Drosophila Heat Shock System as a General Model to Investigate Transcriptional Regulation

M.J. Guertin, S.J. Petesch, K.L. Zobeck, I.M. Min, and J.T. Lis
Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853
Correspondence: johnlis@cornell.edu

Whereas the regulation of a gene is uniquely tailored to respond to specific biological needs, general transcriptional mechanisms are used by diversely regulated genes within and across species. The primary mode of regulation is achieved by modulating specific steps in the transcription cycle of RNA polymerase II (Pol II). Pol II "pausing" has recently been identified as a prevalent rate-limiting and regulated step in the transcription cycle. Many sequence-specific transcription factors (TFs) modulate the duration of the pause by directly or indirectly recruiting positive transcription elongation factor b (P-TEFb) kinase, which promotes escape of Pol II from the pause into productive elongation. These specialized TFs find their target-binding sites by discriminating between DNA sequence elements based on the chromatin context in which these elements reside and can result in productive changes in gene expression or nonfunctional "promiscuous" binding. The binding of a TF can precipitate drastic changes in chromatin architecture that can be both dependent and independent of active Pol II transcription. Here, we highlight heat-shock-mediated gene transcription as a model system in which to study common mechanistic features of gene regulation.

The nucleus is a massive and dynamic cellular organelle that packages, replicates, and regulates the expression of genes. In a typical metazoan, the nucleus is home to ~20,000 genes. Remarkably, each gene is subject to distinct regulation of its level, location, and timing of expression in response to a variety of developmental, environmental, and nutritional cues. This requires an extraordinary amount of information to be packaged into DNA.

The genome is a polymer that is unlike many repeating structures in the cell. Genes arranged along the polymer must each take on a distinctive character and execute regulatory programs that are tailored to the cellular or organismal need for expression of its protein or RNA product. Organisms have evolved a large spectrum of regulatory proteins that respond to a correspondingly complex spectrum of cellular signals. The ability of regulatory machinery to bind and execute specific tasks to control gene expression is dependent on their access to corresponding DNA regulatory elements, which is strongly influenced by chromatin structure and composition. The structure of the chromatin depends not only on the level of occupancy position of nucleosomes along the DNA fiber but also the extent to which nucleosomes are locally compacted into chromatin. The composition of chromatin includes not only the histones (both canonical and variants thereof) that bind to the DNA, but also nonhistone chromosomal proteins and the myriad of posttranslational modifications to the histones and the nonhistone proteins. Just as chromatin structure can influence the relative composition, so too can the composition of chromatin affect its structure. Therefore, it is not surprising that changes in chromatin structure and composition need to occur to effectively regulate gene transcription.

Although individual gene expression requires specific protein–nucleic acid interactions, many of the complex molecular machines that execute transcription are common and shared in various combinations among genes that are each distinctly regulated. At the core of expression is the machine that transcribes all precursor messenger RNA (pre-mRNA): RNA Pol II. Pol II is not sufficient to execute this task and requires a collection of common and some less frequently used general TFs to discern where on the genome to begin transcription (Orphanides et al. 1996; Roeder 1996). Still more complex sets of proteins act commonly to execute a variety of other steps in the nuclear expression of genes including (1) necessary remodeling and modification of chromatin for TF accessibility and Pol II recruitment and initiation, (2) promoter-proximal pausing, (3) transcription elongation, and (4) termination and recycling of the machinery (Fuda et al. 2009).

Here, we summarize focused studies characterizing the architecture and transcription regulation of inducible heat shock (HS) genes in *Drosophila melanogaster*, particularly *Hsp70*. The first section defines characteristics of promoter-proximal paused Pol II and its general role in transcription regulation. The second section describes how activators are targeted to DNA and mediate initiation, release of paused Pol II, and efficient reinitiation at a subset of local genes. The third section describes how gene activation triggers dramatic changes in chromatin architecture. In each, we evaluate how the principles of this defined HS regulatory system are shared by a large fraction of genes across genomes and species.

DEFINING A RATE-LIMITING AND REGULATORY STEP OF TRANSCRIPTION: PROMOTER-PROXIMAL PAUSING

Characterization of Paused RNA Pol II

Recruitment of Pol II to a gene's promoter can be a rate-limiting and regulated step in transcription, and early studies suggested that this is the predominant mode of transcription

regulation (Tjian and Maniatis 1994; Ptashne and Gann 1997). Nonetheless, there existed counterexamples, whereby Pol II occupied the promoters of uninduced genes. For instance, nuclear–run-on (NRO) assays demonstrated that transcriptionally engaged Pol II was enriched at the 5′ end of the inactive β-globin gene of chicken erythrocytes, indicating that the elongation of Pol II could be a controlled regulatory step in eukaryotic transcription (Gariglio et al. 1981). More in-depth studies of the *Drosophila Hsp70* gene were critical in defining key features of promoter-proximal-enriched Pol II. In vivo cross-linking and chromatin immunoprecipitation (ChIP) studies of Pol II at *Hsp70* demonstrated that in unstressed conditions, Pol II is enriched at the 5′ end of *Drosophila Hsp70* (Gilmour and Lis 1985). Subsequently, NRO experiments determined that the promoter-proximal Pol II was transcriptionally engaged but "paused" because the Pol II could not transcribe efficiently in NRO assays unless negative factors were extracted by sarkosyl or high salt (Rougvie and Lis 1988). The precise position of the pause was revealed at near-nucleotide resolution on HS genes by both mapping the single-strand transcription bubble associated with the paused Pol II using permanganate footprinting in living cells (Giardina et al. 1992) and purifying and sizing chain-terminated NRO RNAs (Rasmussen and Lis 1993). Moreover, the latter study also demonstrated that 5′ 7-methylguanosine pre-mRNA capping occurs in vivo on HS genes as Pol II passes through the pause region (Rasmussen and Lis 1993). Sarkosyl-dependent NRO signals of several genes indicated that pausing may be a general regulatory mechanism in *Drosophila* (Rougvie and Lis 1990). These early studies defining the properties of Pol II on *Drosophila* genes were complemented by additional studies of key mammalian genes, notably c-*myc* and c-*fos* (Krumm et al. 1992; Strobl and Eick 1992; Plet et al. 1995).

Pausing Is a General Regulatory Step in Transcription

The first unbiased attempt to gauge the generality of pausing within *Drosophila* used ChIP to purify and clone DNAs associated with Pol II, and these were then probed with NRO RNA (Law et al. 1998). Differential hybridization with standard run-on and sarkosyl-treated run-ons showed that a large fraction of these Pol II–associated fragments contained sarkysol-dependent RNA transcripts. Even with the limited sensitivity of this method to detect pausing, it was estimated that 21% of promoter-proximal–detected Pol II was stimulated in NRO experiments by sarkosyl (Law et al. 1998). The evolution of powerful genome-wide methods made a particularly compelling case that Pol II is enriched at the 5′ end of 10%–30% of genes (Kim et al. 2005; Guenther et al. 2007; Muse et al. 2007; Zeitlinger et al. 2007; Core et al. 2008). Although promoter-proximal pausing was first characterized in eukaryotic cells at signal-regulated genes such as HS- and cell-cycle-regulated genes, pausing is the hallmark of many classes of genes, including those involved in embryonic development, housekeeping processes, the unfolded protein response, DNA damage pathways, and response to stimuli (Zeitlinger et al. 2007; Core et al. 2008). Paused Pol II was proposed to allow rapid and synchronous responses to stimuli, and studies in the developing *Drosophila* embryo have shown that synchronously expressed genes are enriched for hallmarks of paused Pol II, compared with stochastically regulated genes (Boettiger and Levine 2009). Genome-wide data continue to demonstrate the generality of Pol II pausing as a rate-limiting step and likely target of regulation in metazoans.

Release from the Pause

DRB (Dichloro-1-β-D-ribofuranosylbenzimidazole) sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) complexes stabilize pausing of Pol II; conversely, P-TEFb is the critical factor that drives the escape of Pol II from the pause and into productive elongation (Peterlin and Price 2006). In the case of *Hsp70* gene expression, HS induces the binding of heat shock factor (HSF) to upstream target DNA elements, resulting in P-TEFb recruitment to the paused Pol II (Boehm et al. 2003). P-TEFb is a serine/theonine kinase that phosphorylates serine 2 of tandem repeats in the carboxy-terminal domain of the largest subunit of Pol II as well as DSIF and NELF. The phosphorylation of P-TEFb target proteins allows Pol II to be released from the pause and productively transcribe through the body of the gene (Peterlin and Price 2006). The recruitment of P-TEFb to the *Hsp70* promoter is critical for stimulating transcription because artificially tethering P-TEFb to the promoter resulted in a high level of transcription, even in the absence of HS (Lis et al. 2000). Moreover, the P-TEFb–specific kinase inhibitor flavopiridol was used to demonstrate that the enzymatic activity is necessary for the transition of Pol II from the paused state into productive elongation in vivo (Ni et al. 2008). Specific targeting of P-TEFb to the *Hsp70* gene requires HSF; however, HSF does not appear to directly recruit P-TEFb because molecular, cytological, and biochemical experiments have shown that P-TEFb does not localize to upstream heat-shock-factor sequence elements (HSEs), and P-TEFb does not directly bind to HSF (Lis et al. 2000; Boehm et al. 2003). Currently, the mechanism of P-TEFb recruitment to the *Hsp70* gene remains elusive.

Recently, it has been shown that c-*myc* interacts with P-TEFb and is responsible for regulating Pol II release from the pause for 30% of the paused genes in embryonic stem cells (Rahl et al. 2010). This study indicates that paused Pol II is a general feature for controlling transcription in mammalian cells and that upstream TFs have a role in recruiting P-TEFb for the release of paused Pol II.

MASTER HS REGULATORY PROTEIN: HSF

HSF Targets HS DNA Elements Based on Chromatin Structure

HSF is the master regulator of stress-responsive genes among organisms as distantly related as yeast and humans. HSF binds *cis*-regulatory elements to modify gene expression. The DNA-binding domain of HSF and the HSF-targeted DNA consensus sequence element are highly conserved between distantly related organisms (Fig. 1).

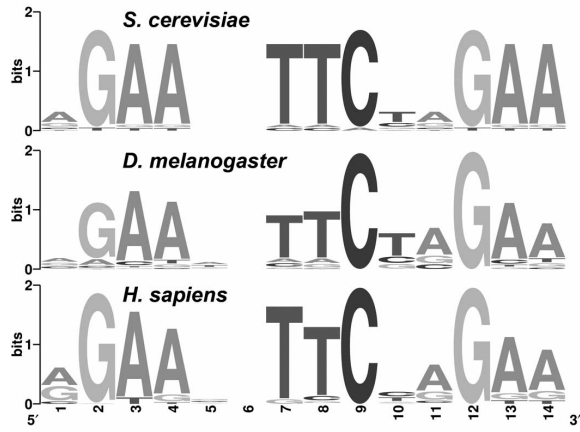

Figure 1. Conservation of HSF-targeted DNA sequence element. The consensus HSE is remarkably conserved among *Saccharomyces cerevisiae*, *D. melanogaster*, and *Homo sapiens* (Fernandes et al. 1994; Trinklein et al. 2004; Guertin and Lis 2010).

The consensus HSE is a tandem array of three 5-mer sites (AGAAN) arranged head to tail, consistent with HSF binding to HSEs as a trimer (Perisic et al. 1989). Monomeric HSF is localized mainly in the nucleoplasm during unstressed conditions (Orosz et al. 1996), and following stress, HSF trimerizes and binds to cognate DNA elements. Trimerization of *Drosophila* HSF is mediated by a trimerization domain, characterized by three blocks of leucine-zipper-like heptad repeats (Rabindran et al. 1993; Wu 1995). Following binding to DNA, HSF affects chromatin structure and composition by directly or indirectly recruiting coactivators (Park et al. 2001; Kim et al. 2004), elongation factors (Lis et al. 2000; Ni et al. 2008; Ardehali et al. 2009), histone-modifying enzymes (Hong et al. 2004; Smith et al. 2004), and chaperones (Saunders et al. 2003). The inducible nature of HSF binding provides an opportunity to evaluate the roles chromatin environment and locally bound factors have in determining how HSF discriminates among HSEs and which local genes will be activated by HSF.

The main predictor of whether a site will be bound by HSF is the presence of an HSE; however, HSF preferentially binds to HSEs that reside within accessible regions of the genome, which are marked by acetylated histones (Fig. 2). The degenerate HSE consensus sequence predicts that HSF is capable of binding to thousands of sites within the *Drosophila* genome (Fig. 1). Nonetheless, in vivo, chromatin structure restricts HSF binding to only a subset of HSEs (Westwood et al. 1991; Guertin and Lis 2010; Gonsalves et al. 2011). This fits a pattern that is well characterized in eukaryotes; some of the best homeoprotein, p53, and estrogen receptor–binding sequences are not occupied by their cognate factor in vivo (Carr and Biggin 2000; Wei et al. 2006; Welboren et al. 2009).

There are two general models in which TFs are thought to discriminate between equivalent binding sites at different places in the genome. TF binding is either facilitated by cooperative interactions with chromatin-bound cofactors or occluded by chromatin structure (Biggin and McGinnis 1997). Trimerized HSF clearly cooperates with other HSF trimers to enhance binding at cognate elements (Xiao et al. 1991); however, HSF primarily binds to HSE motifs based

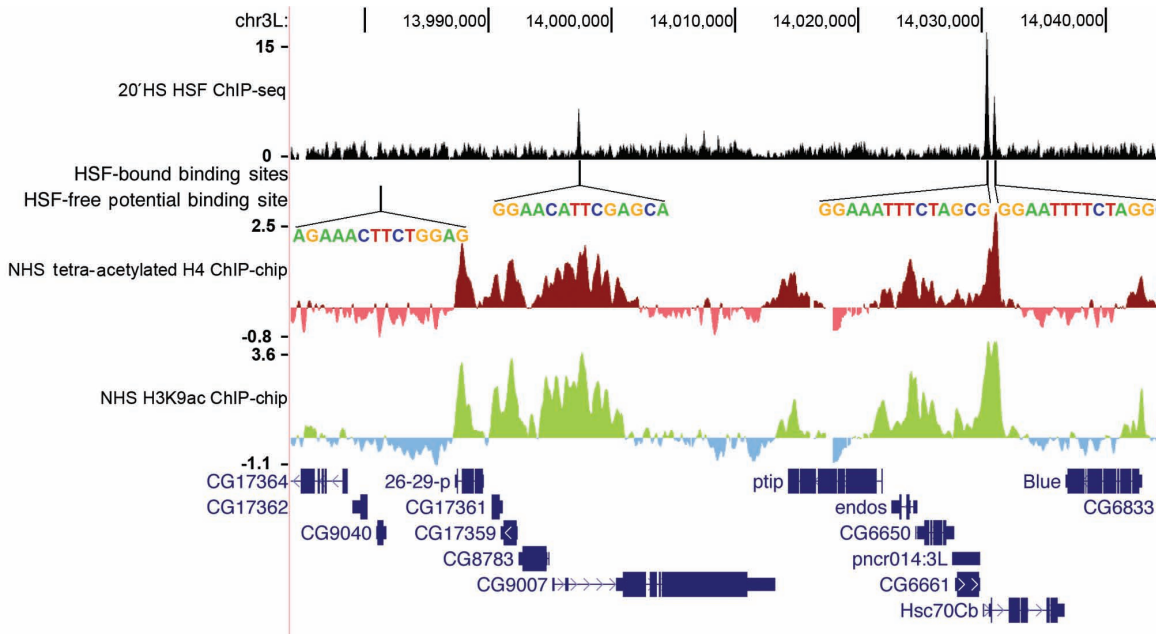

Figure 2. Chromatin landscape dictates HSF binding to HSEs in vivo. This region of chromosome 3L contains four potential HSF-binding sites. HSE sequences conform to the HSE matrix with the following p values, from left to right: 3.9×10^{-6}, 1.8×10^{-5}, 1.4×10^{-5}, and 4.3×10^{-5}. Although the HSF-free motif conforms to the consensus with the lowest p value, chromatin structure restricts HSF occupancy. HSEs that are enriched for H4 and H3 acetylation during non-HS (Kharchenko et al. 2010a) are preferentially bound by HSF in vivo (Guertin and Lis 2010).

on accessibility of the DNA within the context of chromatin (Wu 1980; Guertin and Lis 2010). Although it remains unclear how to formally define accessible chromatin, DNase I hypersensitivity and canonical marks of "active" chromatin (e.g., H4 tetra-acetylation and H3K9 acetylation) can be considered to be an appropriate proxy for chromatin that is accessible for TF binding (Fig. 2) (Guertin and Lis 2010; Li et al. 2010). Moreover, histone modifications, which occur in broad domains and distinct combinations (Heintzman et al. 2007; Wang et al. 2008; Zang et al. 2009; Kharchenko et al. 2010), are a powerful predictor of enhancers and TF-binding sites (TFBSs) (Hon et al. 2009; Guertin and Lis 2010). The generality of this TFBS recognition mechanism is highlighted by the observation that the in vivo occupancy of 16 developmental factors correlates strongly with DNase I hypersensitivity (Li et al. 2010). Furthermore, chromatin accessibility changes that occur throughout development are responsible for differential TF occupancy observed at different stages in development (Li et al. 2010).

Although binding of a TF to DNA is generally a critical step in the activation of target genes, not all TF-binding events are productive—many binding events in higher eukaryotes have no functional consequence on gene expression (Li et al. 2008; MacArthur et al. 2009; Guertin and Lis 2010). It is still unclear how activators discriminate between nearby genes to mediate up-regulation of only a distinct subset of genes.

HSF Functions as a Classic Acidic Activator

Drosophila HSF harbors a conserved transactivation domain at its carboxyl terminus (Wisniewski et al. 1996) that contains hydrophobic and acidic residues characteristic of an acidic activator. Like other acidic activators, HSF has the potential to interact with coactivators such as the Mediator complex and CBP (cAMP response element-binding protein) (Hong et al. 2004; Kim et al. 2004; Smith et al. 2004), and HSF binding to DNA is sufficient to increase local acetylation (Guertin and Lis 2010). This suggests that the activation potential of HSF is at least partly manifested through HSF's ability to recruit chromatin-modifying enzymes that alter chromatin structure. Although HSF's activation domain is sufficient to activate transgenic reporter genes (Wisniewski et al. 1996), HSF binding to a promoter is insufficient to activate local genes in vivo at endogenous binding sites (Guertin and Lis 2010). It is unclear how HSF discriminates among proximal genes to activate; however, directed studies with known HS-regulated genes suggest that GAGA-associated factor (GAF) and paused Pol II may facilitate activation of some genes (H Lee et al. 1992; Tsukiyama et al. 1994; Wilkins and Lis 1997; Leibovitch et al. 2002; C Lee et al. 2008). In contrast, the insulator BEAF (boundary element-associated factor) may be able to insulate local genes from HSF's activation potential (Guertin and Lis 2010; Gonsalves et al. 2011). Although the DNA elements, promoter features, and TFs that are critical for HSF-induced expression of *Hsp70* are well characterized, it remains unclear to what extent these chromatin features dictate the activation potential of HS genes genome wide.

HSF is thought to have a predominant role in releasing paused Pol II into productive elongation (Lis et al. 2000; Ni et al. 2008), but it remains unclear whether HSF is necessary for Pol II initiation or efficient reinitiation following the escape of a paused Pol II into productive elongation, as is seen with other acidic activators (Yudkovsky et al. 2000). In vitro experiments on chromatin templates indicate that HSF has a role in setting up a platform that allows faster initiation (Sandaltzopoulos and Becker 1998), and in vivo studies demonstrate that HSF remains stably bound to promoters during HS over a time frame that allows many dozens of cycles of transcription initiation (Yao et al. 2006). Although the mechanistic role of HSF in transcription after DNA binding is still under investigation, it is clear that HSF is necessary to orchestrate the changes in transcription, chromatin structure, and architecture that are observed following HS (Jedlicka et al. 1997; Petesch and Lis 2008).

HSF PRECIPITATES NUCLEOSOME LOSS AND FORMATION OF A TRANSCRIPTION PUFF WITH COMPARTMENT-LIKE PROPERTIES

Chromatin Dynamics Visualized on Polytene Chromosomes

Drosophila polytene chromosomes have nine inducible HS puffs that form reproducibly and synchronously and are maximal in size by 20 min (Fig. 3) (Ashburner 1970; Lewis et al. 1975). The size of the HS-inducible puffs is roughly correlated with both gene length and strength of the promoter (Simon et al. 1985). HSF is necessary for both gene activation and chromatin decondensation at these nine sites (Jedlicka et al. 1997). In contrast to the evidence that active transcription of these HS genes is responsible for the formation of the HS puffs, other studies show that the two can be uncoupled from each other. Chemical treatment with sodium salicylate induces HSF binding to the same nine target loci and induces puff formation, but active transcription of these genes is blocked (Winegarden et al. 1996). HS genes can thus be used to study how chromatin structure changes during transcriptional activation and how different factors can influence

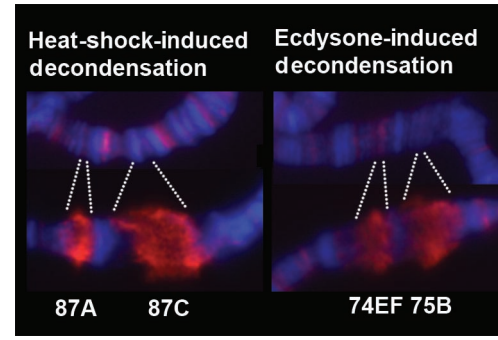

Figure 3. Decondensation of heat-shock- and ecdysone-induced genes. HS genes at cytological loci 87A and 87C and developmental genes at 74EF and 75B decondense dramatically following activation. Hoeschst-stained DNA (blue), serine-5 phosphorylated Pol II (H14 antibody) (red).

the changes in chromatin, either through a transcriptionally dependent or independent pathway. Notably, transcriptional puffs are dynamic structures that are not unique to HS loci because developmental cues such as hormone exposure also induce puff formation (Fig. 3) (Ritossa 1962). As such, polytene chromosomes continue to serve as a visual model for how transcription and TFs can influence chromatin structure and vice versa.

In addition to the many cytological studies performed with the imaging of polytene chromosomes following HS, many in vitro and in vivo studies with *Hsp70* genes have provided evidence for the changes that accompany individual nucleosomes following HS. Just as staining of polytene chromosomes has shown that regions of high levels of chromatin condensation preferentially exclude Pol II, in vitro studies have shown that the presence of nucleosomes impedes the transcription of Pol II (Knezetic and Luse 1986; Lorch et al. 1987). However, the presence of nucleosomes on the gene body of *Hsp70* before HS does not impede the immediate response and efficient transcriptional activation of paused Pol II at *Hsp70* following an instantaneous HS (Boehm et al. 2003). Intriguingly, the first studies mapping nucleosomes at *Hsp70*, both before and after HS, demonstrated that nucleosome structure was severely compromised on the body of the gene following a 20-min HS (Wu et al. 1979). These results corroborate those seen at a lower resolution with the formation of chromosomal puffs at the *Hsp70* loci following HS from polytene chromosomes (Fig. 3), and they also go a step further by demonstrating that individual nucleosomes, and not just the chromatin fiber, are changing during transcriptional activation of the gene.

High Spatial and Temporal Resolution of Changes in Chromatin Dynamics

More recently, the chromatin structure of *Hsp70* has been probed at higher spatial and temporal resolution to understand more precisely how and when nucleosome structures are changing. These results demonstrate that there is a rapid change in nucleosome structure by 30–60 sec of HS that continues throughout the gene and is independent of active transcription (Petesch and Lis 2008). These changes are marked by an increase in the accessibility of micrococcal nuclease to nucleosomal DNA and also a reduction in histone density across the gene. Additionally, between 60 and 120 sec of HS, there is a further transcription-dependent disruption of nucleosomes, seen by another increase in micrococcal nuclease accessibility. These results demonstrate not only the rapid nature of the changes in nucleosomes that occur following HS but also separate transcription-dependent and -independent changes to nucleosome structure that were also documented from those polytene chromosome analyses performed using sodium salicylate (Winegarden et al. 1996).

HSF and PARP Are Critical for Changes in Chromatin Structure

The immediate and widespread changes in chromatin structure at *Hsp70* following HS are critically dependent on HSF and poly(ADP-)ribose polymerase (PARP). PARP is an enzyme that catalyzes the formation of poly(ADP-)ribose (PAR) from NAD^+ molecules onto donor proteins as a posttranslational modification (Krishnakumar and Kraus 2010). As with HSF, PARP has been previously identified to be a critical factor for puffing in response to HS but is also important in puff formation associated with developmental genes that respond to hormones and innate immunity genes that respond to infection (Tulin and Spradling 2003). Both RNA interference (RNAi) knockdown of HSF and PARP as well as catalytic inhibition of PARP result in the failure of *Hsp70* to undergo rapid transcription-independent loss of the chromatin structure following HS (Petesch and Lis 2008). Taken together, these data suggest that PARP activation (presumably directly or indirectly by HSF) and production of long-nucleic-acid–like PAR molecules can directly result in a decondensation event that facilitates the movement of Pol II through nucleosomes.

It is also becoming clear that the formation of PAR at these sites of transcriptional activation has an additional function beyond rapidly decreasing the density of histones and disrupting nucleosome structure along genes. Recent studies using live cell imaging of polytene chromosomes expressing GFP-fusion proteins to PARP and histones have been used to track the dynamic nature of these proteins before and after HS at *Hsp70* loci (Zobeck et al. 2010). Before HS, both PARP and histones are present at *Hsp70*. Following HS, the total levels of both PARP and histones remain unchanged from those levels before HS, although histone density on the gene itself decreases. This indicates that although histones are being evicted from the gene following HS, they are retained locally at *Hsp70* loci. Whether the histones present before HS at the loci are the same as those at the loci after HS has yet to be directly measured. However, these results indicate that in addition to contributing to the decondensation at *Hsp70* loci, the activation of PARP and production of PAR could also aid in the local retention of histones that are evicted following HS.

Compartmentalization of Transcription

Focused studies of *Hsp70* loci of polytene chromosomes indicate that HS activation precipitates de novo assembly, accumulation, and retention of TFs at activated regions. The term “compartment” has been coined to describe the changes in chromatin architecture that facilitate local retention of TFs that were recruited to the loci. Evidence for a transcription compartment arose from multiphoton laser-scanning confocal and spinning-disk microscopy images of native cytological loci in living cells (Yao et al. 2007; Zobeck et al. 2010). Using fluorescence recovery after photobleaching (FRAP) at *Hsp70* loci in polytene nuclei, it was shown that photobleached Pol II was replaced rapidly by unbleached Pol II 10–20 min after HS; however, less recovery was observed following a 40–60-min HS (Fig. 4) (Yao et al. 2007). This change in Pol II exchange occurs in a progressive manner (Yao et al. 2007). Additional data indicate that transcription is still occurring after 40 min at these sites, suggesting that the bleached Pol II molecules are locally retained in a transcription compart-

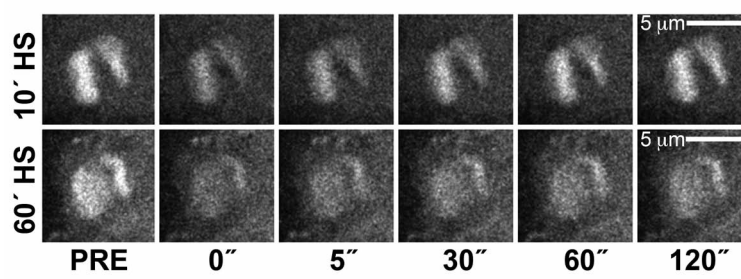

Figure 4. Dynamics of Pol II at *Hsp70* loci in vivo. FRAP (*Fluorescence recovery after photobleaching*) was used to measure dynamics of Pol II (using an Rpb3-GFP [green fluorescent protein] fusion protein) after 10- or 60-min following an instantaneous HS. Nonphotobleached Pol II is rapidly recruited to loci after 10 min but more modestly recruited after 60 min. Time (in seconds) along the *x* axis represents time elapsed after photobleaching.

ment and recycled for transcription observed at late time points (Yao et al. 2007).

Although the compartment was initially characterized for Pol II, this model would predict that other TFs would be retained in the compartment structure. Indeed, the exchange of P-TEFb, Spt6, and Topo I at *Hsp70* loci, as measured by FRAP, decreases at later HS time points. Interestingly, the transcription compartment is reliant on PARP catalytic activity, again implicating PARP as an important player in the chromatin changes that occur following HS (Zobeck et al. 2010). Although compartments were studied in depth at *Hsp70* loci, compartment-like features are found at numerous developmental loci as well, illustrating the generality of this structure (Yao et al. 2007).

Another defining characteristic of compartments is the de novo assembly and accumulation of TFs at activated regions (Fig. 5) (Yao et al. 2007; Zobeck et al. 2010). In contrast to this model, the transcription "factory" model proposes that coregulated genes relocate following activation to preformed foci containing Pol II and the transcription machinery (Iborra et al. 1996). Support for the transcription factory model stems from fluorescence in situ hybridization (FISH) data indicating that coexpressed genes colocalize with Pol II foci more frequently when active compared with when they are inactive (Osborne et al. 2004, 2007). However, the factory model does not hold true for HS-regulated genes because FISH of *Hsp70* and *Hsp83* revealed that the frequency of interactions of these two loci and distribution in the nucleus do not change following gene activation in diploid cells (Yao et al. 2007). Furthermore, live cell imaging in polytene nuclei indicates that *Hsp70* loci are not repositioned on gene activation. Similarly, in mouse 3134 cells, it was shown that a tandem array of mouse mammary tumor virus promoters does not relocate within the nucleus after activation (Becker et al. 2002). Additionally, live cell imaging of MS2-GFP binding to the nascent RNA of a β-globin-gene construct harboring MS2-binding sites revealed that transcriptional activation does not cause movement from the nuclear membrane to the interior of the nucleoplasm (Kumaran and Spector 2008). These data, taken together, suggest that a de novo recruitment mechanism of necessary TFs for building a transcriptionally active gene is conserved beyond *Hsp70* loci of *Drosophila* and is likely a general model of gene regulation.

PERSPECTIVES

We describe a model system that has proven to be extremely useful in elucidating the general mechanisms by which paused Pol II and acidic activators function to direct changes in transcriptional regulation and chromatin structure. However, important unanswered questions remain.

Accumulating evidence shows that Pol II pausing at the promoter-proximal region is a general mechanism to regulate transcription in diverse genes and in organisms that extend beyond the highly studied paused *Drosophila Hsp70* gene; however, the mechanism by which paused Pol II is established is not fully understood. Specific compositions of nucleotides at the promoter are associated with increased levels of pausing (Hendrix et al. 2008; Nechaev et al. 2010), but the nucleotide composition alone is not likely to be sufficient for directing the establishment of the pause. Several factors including DSIF, NELF, and GAF have been shown to be important in establishing or stabilizing the pause, but these are insufficient to dictate all pausing, suggesting that the full array of the sets of factors is not yet identified. Moreover, the factors that are known to affect pausing are not understood completely.

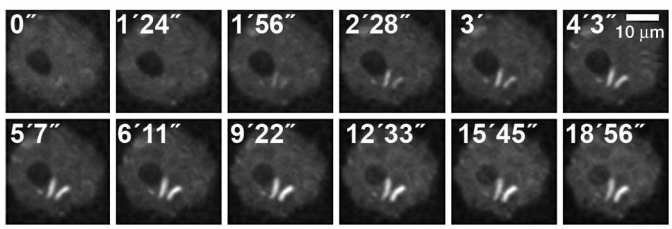

Figure 5. Gradual recruitment of Pol II to the *Hsp70* loci in vivo. Additional Pol II is recruited to *Hsp70* loci (white doublet) at ~2 min post-HS (third panel) and accumulates until ~5-min post-HS (seventh panel). After 5 min of HS, Pol II reaches maximal levels and remains high for the duration of the experiment.

The chromatin environment and chromatin-bound factors may also influence how pausing of Pol II is established at the 5′ end of genes (Gilchrist et al. 2008). Additionally, it remains unclear how a diverse set of upstream activators interfaces with P-TEFb to release the paused Pol II into productive elongation. Furthermore, the full spectrum of P-TEFb phosphorylation targets needs to be identified, as does the mechanistic contribution of their phosphorylation to steps within the transcriptional cycle.

Although HSF is indispensable for the recruitment of P-TEFb and the release of paused Pol II into productive elongation at *Hsp70*, it remains unclear how HSF discriminates among local genes, mediating the activation of only a subset. GAF is an important factor in mediating activation of some, but not all, HS genes (Lee et al. 2008; Ardehali et al. 2009; Guertin and Lis 2010). Additionally, at *Hsp70*, the presence of paused Pol II is a prerequisite for activation by HSF (Lee et al. 1992). In contrast, insulators and other repressive factors may inhibit the communication of HSF and promoters (Guertin and Lis 2010; Gonsalves et al. 2011). To date, the properties of HSF-activated genes have been characterized on a gene-by-gene basis, but genome-wide methodologies allow for an unbiased analysis contrasting the characteristics of genes that are differentially affected by HSF-proximal binding. We anticipate that many features including DNA promoter elements, activator-binding kinetics and proximity to the transcriptional start sites (TSSs), prebound TFs, local histone modifications, and paused Pol II status can act together to influence gene activation and affect chromatin structure.

The dynamic interplay between how chromatin and transcription influence each other needs to be understood quantitatively and in mechanistic detail. It is still unclear as to what extent chromatin structure affects the ability of the transcriptional machinery to bind DNA and the extent to which chromatin structure affects the establishment and duration time of paused Pol II. During the process of transcriptional elongation, the fate of individual histones during the passage of Pol II through the nucleosome remains unresolved. We need to define the intermediate in vivo nucleosome structures generated as Pol II transcribes through a gene and discern the fate of those histones that are evicted from the DNA. Additionally, we need to understand the changes in chromatin structure that occur both dependently and independently of actively transcribing Pol II, in particular, the general role that activators and PARP have in mediating transcription-independent changes in chromatin. Elucidating these mechanistic details will provide critical insights into the ways that chromatin structure influences transcription and vice versa.

Live cell imaging of *Drosophila* polytene chromosomes has provided novel insights into the dynamics of TFs at activated genes. These general features may be conserved in diploid cells, which exhibit some of the characteristics of a transcription compartment. However, the retention of TFs in a transcription-activated compartment has only been observed in *Drosophila* polytene nuclei. To address the generality of compartments, new imaging methods that improve resolution and sensitivity need to be developed.

The HS genes of *Drosophila* have been integral in defining and characterizing general mechanisms by which signal-regulated transcription occurs. The recent emergence of genomic expression and TF-binding data, live cell imaging techniques, and molecular tools promises that the HS system will continue to provide novel insights into transcriptional regulation and chromatin structure.

ACKNOWLEDGMENTS

We thank J. Werner for providing the immunofluorescence images used in Figure 3.

REFERENCES

Ardehali MB, Yao J, Adelman K, Fuda NJ, Petesch SJ, Webb WW, Lis JT. 2009. Spt6 enhances the elongation rate of RNA polymerase II in vivo. *EMBO J* **28:** 1067–1077.

Ashburner M. 1970. Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*. V. Responses to environmental treatments. *Chromosoma* **31:** 356–376.

Becker M, Baumann C, John S, Walker DA, Vigneron M, McNally JG, Hager GL. 2002. Dynamic behavior of transcription factors on a natural promoter in living cells. *EMBO Rep* **3:** 1188–1194.

Biggin MD, McGinnis W. 1997. Regulation of segmentation and segmental identity by *Drosophila* homeoproteins: The role of DNA binding in functional activity and specificity. *Development* **124:** 4425–4433.

Boehm AK, Saunders A, Werner J, Lis JT. 2003. Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock. *Mol Cell Biol* **23:** 7628–7637.

Boettiger AN, Levine M. 2009. Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science* **325:** 471–473.

Carr A, Biggin MD. 2000. Accessibility of transcriptionally inactive genes is specifically reduced at homeoprotein-DNA binding sites in *Drosophila*. *Nucleic Acids Res* **28:** 2839–2846.

Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322:** 1845–1848.

Fernandes M, Xiao H, Lis JT. 1994. Fine structure analyses of the *Drosophila* and *Saccharomyces* heat shock factor: Heat shock element interactions. *Nucleic Acids Res.* **22:** 167–173.

Fuda NJ, Ardehali MB, Lis JT. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461:** 186–192.

Gariglio P, Bellard M, Chambon P. 1981. Clustering of RNA polymerase B molecules in the 5′ moiety of the adult β-globin gene of hen erythrocytes. *Nucleic Acids Res* **9:** 2589–2598.

Giardina C, Perez-Riba M, Lis JT. 1992. Promoter melting and TFIID complexes on *Drosophila* genes *in vivo*. *Genes Dev* **6:** 2190–2200.

Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, 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 DS, Lis JT. 1985. In vivo interactions of RNA polymerase II with genes of *Drosophila melanogaster*. *Mol Cell Biol* **5:** 2009-2018.

Gonsalves SE, Moses AM, Razak Z, Robert F, Westwood JT. 2011. Whole-genome analysis reveals that active heat shock factor binding sites are mostly associated with non-heat shock genes in *Drosophila melanogaster*. *PLoS ONE* **6:** e15934.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130:** 77–88.

Guertin MJ, Lis JT. 2010. Chromatin landscape dictates HSF binding to target DNA elements. *PLoS Genet.* **6:** e1001114.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD,

Barrera LO, Van Calcar S, Qu C, Ching KA, et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39:** 311–318.

Hendrix DA, Hong JW, Zeitlinger J, Rokhsar DS, Levine MS. 2008. Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proc Natl Acad Sci* **105:** 7762–7767.

Hon G, Wang W, Ren B. 2009. Discovery and annotation of functional chromatin signatures in the human genome. *PLoS Comput Biol* **5:** e1000566.

Hong S, Kim SH, Heo MA, Choi YH, Park MJ, Yoo MA, Kim HD, Kang HS, Cheong J. 2004. Coactivator ASC-2 mediates heat shock factor 1-mediated transactivation dependent on heat shock. *FEBS Lett* **559:** 165–170.

Iborra FJ, Pombo A, McManus J, Jackson DA, Cook PR. 1996. The topology of transcription by immobilized polymerases. *Exp Cell Res* **229:** 167–173.

Jedlicka P, Mortin MA, Wu C. 1997. Multiple functions of *Drosophila* heat shock transcription factor *in vivo*. *EMBO J* **16:** 2452–2462.

Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Larschan E, Gorchakov AA, Gu T, Linder-Basso D, et al. 2010. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature* doi:10.1038/nature09725.

Kim TW, Kwon YJ, Kim JM, Song YH, Kim SN, Kim YJ. 2004. MED16 and MED23 of Mediator are coactivators of lipopolysaccharide- and heat-shock-induced transcriptional activators. *Proc Natl Acad Sci* **101:** 12153–12158.

Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B. 2005. A high-resolution map of active promoters in the human genome. *Nature* **436:** 876–880.

Knezetic JA, Luse DS. 1986. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* **45:** 95–104.

Krishnakumar R, Kraus WL. 2010. PARP-1 regulates chromatin structure and transcription through a KDM5B-dependent pathway. *Mol Cell* **39:** 736–749.

Krumm A, Meulia T, Brunvand M, Groudine M. 1992. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. *Genes Dev* **6:** 2201–2213.

Kumaran RI, Spector DL. 2008. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol* **180:** 51–65.

Law A, Hirayoshi K, O'Brien T, Lis JT. 1998. Direct cloning of DNA that interacts in vivo with a specific protein: Application to RNA polymerase II and sites of pausing in *Drosophila. Nucleic Acids Res* **26:** 919–924.

Lee H, Kraus KW, Wolfner MF, Lis JT. 1992. DNA sequence requirements for generating paused polymerase at the start of hsp70. *Genes Dev* **6:** 284–295.

Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS. 2008. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila. Mol Cell Biol* **28:** 3290–3300.

Leibovitch BA, Lu Q, Benjamin LR, Liu Y, Gilmour DS, Elgin SC. 2002. GAGA factor and the TFIID complex collaborate in generating an open chromatin structure at the *Drosophila melanogaster hsp26* promoter. *Mol Cell Biol* **22:** 6148–6157.

Lewis M, Helmsing PJ, Ashburner M. 1975. Parallel changes in puffing activity and patterns of protein synthesis in salivary glands of *Drosophila*. *Proc Natl Acad Sci* **72:** 3604–3608.

Li XY, MacArthur S, Bourgon R, Nix D, Pollard DA, Iyer VN, Hechmer A, Simirenko L, Stapleton M, Luengo Hendriks CL, et al. 2008. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol* **6:** e27.

Li XY, Thomas S, Sabo PJ, Eisen MB, Stamatoyannopoulos JA, Biggin MD. 2011. The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding. *Genome Biol* (in press).

Lis JT, Mason P, Peng J, Price DH, Werner J. 2000. P-TEFb kinase recruitment and function at heat shock loci. *Genes Dev* **14:** 792–803.

Lorch Y, LaPointe JW, Kornberg RD. 1987. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* **49:** 203–210.

MacArthur S, Li XY, Li J, Brown JB, Chu HC, Zeng L, Grondona BP, Hechmer A, Simirenko L, Keranen SV, et al. 2009. Developmental roles of 21 *Drosophila* transcription factors are determined by quantitative differences in binding to an overlapping set of thousands of genomic regions. *Genome Biol* **10:** R80.

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. 2007. RNA polymerase is poised for activation across the genome. *Nat Genet* **39:** 1507–1511.

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

Ni Z, Saunders A, Fuda NJ, Yao J, Suarez JR, Webb WW, Lis JT. 2008. P-TEFb is critical for the maturation of RNA polymerase II into productive elongation *in vivo*. *Mol Cell Biol* **28:** 1161–1170.

Orosz A, Wisniewski J, Wu C. 1996. Regulation of *Drosophila* heat shock factor trimerization: Global sequence requirements and independence of nuclear localization. *Mol Cell Biol* **16:** 7018–7030.

Orphanides G, Lagrange T, Reinberg D. 1996. The general transcription factors of RNA polymerase II. *Genes Dev* **10:** 2657–2683.

Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, et al. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* **36:** 1065–1071.

Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, Corcoran AE, Fraser P. 2007. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol* **5:** e192.

Park JM, Werner J, Kim JM, Lis JT, Kim YJ. 2001. Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock. *Mol Cell* **8:** 9–19.

Perisic O, Xiao H, Lis JT. 1989. Stable binding of *Drosophila* heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. *Cell* **59:** 797–806.

Peterlin BM, Price DH. 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* **23:** 297–305.

Petesch SJ, Lis JT. 2008. Rapid, transcription-independent loss of nucleosomes over a large chromatin domain at Hsp70 loci. *Cell* **134:** 74–84.

Plet A, Eick D, Blanchard JM. 1995. Elongation and premature termination of transcripts initiated from c-fos and c-myc promoters show dissimilar patterns. *Oncogene* **10:** 319–328.

Ptashne M, Gann A. 1997. Transcriptional activation by recruitment. *Nature* **386:** 569–577.

Rabindran SK, Haroun RI, Clos J, Wisniewski J, Wu C. 1993. Regulation of heat shock factor trimer formation: Role of a conserved leucine zipper. *Science* **259:** 230–234.

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. 2010. c-Myc regulates transcriptional pause release. *Cell* **141:** 432–445.

Rasmussen EB, Lis JT. 1993. In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes. *Proc Natl Acad Sci* **90:** 7923–7927.

Ritossa FM. 1962. A new puffing pattern induced by temperature shock and DNP in *Drosophila*. *Experientia* **18:** 571–573.

Roeder RG. 1996. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* **21:** 327–335.

Rougvie AE, Lis JT. 1988. The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* **54:** 795–804.

Rougvie AE, Lis JT. 1990. Postinitiation transcriptional control in *Drosophila melanogaster*. Mol Cell Biol **10:** 6041–6045.

Sandaltzopoulos R, Becker PB. 1998. Heat shock factor increases the reinitiation rate from potentiated chromatin templates. *Mol Cell Biol* **18:** 361–367.

Saunders A, Werner J, Andrulis ED, Nakayama T, Hirose S, Reinberg D, Lis JT. 2003. Tracking FACT and the RNA polymerase

II elongation complex through chromatin *in vivo*. *Science* **301:** 1094–1096.

Simon JA, Sutton CA, Lobell RB, Glaser RL, Lis JT. 1985. Determinants of heat shock-induced chromosome puffing. *Cell* **40:** 805–817.

Smith ST, Petruk S, Sedkov Y, Cho E, Tillib S, Canaani E, Mazo A. 2004. Modulation of heat shock gene expression by the TAC1 chromatin-modifying complex. *Nat Cell Biol* **6:** 162–167.

Strobl LJ, Eick D. 1992. Hold back of RNA polymerase II at the transcription start site mediates down-regulation of *c-myc in vivo*. *EMBO J* **11:** 3307–3314.

Tjian R, Maniatis T. 1994. Transcriptional activation: A complex puzzle with few easy pieces. *Cell* **77:** 5–8.

Trinklein ND, Murray JI, Hartman SJ, Botstein D, Myers RM. 2004. The role of heat shock transcription factor 1 in the genome-wide regulation of the mammalian heat shock response. *Mol Biol Cell* **15:** 1254–1261.

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

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

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, Zhao K. 2008. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40:** 897–903.

Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, et al. 2006. A global map of p53 transcription-factor binding sites in the human genome. *Cell* **124:** 207–219.

Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG. 2009. ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J* **28:** 1418–1428.

Westwood JT, Clos J, Wu C. 1991. Stress-induced oligomerization and chromosomal relocalization of heat-shock factor. *Nature* **353:** 822–827.

Wilkins RC, Lis JT. 1997. Dynamics of potentiation and activation: GAGA factor and its role in heat shock gene regulation. *Nucleic Acids Res* **25:** 3963–3968.

Winegarden NA, Wong KS, Sopta M, Westwood JT. 1996. Sodium salicylate decreases intracellular ATP, induces both heat shock factor binding and chromosomal puffing, but does not induce *hsp 70* gene transcription in *Drosophila*. *J Biol Chem* **271:** 26971–26980.

Wisniewski J, Orosz A, Allada R, Wu C. 1996. The C-terminal region of *Drosophila* heat shock factor (HSF) contains a constitutively functional transactivation domain. *Nucleic Acids Res.* **24:** 367–374.

Wu C. 1980. The 5′ ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* **286:** 854–860.

Wu C. 1995. Heat shock transcription factors: Structure and regulation. *Annu Rev Cell Dev Biol* **11:** 441–469.

Wu C, Wong YC, Elgin SC. 1979. The chromatin structure of specific genes. II. Disruption of chromatin structure during gene activity. *Cell* **16:** 807–814.

Xiao H, Perisic O, Lis JT. 1991. Cooperative binding of *Drosophila* heat shock factor to arrays of a conserved 5 bp unit. *Cell* **64:** 585–593.

Yao J, Munson KM, Webb WW, Lis JT. 2006. Dynamics of heat shock factor association with native gene loci in living cells. *Nature* **442:** 1050–1053.

Yao J, Ardehali MB, Fecko CJ, Webb WW, Lis JT. 2007. Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. *Mol Cell* **28:** 978–990.

Yudkovsky N, Ranish JA, Hahn S. 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408:** 225–229.

Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W. 2009. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25:** 1952–1958.

Zeitlinger J, Stark A, Kellis M, Hong JW, Nechaev S, Adelman K, Levine M, Young RA. 2007. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat Genet* **39:** 1512–1516.

Zobeck KL, Buckley MS, Zipfel WR, Lis J. 2010. Recruitment timing and dynamics of transcription factors at the *Hsp70* loci in living cells. *Mol Cell* **40:** 965–975.

Cold Spring Harbor Symposia on Quantitative Biology

Drosophila Heat Shock System as a General Model to Investigate Transcriptional Regulation

M.J. Guertin, S.J. Petesch, K.L. Zobeck, et al.

Cold Spring Harb Symp Quant Biol published online April 5, 2011
Access the most recent version at doi:10.1101/sqb.2010.75.039