IN VIVO ANALYSIS OF THE MOLECULAR MECHANISMS OF LONG- AND SHORT-RANGE TRANSCRIPTIONAL REPRESSION

By

Li Li

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Microbiology and Molecular Genetics

2010

ABSTRACT

IN VIVO ANALYSIS OF THE MOLECULAR MECHANISMS OF LONG- AND SHORT-RANGE TRANSCRIPTIONAL REPRESSION

By

Li Li

Transcription repression is essential for establishing precise patterns of gene expression during development. Studies of Drosophila transcriptional regulation reveal that transcriptional repressors fall into two classes: those acting locally as “short-range” repressors, and those acting dominantly as “long-range” repressors. The types of transcription factors and cofactors involved are highly conserved in all metazoans. One area in which our knowledge has lagged is in understanding how these different classes of repressors function at a mechanistic level. Understanding of repression mechanisms in the developmental circuits will shed light on key biological processes including stem cell reprogramming, cancer, and development.

In this work, I describe studies that reveal molecular mechanisms of long- and short-range repression during early *Drosophila* development. My chromatin studies of the prototypic long-range repressor Hairy and the short-range repressor Knirps suggest that these two modes of repression induce distinct chromatin states on the repressed genes. These results indicate that transcriptoinal corepressors can be deployed in a context-dependent manner to effect transcription. To discern the importance of individual corepressors, my studies have also examined the role of the conserved Groucho corepressor in Hairy-mediated long-range repression.

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who contributed to my thesis work. First of all, I would like to express my deepest gratitude to my mentor, Dr. David N. Arnosti, for his enthusiasm, patience, support, and excellent guidance during my graduate study. David invested an incredible amount of time in my research projects, and helped me tremendously in becoming a better scientist. I would also like to thank all my lab members for their insight and discussions, with special thanks to Dr. Sandhya Payankaulam for being a supportive colleague and a great friend. My research project was substantially enriched by the thoughtful contribution of my committee members: Dr. Andy Amalfitano, Dr. Lee Kroos, Dr. Richard Schwartz, and Dr. Michael Thomashow. My deepest gratitude goes to my parents in China, whose supports and encouragements inspired me during the course of this work, and made it all worth it.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

aa: amino acid

bp: base pair (s)

cDNA: complementary deoxyribonucleic acid

ChIP: chromatin immunoprecipitation

ChIP-chip: chromatin immunoprecipitation followed by microarray analysis

ChIP-seq: chromatin immunoprecipitation followed by sequencing analysis

CtBP: C-terminal binding protein

DBD: DNA binding domain

DNA: deoxyribonucleic acid

EDTA: ethylenedinitrilo tetraacetic acid

eve: *even-skipped*

ftz: *fushi tarazu*

HAT: histone acetyltransferase

HDAC: histone deacetylase

HEPES: N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid

HMT: histone methyltransferase

hsp: heat shock promoter

PcG: polycomb group

PCR: polymerase chain reaction

Pol II: polymerase II

Rb: Retinoblastoma tumor suppressor protein

RNA: ribonucleic acid

TBP: TATA box binding protein

Chapter I
Introduction

Transcriptional repression is essential for establishing precise patterns of gene expression during development. One of the first major findings stemming from analysis of early *Drosophila* transcriptional regulation was the realization that transcriptional repressors appear to fall into two classes: those acting locally as "short-range" repressors, and those acting globally as "long-range" repressors. The types of transcription factors and cofactors involved point strongly to conservation of these pathways in all metazoans. One area in which our knowledge has lagged is in understanding how these different classes of repressors function at a mechanistic level. The interest in this topic is huge, because the interaction of transcriptional corepressors with developmental circuits is central to key areas of biology including stem cell reprogramming, cancer, and development.

In this work, I describe studies that shed light on the mechanisms of long- and short-range repression during early *Drosophila* development. My chromatin studies of the prototypic long-range repressor Hairy and the short-range repressor Knirps provide mechanistic information about short- and long-range repression at the chromatin level. My work indicates that transcriptional corepressors can be differentially deployed to mediate chromatin changes in a context-dependent manner. To discern the importance of individual corepressors, my studies have also examined the role of the conserved Groucho corepressor in Hairy-mediated long-range repression.

In section I of this chapter, I will first review the literature regarding known general transcriptional repression mechanisms from an evolutionary perspective. Then in section II, I

will introduce the distinctions between long- and short-range transcriptional repressors, with focus on the prototypic long-range repressor Hairy and short-range repressor Knirps, as well as their corepressors. Finally in section III, I will discuss the previous knowledge on the mechanisms of long- and short-range repression, which serve as a background to my own discoveries.

Section I: Transcriptional repression: conserved and evolved features[1]

Transcriptional repression embodies the original mechanism discovered for gene regulation, dating to the pioneering work of Jacob and Monod. Subsequent discoveries filled out a picture showing that bacterial genes are regulated by sophisticated layers of activation and repression, as well as complex posttranscriptional mechanisms. From basic biophysical considerations, simple bacterial *cis* regulatory transcriptional elements should be capable of generating complex outputs of the kind seen with eukaryotic genes [1]. Nature is not so parsimonious, however; the evolution of more sophisticated gene expression programs in eukaryotes has been accompanied by the appearance of more complex machinery and mechanisms of transcriptional control, including repression. A multitude of cellular processes control transcriptional activity of genes, and include negative regulatory pathways that limit production and activity of activator proteins. Dedicated DNA-binding transcription repressors that actively block transcription play equally crucial roles, however. Focusing on this latter aspect, here we survey our current understanding of their repression mechanisms, and suggest

[1] Section I has been published as part of a review article for which I was a coauther: **Payankaulam, S., Li, L.M. and Arnosti, D.N.** (2010) Transcriptional repression: conserved and evolved features. Current Biology 20, R764-R771.

how evolving complexity is matched by adoption of multitiered repressor/corepressor systems that themselves are subject to considerable elaboration and modification.

Transcriptional repression in bacteria: complex responses from streamlined systems

Bacterial RNA polymerase action has been shown to involve binding of the enzyme to the promoter to form a closed complex, melting of the DNA to form the open complex, promoter escape after formation of the first few phosphodiester bonds, and elongation and termination [2, 3]. All of these steps are possible points for control (Figure 1-1). With the exception of antitermination mechanisms that involve proteins complexed to the nascent transcript, bacterial repression pathways invariably involve the direct action of DNA-binding repressor proteins on the transcription machinery to block the activity of RNA polymerase at various steps, including interfering with recruitment, initiation, or elongation. A common mechanism involves an either/or occupancy of promoter regions by inhibitors and the polymerase. The well-studied LacI repressor, for example, utilizes a primary binding site close to the transcription initiation site to block RNA polymerase access to the promoter; this mechanism is greatly enhanced by loop formation involving distal auxiliary sites [4]. Interestingly, placing binding sites just 3′ of the initiation site can switch the mode of repression to blockage of promoter escape, suggesting a flexibility in possible effects mediated by these simple components [5]. Other bacterial repressors, including phage lambda cI repressor and the LexA repressor also target promoter proximal sites to block RNA polymerase binding [6, 7]. In cases where the repressor binding site does not overlap with the promoter, repressor binding can nucleate the binding of additional repressor molecules to occlude RNA polymerase, as seen

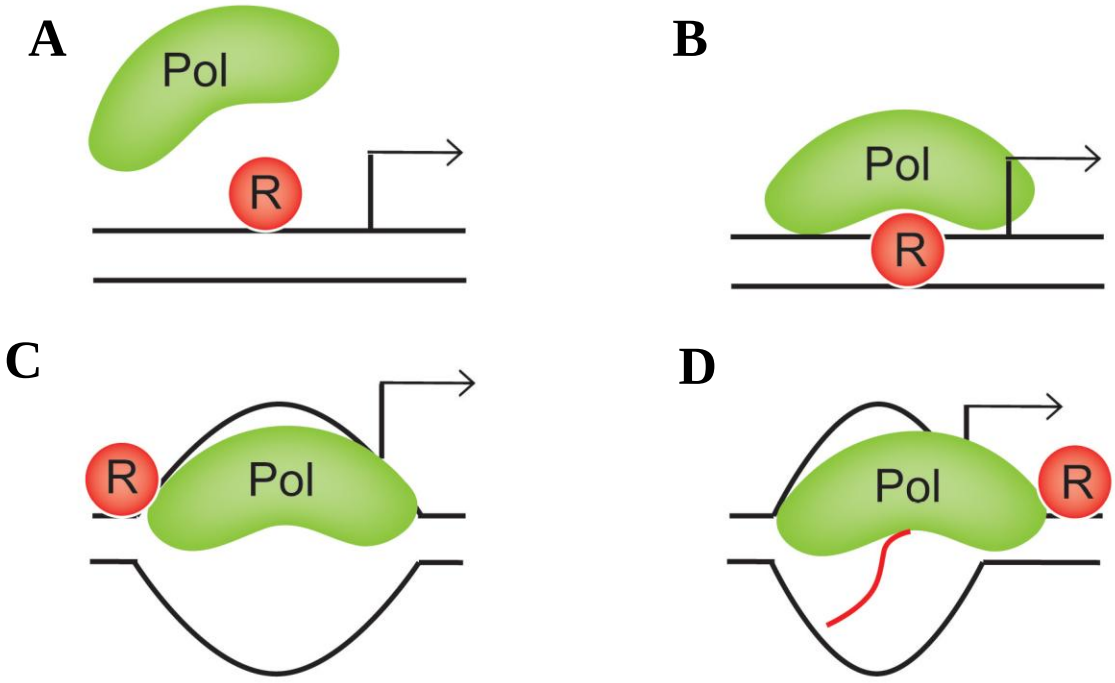


Figure 1-1. Distinct repression mechanisms used by prokaryotic repressors. A bacterial repressor (R) directly targets the core transcription machinery in the following stages: (A) repressor binds to the promoter proximal sites to block the binding of RNA polymerase (Pol); (B) and (C) repressor binds simultaneously to the promoter with RNA polymerase, inhibiting its transition from closed to open complex, or preventing promoter escape; (D) inhibition of RNA polymerase promoter escape can also occur when repressor binds 3′ of the initiation site. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

with DnaA binding to the *Escherichia coli dnaA* gene [8]. Binding of this protein within the gene can also induce premature termination [9].

A second common mechanism involves simultaneous occupancy of RNA polymerase and repressor, with inhibitory consequences. The phage ϕ29 protein p4 binds to sites 5′ of the viral A2b promoter and interacts with the α subunit of the RNA polymerase to prevent promoter escape, although abortive initiation can occur [10]. The MerR repressor acts at an earlier step; this repressor and the polymerase bind simultaneously to the *merT* promoter, resulting in a complex that is blocked in the transition from closed to open complex in the absence of mercury [11, 12]. Upon binding of the metal, allosteric alterations in MerR change DNA bending to permit open complex formation [13]. Similar allosteric controls of RNA polymerase closed to open complex formation have been reported for the *E. coli* GalR repressor and the plasmid-borne KorB repressor. The latter protein occupies promoter proximal sites to exert a moderate level of repression, but strongly synergizes with two other plasmid encoded repressors, KorA and TrbA, that can functionally interact with KorB at a distance, suggesting that distally-acting multiprotein complexes can also be utilized to repress bacterial genes [14].

These examples show that a major theme of transcriptional regulation in bacteria is the direct physical interactions of repressors (and activators) with the basal transcription machinery, programmed by the structure of *cis* regulatory regions. More complex responses are effected by utilization of additional auxiliary sites, as with the

LacI repressor, but there are relatively few instances known where two distinct DNA binding proteins are jointly employed to effect repression (or activation), and in no cases do the repressors need to mediate their control of polymerase through intermediate protein

complexes. This direct interaction contrasts strongly with the situation in eukaryotes, as we discuss below.

Archaeal transcription — complex core machinery, simple repression mechanisms

Recent studies of the third domain of life, the archaea, indicate that the transcriptional machinery of these organisms exhibits similarities with the machineries of both bacteria and eukaryotes. The archaeal RNA polymerase of *Methanococcus jannaschi* has twelve subunits homologous to those of eukaryotic RNA polymerase II; a core complex containing just the homologs to bacterial ββ′α is sufficient for *in vitro* initiation [15]. The holoenzyme is positioned by basal factors TFB and TBP that bind to the basal promoter, similar to the case in eukaryotes [16].

Some Euryarchaea contain H3 and H4-like histones that serve to compact DNA, but these histones lack the tail regions frequently modified for gene regulatory purposes in eukaryotes, and they bind less tightly to DNA than eukaryotic histones [17]. Other archaeal genomes lack these eukaryotic-like histone genes, but appear to have basic DNA-binding proteins that serve a similar purpose [18, 19]. It is not clear if archaeal histones are as inherently inhibitory to transcription as those present in eukaryotes, and whether all types of histone analogs in archea have similar repressive potential.

Archaeal genomes contain numerous homologs to bacterial transcription factors, and insight into transcriptional repression in archaea was first obtained through characterization of one of these, the metal-dependent regulator Mdr1 of *Archaeoglobus fulgidus*. In a similar manner bacterial repressors, the Mdr1 protein regulates transcription of its own gene by blocking recruitment of the RNA polymerase to the promoter [20, 21]. Another transcriptional

repressor, the Phr protein of *Pyrococcus furiosus,* acts through promoter-proximal binding sites to similarly prevent RNA polymerase from accessing the promoter [22]. The basal factors bound at the transcription start site can also be targeted; Lrs14, an archaeal homolog of the bacterial leucine responsive regulatory protein family, inhibits the promoter binding of archaeal general transcription factors TBP and TFB, thus interfering with early steps of transcription initiation [23, 24]. *In vitro* experiments have revealed that NpR, a regulator of nitrogen metabolism in *Methanococcus maripaludis* can also inhibit binding of TFB and TBP.

Thus, even though the archaeal core transcription machinery bears a strong resemblance to that of eukaryotes, transcriptional repression appears to follow the simple scenario seen in bacteria, whereby DNA-binding proteins directly target basal machinery at the promoter. These mechanistic studies, however, have relied largely on *in vitro* assays in the absence of archaeal nucleosome-like proteins, where possible effects of covalent modification of nucleosome-like proteins are neglected [25]. A more complete picture is likely to emerge as biochemical studies progress. One potentially unifying notion is that the small bacterial-sized genomes found in archaea may demand no more highly sophisticated repression mechanisms than those present in bacteria. However, it is striking that the bacterial-sized genome of the budding yeast *Saccharomyces cerevisiae* does feature a full panoply of chromatin-mediated repression mechanisms. It is possible that other features of the eukaryotic lifestyle dictate elaboration of more complex mechanisms for transcriptional control.

Repression features layers of regulation in eukaryotes

To consider how eukaryotic repression mechanisms compare with these bacterial and archaeal mechanisms, we should note key general features of eukaryotic transcription control.

The distinctly different set of proteins and mechanisms involved in eukaryotic repression correlates with enhanced structural and functional demands in eukaryotes (Figure 1-2). Especially in multicellular eukaryotes, achieving correct temporal- and tissue- specific gene expression poses enormous functional challenges. In addition, the large size of many eukaryotic genomes dilutes regulatory sequences in a sea of non-coding elements, making establishment of the correct interactions between regulatory proteins and target genes problematic. Presumably as an adaptation to these demands, the chromatin structure of eukaryotes is profoundly different from that of bacteria and archaea. Eukaryotic histone proteins, assembled on DNA as nucleosomes, pose a formidable intrinsic barrier to access of transcription machinery, reducing background transcriptional noise that fortuitous interactions may generate [26]. Eukaryotic histones, with flexible 'tail' regions, also provide a platform for elaboration of various modification marks of regulatory significance.

An additional, recently appreciated feature of eukaryotic transcriptional regulation is the presence of RNA polymerase at many promoters even when the genes are in a quiescent state, suggesting that promoter escape may be the rate-limiting step [27, 28]. In bacteria, such stalled complexes reflect either a deficient closed-to-open complex transition, a function of the DNA template that is compensated for by activators, or the active intervention by repressors that prevent an otherwise favorable reaction. In eukaryotes, it is not clear whether stalled polymerases are influenced by DNA binding repressors, or if such stalling is solely a consequence of basal promoter sequences and chromatin composition [29].

Eukaryotic transcriptional systems may have developed a number of specific features to provide additional complexity necessary for gene regulatory demands in these organisms. First,

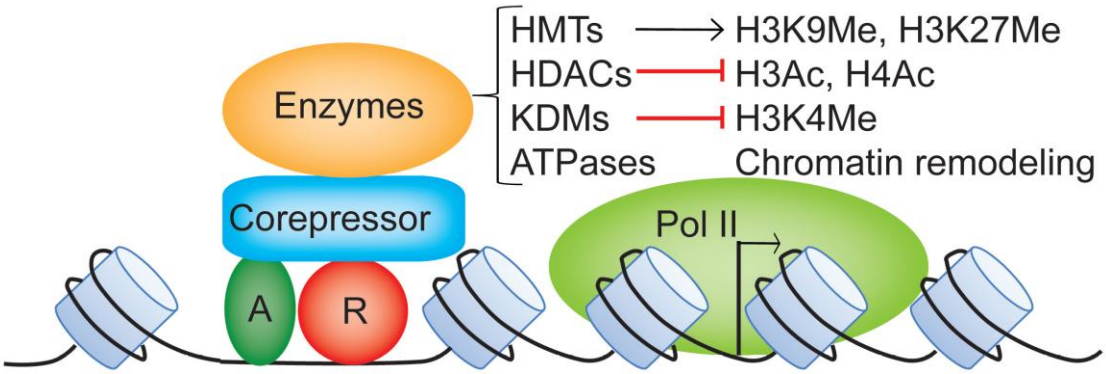


Figure 1-2. Distinct repression mechanisms used by eukaryotic repressors. Because of the increased genome complexity and presence of chromatin, eukaryotic repressors (R) rely largely on recruitment of corepressors and chromatin modifying enzymes, including HMTs (histone methyltransferases), HDACs (histone deacetylases), KDMs (lysine demethylases) and chromatin remodeling factors with ATPase activity. In some cases these corepressors interact directly with activators (A) or the basal machinery.

increased complexity in a transcriptional system may be generated by increasing the number of transcription factors, a trend also noted in bacteria, where those species with complex developmental life cycles tend to possess a large number of RNA polymerase binding sigma factors. Genomic studies have indeed revealed lineage-specific expansions of specific eukaryotic transcription families, suggesting that there are advantages in allocating functions among more actors [30]. When viewed at a global level, however, the percentage of genes annotated as encoding DNA-binding transcription factors is similar in eukaryotes and bacteria, ranging from 7% in *E. coli* to 12% in *Streptomyces coelicolor*, compared with 5.9% in the plant *Arabidopsis thaliana* and 10% in humans [31, 32]. Some eukaryotes, such as the nematode *Caenorhabditis elegans,* have a smaller percentage of genes devoted to DNA-binding transcription factors (5%) than do many bacteria [33]. Alternative splicing and posttranslational modifications can of course increase the number of isoforms present in a cell, but this is not the sole source of complexity.

A second layer of complexity is the nature of eukaryotic *cis* regulatory elements; many bacterial elements consist of only a few binding sites, whereas eukaryotic promoters and enhancers commonly have dozens of motifs responsible for assembling core and regulatory machinery. Such combinatorial interactions provide an effective way to differentiate real from spurious cis elements in large genomes, facilitate cooperative mechanisms to overcome the barrier of histones, and offer more complex switch responses for fine-scale control. Yet with respect to transcriptional repressors, these proteins are generally capable of mediating effective transcriptional responses from rather simple cis elements; often from one or a few

binding sites for transcription factors. Hwa's predictions that *cis* element structures need not be elaborate to generate proper readouts appear to be accurate in this context [34].

A third major area of difference between bacterial and eukaryotic systems is the elaboration of core transcriptional machinery, which is defined as the set of proteins required for assembly and initiation at a basal promoter. In addition to the twelve polypeptides associated with the RNA polymerase II responsible for transcription of protein coding genes, at least 60–80 other polypeptides are considered critical for the essential activity of transcriptional initiation [35]. Clearly, this complexity should offer additional levels of regulation. In fact, recent discoveries of tissue-specific forms of this basal transcriptional machinery suggest that developmental gene regulation does involve alterations to this core set of factors [36]. It is less clear that with a larger complex of proteins assembling at promoters, a proportionally greater number of control points or available protein surfaces are directly contacted by different transcription factors. Many activator effects funnel through a limited number of targets in the basal machinery, and as we note below, repressors generally appear to predominantly interact with the chromatin.

The nature of the interaction between transcription factor and transcription machinery/DNA template is where bacterial and eukaryotic regulation shows the greatest divergence. Eukaryotic transcriptional activators appear to function through two general levels, namely, recruiting coactivator complexes that can modify chromatin, and interacting with core transcriptional machinery [37]. In the latter category, some transcriptional activators directly target components of the basal machinery, such as TAF proteins of the TFIID complex, to recruit or allosterically affect these proteins, while other transcriptional factors interact with core

machinery by means of coactivator complexes. DNA-binding eukaryotic transcriptional repressors, on the other hand, have only in rare instances been documented to directly contact and modify the activity of core machinery [38]. Some *in vitro* systems have documented competitive interactions between repressors and the basal transcriptional machinery, but the *in vivo* relevance of such studies remains uncertain [39]. Physical competition may provide a mechanism, as DNA binding repressors have also been shown to compete for overlapping activator binding sites, but the generality of this model is uncertain, for most repressor binding sites do not directly overlap activator sites [40]. Most DNA-binding transcriptional repressors appear to rely instead on indirect interactions, mediated by corepressors.

Corepressors — definition, function and conservation

The term corepressor refers to either a single protein or a scaffolding protein mediating the assembly of a multi-subunit complex that is recruited to a gene by a transcription factor. Repressors have discreet peptide motifs often contained within transcriptional repression domains which permit interaction with specific corepressors. It has been assumed that the transcription factor–corepressor complex is modular, because direct tethering of corepressors to DNA via a heterologous DNA-binding domain elicits transcriptional repression [40-42]. This assay neglects important context effects, however, as we discuss below. Corepressor complexes can exhibit a range of activities; they frequently contain histone deacetylase (HDAC) activity essential for stripping activation marks from histone tails. Corepressor complexes can contain other enzymatic activities as well, including histone methyl transferase or demethylase activities that can reset chromatin marks (Figure1-2).

Another conserved group of cofactors are SWI/SNF nucleosome remodeling complexes which generally serve an activating role, but in some cases are recruited by transcription factors to induce gene repression and thus can be considered a special type of corepressor. The Mi2–NuRD corepressor complex combines nucleosome remodeling activity with histone deacetylase activity, and does appear to play a dedicated repressive role [43-45]. In addition to nucleosome modification, corepressors have been shown to directly interfere with coactivators such as the CBP/p300 histone acetyltransferase or the mediator complex [46, 47].

These latter examples appear to be exceptions involving nonhistone targets, but in general, histone modifiying enzymes are capable of modifying transcription factors with the same posttranscriptional modifications as seen on histone tails. Nonetheless, our current understanding is that corepressors predominantly target chromatin structures. With hundreds of transcription factors in the cell, but only a dozen or so corepressor complexes, it may appear that implementing transcriptional repression through these cofactors produces a bottleneck of limited, generic activities. However, corepressors are not simply interchangeable levers to lower the nucleosome barrier, but sensitive control points that provide additional regulatory modulations through structural diversification, by sensing of signaling cascades, and context-dependent responses. These elaborations tune cell/tissue specific gene responses that are keys for development.

Corepressors link signal pathways with transcription activity

Signaling pathways link environmental and physiological stimuli to transcription activity through posttranslational modifications of transcription factors, for example phosphorylation, acetylation, hydroxylation, and sumoylation. Corepressor complexes are similarly targets of

signaling pathways and may provide unique channels of communication with the cellular environment. Posttranslational modifications of corepressors can either change their subcellular localization or regulate the binding to transcription factors. In addition, some corepressors can bind to small molecule compounds in the environment, which can modulate repression potencies (Figure 1-3).

A number of posttranslational modifications affect subcellular localization of corepressors. Mammalian CtBP1 is SUMOylated, a modification that is essential for its accumulation in the nucleus and repression of the E-cadherin promoter [48]. Acetylation of CtBP2 by p300 can play a similar role; this modification is required for nuclear retention and repression activity [49]. The SMRT corepressor is phosphorylated by kinases of the MAP kinase pathway, leading to subcellular redistribution [50].

Posttranslational modifications of corepressors also affect the recruitment of corepressor complexes by transcription factors. Phosphorylation is one of the most prevalent post-translational modifications that modulate corepressor activities. During G1–S phase transition, the Rb corepressor is phosphorylated at multiple residues by the combinatorial effects of several cyclin-dependent kinases (CDKs), which leads to the disruption of the Rb–E2F complex [51]. Groucho activity is also modulated by phosphorylation, in response to signaling pathways. MAPK pathway phosphorylation attenuates Groucho's binding to the Eyeless transcription factor, while in contrast phosphorylation of the Groucho homolog TLE1 by casein kinase II enhances its association with the Hairy homolog Hes1 and with chromatin [52-54]. This mechanism is suggested to allow Groucho to play an integrating role between the EGF receptor and Notch signaling pathways [55]. A distinct modification, polyADP ribosylation by PARP1, is

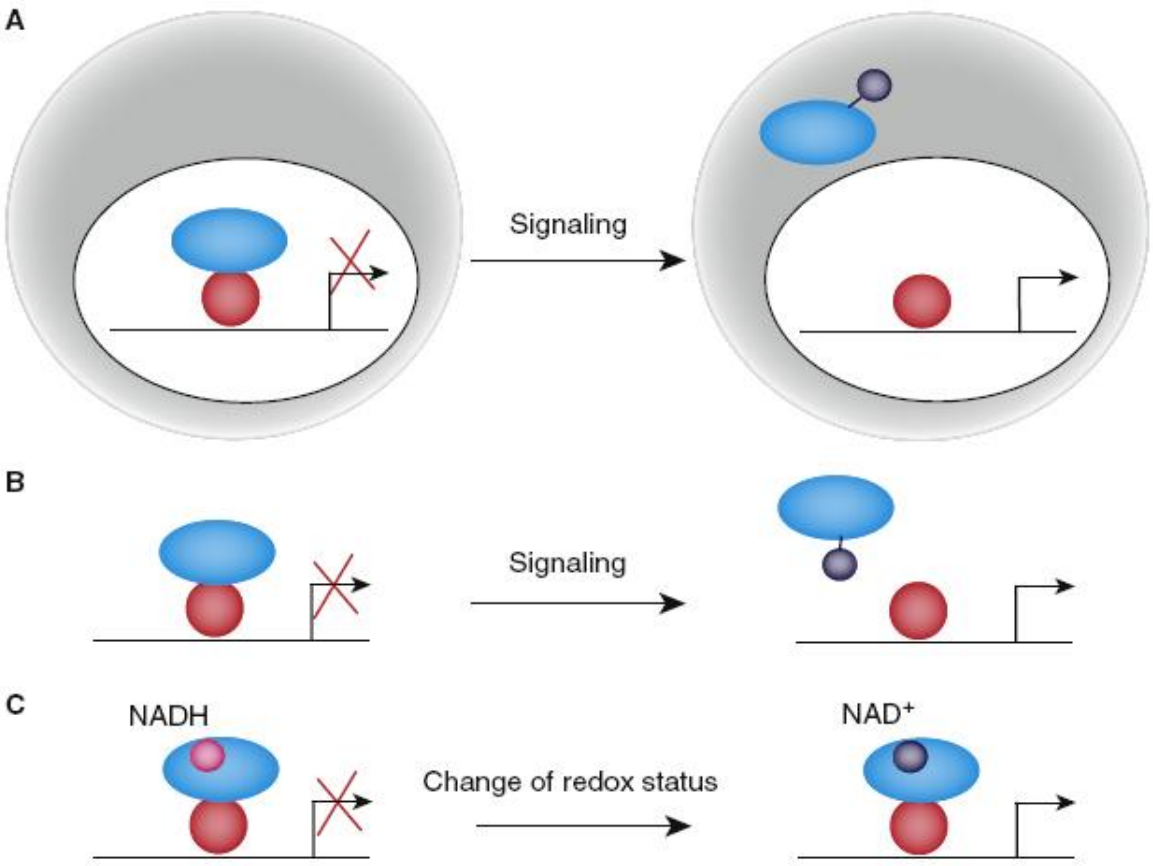


Figure 1-3. Modulation of corepressor activity through posttranslational modification and small molecules. (A) Subcellular localization of corepressors is affected by posttranslational modifications, such as phosphorylation, SUMOylation or acetylation. (B) Posttranslational modifications of corepressors alter a protein-protein interaction surface, either inhibiting or enhancing the recruitment of corepressors. (C) Binding to cellular compounds may impact the corepressor activity: CtBP is suggested to be a more potent corepressor when bound to NADH compared to NAD+, thus allowing it to function as a cellular redox state sensor.

suggested to mediate the dismissal of the TLE1 corepressor complex from HES1-regulated promoter during neural stem cell differentiation [56].

In addition to these covalent modifications, binding of ligands has been suggested to provide a way for corepressors to respond to the changes in the cellular environment. CtBP exhibits remarkable sequence similarity to dehydrogenases and has a Rossman fold for binding to NAD+/NADH [57]. Binding of the cofactor is essential for CtBP's interaction with cellular and viral transcriptional repressors, and it was reported that NADH has a much higher affinity than NAD^+ for CtBP [58]. CtBP might serve as a redox sensor to regulate transcription according to changes in nuclear NAD^+/NADH ratios. These modifications provide a sampling of some of the ways that corepressor activity can be modulated by posttranslational effects that link corepressor function to cellular physiology.

Section II: Long- and short-range transcriptional repressors

Analysis of transcriptional regulation in Drosophila indicates a further level of complexity regarding the transcription repression mechanisms in eukaryotes. Studies revealed that transcriptional repressors can be classified into two classes according to their range of action [59]. Long-range repressors act over a range of more than 1 kb to dominantly silence the whole targeted locus. In Drosophila, Hairy and Dorsal repressors have been found to repress enhancers from a distance in transgenic reporter assays in *Drosophila* embryos [60, 61]. Related examples of long-range repression in mammals include: a silencer element contained within

the mouse CD4 gene, a neuron-restrictive silencer element regulating the mouse neuronal growth associated protein *scg-10*, a retinoic acid response element within the mouse *Hoxb-1* locus, and the mammalian E2F-Rb complex binding element [62-66]. Although repressors binding to most of these elements are not very well characterized, the elements have been demonstrated to function in a direction- and orientation- independent manner to mediate dominant repression. On the other hand, short-range repressors work over a distance of less than 100 – 150 bp to inhibit activators in a local fashion [67]. Prototype short-range repressors in Drosophila include Knirps, Kruppel and Snail [59-61, 68, 69] (Figure1-4).

Long- and short-range repressors identified in Drosophila have homologs throughout metazoa, including highly conserved functional motifs, such as cofactor binding surfaces and DNA binding domains. I will introduce the prototypic long-range repressor Hairy and short-range repressor Knirps in more detail, as they were used as models to study mechanisms of long- and short-range repression in the following chapters.

The long-range repressor Hairy

1. The Hairy and Enhancer of Split related proteins

Hairy is the founding member of the Hairy and Enhancer of Split (HES) related protein family. Other proteins within this family include Deadpan, HES, Hey and Stra 13. Most of the family members function as DNA binding transcriptional repressors, which contain a tandem arrangement of the basic helix-loop-helix domain and an adjacent sequence known as the Orange domain. [70]. Except for the Stra 13 proteins, proteins in this family also have a conserved C-terminal tetrapeptide motif, either WRPW or YXXW that mediates interactions with the Groucho corepressor [70].

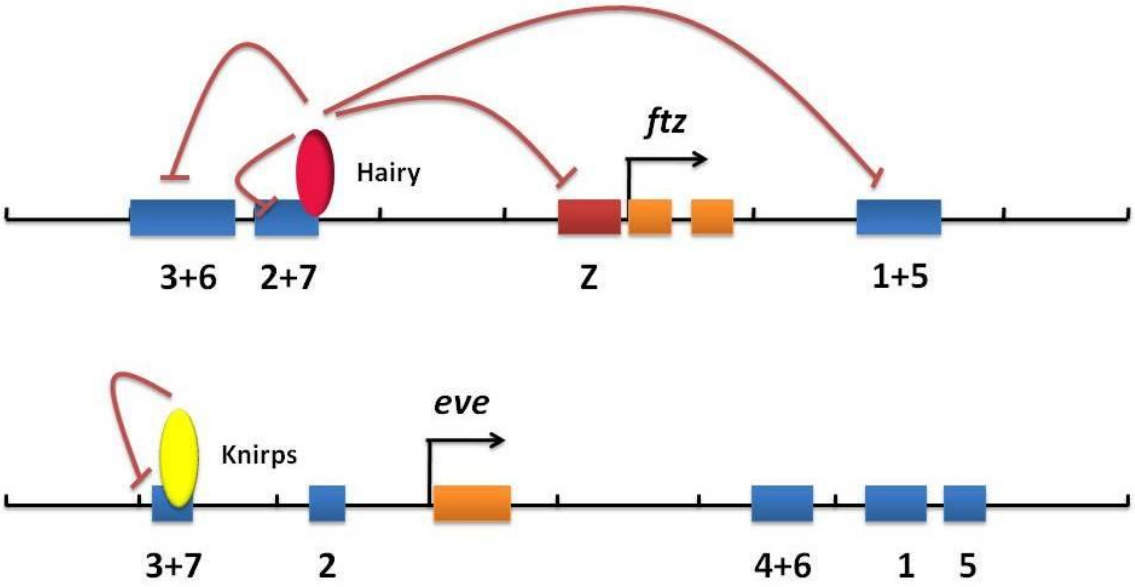


Figure 1-4. Long-and short-range repression. (A) Long-range repressors like Hairy can dominantly repress all enhancers within the targeted locus. Shown in the figure is the schematic representation of Hairy regulation of its canonical target *fushi tarazu (ftz)*. (B) Short-range repressors like Knirps can only locally regulate the enhancers they are located in. Shown in the figure is the schematic representation of Knirps regulation of its canonical target *even skipped (eve)*. The boxes represent stripe-specific enhancers, and the numbers underneath the boxes represent the targeted stripes.

HES proteins regulate multiple biological processes in *Drosophila* as well as vertebrates, including mesoderm segmentation, vertebrate neurogenesis, neural cell fate determination, blood vessel formation, myogenesis, and inhibition of oncogenesis [70]. Understanding Hairy repression will not only shed light on the basic long-range repression mechanisms, but also provide insights into many important biological processes.

2. Hairy –mediated transcriptional repression in *Drosophila*

In *Drosophila*, Hairy directly binds to DNA and functions as a homodimer [71, 72]. *In vivo* transgenic reporter studies revealed that Hairy mediates long-range dominant repression from isolated binding sites, although later studies suggested that this repression can be potentiated by adjacent activators [60, 73]. Like many repressors in eukaryotes, Hairy repression is mainly mediated by corepressors, and the Groucho corepressor has been suggested to account for most of Hairy repression activity [74, 75].

Other than Groucho, Hairy has also been shown to interact with corepressor dSir2 through a motif RRAR in the basic region of Hairy and with corepressor C-terminal binding protein (CtBP) through the PLSLV motif located in close proximity to the Groucho binding motif, only nine amino acids away from the WRPW motif [76, 77] (Figure 1-5, A). The roles of these different corepressors in Hairy-mediated repression are still unclear. One possibility is that Hairy utilizes these corepressors in a context-dependent manner; *in vivo* binding analysis carried out using DamID suggested that the Groucho and CtBP corepressors do not entirely overlap with Hairy binding [78]. It is also possible that these corepressors interact with each other to fine-tune the output of Hairy. dCtBP and Groucho are seen to have antagonistic effects in Hairy-mediated repression [79].

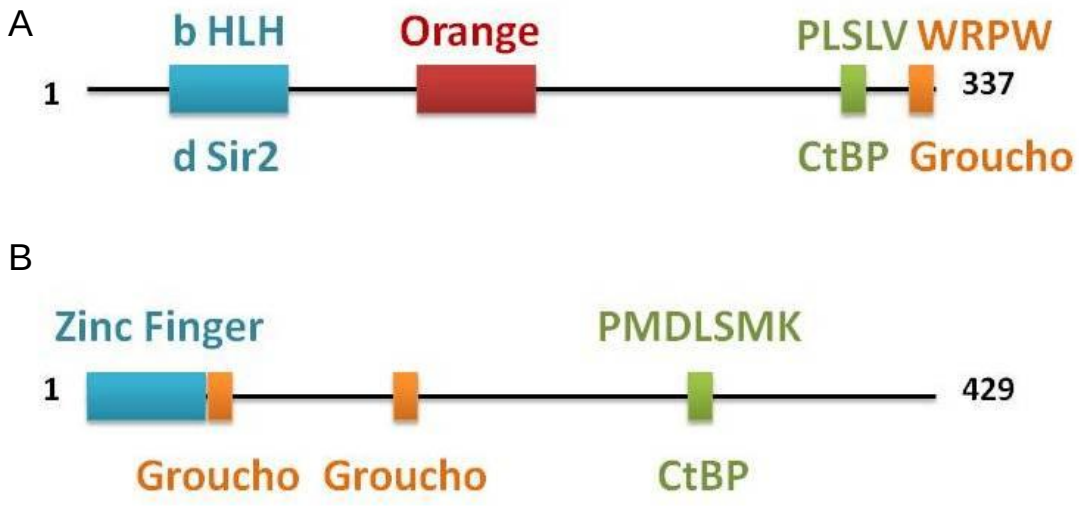


Figure 1-5. Both long-range repressor Hairy and short-range repressor Knirps recruit multiple corepressors. Schematic representations of functional domains of Hairy and Knirps are shown. (A) Hairy binds to DNA through bHLH domain, which also recruits dSir2 corepressor. Orange domain is conserved among Hairy and Enhancer of split family members, and is important for Hairy function. The C-terminus of Hairy protein contains the Groucho recruitment motif 'WRPW' and the CtBP recruitment motif 'PLSLV'. These two motifs are nine amino acids apart. (B) Knirps binds DNA through Zinc Finger at its N-teminus. Groucho is recruited to Knirps through the collaborative action of two motifs located at 85-91aa and 169-189aa. CtBP is brought to the complex through an atypical 'PMDLSMK' motif.

Despite the importance of Hairy and its related family members in Drosophila, only a handful of Hairy targets have been identified so far. Hairy was first identified as a transcriptional repressor through its ability to repress *fushi tarazu* expression upon ectopic expression using a *hsp70* promoter [80]. During development, Hairy is expressed in an anterior-posterior seven stripe segmental pattern that serves to stabilize the striped expression pattern of *ftz*. The initial *ftz* pattern is established by regional cues provided by gap genes [81]. Hairy stabilizes this pattern by repressing this gene in the interstripe regions. Hairy also inhibits sensory organ formation by direct repression of *achaete,* through binding sites located 300 bp upstream of this gene's basal promoter [72]. Recent studies have suggested that Hairy also regulates the expression of TCA cycle enzymes, functioning as a metabolic switch to mediate hypoxic tolerance [82].

The short-range repressor Knirps

Knirps is a member of the nuclear receptor superfamily, possessing this family's characteristic zinc finger DNA binding domain, which is coupled to a unique C-terminal motif 'PMDLSMK' that interacts with CtBP [83, 84]. Knirps is an important regulator for striped patterns of pair-rule gene expression in the *Drosophila* embryo [85]. *even skipped* is a canonical Knirps target, with extensive information about the role of this repressor on the modular *eve* stripe 3+7 and 4+6 enhancers [86]. During later stages of development, Knirps and its related proteins determine the position of the second longitudinal wing vein (L2) [87]. They also control cell fates and branch morphogenesis during tracheal development [88].

Knirps has been shown to mediate both local quenching and direct repression in Drosophila embryos in both a CtBP-dependent and -independent manner [59, 89] (Figure1-5,

B). The Rpd3 deacetylase has also been shown to exist in the Knirps corepressor complex through its interaction with CtBP [90]. Earlier studies suggested that similar to the distinction between transcription repressors, distinct corepressors are required for long- and short-range repression: Groucho is the mediator for long-range repression, and CtBP is required for short-range repression [84, 91]. Although functioning as a dedicated short-range repressor, Knirps also recruits and functionally interacts with the Groucho corepressor, which had previously been suggested to be a dedicated long-range corepressor [92].

Long- and short-range transcriptional corepressors

The "Long-range" corepressor Groucho

1. Groucho and the Transducin-like Enhancer of split (TLE) proteins

The Drosophila *groucho* (*gro*) gene was identified through a viable mutation that results in clumps of extra bristles above the adult eyes, reminiscent of the bushy eyebrows of Groucho Marx [93] . Groucho is the prototype of the transducin-like Enhancer of split (TLE) protein family, whose members are found in most metazoans. Members within this family contain 40 amino acid tandem repeats with characteristically located tryptophan and aspartate residues, so called the WD-40 motif [94]. They function mainly during segmentation, central and peripheral neurogenesis, as well as epithelial differentiation in *Drosophila* and mammals [74, 95].

2. Mechanisms of Groucho/TLE-mediated repression

Groucho's role as a transcriptional corepressor was initially revealed through its interaction with Hairy. Later studies revealed that Groucho interacts with many other sequence-specific repressors including Dorsal, Runt, Brinker and Engrailed [74, 96-99].

Courey and colleagues proposed that Groucho functions by spreading along the chromatin, based on the evidence that Groucho can oligomerize *in vitro* by means of its N-terminal glutamine-rich-domain, which is important for its function. Biochemical studies indicated that Groucho interacts with the Rpd3 histone deacetylase and with deacetylated histones [41, 91, 100, 101]. Our laboratory has demonstrated that Groucho spreads along the transcribed region of a Hairy-repressed reporter *in vivo* [102].

The fact that Groucho interacts with Hairy and Dorsal, two well-characterized long-range repressors, as well as the molecular characteristics of Groucho, led to the prevalent view that Groucho is a long-range corepressor [91]. However, accumulating evidence suggests that Groucho also is involved in short-range repression mediated by Slp1 and Knirps [92, 103]. Thus, rather than functioning as a dedicated long-range repressor, Groucho is more likely to function in a context-dependent manner to mediate repression.

The short-range corepressor CtBP

1. CtBP family protein

CtBP was discovered during a search for cellular proteins that complex with the C-terminal region of E1A implicated in negative modulation of oncogenic transformation. CtBP interacts with a short sequence motif (PLDLS) conserved among the E1A proteins of all human and primate adenoviruses [104]. The Drosophila CtBP has been shown to be essential for wild-type function of short-range repressors, and it binds to the short-range repressors Knirps, Kruppel and Snail [84]. Although CtBP has also been shown to interact with the long-range repressor Hairy, this interaction appears to antagonize with Hairy repression, through competition with Groucho corepressor [79].

2. Mechanisms of CtBP repression

Proteomic analysis of the mammalian CtBP1 complex revealed that CtBP1 recruits multiple enzymatic components, including the histone deacetylase Rpd3, histone methyltransferases G9a and GLP which methylate histone H3 at lysine 9, and the histone demethylase LSD1, which demethylates di-methylated histone H3 lysine 4 [105, 106]. It is possible that CtBP-mediated repression is a combinatorial effect of decreased activating chromatin marks and elevated repressive chromatin marks.

Context-dependent activity of corepressors[2]

Assuming a cofactor is in the correct subcellular location and in an active state, rather than simply switching a gene off, corepressors can be deployed in a context-dependent manner to mediate distinct types of repression. Recent studies of the Groucho corepressor have provided a tantilizing insight into yet another level of functional complexity affecting repression systems. Although considered as a long-range corepressor for a long time, recent evidence indicates that Groucho is also an essential corepressor for the Slp1 and Knirps short-range repressors in the *Drosophila* embryo [78,79]. The same component of the cellular corepression machinery can thus evince distinct functions —to what may we attribute this flexibility? Several molecular features of the protein and its complexes provide clues.

In some cases, transcription factors interact with multiple corepressors to elicit context-dependent effects. Different corepressor complexes are recruited to generate developmental stage-specific, or gene- or enhancer- specific effects. The *Drosophila* Runt repressor requires different corepressors for repression of *engrailed* at different developmental stages: initially

[2] This section has been published as part of a review article for which I was a coauther: **Payankaulam, S., Li, L.M. and Arnosti, D.N.** (2010) Transcriptional repression: conserved and evolved features. Current Biology 20, R764-R771.

repression is dependent on Tramtrack, and later Groucho, CtBP and Rpd3 are required for the maintenance of repression after blastoderm stage [80]. The *Drosophila* repressor Brinker protein displays gene specific effects, relying on either CtBP or Groucho complexes to repress different target genes [81].

Even distinct enhancers from the same gene can show cofactor specific effects. The *even skipped* stripe 3/7 enhancer is repressed by Knirps even in the absence of CtBP, while this same repressor requires CtBP to inactivate the stripe 4/6 enhancer during the same period of development. Knirps relies on Groucho for repression of the stripe 3/7 enhancer [79,82]. Such combinatorial usage of corepressor complexes is not restricted to transcription factors in Drosophila but is also observed in higher eukaryotes. For example, the *Xenopus* Tcf3 (XtCF3) interacts with both the Groucho and CtBP corepressors to regulate target gene expression [83].

Section III. Long- and short-range repression: two sides of the same coin?

The fact that both long- and short-range repressors can recruit CtBP and Groucho raises several interesting questions. Is there indeed a distinction between long- and short-range repression? What are the underlying differences between long- and short-range repression? In order to better dissect the two modes of repression during early *Drosophila* development, it is important to review previous knowledge of long- and short-range repression mechanisms.

Mechanisms of short-range repression

Transgenic reporter assays performed in *Drosophila* embryos have revealed several possible mechanisms of short-range repression. Short-range repression can involve direct inhibition of activator binding as seen for example by Snail repressor binding sites that overlap with

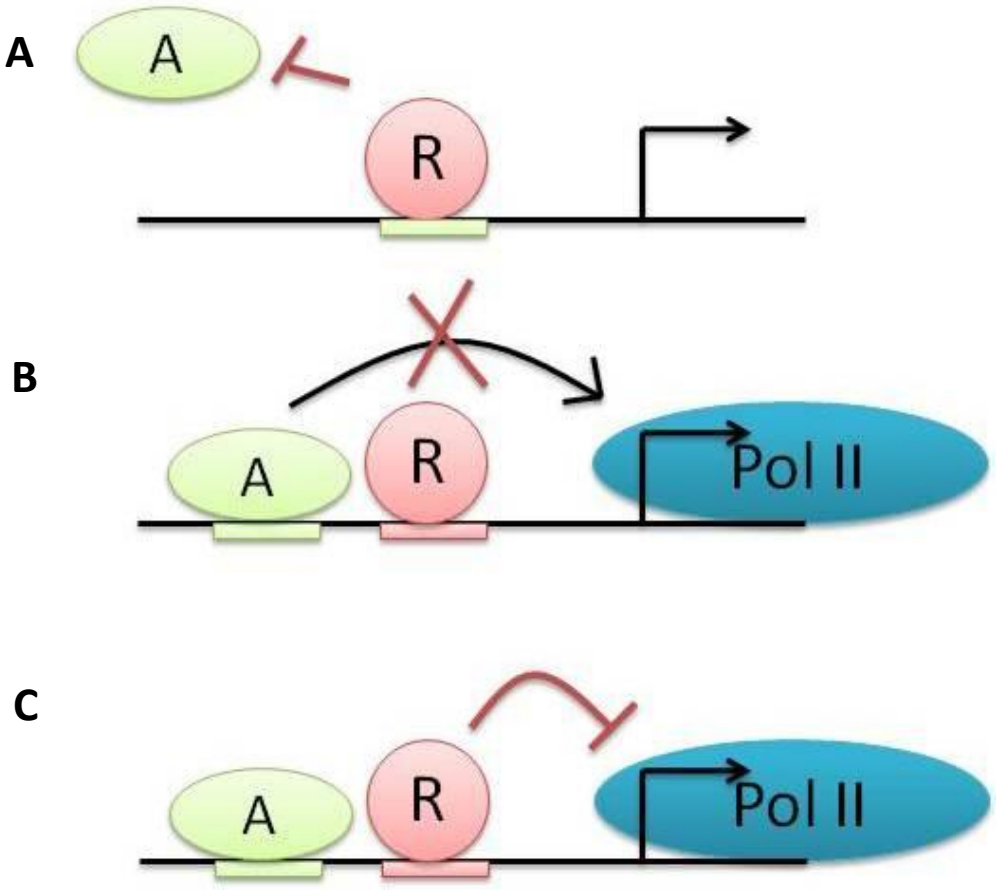


Figure 1-6. Possible mechanisms of short-range repression. Analysis of short-range repression in Drosophila embryos suggests that there might be three different mechanisms. (A) Short-range repressors (R) compete for the same binding sites with activators (A). (B) Short-range repressors (R) disrupt the crosstalk between the activators (A) and the basal transcription machinery (Pol II). (C) Short-range repressors (R) can directly interact and interfere with the basal transcription machinery (Pol II).

Twist activator binding sites in the dorsal-ventral enhancers regulating genes expressed in the neuroectodem of the embryo (Figure 1-6). Short-range repression can also involve a "quenching" mechanism, which need not involve binding sites that overlap with activator sites. Alternatively, when suitably positioned proximal to the transcription start site, short-range repressors can directly interact with the basal transcriptional machinery [59, 68].

These conclusions are largely based on experimental approaches that involve manipulation of repressor binding sites and measurement of the associated transcriptional readouts. Many mechanistic questions remain to be answered. Do activators and components of the basal transcriptional machinery stay bound upon repression? How might the repressors inhibit the crosstalk between activators and the basal transcriptional machinery? How do the repressors regulate the basal transcriptional machinery and affect chromatin structure? Genome-wide ChIP-chip studies reveal that the Snail short-range repressor co-occupied repressed targets with Twist activator and RNA polymerase II, providing direct evidence for the quenching model in which the crosstalk between activators and the basal transcriptional machinery after recruitment was interrupted by the short-range repressor [28, 107].

Mechanisms of long-range repression

The two most popular models to explain long-range repression are mainly through based on studies of Polycomb group proteins and Sir family proteins. The first model is the 'spreading' mechanism, in which binding of a repressor at one site facilitates the adjacent binding of additional repressors or corepressors until the resulting protein complex extends throughout the targeted locus, somehow interfering with gene expression (Figure 1-7, A). The best example for this mechanism by the Polycomb group proteins-mediated repression,

occupancy of a locus by the Polycomb group complex is associated with transcription silencing, despite continued occupancy of the basal transcriptional machinery, as well as wide-spread H3K27, H3K9 and H4K20 methylation. Though the exact roles of these epigenetic marks are still unclear [108, 109]. Another well-studied example for a spreading mechanism is exemplified by SIR-dependent repression in budding yeast, in which the Sir2/Sir3/Sir4 proteins assemble a silent chromatin region of around 3 kb at telomeres and the silent mating type loci [110]. Similarly, spreading of *Drosophila* Groucho protein has been observed on a reporter gene regulated by Hairy long-range repression, which is thought to deliver the histone deactylase Rpd3 along the chromatin to produce an inaccessible chromatin environment [102].

An alternative model for long-range repression involves a looping mechanism, in which the intervening DNA longs out to allow contact between repressors and their targeted enhancers and promoters (Figure 1-7, B). One alternative model for the role of H3K27 methylation in PRC1-mediated repression is that H3K27 facilitates intralocus looping to bring PRE (polycomb response element)-bound complexes into contact with the body of the silenced gene [111]. This model is supported by the studies showing that one insulator between the PRE and promoter blocks silencing, but two insulators in tandem allow PRC repression to bypass this blockage [112]. The observation suggests the continuous chromatin context is not necessary for Polycomb repression, and the PRC can directly repress the target promoter through looping out of the DNA in-between (Figure 1-7, B).

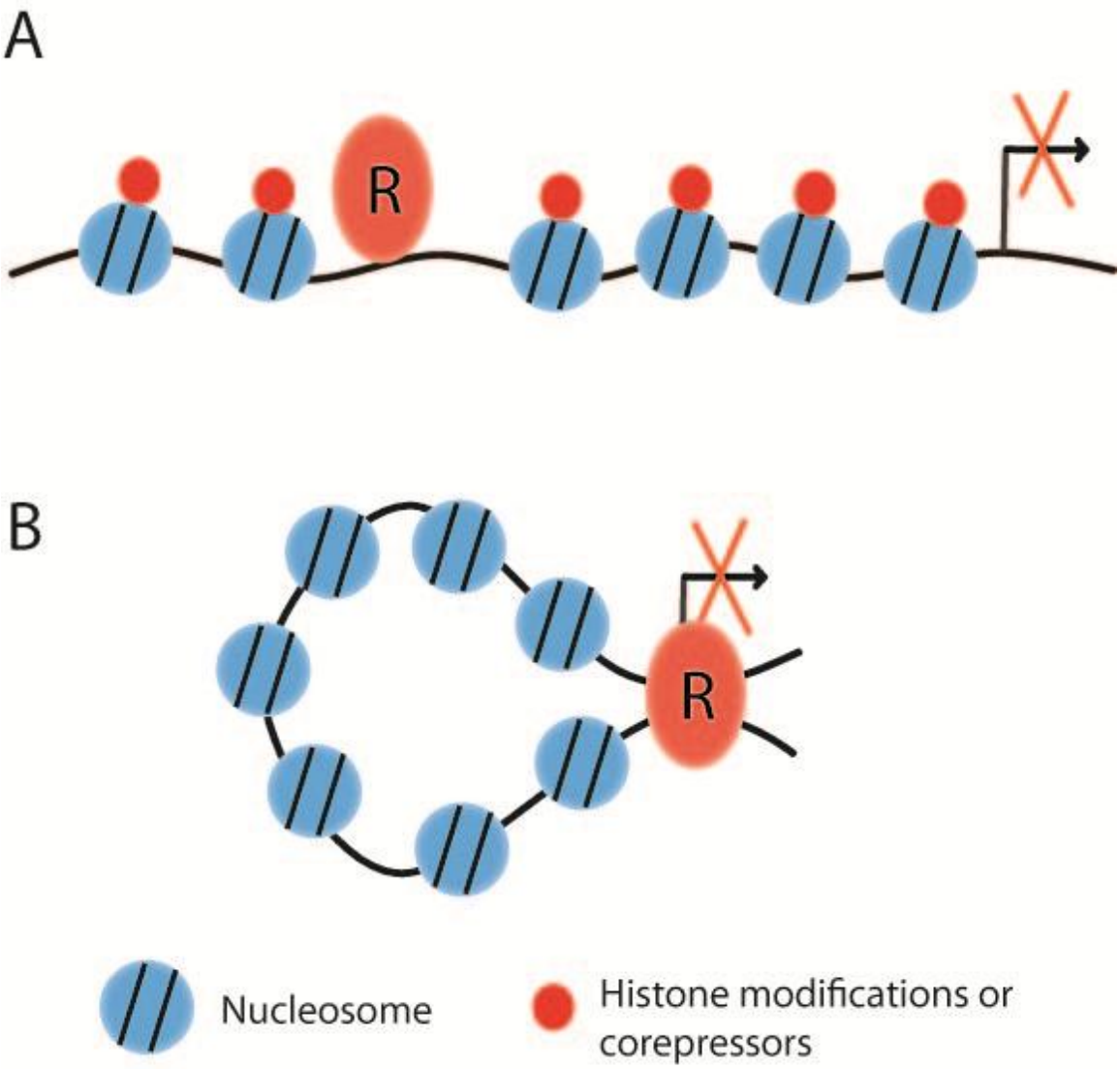


Figure 1-7. Mechanisms of long-range transcriptional repression. There are two mechanisms by which long-range repressors (R) can reach distal promoters. (A) **Spreading Mechanism**: Binding of the long-range repressors might facilitate the binding of additional repressors or corepressors to adjacent sites, or modifications of chromatin nearby, until the resulting protein complex or modified chromatin extend throughout the targeted locus. (B) **Looping Mechanism:** The repressors (R) might interact with proteins binding to the basal promoters, allowing direct contacts between the enhancers and promoters, and looping out the chromatin in-between.

Conclusions

Examination of the mechanisms used by DNA-binding transcriptional repressors indicates that a relatively limited set of direct interactions is sufficient to regulate the activity of bacterial and archaeal promoters, which feature compact *cis* regulatory elements. Eukaryotic genes, whether because of their more complex *cis* regulatory structure, chromatin packaging, or regulatory demands tend to require the action of corepressors, proteins that directly or indirectly mediate histone modifications. Other than simply dictating an 'OFF' state of the targeted genes, transcriptional repressors can be classified into long-and short-range repressors based on their ranges of action. To understand the molecular distinctions between long-and short-range transcriptional repressors, it is necessary to determine how they regulate the chromatin environment, activators and the basal transcriptional machinery. A single corepressor can be recruited by multiple transcriptional repressors. On the other hand, a single transcription repressor is usually capable of recruiting multiple corepressors to differentially repress their target genes. It is important to understand the roles of different corepressors within a single repressor complex, in order to better predict the repression output mediated by the particular repressor.

References

References

1. Bintu, L., et al., Transcriptional regulation by the numbers: applications. Curr Opin Genet Dev, 2005. **15**(2): p. 125-35.

2. Chamberlin, M.J., The selectivity of transcription. Annu Rev Biochem, 1974. **43**(0): p. 721-75.

3. Krummel, B. and M.J. Chamberlin, RNA chain initiation by Escherichia coli RNA polymerase. Structural transitions of the enzyme in early ternary complexes. Biochemistry, 1989. **28**(19): p. 7829-42.

4. Lewis, M., Response: DNA Looping and Lac Repressor--CAP Interaction. Science, 1996. **274**(5294): p. 1931-2.

5. Lopez, P.J., et al., On the mechanism of inhibition of phage T7 RNA polymerase by lac repressor. J Mol Biol, 1998. **276**(5): p. 861-75.

6. Hawley, D.K., A.D. Johnson, and W.R. McClure, Functional and physical characterization of transcription initiation complexes in the bacteriophage lambda OR region. J Biol Chem, 1985. **260**(14): p. 8618-26.

7. Bertrand-Burggraf, E., et al., Promoter properties and negative regulation of the uvrA gene by the LexA repressor and its amino-terminal DNA binding domain. J Mol Biol, 1987. **193**(2): p. 293-302.

8. Lee, Y.S. and D.S. Hwang, Occlusion of RNA polymerase by oligomerization of DnaA protein over the dnaA promoter of Escherichia coli. J Biol Chem, 1997. **272**(1): p. 83-8.

9. Tesfa-Selase, F. and W.T. Drabble, Specific binding of DnaA protein to a DnaA box in the guaB gene of Escherichia coli K12. Eur J Biochem, 1996. **241**(2): p. 411-6.

10. Monsalve, M., et al., Activation and repression of transcription at two different phage phi29 promoters are mediated by interaction of the same residues of regulatory protein p4 with RNA polymerase. EMBO J, 1996. **15**(2): p. 383-91.

11. Heltzel, A., et al., Activator-dependent preinduction binding of sigma-70 RNA polymerase at the metal-regulated mer promoter. Biochemistry, 1990. **29**(41): p. 9572-84.

12. Summers, A.O., Untwist and shout: a heavy metal-responsive transcriptional regulator. J Bacteriol, 1992. **174**(10): p. 3097-101.

13. Ansari, A.Z., J.E. Bradner, and T.V. O'Halloran, DNA-bend modulation in a repressor-to-activator switching mechanism. Nature, 1995. **374**(6520): p. 371-5.
14. Williams, D.R., M. Motallebi-Veshareh, and C.M. Thomas, Multifunctional repressor KorB can block transcription by preventing isomerization of RNA polymerase-promoter complexes. Nucleic Acids Res, 1993. **21**(5): p. 1141-8.

15. Korkhin, Y., et al., Evolution of Complex RNA Polymerases: The Complete Archaeal RNA Polymerase Structure. PLoS Biol, 2009. **7**(5): p. e102.

16. Zhang, G., et al., Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell, 1999. **98**(6): p. 811-24.

17. Wilkinson, S.P., M. Ouhammouch, and E.P. Geiduschek, Transcriptional activation in the context of repression mediated by archaeal histones. Proc Natl Acad Sci U S A. **107**(15): p. 6777-81.

18. Li, W.T., et al., MJ1647, an open reading frame in the genome of the hyperthermophile Methanococcus jannaschii, encodes a very thermostable archaeal histone with a C-terminal extension. Extremophiles, 2000. **4**(1): p. 43-51.

19. Fahrner, R.L., et al., An ancestral nuclear protein assembly: crystal structure of the Methanopyrus kandleri histone. Protein Sci, 2001. **10**(10): p. 2002-7.

20. Bell, S.D. and S.P. Jackson, Mechanism of autoregulation by an archaeal transcriptional repressor. J Biol Chem, 2000. **275**(41): p. 31624-9.

21. Napoli, A., et al., An Lrp-like protein of the hyperthermophilic archaeon Sulfolobus solfataricus which binds to its own promoter. J Bacteriol, 1999. **181**(5): p. 1474-80.

22. Vierke, G., et al., A novel archaeal transcriptional regulator of heat shock response. J Biol Chem, 2003. **278**(1): p. 18-26.

23. Lie, T.J., G.E. Wood, and J.A. Leigh, Regulation of nif expression in Methanococcus maripaludis: roles of the euryarchaeal repressor NrpR, 2-oxoglutarate, and two operators. J Biol Chem, 2005. **280**(7): p. 5236-41.

24. Lie, T.J. and J.A. Leigh, A novel repressor of nif and glnA expression in the methanogenic archaeon Methanococcus maripaludis. Mol Microbiol, 2003. **47**(1): p. 235-46.

25. Bell, S.D., et al., The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. Science, 2002. **296**(5565): p. 148-51.

26. Struhl, K., Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell, 1999. **98**(1): p. 1-4.

27. Fuda, N.J., M.B. Ardehali, and J.T. Lis, Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature, 2009. **461**(7261): p. 186-92.

28. Zeitlinger, J., et al., RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat Genet, 2007. **39**(12): p. 1512-6.

29. Hendrix, D.A., et al., Promoter elements associated with RNA Pol II stalling in the Drosophila embryo. Proc Natl Acad Sci U S A, 2008. **105**(22): p. 7762-7.

30. Madan Babu, M. and S.A. Teichmann, Evolution of transcription factors and the gene regulatory network in Escherichia coli. Nucleic Acids Res, 2003. **31**(4): p. 1234-44.

31. Bentley, S.D., et al., Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature, 2002. **417**(6885): p. 141-7.

32. Vaquerizas, J.M., et al., A census of human transcription factors: function, expression and evolution. Nat Rev Genet, 2009. **10**(4): p. 252-63.

33. Riechmann, J.L., et al., Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science, 2000. **290**(5499): p. 2105-10.

34. Buchler, N.E., U. Gerland, and T. Hwa, On schemes of combinatorial transcription logic. Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5136-41.

35. Sikorski, T.W. and S. Buratowski, The basal initiation machinery: beyond the general transcription factors. Curr Opin Cell Biol, 2009. **21**(3): p. 344-51.

36. D'Alessio, J.A., K.J. Wright, and R. Tjian, Shifting players and paradigms in cell-specific transcription. Mol Cell, 2009. **36**(6): p. 924-31.

37. Weake, V.M. and J.L. Workman, Inducible gene expression: diverse regulatory mechanisms. Nat Rev Genet. **11**(6): p. 426-37.

38. Kim, J., G. de Haan, and D.J. Shapiro, DNA bending between upstream activator sequences increases transcriptional synergy. Biochem Biophys Res Commun, 1996. **226**(3): p. 638-44.

39. Austin, R.J. and M.D. Biggin, A domain of the even-skipped protein represses transcription by preventing TFIID binding to a promoter: repression by cooperative blocking. Mol Cell Biol, 1995. **15**(9): p. 4683-93.

40. Wang, H. and D.J. Stillman, Transcriptional repression in Saccharomyces cerevisiae by a SIN3-LexA fusion protein. Mol Cell Biol, 1993. **13**(3): p. 1805-14.

41. Chen, G., P.H. Nguyen, and A.J. Courey, A role for Groucho tetramerization in transcriptional repression. Mol Cell Biol, 1998. **18**(12): p. 7259-68.

42. Sutrias-Grau, M. and D.N. Arnosti, CtBP contributes quantitatively to Knirps repression activity in an NAD binding-dependent manner. Mol Cell Biol, 2004. **24**(13): p. 5953-66.

43. Zamir, I., J. Zhang, and M.A. Lazar, Stoichiometric and steric principles governing repression by nuclear hormone receptors. Genes Dev, 1997. **11**(7): p. 835-46.

44. Burke, L.J. and A. Baniahmad, Co-repressors 2000. FASEB J, 2000. **14**(13): p. 1876-88.

45. Perissi, V., et al., Deconstructing repression: evolving models of co-repressor action. Nat Rev Genet. **11**(2): p. 109-23.

46. Kim, J.H., et al., CtBP represses p300-mediated transcriptional activation by direct association with its bromodomain. Nat Struct Mol Biol, 2005. **12**(5): p. 423-8.

47. Gromoller, A. and N. Lehming, Srb7p is a physical and physiological target of Tup1p. EMBO J, 2000. **19**(24): p. 6845-52.

48. Lin, X., et al., Opposed regulation of corepressor CtBP by SUMOylation and PDZ binding. Mol Cell, 2003. **11**(5): p. 1389-96.

49. Zhao, L.J., et al., Acetylation by p300 regulates nuclear localization and function of the transcriptional corepressor CtBP2. J Biol Chem, 2006. **281**(7): p. 4183-9.

50. Hong, S.H. and M.L. Privalsky, The SMRT corepressor is regulated by a MEK-1 kinase pathway: inhibition of corepressor function is associated with SMRT phosphorylation and nuclear export. Mol Cell Biol, 2000. **20**(17): p. 6612-25.

51. Poznic, M., Retinoblastoma protein: a central processing unit. J Biosci, 2009. **34**(2): p. 305-12.

52. Choi, C.Y., et al., Phosphorylation by the DHIPK2 protein kinase modulates the corepressor activity of Groucho. J Biol Chem, 2005. **280**(22): p. 21427-36.

53. Nuthall, H.N., et al., Role for Hes1-induced phosphorylation in Groucho-mediated transcriptional repression. Mol Cell Biol, 2002. **22**(2): p. 389-99.

54. Hasson, P., et al., EGFR signaling attenuates Groucho-dependent repression to antagonize Notch transcriptional output. Nat Genet, 2005. **37**(1): p. 101-5.

55. Hasson, P. and Z. Paroush, Crosstalk between the EGFR and other signalling pathways at the level of the global transcriptional corepressor Groucho/TLE. Br J Cancer, 2006. **94**(6): p. 771-5.

56. Ju, B.G., et al., Activating the PARP-1 sensor component of the groucho/ TLE1 corepressor complex mediates a CaMKinase IIdelta-dependent neurogenic gene activation pathway. Cell, 2004. **119**(6): p. 815-29.

57. Schaeper, U., et al., Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation. Proc Natl Acad Sci U S A, 1995. **92**(23): p. 10467-71.

58. Zhang, Q., D.W. Piston, and R.H. Goodman, Regulation of corepressor function by nuclear NADH. Science, 2002. **295**(5561): p. 1895-7.

59. Arnosti, D.N., et al., The gap protein knirps mediates both quenching and direct repression in the Drosophila embryo. EMBO J, 1996. **15**(14): p. 3659-66.

60. Barolo, S. and M. Levine, hairy mediates dominant repression in the Drosophila embryo. EMBO J, 1997. **16**(10): p. 2883-91.

61. Ip, Y.T., et al., The dorsal morphogen is a sequence-specific DNA-binding protein that interacts with a long-range repression element in Drosophila. Cell, 1991. **64**(2): p. 439-46.

62. Sawada, S., et al., A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. Cell, 1994. **77**(6): p. 917-29.

63. Siu, G., et al., A transcriptional silencer controls the developmental expression of the CD4 gene. EMBO J, 1994. **13**(15): p. 3570-9.

64. Schoenherr, C.J. and D.J. Anderson, The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science, 1995. **267**(5202): p. 1360-3.

65. Studer, M., et al., Role of a conserved retinoic acid response element in rhombomere restriction of Hoxb-1. Science, 1994. **265**(5179): p. 1728-32.

66. Weintraub, S.J., et al., Mechanism of active transcriptional repression by the retinoblastoma protein. Nature, 1995. **375**(6534): p. 812-5.

67. Gray, S. and M. Levine, Transcriptional repression in development. Curr Opin Cell Biol, 1996. **8**(3): p. 358-64.

68. Gray, S., P. Szymanski, and M. Levine, Short-range repression permits multiple enhancers to function autonomously within a complex promoter. Genes Dev, 1994. **8**(15): p. 1829-38.

69. Gray, S. and M. Levine, Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in Drosophila. Genes Dev, 1996. **10**(6): p. 700-10.

70. Davis, R.L. and D.L. Turner, Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning. Oncogene, 2001. **20**(58): p. 8342-57.

71. Van Doren, M., et al., Negative regulation of proneural gene activity: hairy is a direct transcriptional repressor of achaete. Genes Dev, 1994. **8**(22): p. 2729-42.

72. Ohsako, S., et al., Hairy function as a DNA-binding helix-loop-helix repressor of Drosophila sensory organ formation. Genes Dev, 1994. **8**(22): p. 2743-55.

73. Nibu, Y., H. Zhang, and M. Levine, Local action of long-range repressors in the Drosophila embryo. EMBO J, 2001. **20**(9): p. 2246-53.

74. Paroush, Z., et al., Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell, 1994. **79**(5): p. 805-15.

75. Fisher, A.L., S. Ohsako, and M. Caudy, The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. Mol Cell Biol, 1996. **16**(6): p. 2670-7.

76. Rosenberg, M.I. and S.M. Parkhurst, Drosophila Sir2 is required for heterochromatic silencing and by euchromatic Hairy/E(Spl) bHLH repressors in segmentation and sex determination. Cell, 2002. **109**(4): p. 447-58.

77. Poortinga, G., M. Watanabe, and S.M. Parkhurst, Drosophila CtBP: a Hairy-interacting protein required for embryonic segmentation and hairy-mediated transcriptional repression. EMBO J, 1998. **17**(7): p. 2067-78.

78. Bianchi-Frias, D., et al., Hairy transcriptional repression targets and cofactor recruitment in Drosophila. PLoS Biol, 2004. **2**(7): p. E178.

79. Zhang, H. and M. Levine, Groucho and dCtBP mediate separate pathways of transcriptional repression in the Drosophila embryo. Proc Natl Acad Sci U S A, 1999. **96**(2): p. 535-40.

80. Ish-Horowicz, D. and S.M. Pinchin, Pattern abnormalities induced by ectopic expression of the Drosophila gene hairy are associated with repression of ftz transcription. Cell, 1987. **51**(3): p. 405-15.

81. Yu, Y. and L. Pick, Non-periodic cues generate seven ftz stripes in the Drosophila embryo. Mech Dev, 1995. **50**(2-3): p. 163-75.

82. Zhou, D., et al., Mechanisms underlying hypoxia tolerance in Drosophila melanogaster: hairy as a metabolic switch. PLoS Genet, 2008. **4**(10): p. e1000221.

83. Oro, A.E., et al., The Drosophila gene knirps-related is a member of the steroid-receptor gene superfamily. Nature, 1988. **336**(6198): p. 493-6.

84. Nibu, Y., H. Zhang, and M. Levine, Interaction of short-range repressors with Drosophila CtBP in the embryo. Science, 1998. **280**(5360): p. 101-4.

85. Kosman, D. and S. Small, Concentration-dependent patterning by an ectopic expression domain of the Drosophila gap gene knirps. Development, 1997. **124**(7): p. 1343-54.

86. Clyde, D.E., et al., A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. Nature, 2003. **426**(6968): p. 849-53.

87. Lunde, K., et al., The knirps and knirps-related genes organize development of the second wing vein in Drosophila. Development, 1998. **125**(21): p. 4145-54.

88. Chen, C.K., et al., The transcription factors KNIRPS and KNIRPS RELATED control cell migration and branch morphogenesis during Drosophila tracheal development. Development, 1998. **125**(24): p. 4959-68.

89. Struffi, P., et al., Quantitative contributions of CtBP-dependent and -independent repression activities of Knirps. Development, 2004. **131**(10): p. 2419-29.

90. Struffi, P. and D.N. Arnosti, Functional interaction between the Drosophila knirps short range transcriptional repressor and RPD3 histone deacetylase. J Biol Chem, 2005. **280**(49): p. 40757-65.

91. Courey, A.J. and S. Jia, Transcriptional repression: the long and the short of it. Genes Dev, 2001. **15**(21): p. 2786-96.

92. Payankaulam, S. and D.N. Arnosti, Groucho corepressor functions as a cofactor for the Knirps short-range transcriptional repressor. Proc Natl Acad Sci U S A, 2009. **106**(41): p. 17314-9.

93. Chen, G. and A.J. Courey, Groucho/TLE family proteins and transcriptional repression. Gene, 2000. **249**(1-2): p. 1-16.

94. Hartley, D.A., A. Preiss, and S. Artavanis-Tsakonas, A deduced gene product from the Drosophila neurogenic locus, enhancer of split, shows homology to mammalian G-protein beta subunit. Cell, 1988. **55**(5): p. 785-95.

95. Dehni, G., et al., TLE expression correlates with mouse embryonic segmentation, neurogenesis, and epithelial determination. Mech Dev, 1995. **53**(3): p. 369-81.

96. Flores-Saaib, R.D., S. Jia, and A.J. Courey, Activation and repression by the C-terminal domain of Dorsal. Development, 2001. **128**(10): p. 1869-79.

97. Jimenez, G., Z. Paroush, and D. Ish-Horowicz, Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. Genes Dev, 1997. **11**(22): p. 3072-82.

98. Aronson, B.D., et al., Groucho-dependent and -independent repression activities of Runt domain proteins. Mol Cell Biol, 1997. **17**(9): p. 5581-7.

99. Hasson, P., et al., Brinker requires two corepressors for maximal and versatile repression in Dpp signalling. EMBO J, 2001. **20**(20): p. 5725-36.

100. Chen, G., et al., A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in Drosophila development. Genes Dev, 1999. **13**(17): p. 2218-30.

101. Flores-Saaib, R.D. and A.J. Courey, Analysis of Groucho-histone interactions suggests mechanistic similarities between Groucho- and Tup1-mediated repression. Nucleic Acids Res, 2000. **28**(21): p. 4189-96.

102. Martinez, C.A. and D.N. Arnosti, Spreading of a corepressor linked to action of long-range repressor hairy. Mol Cell Biol, 2008. **28**(8): p. 2792-802.

103. Andrioli, L.P., et al., Groucho-dependent repression by sloppy-paired 1 differentially positions anterior pair-rule stripes in the Drosophila embryo. Dev Biol, 2004. **276**(2): p. 541-51.

104. Schaeper, U., et al., Interaction between a cellular protein that binds to the C-terminal region of adenovirus E1A (CtBP) and a novel cellular protein is disrupted by E1A through a conserved PLDLS motif. J Biol Chem, 1998. **273**(15): p. 8549-52.

105. Shi, Y., et al., Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature, 2003. **422**(6933): p. 735-8.

106. Shi, Y., et al., Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell, 2004. **119**(7): p. 941-53.

107. Zeitlinger, J., et al., Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests integration of diverse patterning processes in the Drosophila embryo. Genes Dev, 2007. **21**(4): p. 385-90.

108. Schwartz, Y.B., et al., Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet, 2006. **38**(6): p. 700-5.

109. Papp, B. and J. Muller, Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev, 2006. **20**(15): p. 2041-54.

110. Strahl-Bolsinger, S., et al., SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. Genes Dev, 1997. **11**(1): p. 83-93.

111. Schwartz, Y.B. and V. Pirrotta, Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet, 2007. **8**(1): p. 9-22.

112. Comet, I., et al., PRE-mediated bypass of two Su(Hw) insulators targets PcG proteins to a downstream promoter. Dev Cell, 2006. **11**(1): p. 117-24.

Chapter II
Long- and Short-Range Transcriptional Repressors induce Distinct Chromatin States on Repressed Genes[3]

Abstract

Transcriptional repression is essential for establishing precise patterns of gene expression during development. Repressors governing early *Drosophila* segmentation can be classified into short- or long-range factors based on their ranges of action, acting either locally to quench adjacent activators, or broadly to silence an entire locus. Paradoxically, these repressors recruit common corepressors, Groucho and dCtBP, despite their different ranges of repression. To reveal the mechanisms underlying these two distinct modes of repression, we performed chromatin analysis, using the prototypical long-range repressor Hairy and the short-range repressor Knirps. Chromatin immunoprecipitation and micrococcal nuclease mapping studies reveal that Knirps causes local changes of histone density and acetylation, and the inhibition of activator recruitment, without affecting the recruitment of basal transcriptional machinery. In contrast, Hairy induces wide-spread histone deacetylation and inhibits the recruitment of basal machinery without inducing chromatin compaction. Our study provides unprecedented mechanistic information into short- and long-range repression at the chromatin level, and suggests that the transcriptional corepressors can be differentially deployed to mediate chromatin changes in a context-dependent manner.

Key words

chromatin/ Hairy/Knirps/repression/transcription

[3] Chapter II is presented in the form of a manuscript to be submitted for publication as: Li, L.M., and Arnosti, D.N Long- and short-range transcription repressors induce distinct chromatin states on repressed genes.

Introduction

Eukaryotic development involves the establishment of precise temporal and spatial patterns of gene expression. The segmentation process in *Drosophila* is a paradigm for developmental regulation, and is governed by a cascade of transcription factors, including maternal, gap, pair-rule, and segment-polarity genes [1]. Among these factors, numerous transcriptional repressors play central roles, setting the boundaries for gene expression [2].

A unifying principle emerging from molecular studies of these disparate proteins relates to their ranges of action [3]. Long-range repressors such as Hairy and Dorsal cause dominant inactivation of activators over a range of more than 1000 bp, whereas short-range repressors such as Knirps, Krüppel, and Snail interfere with locally bound activators, leaving untouched the function of more distal enhancers [3, 4]. Short-range repression underlies enhancer autonomy, in which different enhancers work independently of one another within a complex cis-regulatory region [5]. It is not known if the special attributes of long-range repressors are uniquely harnessed to specific gene structures, however.

The different activities of long- and short-range repressors have been proposed to reflect the recruitment of distinct corepressors, namely Groucho and the C-terminal binding protein (CtBP). Groucho, a conserved WD-40 motif protein related to the human TLE and the mouse Grg protein, was initially identified as a corepressor of Hairy and later other repressors such as Deadpan and Engrailed [6, 7]. Groucho is capable of oligomerization, and interacts with histones and the histone deacetylase Rpd3 [8-10]. Based on these observations, it has been proposed that long-range repression involves spreading of the Groucho corepressor along the chromatin [11]. Indeed, Groucho spreading has been observed on a Hairy target reporter gene

[12]. CtBP, originally identified as an E1a-interacting protein in vertebrate cells, was found to bind the Knirps, Krüppel and Snail short-range repressors in *Drosophila* and mediate their activity [13]. CtBP can interact with histone deacetylases and demethylases, suggesting that it acts via chromatin modifications [14, 15].

While extensive data points to Groucho's and CtBP's involvement in numerous repression systems, recent studies have shown that the distinction in repression range is not a simple function of recruiting different cofactors. For example, the short-range repressor Knirps utilizes both CtBP and Groucho to repress *even skipped* [16]. This suggests that different types of repressors can recruit the same cofactors to effect short- or long-range repression. Mechanistic insights into how these processes differ have been lacking: long-range repression has been associated with Groucho cofactor spreading, but almost no mechanistic information is available for short-range repressors, which have been alternately suggested to act via competition for the binding sites with the activators, interfering with the basal transcriptional machinery, or quenching the interaction between activators and the basal transcriptional machinery [4, 17].

In this study, we investigated chromatin mechanisms underlying the functional differences between long- and short-range repression. Using the long-range repressor Hairy and short-range repressor Knirps as models, we provide evidence that the two classes of repressors have distinct effects on chromatin structure, as well as activator and basal transcriptional machinery recruitment, despite the common corepressors they recruit. These results suggest that these transcriptional corepressors cause context-specific (in this case,

repressor-specific) chromatin changes, and also provide insights into how developmental transcription factors regulate RNA polymerase II.

Results

Local and global repression

Hairy directly represses *fushi tarazu* (*ftz*), a secondary pair-rule gene expressed in the blastoderm embryo in a seven-stripe pattern [18]. ftz is regulated by both regionally acting gap genes, as well as the *hairy* pair-rule gene [19]. Two dense clusters of peaks for Caudal, Hunchback, Knirps, Giant, Huckebein, Kruppel and Tailless transcription factors mapping on the 5' regulatory region of the ftz locus represent presumptive regulatory elements, in addition to the promoter-proximal Zebra, and the stripe 1+5 enhancer located 3' of the gene [20-22](Figure 2-1A). Hairy is found to bind *in vivo* to several of these regions. In the blastoderm embryo, the *ftz* gene is active in some nuclei and repressed in others by Hairy. In order to obtain a population of entirely repressed nuclei for chromatin studies, we overexpressed Hairy protein in embryos using a heat shock driver, which results in complete repression of *ftz* (Figure 2-1A). Hairy repression of *ftz* relies on the Groucho corepressor, as a mutant version of Hairy that does not bind to Groucho failed to repress *ftz* (Figure 2-1E, F).

Interestingly, a titration of heat shock induction times resulted in a non-uniform, progressive loss of specific *ftz* stripes, with stripe 4 being the most sensitive and stripe 1+5 the least (Figure 2-1C; differential repression quantified in Supplemental Table 2-1). This result points to the intriguing possibility that Hairy can act locally, at least very transiently, although the end result of Hairy repression is complete silencing of all enhancer elements. This

asynchronous repression of the *ftz* locus also suggests that Hairy-mediated long-range repression does not act solely by direct targeting the basal promoter, a previous model for this class of repressor, because this mechanism would cause uniform inhibition of stripe elements [17].

Similar to *ftz*, the pair-rule gene *even-skipped* (*eve*) is also expressed in a seven-stripe pattern and is regulated by multiple modular enhancers. *eve* is a well-characterized target of the short-range repressor Knirps, which sets posterior boundaries of *eve* stripe 3 and 4 and anterior borders of *eve* stripe 6 and 7 [23, 24]. After substantial overexpression of Knirps (20 minute heat shock induction), the repressor is able to repress all of the *eve* stripe enhancers except for stripe 5 enhancer (Figure 2-1B). When the induction is titrated, Knirps represses individual enhancers in a step-wise manner, with the most sensitive enhancers down-regulated earliest, at a low dose of Knirps [25]. Together, these experiments indicate that Hairy can initially act locally, but ultimately in a globally dominant fashion, while Knirps acts in a restricted manner (Figure 2-1D).

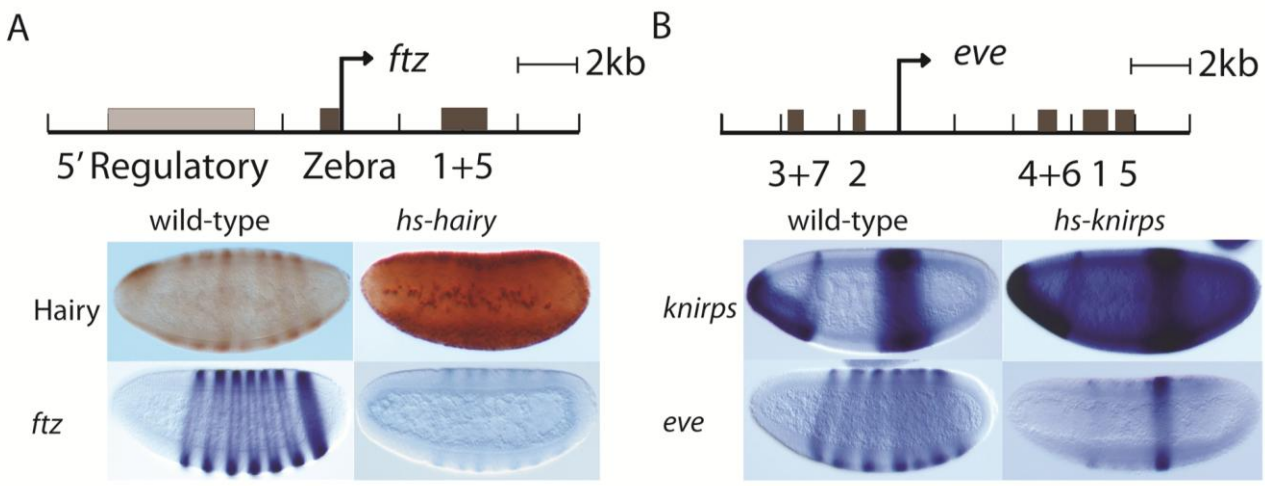


Figure 2-1. Repression of *ftz* by Hairy and *eve* by Knirps. (A) Overexpression of Hairy long-range repressor by 20 minute heat shock represses *ftz* expression. Hairy protein (top panel) and *ftz* mRNA (lower panel) expression pattern in wild-type and *hs-hairy* transgenic embryos after 20 minute heat shock are shown. (B) Substantial heat shock induction (20 minute) of Knirps represses all *eve* enhancers except the stripe 5 enhancer. *knirps* (top panel) and *eve* (lower panel) mRNA expression pattern in wild-type and *hs-knirps* transgenic embryos after 20 minute heat shock are shown.

Figure 2-1 continued

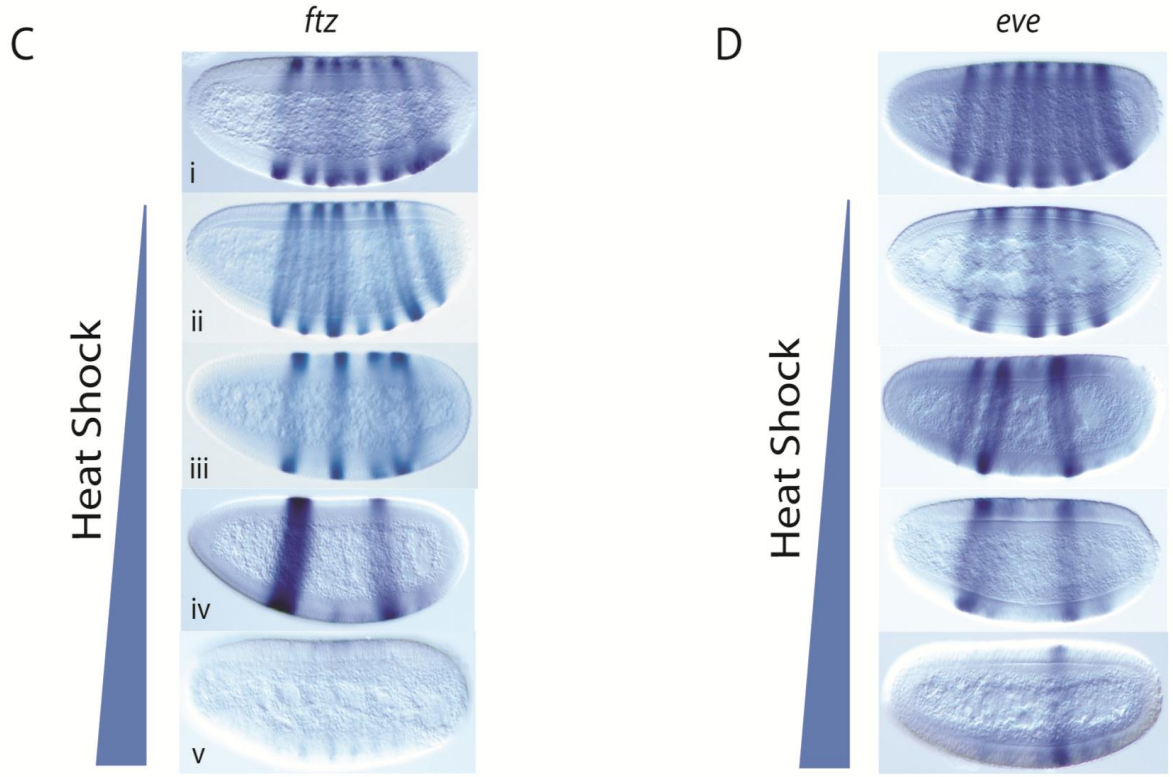
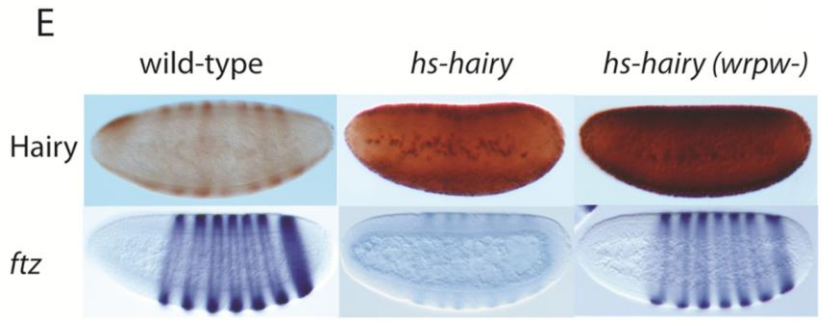


Figure 2-1. Repression of *ftz* by Hairy and *eve* by Knirps. (C) Overexpression of the long-range Hairy repressor by titrated heat shock results in progressive repression of *ftz* enhancers shown by *in situ* hybridization. The order of repression is stripe 4, stripe 2+7, stripe 3+6, stripe 1+5 (from top to bottom). Details of the heat shock titration experiment are listed in Supplemental Table 2-1. (D) Heat shock titration of Knirps represses *eve* in a step-wise manner, with stripe 3+7 enhancer the most sensitive and stripe 1 enhancer the least sensitive (from top to bottom). Details of the heat shock titration experiments have been published and can be found in Struffi et al. 2004.

Figure 2-1 continued



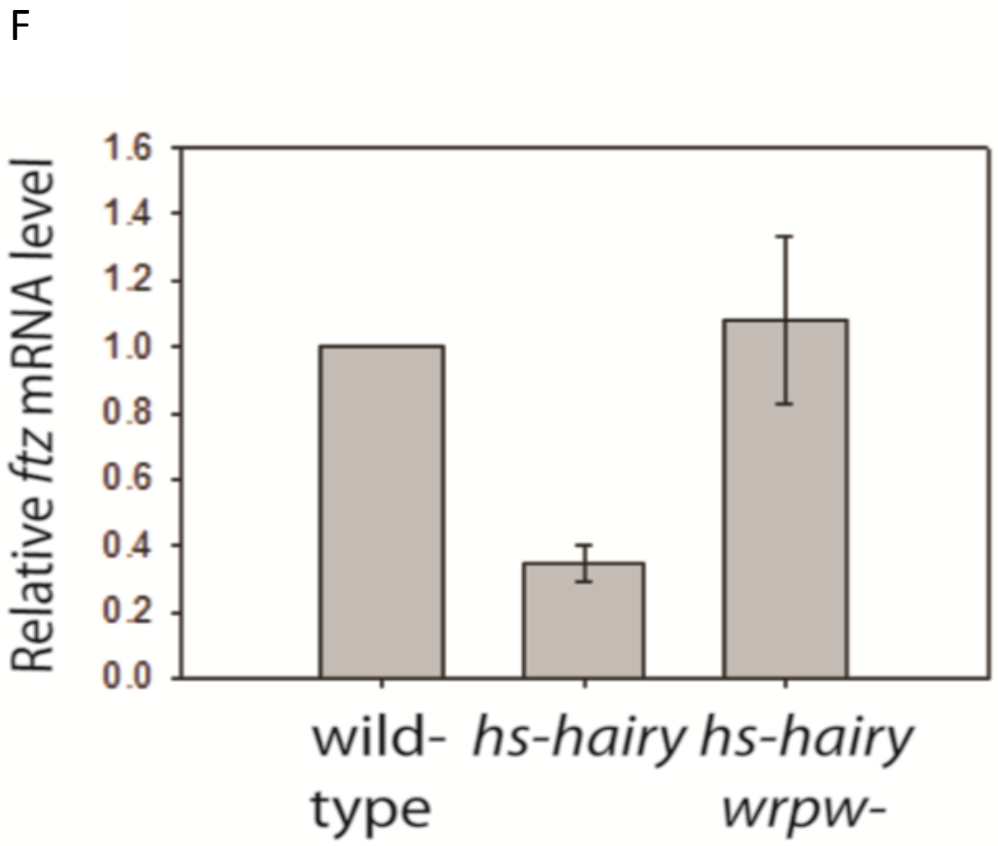


Figure 2-1. Repression of *ftz* by Hairy and *eve* by Knirps. (E) Groucho binding deficient Hairy fails to repress *ftz*. Hairy protein (top panel) and *ftz* mRNA (lower panel) expression pattern in wild-type , *hs-hairy* or *hs-hairy (wrpw-)* (Groucho binding deficiency version of Hairy) transgenic embryos after 20 minute heat shock are shown. (F) *ftz* mRNA levels in wild-type, *hs-hairy* or *hs-hairy (wrpw-)* embryos are quantified by real-time PCR. *ftz* mRNA level is lower only in embryos containing the *hs-hairy* transgene. Heat shock alone, in the absence of the *hs-hairy* or *hs-knirps* transgenes, had no effect on *ftz* and *eve*.

Hairy and Knirps differentially affect chromatin structure

To compare the effects of repression by Hairy and Knirps, we studied chromatin changes associated with repression of *ftz* and *eve* using chromatin immunoprecipitation. We observed little or no significant change of histone H3 occupancy at most regions sampled throughout the *ftz* locus after Hairy overexpression with one possible exception at -5 kbp (Figure 2-2A; Supplemental Table 2-2). In contrast, Knirps repression of *eve* resulted in increased histone H3 density in two of the three regions corresponding to the Knirps-sensitive enhancers, namely stripe 4+6 and stripe 2 (Figure 2-2B; Supplemental Table 2-2). Little change was noted in the promoter region, transcribed region, or the stripe 1 and 5 enhancers, which are not readily repressed by Knirps. An apparent increase in histone H3 density on the repressed stripe 3+7 enhancer, although of low statistical significance, correlates with other alterations common to repressed enhancers, noted below (Figure 2-2B; Supplemental Table 2-2).

To provide a more detailed picture of chromatin structure, we adapted a micrococcal nuclease (MNase) mapping protocol used in yeast and cultured cells for *Drosophila* embryos (Sekinger et al., 2005; Petesch and Lis, 2008; Li and Arnosti, 2010). MNase mapping showed that Hairy repression had little effect on chromatin accessibility throughout the *ftz* locus (Figure 2-2C; Supplemental Table 2-2), whereas Knirps induced a significant increase in MNase insensitivity specifically at *eve* stripe 3+7, 2, 4+6 enhancers, and a minor increase in stripe 1 protection (Figure 2-2D; Supplemental Table 2-2). The promoter and the *eve* stripe 5 enhancer were not much changed, mirroring the patterns noted for overall histone H3 occupancy. The changes noted for the *eve* locus appear to be specific; as Knirps did not induce any change of a non-targeted site on the third chromosome. Hairy also had no effect at this locus (Figure 2-2C).

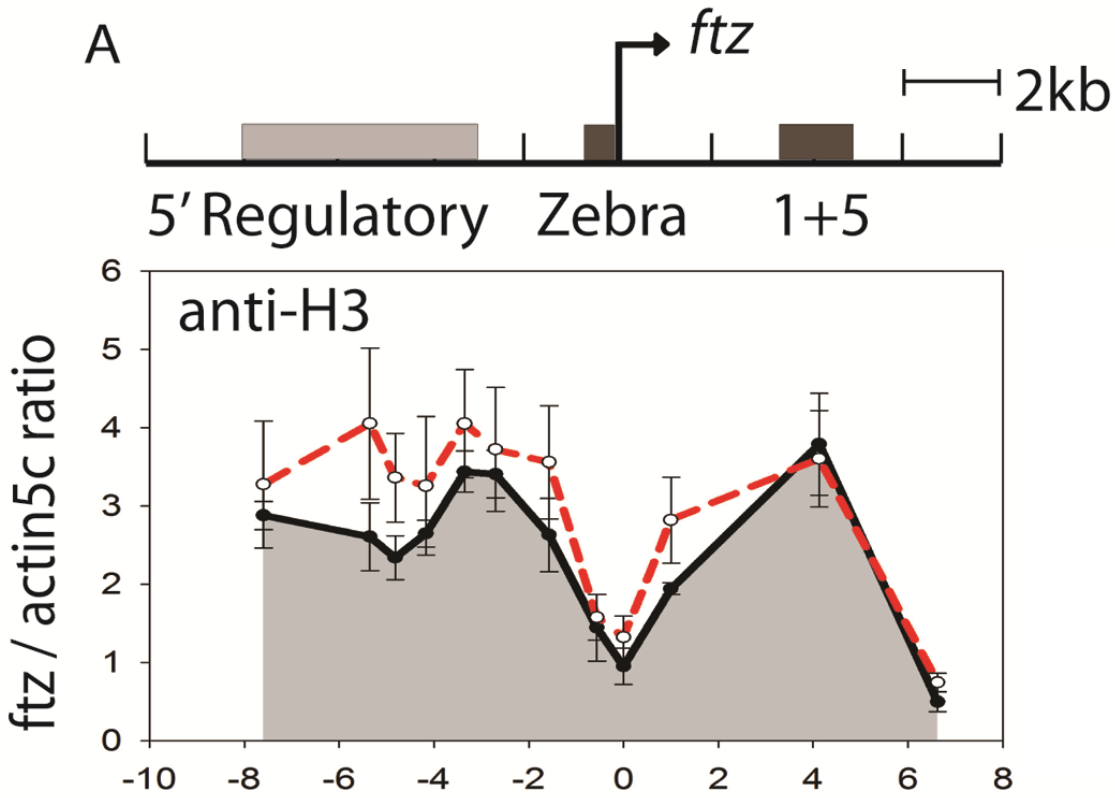


Figure 2-2. Hairy and Knirps-mediated transcriptional repression result in differential changes in histone H3 occupancy and micrococcal nuclease sensitivity. (A, B) Histone H3 occupancy was measured by chromatin immunoprecipitation in regions of *ftz* and *eve* before (solid line) and after repression (dashed line). Relative immunoprecipitation signal was obtained by normalization to the *actin5C* promoter region, which was not affected by Hairy or Knirps repression.

Figure 2-2 continued

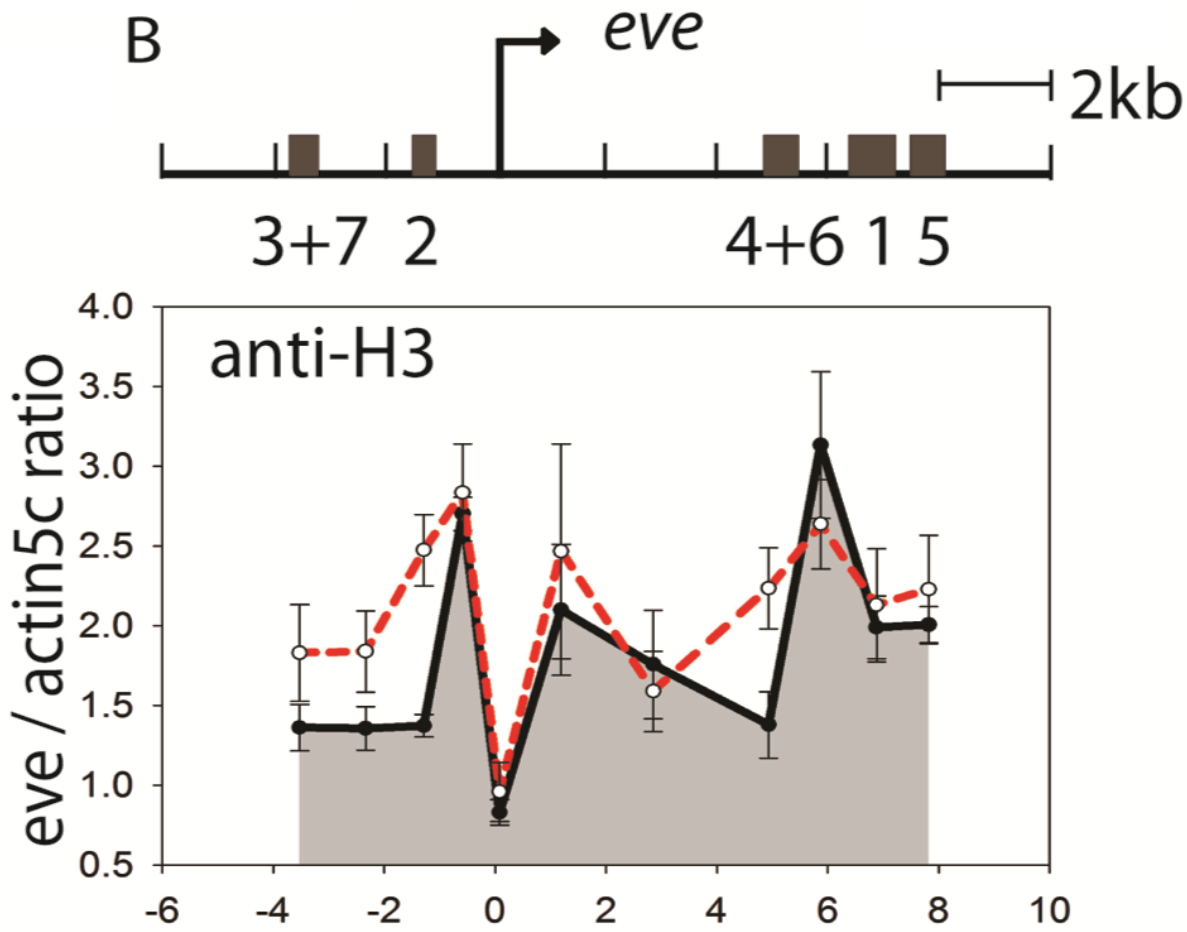


Figure 2-2. Hairy and Knirps-mediated transcriptional repression result in differential changes in histone H3 occupancy and micrococcal nuclease sensitivity. (A, B) Histone H3 occupancy was measured by chromatin immunoprecipitation in regions of *ftz* and *eve* before (solid line) and after repression (dashed line). Relative immunoprecipitation signal was obtained by normalization to the *actin5C* promoter region, which was not affected by Hairy or Knirps repression.

Figure 2-2 continued

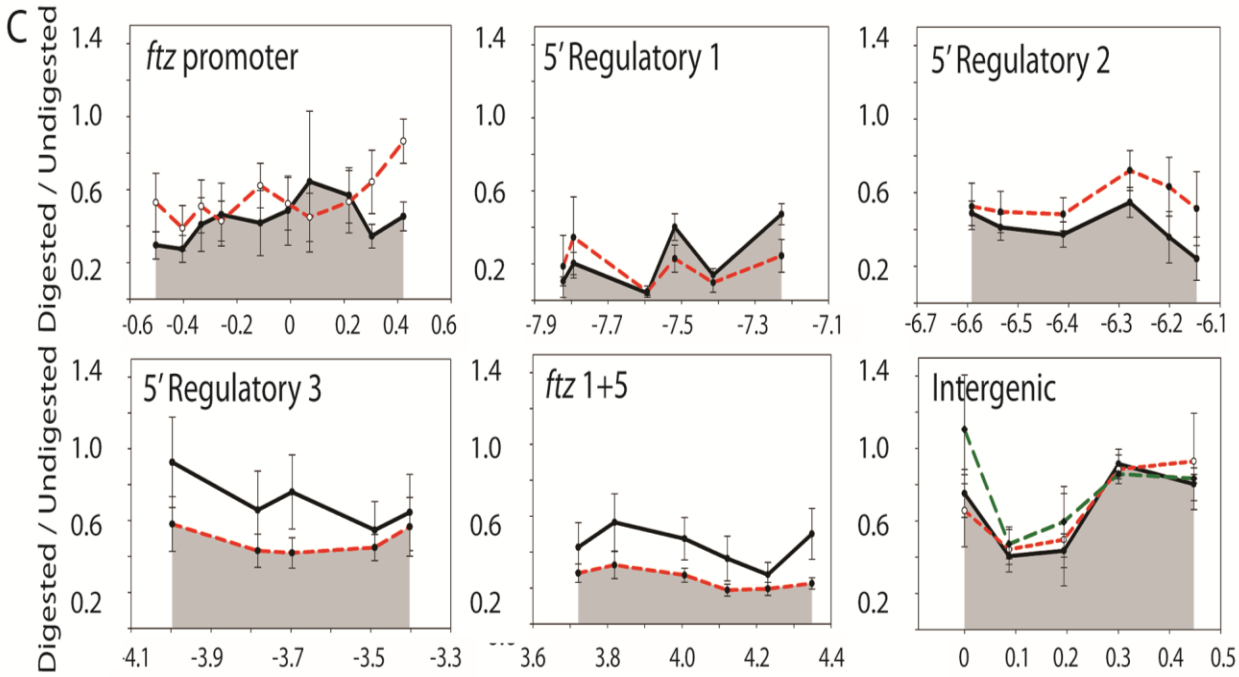


Figure 2-2. Hairy and Knirps-mediated transcriptional repression result in differential changes in histone H3 occupancy and micrococcal nuclease sensitivity. (C, D) MNase sensitivity of the *ftz* and *eve* loci. Hairy induced repression does not significantly change the overall MNase protection pattern in any of the regions tested in the *ftz* locus (solid line before, dashed line after repression). Knirps repression is associated with increased resistance to MNase digestion at the *eve* 3+7, 2 and 4+6 enhancers. Little or no change is observed at the promoter, stripe 1 enhancer and stripe 5 enhancer. The specificity of MNase digestion was also shown by digestion pattern of a 450 bp intergenic region on the 3rd chromosome as shown in (C), for embryos with no repressor overexpression (solid line), Hairy overexpression (long dashed line), and Knirps overexpression (short dashed line). All results in A-D and in following figures represent at least three biological replicates; error bars show standard errors. Areas under the lower plots are shadowed for the clarity of presentation. For this and later figures, the statistical significance of the differences between each pair of points is shown in Supplemental Table 1; for histone H3 occupancy, the p values are < 0.05 for *eve* stripes 2 and 4+6; for *ftz*, no points reached this level of significance. Nonrepressed embryos (lacking the heat shock-inducible repressor gene) were heat shocked to control for any effect of the heat induction itself. In this and subsequent figures, the Y-axis shows the amount of immunoprecipitate as percentage of the input.

Figure 2-2 continued

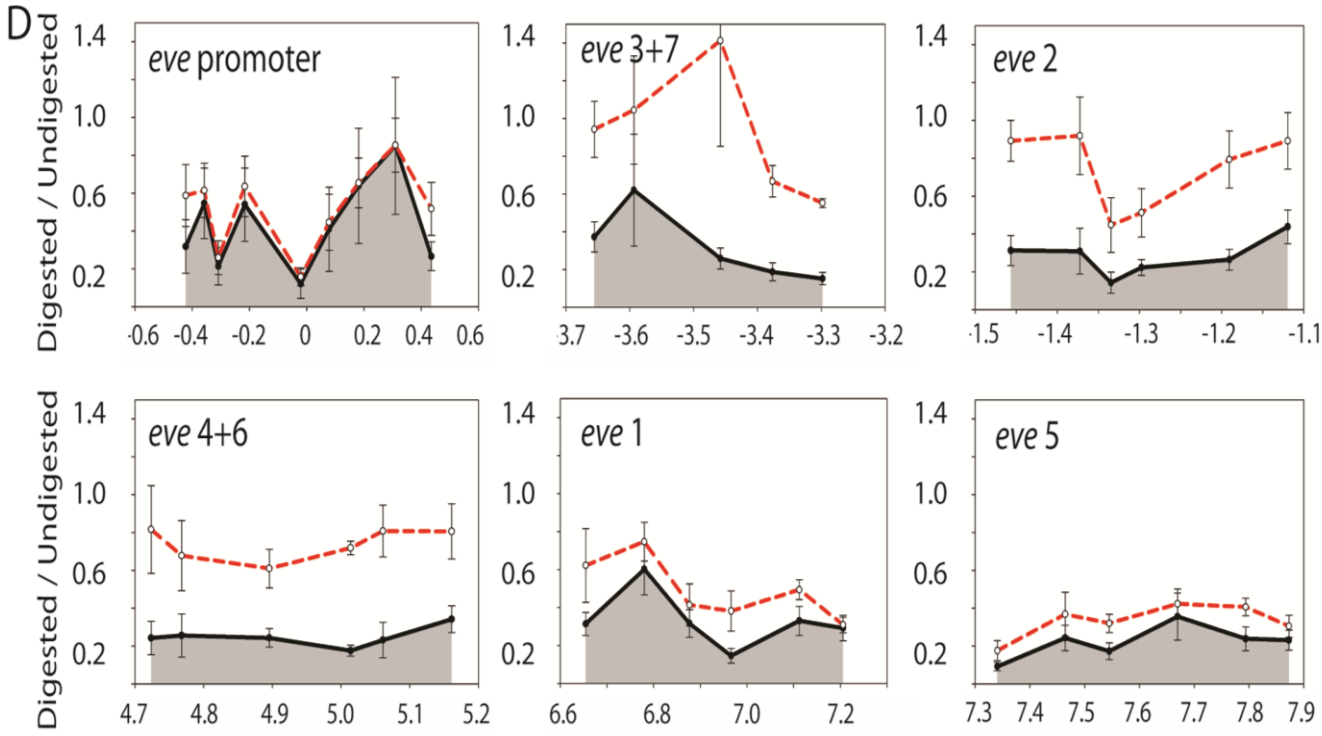


Figure 2-2. Hairy and Knirps-mediated transcriptional repression result in differential changes in histone H3 occupancy and micrococcal nuclease sensitivity. (C, D) MNase sensitivity of the *ftz* and *eve* loci. Hairy induced repression does not significantly change the overall MNase protection pattern in any of the regions tested in the *ftz* locus (solid line before, dashed line after repression). Knirps repression is associated with increased resistance to MNase digestion at the *eve* 3+7, 2 and 4+6 enhancers. Little or no change is observed at the promoter, stripe 1 enhancer and stripe 5 enhancer. The specificity of MNase digestion was also shown by digestion pattern of a 450 bp intergenic region on the 3rd chromosome as shown in (C), for embryos with no repressor overexpression (solid line), Hairy overexpression (long dashed line), and Knirps overexpression (short dashed line). All results in A-D and in following figures represent at least three biological replicates; error bars show standard errors. Areas under the lower plots are shadowed for the clarity of presentation. For this and later figures, the statistical significance of the differences between each pair of points is shown in Supplemental Table 1; for histone H3 occupancy, the p values are < 0.05 for *eve* stripes 2 and 4+6; for *ftz*, no points reached this level of significance. Nonrepressed embryos (lacking the heat shock-inducible repressor gene) were heat shocked to control for any effect of the heat induction itself. In this and subsequent figures, the Y-axis shows the amount of immunoprecipitate as percentage of the input.

The consistent results from overall histone H3 density and MNase mapping suggest that Hairy-mediated long-range repression does not involve compaction of chromatin on the *ftz* locus. Previous studies showed that repression of *slp1* locus by Runt, another Groucho-dependent repressor, does not involve changes in H3 density, thus it is possible that Groucho-dependent long-range repression does not involve chromatin compaction in general [26]. In contrast, these findings indicate that repression by Knirps is associated with an increase in the histone density of targeted enhancer regions, which may result either from Knirps recruitment of factors that mediate chromatin condensation, or the blocking of proteins responsible for loosening of chromatin.

Hairy mediates wide-spread histone deacetylation, whereas Knirps only causes local histone deacetylation

Histone acetylation is dynamically regulated on actively transcribed genes in eukaryotes, with histone acetylation generally correlated with active loci [27]. The histone deacetylase Rpd3 is a component of both Hairy and Knirps corepressor complexes [9, 28]. We assayed histone acetylation levels across the *eve* and *ftz* genes before and after repression. Hairy repression resulted in wide-spread histone H4 deacetylation throughout the *ftz* locus (Figure 2-3A; Supplemental Table 2-2). Using anti H3-acetylation antibodies, similar widespread H3 deacetylation was also noted (data not shown). This distributed effect on the *ftz* locus correlates with the prior observations that Hairy-mediated long-range repression might involve a Groucho-mediated “spreading” mechanism [12]. By this means, Rpd3 may be delivered to extensive areas of a gene.

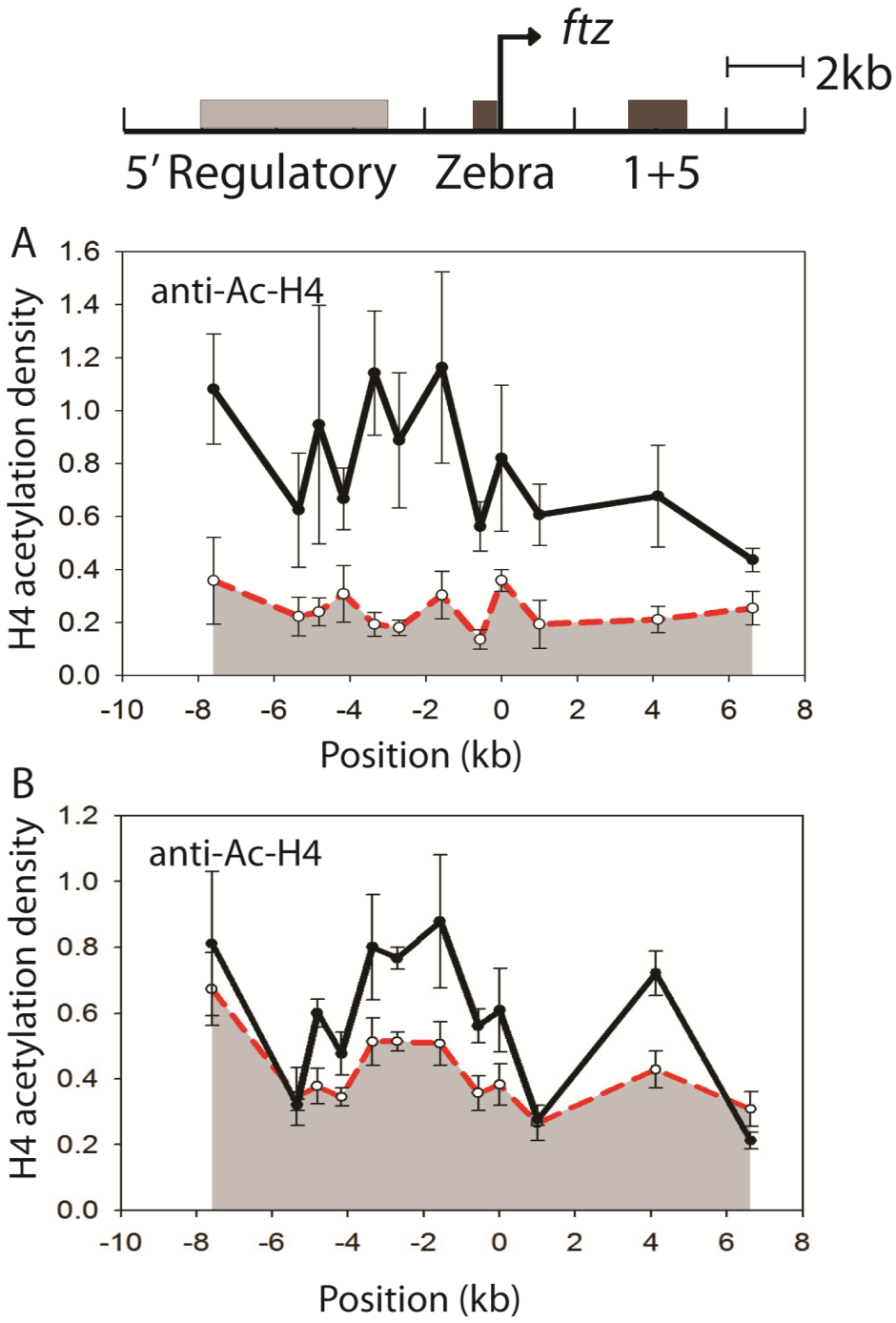


Figure 2-3. Hairy induces global, while Knirps induces local changes in histone acetylation levels. For both repressors, H3K27 tri-methylation are little changed overall. (A, C) H4K-5,8,12,16 acetylation was assayed by chromatin immunoprecipitation on the *ftz* and *eve* loci before (solid line) and after (dashed line) repression by Hairy and Knirps respectively. Significant reduction in H4 acetylation was observed at all loci tested in *ftz*, while reductions in H4 acetylation on *eve* was limited to repressed enhancers. (B) H4 deacetylation is localized around Hairy binding sites after brief 5 minute heatshock induction.

Figure 2-3 continued

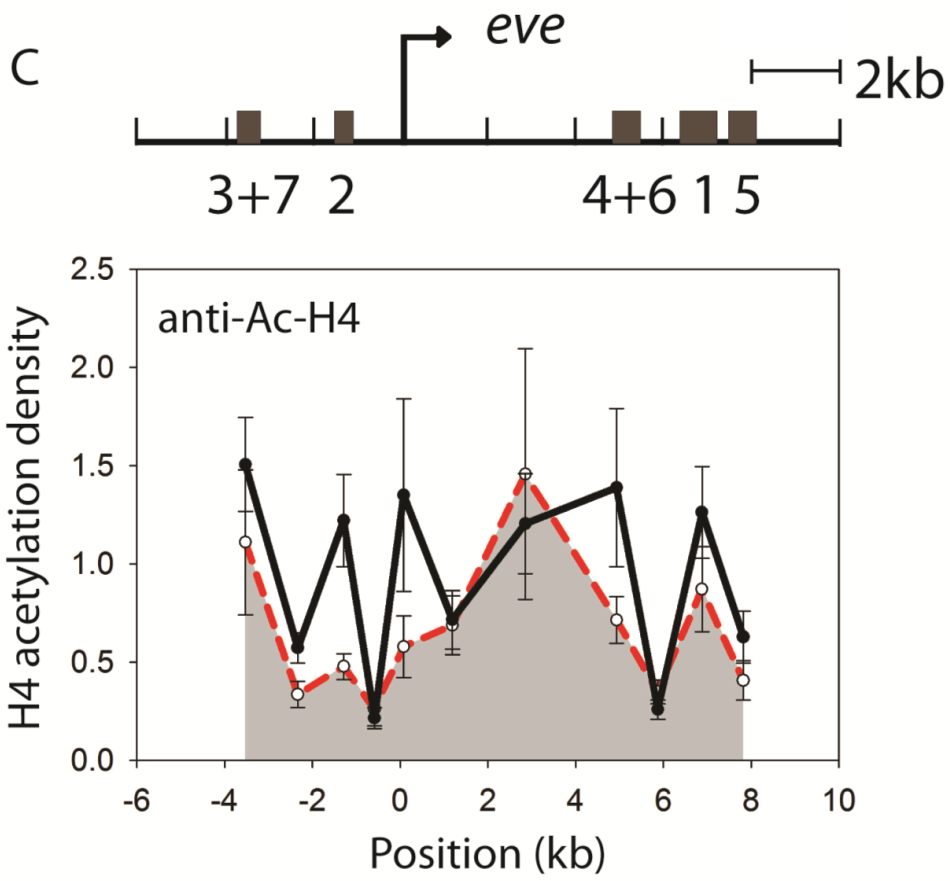


Figure 2-3. Hairy induces global, while Knirps induces local changes in histone acetylation levels. For both repressors, H3K27 tri-methylation are little changed overall. (A, C) H4K-5,8,12,16 acetylation was assayed by chromatin immunoprecipitation on the *ftz* and *eve* loci before (solid line) and after (dashed line) repression by Hairy and Knirps respectively. Significant reduction in H4 acetylation was observed at all loci tested in *ftz*, while reductions in H4 acetylation on *eve* was limited to repressed enhancers. (B) H4 deacetylation is localized around Hairy binding sites after brief 5 minute heatshock induction.

Figure 2-3 continued

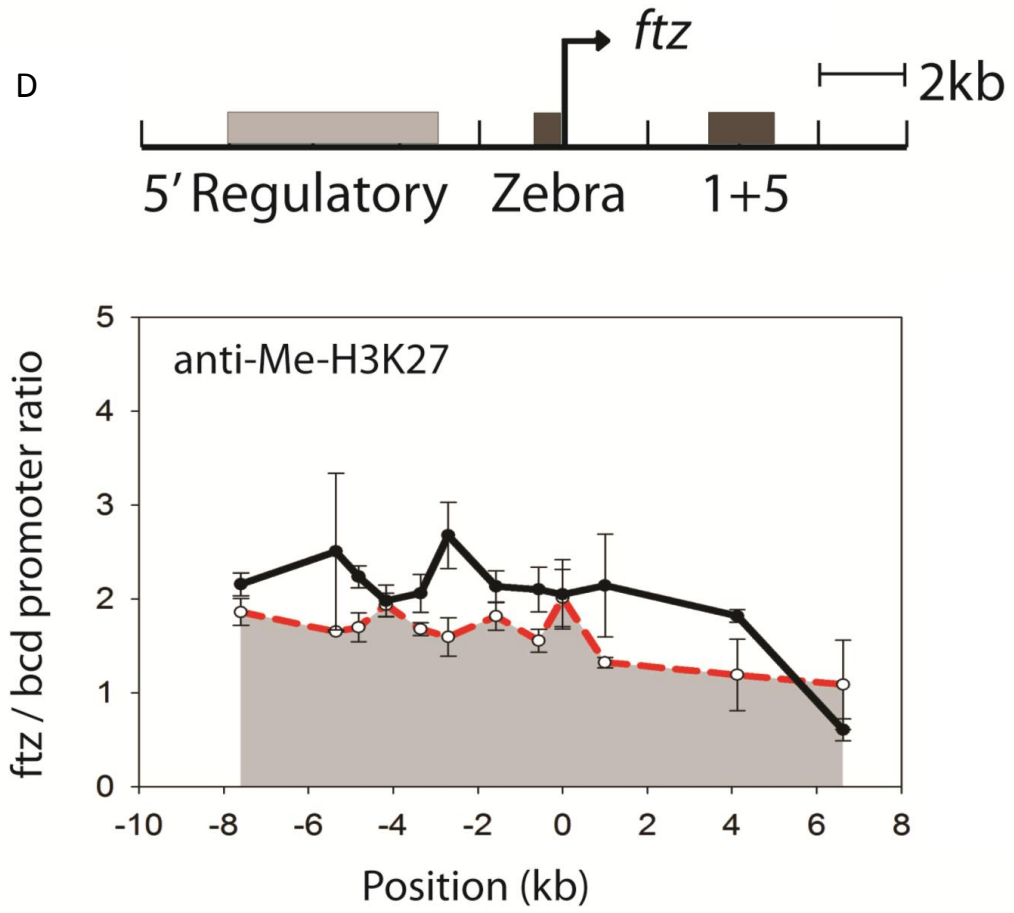


Figure 2-3. Hairy induces global, while Knirps induces local changes in histone acetylation levels. For both repressors, H3K27 tri-methylation are little changed overall. (D, E) Tri-methylated H3 lysine 27 levels are around 5-10 times higher on *eve* and *ftz* than *tubulin*, *actin5C* and *hsp70*. In a repressed state, the levels are only modestly affected in portions of the *ftz* gene, and no significant reduction was observed for the *eve* gene.

Figure 2-3 continued

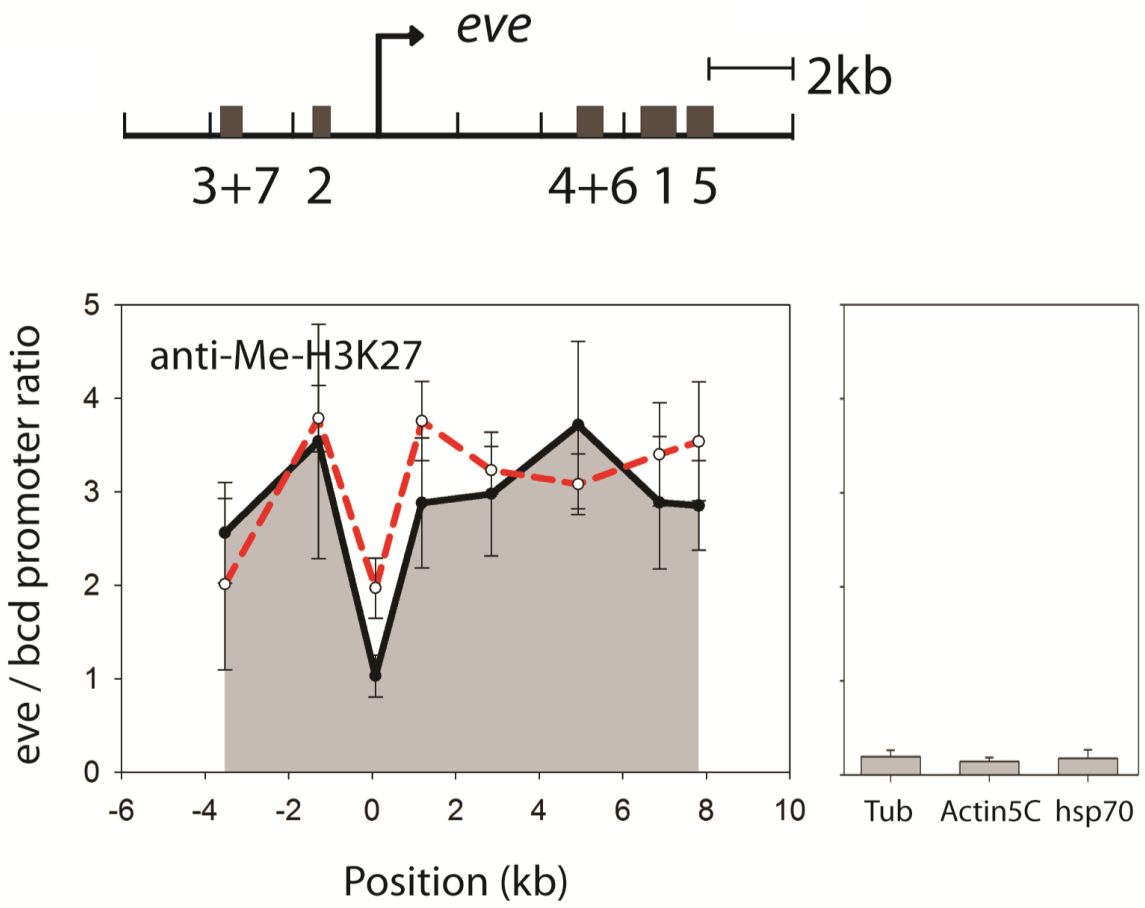


Figure 2-3. Hairy induces global, while Knirps induces local changes in histone acetylation levels. For both repressors, H3K27 tri-methylation are little changed overall. (D, E) Tri-methylated H3 lysine 27 levels are around 5-10 times higher on *eve* and *ftz* than *tubulin*, *actin5C* and *hsp70*. In a repressed state, the levels are only modestly affected in portions of the *ftz* gene, and no significant reduction was observed for the *eve* gene.

To test whether a spreading of histone deacetylation might correlate with the successive inhibition of *ftz* enhancers we noted in , we investigated histone acetylation levels across *ftz* after a short 5 minute heat shock followed by immediate fixing. In this setting, deacetylation was mostly concentrated around Hairy binding sites, consistent with a spreading action of this repressor (Figure 2-3B).

A different picture emerged from studies of Knirps acting on *eve*. Here, repression led to selective decreases in H3 and H4 acetylation levels, concentrated over the *eve* stripe 4+6 and stripe 2 enhancers, with lesser decreases noted at stripe 3+7 and stripe 1 enhancers (Figure 2-3C; Supplemental Table 2-2). A local change in acetylation was also noted near the transcriptional initiation site, but not immediately 5' and 3' of this area. The reductions in histone acetylation levels seen on both *eve* and *ftz* are consistent with Hairy and Knirps recruiting deacetylases to their target genes. However, it is striking that the broad deacetylation mediated by Hairy on *ftz* is not associated with changes in histone density or resistance to nuclease accessibility, whereas increased histone density is associated with Knirps repression on *eve*. It is possible that in addition to inducing deacetylation, Knirps triggers additional histone modifications or interacts with nucleosome remodeling complexes to increase the density of histones at the enhancers.

Repression does not involve changes in H3 lysine 27 trimethylation

The H3 lysine 27 trimethylation mark is established by the Enhancer of Zeste methyltransferase, and is an essential repression mark for establishing and maintaining polycomb-mediated long-term developmental regulation of *Hox* genes [29, 30]. Therefore, we

investigated whether repression mediated by Hairy and Knirps in the blastoderm embryo involves modulation of H3K27 trimethylation level. Before induction of the repressors, we detected robust levels of H3K27 trimethylation on the *ftz* and *eve* loci, at least 5-10 times higher than levels on housekeeping genes such as *tubulin* and *actin5C* (Figure 2-3E). However, only modest changes were observed on the *ftz* locus in response to Hairy. Similarly, the levels of H3 lysine 27 tri-methylation on the *eve* locus were not significantly affected by Knirps (Figure 2-3D, 2-3E; Supplemental Table 2-2). Although, a change in H3 lysine 27 tri-methylation is not required for repression by these early developmental repressors, the enriched levels of H3K27 trimethylation observed on the *eve* and *ftz* genes do suggest the possible involvement of Polycomb-group proteins in the regulation of these pair-rule genes; this chromatin modification might prime these loci for later, permanent inactivation.

Differential effects of Hairy and Knirps on activator recruitment

Our previous studies indicated that Hairy can effectively repress a reporter gene without displacing the activators [12]. We sought to test whether this was the case on an endogenous gene, *ftz*, by examining occupancy by Caudal, a transcription factor that also activates *eve*. Caudal activates the posterior stripes of both *ftz* and *eve*, and we found that Caudal binds the *ftz* 5' regulatory region and the promoter-proximal *Zebra* element, consistent with a recent global study [31-33]. Repression of the locus by Hairy did not affect Caudal binding pattern (Figure 2-4A; Supplemental Table 2-2), similar to the results obtained with a Hairy regulated reporter gene [12]. In contrast, Knirps repression decreased Caudal occupancy specifically at *eve* 3+7 and 4+6

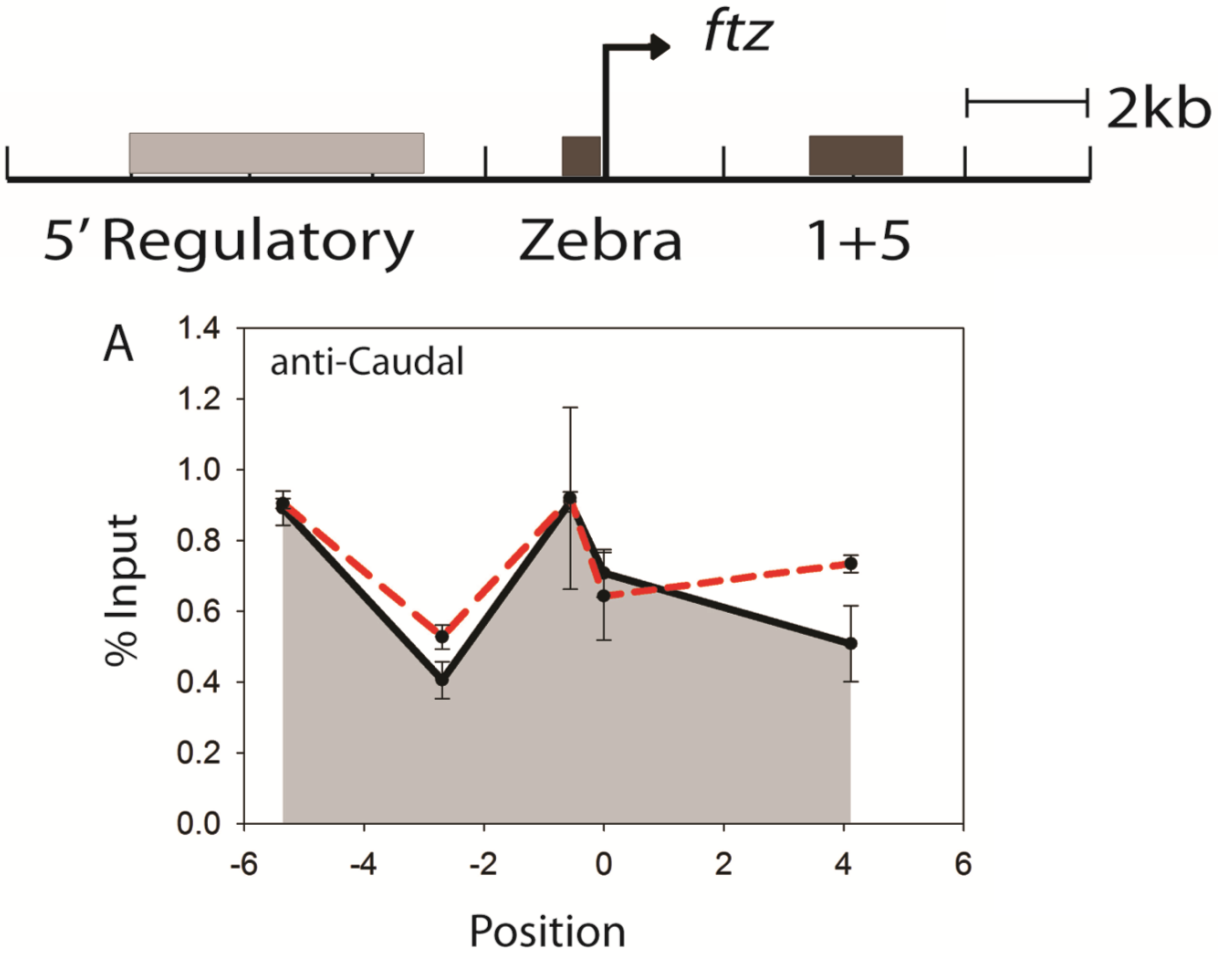


Figure 2-4. Activator occupancy of *ftz* and *eve* before (solid line) and after (dashed line) repression. (A) Caudal protein occupancy of *ftz* measured by chromatin immunoprecipitation. Caudal activator levels do not decrease on the *ftz* locus in reponse to repression by Hairy.

Figure 2-4 continued

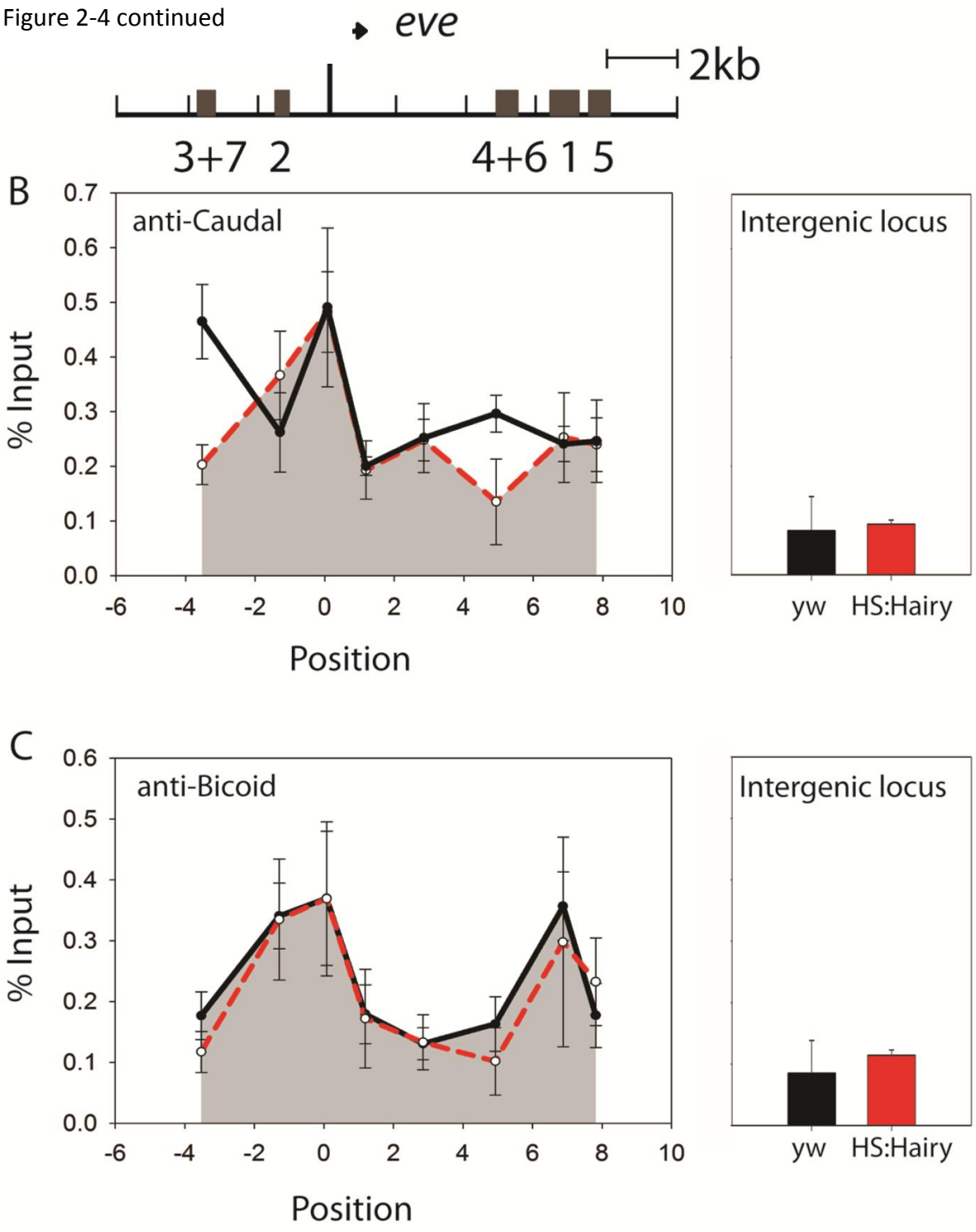


Figure 2-4. Activator occupancy of *ftz* and *eve* before (solid line) and after (dashed line) repression. (B) Caudal occupancy decreases at *eve* stripe 3+7 and 4+6 enhancers after repression by Knirps. (C) Bicoid occupancy of *eve* is unchanged after Knirps repression. No Bicoid binding was detected on the *ftz* locus.

in occupancy down to near-baseline levels. This decrease is not an effect of global decrease of Caudal occupancy, as the Caudal binding peak at the *eve* promoter is not affected. A similar decrease in Caudal occupancy was also observed on a *hunchback* enhancer after repression by Knirps (data not shown). Interestingly, Bicoid occupancy of the *eve* stripe 2 and stripe 1 enhancers was not altered by Knirps, although these enhancers were repressed (Figure 2-4C). It is possible that Knirps utilizes multiple mechanisms for repression: on the *eve* stripe 3+7 enhancer, the high density of Knirps binding sites may permit competitive binding with activators such as Caudal [24, 34]. Interestingly, this enhancer exhibited a relatively modest decrease in histone acetylation upon repression, thus acetylation levels may not directly reflect the abundance of recruited repressors. In the case of the stripe 2 enhancer, which exhibited more significant changes in H4 acetylation, Knirps might mainly repress by altering chromatin structure via deacetylation, without reducing Bicoid activator binding.

Distinct effects on RNA polymerase II by long- and short-range repressors

New insights have suggested many developmental genes, including those regulated by short-range repressors such as Snail, feature RNA polymerase paused in the promoter region even in their inactive state, suggesting post-recruitment levels of regulation [35]. We analyzed components of the core machinery before and after repression by Hairy and Knirps. Upon Hairy repression, a marked decrease of RNA polymerase II occupancy was observed at the *ftz* locus. The same trend is observed for the pre-initiating, initiating, and elongation forms of RNA polymerase II (Fig 2-5A-C). These results suggest that Hairy directly or

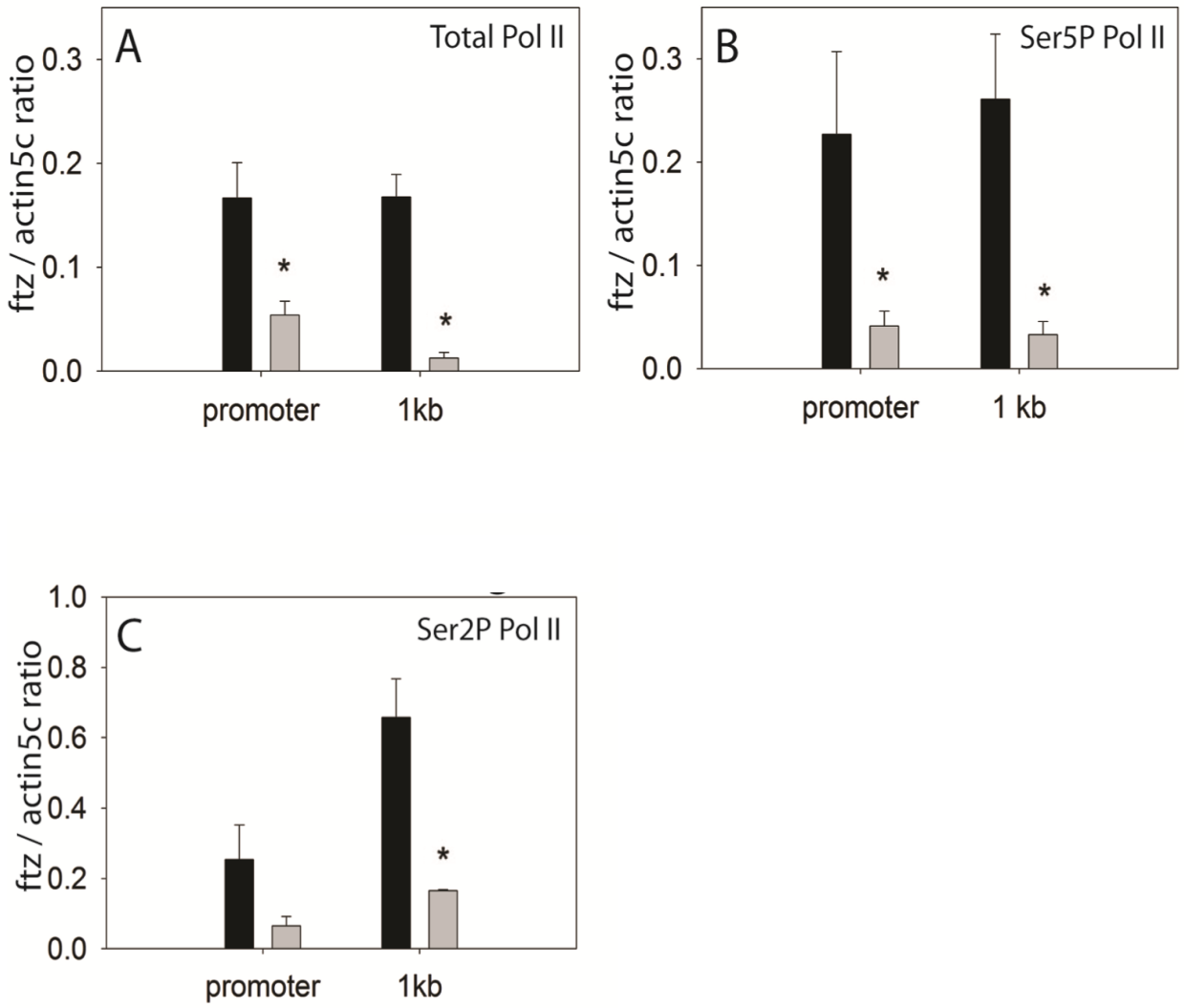


Figure 2-5. Pol II and TATA binding protein occupancy at *ftz* and *eve* promoter and transcribed regions. (A-C) ChIPs were performed using antibodies against unphosphorylated, initiating (Ser5P) and elongating (Ser2P) Pol II on *ftz* before (black bars) and after repression (gray bars). Strong decreases in all forms of Pol II were noted. (Asterisk denotes where $p<0.05$).

Figure 2-5 continued

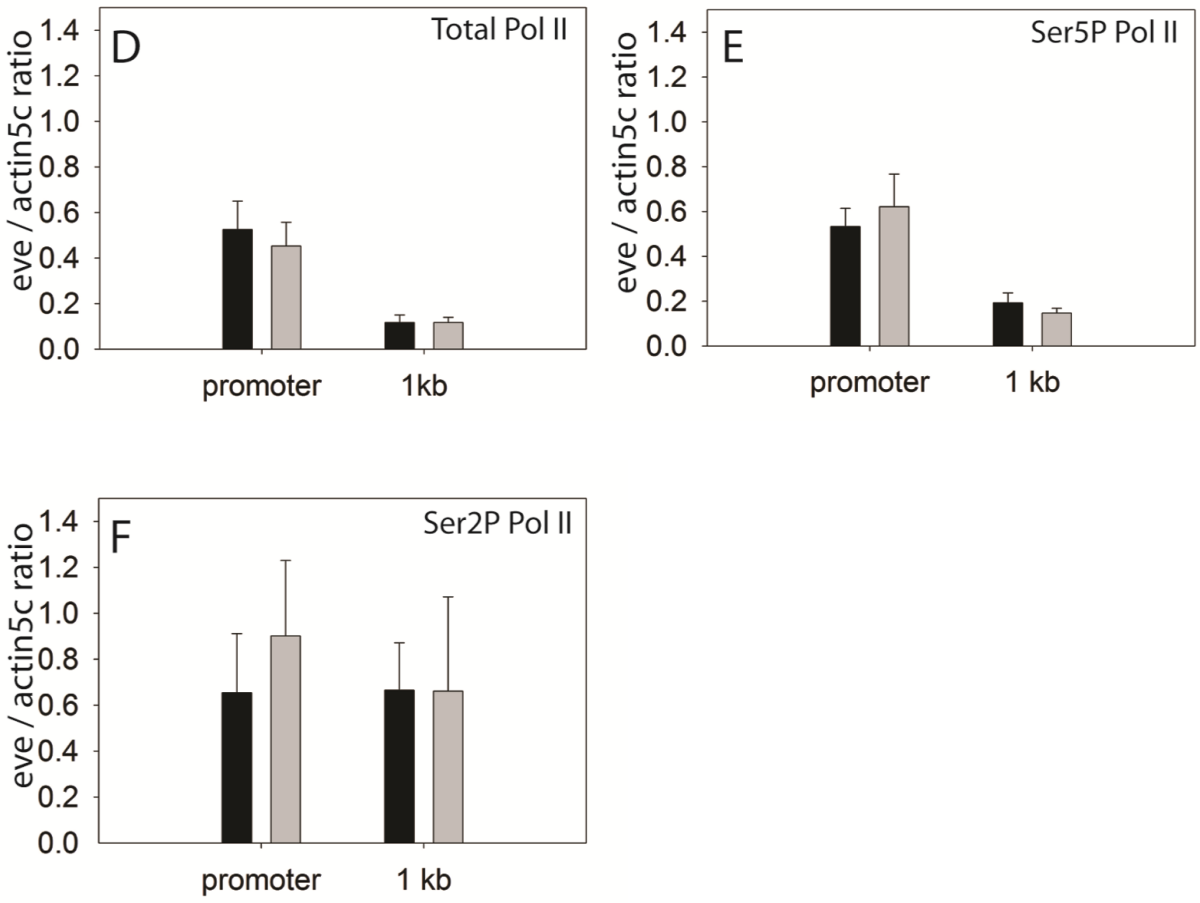


Figure 2-5. Pol II and TATA binding protein occupancy at *ftz* and *eve* promoter and transcribed regions. (D-F) A similar analysis of *eve* did not show significant changes in Pol II occupancy after Knirps repression.

Figure 2-5 continued

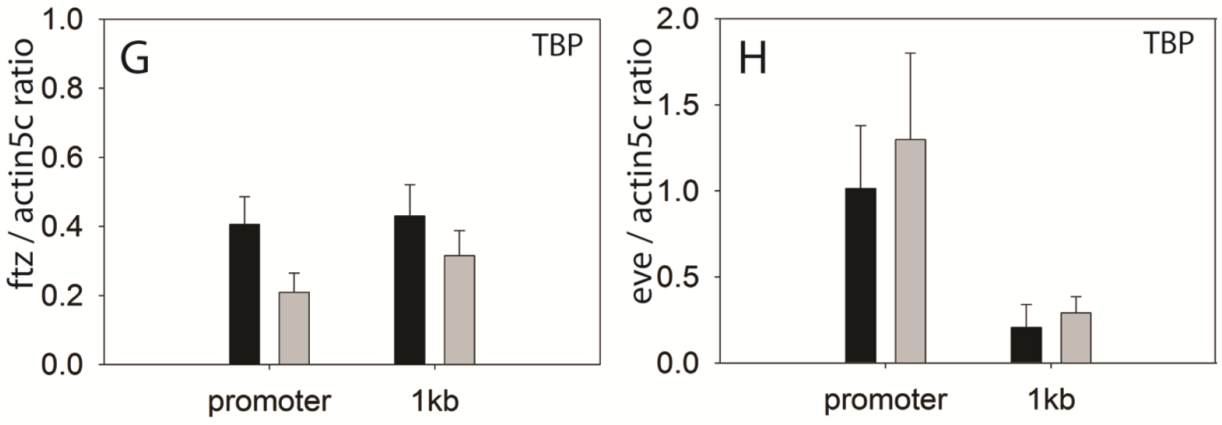


Figure 2-5. Pol II and TATA binding protein occupancy at *ftz* and *eve* promoter and transcribed regions. (G, H) TBP occupancy of the *ftz* promoter decreased after repression by Hairy, but no significant changes were noted for TBP occupancy of the *eve* gene after Knirps repression.

indirectly blocks recruitment of RNA polymerase II. Similar decreases were noted with levels of TATA-box binding protein (TBP) at the promoter (Fig. 2-5G).

In contrast, induction of Knirps does not change Pol II occupancy at the *eve* transcription unit, even under condition where most enhancers were repressed (Fig. 2-5D-F). (Under conditions tested here, over three-quarters of the embryos have shut down expression of all but stripe 1 and/or 5). Similarly, TBP occupancy remains at the comparable level before and after Knirps repression (Fig. 2-5H). The constant level of RNA polymerase on the *eve* transcription unit was a surprise, in light of the sharp reduction in mRNA production, as measured by in situ hybridization. However, there is precedence for this effect: Runt repression of *Slp1* appears to act through elongation control, which causes no change of the concentration of RNA polymerase II on *slp1* [26]. Knirps may produce a similar effect by inducing a slower transit rate of RNA polymerase II on the repressed *eve* locus. Similar observations have been made at the *hsp70* gene upon depletion of elongation factors such as Spt6 or Paf1 [36, 37].

Discussion

Our study provides molecular insights into general mechanisms mediated by two major classes of repressors vital for early *Drosophila* development. Long-range repression mediated by Hairy is dependent upon Groucho, whose extensive interaction with large tracts of chromatin have led to a "spreading" model of repression [11]. Widespread histone deacetylation is consistent with a picture of a globally repressed locus, but interestingly, rapid kinetic "snapshots" of partially repressed states indicate that chromatin modification is nucleated around Hairy binding sites, with initial inhibition of separate cis-elements. The

progressive repression of *ftz* by Hairy suggests the Hairy-mediated repression is not simply directly targeting the basal promoter, because in this case all *ftz* enhancers would be repressed at the same time. Pol II and TBP occupancy is inhibited when the whole locus is repressed, which may either reflect a complete loss of cis-regulatory signaling with the silencing of all the enhancers or Groucho-mediated chromatin modification may in the end also target the basal promoter blocking access by the preinitiation complex.

The global mechanism of Hairy-mediated repression resembles the deacetylated chromatin domain mediated by Sir2 in both the telomeres and mating type loci in yeast, although the latter system farms a stable state that is propagated for generations [38]. Hairy in contrast appears to generate a relatively labile state that is reset as soon as the nucleating transcription factor is turned over (Martinez & Arnosti, 2008 and data not shown). This transient long-range repression might fit into the role of Hairy during early development, which allows for effective yet dynamic regulation of its targets. Polycomb protein-mediated long-range repression also involves similar wide-spread chromatin modifications, including H3K27, K9 and K20 methylation chromatin marks [39]. We did note an enrichment of H3K27 methylation of *ftz* (as well as *eve*), but Hairy-mediated repression did not involve dynamic modulation of this chromatin mark.

There is extensive information about cis-regulatory requirements for short-range repressors such as Knirps, but we have had limited insight into the molecular machanisms of these transcriptional regulators, other than indications that chromatin deacetylation should play a role, above and beyond the simple competitive binding models first proposed for those repressors [4, 28]. Our results strongly point to important local changes occurring on cis-

regulatory elements that coincide with silencing of gene function, including increases in histone density, local resistance of chromatin to nuclease attack, deacetylation and loss of activator binding. This picture strongly supports a local 'quenching' model of short-range repression, and is at odds with the alternative 'hitchhiking' model that posited long-range, albeit transient, interactions with the basal transcription machinery [4, 17]. At the same time, the lack of change in Pol II occupancy indicates that these local modifications can influence polymerase elongation, a feature that needs further investigation. Other transcriptional repressors have also been linked to repression of Pol II elongation [26, 37].

The local increase in histone density may indicate an inhibition of nucleosome disassembly or recruitment of cofactors that deposit nucleosomes at the enhancer regions by Knirps. These mechanistic insights may provide key discoveries to facilitate mathematical modeling of cis-regulatory elements controlled by these short-range repressors. Modeling can in turn provide a basis for understanding the evolution of the developmental enhancers [40].

With the recent demonstration that transcriptional factors considered to be short- and long-range repressors utilize shared cofactors, namely CtBP and Groucho, there has been a question of whether long-range repression is actually functionally distinct from short-range repression [16]. Our study provides evidence that the chromatin states associated with long- and short-range repressors are distinct in several ways. We do not yet know whether the effects seen on *ftz* are observed for all Hairy targets, although the similarity of changes observed with a *lacZ* reporter subject to Hairy repression suggests they are conserved [12]. Similarly, the reproducibility of Knirps-induced changes at different *eve* enhancers indicates that this protein can effect related chromatin changes a cis regulatory modules bound by

different activators. Snail, another short-range repressor, also appears to mediate localized deacetylation and activator displacement, thus this mechanism may be a common feature of this entire class of proteins (Yutaka Nibu, personal communication). The highly divergent activities of Knirps and Hairy demonstrated in this study not only underscore the fact that these proteins do mediate biochemically divergent events, but also raise interesting questions about how similar cofactors can mediate such distinct effects in a context-dependent manner. It is possible that the corepressors adopt distinct conformations when recruited by different repressors, or the corepressor may form distinct complexes with unique activities [41]. In addition to determining how cis- and trans-acting factors affect repression pathways, these mechanistic insights will provide important contextual information for interpretation of genome-wide transcription factor binding and chromatin modifications.

Materials and Methods

Plasmid construction

Transgenic flies carrying inducible Hairy genes were generated using pCasper-hs transformation vector [25]. The genes were created by joining a EcoRI/XbaI fragment containing a Kozak sequence, initiator ATG, and coding sequence for either wildtype or the mutant (WRPW/AAAA) Hairy protein (primer sequences are available upon request) amplified from a pGEX-2T vector containing *hairy* cDNA [6].

In situ hybridization and antibody staining of Drosophila embryos

Embryos were fixed for in situ hybridization and stained using anti-digoxigenin-UTP-label RNA probe to *ftz* or *eve* as previously described [25].

Embryo Collection

Embryos used for chromatin immunoprecipitation and micrococcal nuclease protection experiments were 2-3 hour old, exposed to 20 minute heat shock to induce maximal repression, and allowed no recovery period after heat shock treatment. To control for possible nonspecific effects of heat shock, wild-type embryos without *hs-trangenes* were similarly treated to generate the chromatin profiles of *ftz* and *eve* in the unrepressed state (heat shock alone has no effect on the expression patterns of *eve* or *ftz*; Struffi et al., 2004 and data not shown). Embryos containing either the *hs-hairy* or *hs-knirps* transgene were used to generate the “after repression”chromatin.

Chromatin immunoprecipitation

Heat shocks and chromatin immunoprecipitations were performed as previously described [12], with the exceptions that embryos were sonicated for 20s (60% duty cycle) and

cooled on ice for 30s 15 times; beads were washed twice with the low-salt buffer, twice with the high-salt buffer, and twice with Tris-EDTA. We used the following antibodies: mouse monoclonal IgG (10 ul; Upstate), rabbit anti-H3 (1ul; Abcam), rabbit anti-acetyl H4 (1ul; Upstate); rabbit anti-trimethyl H3K27 (2ul; Abcam); 8WG16 (10ul; Covance; anti-unphosphorylated Pol II CTD); H5 (10ul; Covance; anti-Ser2-phosphorylated CTD); H14 (10ul; Covance; anti-Ser5-phosphorylated CTD); TBP (2ul; Abcam); rabbit anti-mouse IgM (10ul; Abcam); rabbit anti-Bicoid serum (10 ul; gift from M. Biggin); rabbit anti-Caudal serum (10ul; gift from M. Biggin).

Micrococcal nuclease mapping in *Drosophila* embryos

Micrococcal nuclease mapping in *Drosophila* embryos was performed as previously described [42].

Quantitative PCR analysis

The samples from ChIP and MNase mapping were analyzed using real-time PCR (Applied Biosystem 7500). Primer pairs had a Tm in the range of 58-60^{o}C, and amplicons ranged from 50 to 150 bp. Primer sequences are listed in Supplemental Table 3. For ChIP samples, a standard curve was generated by serially diluting input samples to quantify IP samples. For MNase digests, a ratio was calculated between MNase digested and undigested samples. All values used were collected from the linear range of amplification.

Acknowledgments

We thank Xiaoyong Li and Mark Biggin for the Bicoid and Caudal antisera; Zachary Burton and Mark Schroeder for critical reading of the manuscript; Sandhya Payankaulam for

valuable advice on Knirps-related experiments and the manuscript; and other Arnosti lab members for useful discussions. This work is supported by NIH grant GM56976 to D.N.A.

APPENDICES

Appendix A-1

Supplementary Tables for Chapter II

Supplemental Table 2.1 Numbers of embryos within each category after heat shock titration of Hairy protein are summarized. Staining patterns for embryos in each category can be referred to Figure 1C. For each experiment, total n=100 .

Induction(min)	Recovery (min)	i	ii	iii	iv	v
5	0	71	29	0	0	0
10	0	52	21	11	1	15
20	0	6	13	4	3	74
5	30	2	35	26	15	24

Supplemental Table 2.1 p-values for Student's t-test comparing before repression and after repression chromatin immunoprecipitation signals of each data point in Figure 2-2,2-3 and 2-4 (n ≥ 3).

Hairy repression on *ftz*								Zebra	+1	1kb	1+5	
Position	-7593	-5350	-4813	-4171	-3352	-2697	-1570	-562	0	1000	4121	6624
Histone H3 (Figure 2-2A)	0.68	0.21	0.14	0.52	0.43	0.72	0.31	0.81	0.33	0.15	0.84	0.19
Histone H4 acetylation (Figure 2-3A)	0.034	0.058	0.17	0.064	0.007	0.033	0.06	0.013	0.069	0.031	0.057	0.058
Histone H4 acetylation (Figure2- 3B)	0.59	0.41	0.009	0.091	0.13	0.0002	0.11	0.019	0.14	0.87	0.007	0.10
Histone H3 K27 me (Figure 2-3D)	0.19	0.36	0.049	0.85	0.15	0.056	0.23	0.11	0.38	0.94	0.18	0.21
Caudal (Figure 2-4A)		0.80				0.12		0.97	0.67		0.11	

Knirps repression on *eve*	3+7		2		+1	1kb		4+6		1	5
Position	-3522	-2330	-1281	-583	83	1194	2856	4934	5876	6882	7821
Histone H3 (Figure 2-2B)	0.20	0.17	0.009	0.70	0.53	0.65	0.70	0.031	0.99	0.74	0.57
Histone H4 acetylation (Figure 2-3C)	0.39	0.045	0.038	0.66	0.16	0.90	0.15	0.15	0.28	0.29	0.16
Histone H3 K27 me (Figure 2-3D)	0.63		0.86		0.07	0.34	0.74	0.54		0.60	0.44
Caudal (Figure 2-4B)	0.015		0.38		0.96	0.90	0.96	0.11		0.90	0.95
Bicoid (Figure 2-4C)	0.29		0.96		0.99	0.94	0.97	0.42		0.76	0.56

Supplemental Table 2.2: p-values for Student's t-test comparing before repression and after repression Micrococcal Nuclease Digested/Undigested values of each data point in Figure 2C (n ≥ 3).

ftz promoter										
Position	-503	-404	-334	-259	-113	-10	72	219	304	422
p-value	0.29	0.50	0.66	0.86	0.37	0.88	0.61	0.89	0.91	0.048
5' regulatory_2										
Position	-6592	-6535	-6411	-6279	-6201	-6147				
p-value	0.79	0.54	0.35	0.22	0.22	0.27				
5' regulatory_1										
Position	-7822	-7794	-7593	-7519	-7414	-7228				
p-value	0.50	0.49	0.82	0.17	0.52	0.076				
5'_regulatory_3										
Position	-3997	-3783	-3697	-3490	-3403					
p-value	0.27	0.33	0.15	0.59	0.77					
*ftz*_1+5										
Position	3722	3819	4007	4122	4231	4349				
p-value	0.36	0.18	0.13	0.23	0.31	0.082				

Supplemental Table 2.3: p-values for Student's t-test comparing before repression and after repression Micrococcal Nuclease Digested/Undigested values of each data point in Figure 2D (n ≥ 3).

eve 2									
Position	-1456	-1372	-1334	-1297	-1190	-1119			
p-value	0.005	0.041	0.097	0.074	0.016	0.04			
*eve*4+6									
Position	4723	4768	4895	5014	5061	5161			
p-value	0.06	0.099	0.018	0.00002	0.015	0.025			
*eve*3+7									
Position	-3655	-3593	-3458	-3376	-3298				
p-value	0.015	0.34	0.031	0.002	0.00005				
eve 1									
Position	6654	6780	6877	6966	7112	7206			
p-value	0.20	0.41	0.50	0.086	0.10	0.83			
*eve*5									
Position	7340	7464	7545	7669	7794	7873			
p-value	0.21	0.40	0.29	0.65	0.13	0.38			
eve p									
Position	-423	-358	-309	-216	-21	78	182	310	436
p-value	0.26	0.78	0.75	0.72	0.70	0.90	0.97	0.99	0.22

Supplemental Table 2-3. List of primer pairs used in chromatin immunoprecipitation and micrococcal nuclease sensitivity analysis

Primer pairs for chromatin immunoprecipitation				
ftz locus				
Location	FWR	Sequence	Rev	Sequence
-7593	DA 2156	CGAGTTTTTTAAGCGACAATATGTCT	DA 2157	TCCGTCTGTCCATATGAACGAT
-5350	DA 1956	AAGGAAAGTCTGTTTTGGGTAAACA	DA 1957	CCTTGATTCGTCCGTGGAAA
-4813	DA 2158	TCAGCGGAATCGCATTGAC	DA 2159	TCGTGGGACATCAACATCTAATAAG
-4171	DA 1958	CAGGAGCAATTACAGCCTTATCCT	DA 1959	CGAGGTCCTGTGCGTCCTA
-3352	DA 2160	GAAAACCGCCGGAGGAGAT	DA 2161	CTCGGGATTTGTTTGGTTTTCT
-2697	DA 1960	TCCTAATGATCTTGGTCACTTTGG	DA 1961	TTCTGGCTGCTGTTGAAGTCA
-1570	DA 2162	ACATCGGAGAGATGCGAACTG	DA 2163	CCTTACCCCATCCGCATACTC
-562	DA 1962	CCAGGGATCGGACGTAATGT	DA 1963	AAGACTGGCTGCCAGTTGTTC
0	DA 1882	TCTCCGTGCCTGCAAGGACATTT	DA 1883	TAGCTGTAGTGGCTCTGGCTGTTT
1000	DA 2013	TCCAAAACCGACGCATGAA	DA 2014	GTGTAGCCGGCACCACAGT
4121	DA 2164	GATCGTAAAGTTATCCTTTCGCAAA	DA 2165	TGCGCGGACAAAGGTGTT
6624	DA 1964	GGGTTCTGAGGCCATTATGATTT	DA 1965	TCGATTAGGTGCATGTCTTATGTAGA
eve locus				
-3523	DA2205	TTTGTTCCGGGCTCAGAAAT	DA2206	CTCTCGCGGCAGGACACT
-2330	DA2782	TGTAAAAAACTGAACTGGCACTCTTC	DA2783	GTTGGAATGGGTTGAGTAAGTGATT
-1282	DA2207	AGTTGCAGCGTTTCGCTTTC	DA2208	AGTATGGCACAGCGTGTTACCA
-583	DA2784	CAGCAGTTCCCAAATGGTTATG	DA2785	CCATGGCTGCCATGACTTT
84	DA2209	GCAGAGCGCAGCGGTATAA	DA2210	AGGATATTAACGAAGGCAGTTAGTTGT
1194	DA2211	AGCGGGAGCAGGAACTGAT	DA2212	TCACGCCAAGGTGTTCGA
2857	DA2786	CGCGATCTTAGCCGGAATC	DA2787	AGCGCTTAGTGTTTTATGGAGGAT
4935	DA2213	CCAGGCAATTGTCCTTTTTTG	DA2214	CGATCCCAGGCGAATATCAT

Table continued next page

Supplemental Table 2-3. List of primer pairs used in chromatin immunoprecipitation and micrococcal nuclease sensitivity analysis

5877	DA2215	CGTGGCGCAGGTTGATAAT	DA2216	CAGCTAGTCGCATAATTTTGCAA
6883	DA2501	CCCAGAGAATGGGCAACAAG	DA2502	GTCCCGCTTCGGATGAACT
7822	DA2503	ATAACCCTACGAGCGGCTCTT	DA2504	GCAGTTACGCGGAAGTCGAA
actin5c				
p	DA2095	CCACTTTCAGTCGGTTTATTCCA	DA2096	AAATTCAAGGCGTGAAAACTAC TGT
1kb	DA2217	AGTCCGGCCCCTCCATT	DA2218	CTGATCCTCTTGCCCAGACAA
Mnase footprinting primers				
ftz promoter				
-503	DA2266	CGGGTTCCGGGAACAAC	DA2267	GGTTCCCTAAAAATTTCTCTGGA TT
-404	DA2268	CCTCTGCCCCGAAGGAA	DA2269	CCGATCTCCCTCTAATTGCAAT
-334	DA2270	AATTAGAGGGAGATCGGCTGA GA	DA2271	TGAGACGGTGCTGCAGTAATTA C
-259	DA2272	CGTCTCAAGGTCGCCGAGTA	DA2273	TACCCATCGCGCCCATTAT
-113	DA2274	TGCCTGCAAGGACATTTCG	DA2275	ATAAATATCCCTGCGCATGACA
-10	DA2276	CTCTGATTTTGCTATATATGCAG GATCT	DA2277	CACGCAACGCTGGTGAGTT
72	DA2278	ACATCGCAGAGTTAGAGAAGA AATCTAG	DA2279	TGTTGTCGGCGTAGCTGTAGTG
219	DA2280	GCAGCAATGCCTACTATCAGAA CA	DA2281	TTGTAGTAGTAGCAGCTCTCCG AGTAAC
304	DA2282	GGAGAGCTGCTACTACTACAAC AATCAG	DA2283	GCGATGATGGAAGCAGCAT
422	DA2284	AGGAGCGACCCAGCACACT	DA2285	GCTCGACGGTTGTGTAGAAATA GTC
ftz 3+6 enhancer				
-7822	DA2905	GGCACATTTTTTGTGTGCTTTAT T	DA2906	GAGTGTATTCTTTGTAAAATTGT CAATGG
-7794	DA2907	ACGATTTTTAGCGGTATCATAC GAT	DA2908	TGTAGGAGTGTATTCTTTGTAAA ATTGTCA
-7593	DA2909	CGAGTTTTTTAAGCGACAATAT GTCT	DA2910	TCCGTCTGTCCATATGAACGAT
-7519	DA2911	TCCCGAGATCACATCGTTCA	DA2912	CGGGATTACTGGCCAATTCA
-7414	DA2913	CACTTTCATCCTTTGATACATAC TTTTCAT	DA2914	GCAATAACTGGGAGTGGTGAAT AA
-7228	DA2915	CTCCCATCGGGTTTTGACTGT	DA2916	GCCTCAAGCGGCGTAGATAA
ftz 2+7 enhancer				
-6592	DA2770	GCCTCCTTTGGTAGCCTCAA	DA2771	TCGCAGGATACGCGCTTAT

Table continued next page

Supplemental Table 2-3. List of primer pairs used in chromatin immunoprecipitation and micrococcal nuclease sensitivity analysis

-6535	DA2772	GCGCGTATCCTGCGAATTAA	DA2773	AAACCCATACGTTGCATAGACTTG
-6411	DA2774	TTGTCTTGATGGCATTGTTTCTG	DA2775	GAAGGTGACGTTTTATTGACTGTCTTC
-6279	DA2776	AACAATTCCGGCGTTCTCAA	DA2777	TTAAATTAGAGCAAACGGCGACTA
-6201	DA2778	TGCATTGTTTTGGAATGTTTCC	DA2779	TGGGTCATAAAAACGAAAGACTTG
-6147	DA2780	CGCCACACACTTTTCGTTACAA	DA2781	GCAACATCTATTCAAAAAATCGTTCA
ftz inter enhancer				
-3997	DA2875	TCGGCATCCTTGACTTTGATT	DA2876	GGTCGTCTCAATGGGTGGATA
-3783	DA2879	TTTCCCCAGCGCAGACTTAA	DA2880	TGCGTTATGGGATTTGTTTGAGT
-3697	DA2881	GAATAATTCTCGAGCATCGCTACTAA	DA2882	GTTCGGATCGCAGGCAAT
-3490	DA2883	TAATTTTTCTCGGCCCTTTACTATTT	DA2884	CAATCGAGTATATCACACTCATTATCACAGT
-3403	DA2885	CACCCCTTCCCCAATTATGTG	DA2886	TCCTTTGAAGTTGCTGCTGATC
ftz 1+5 enhancer				
3722	DA2887	CGTTAACTCTTCCTGTGTCCTTCTG	DA2888	CCGACTCCGAGGACCTCTAAT
3819	DA2889	GAGCCATAAACCCCGAGATTC	DA2890	GACTGAAACCGGGTTTGCA
4007	DA2891	CGAATAGGCGTGCCATTGT	DA2892	GCTGGTGGAAGGGATAGAGATG
4122	DA2893	GATCGTAAAGTTATCCTTTCGCAAA	DA2894	TGCGCGGACAAAGGTGTT
4231	DA2895	AGTTGGCGAATTATCCGTTCA	DA2896	GCCCATCGCGAATTTCCTA
4349	DA2897	GGAAGGAACACGACAGGCTAGA	DA2898	CCGCTCCGTGTTTATCTTTAACA
eve promoter				
-423	DA2290	GCTGGCAGGCGATCGATA	DA2291	TGGCATTATGCCGCTCAGT
-358	DA2292	ACCCGACTGAGCGGCATA	DA2293	GGCGGGCCCCCTTAA
-309	DA2294	GATGAAGCCGATAAAATCCCATTAT	DA2295	CGGATGAGCGAATCTTTTGTCT
-216	DA2296	CGCTCATCCGCTATGAATACC	DA2297	TCTGCTCGGCAGGACCAT
-21	DA2298	ACTCTCAGCACCGCACGATTAG	DA2299	AGGCAGTTAGTTGTTGACTGTGCG

Table continued next page

Supplemental Table 2-3. List of primer pairs used in chromatin immunoprecipitation and micrococcal nuclease sensitivity analysis

78	DA2300	TGCAGAGCGCAGCGGTATAAA	DA2301	CGAAGGCAGTTAGTTGTTGACTGTGC
182	DA2302	TTGAATCACAAGACGCATACCAA	DA2303	CATGGTGGCTCTCCATGTTG
310	DA2304	CGGCAAGCCCCAGACA	DA2305	CGTTTGCTCGGCTTTATCTTTAA
436	DA2306	ACACCCAATCTTTCCTCTGTCCA	DA2307	TGTAGAACTCCTTCTCCAAGCGAC
eve 3+7				
-3656	DA2694	TGTTTTGGCCGACCGATTT	DA2695	TGGGAACGCGGCCATA
-3594	DA2696	GGCCGCGTTCCCATTT	DA2697	CGCCGGGACCCATAAAA
-3458	DA2698	TGTCGCCCGTATTAGGAAAGTAG	DA2699	AAAAACTAGCGCAGCGAAAAA
-3377	DA2700	CCAGCGAACTGCTCTAATTTTTTAA	DA2701	ACCTTGTCCGCGTTTTTCC
-3298	DA2702	CGCGGACAAGGTTATAACGCTCTACTT	DA2703	GTGAATGCCATCGCGGACAAACAA
eve 2				
-1457	DA2654	TAACTGGCAGGAGCGAGGTATC	DA2655	GAGAAACCAGGCCAGCTTTTC
-1373	DA2656	CCGGTACTGCATAACAATGGAA	DA2657	CTGATGGCAAACGGATTAACAC
-1335	DA2658	CGTTTGCCATCAGCGAGATT	DA2659	TTGTTCAAGATGCTGCAATAAAGTC
-1298	DA2660	GCAGCGTTTCGCTTTCGT	DA2661	TGAAAGTATGGCACAGCGTGTT
-1191	DA2662	GGACCCTGGACTATAATCGCACAA	DA2663	GCAGAAGATGGCGATGGCTAGAT
-1120	DA2664	CATCTTCTGCGGGCGTTTGTTTGT	DA2665	GGGCTAGGGATCGGGATTGGATT
eve 4+6				
4724	DA2666	TCCGCCGACCCTCCATA	DA2667	CGGTCATTAAAGCCCCTAAAAA
4768	DA2668	TCTCGTTTTTTTCGCGTTATTTT	DA2669	AATCAGTGCGCGAAATGTGA
4896	DA2670	GCCTCGAGCAGGACTCTTTG	DA2671	CGCTGGACGAAAAAACTGAGA
5014	DA2672	GGCGTTTTATGGCGGAGAT	DA2673	CCGCCTAATTTTTATGGCCTAA
5061	DA2674	GGCATAAAAAAACTGCATTGGAA	DA2675	CCGCGAAATGCGAATCTAG
5161	DA2676	CAGAATGCCAGAATGGTCAGAA	DA2677	ACCGCAGCGAGTCAATTTTTT
eve 1				
6655	DA2728	CCGTACCTGAGCCCACTGAT	DA2729	AGATTAGAGCCAAAGAGTTGACAATATTT

Table continued next page

Supplemental Table 2-3. List of primer pairs used in chromatin immunoprecipitation and micrococcal nuclease sensitivity analysis

6780	DA2730	TGGCGGAGAGGGAGATGAG	DA2731	GCCAATCTCGACTGCCACTT
6877	DA2732	GCAACAAGTAGCGGCGAATT	DA2733	TGCGCGGGCAAGAGTT
6966	DA2734	CCGCGCATCTCAGTTCATC	DA2735	CGCCTCGAAATTAATGCAAAC
7112	DA2736	AGAAACCGCCGAGGACACT	DA2737	CAGTTGCAAATGCGAAAGGA
7206	DA2738	ACTGGGTCCGCCCCTAATC	DA2739	GGTAATCGCAGATAATCCCATTACC
eve 5				
7341	DA2740	AGCGAAGTCAACTAAATCCAATCC	DA2741	GGGCAACTTAATTGCCAGAAAT
7465	DA2742	CTGTCCTGCACAAATGCTGAA	DA2743	GCGGCCTTGGGATATCTGA
7545	DA2744	CGGCAGCAAATTTCCCTTT	DA2745	CCCAAACAATGCAGCGAGTT
76670	DA2746	GCGTGCAATTATGGACATCCT	DA2747	GGGTTCTGTAGGCGGATGAA
7794	DA2748	CAGCAGCTTGCGTAAAATTTCA	DA2749	AAAAAGGATAATGCCAGGGATTC
7873	DA2750	TCCTTTTTACCTCTTGCCAATCC	DA2751	CTGGAAATCGCCGTCCTTT
Intergenic control				
0	DA2806	CTGCCATGCGCTTTTGTCT	DA2807	GGCGGAGTAGTTGGATTTTACG
87	DA2808	GGCCGTATTACAAACGAGTATGC	DA2809	GCGGAAGTGCTTTATTTAAGCTAGA
194	DA2810	AAGTCGCACTCGGTAGCTCTCT	DA2811	AAGGTCCTGGCACGATCCT
300.5	DA2812	GGCCGGCTGCAACATG	DA2813	CCCCTTCACCATCAACATCCT
447.5	DA2814	TTTGTGTAGGCGACATTGTTGTAA	DA2815	CTGTATGGGCCTCGGTGTTT

Appendix A-2

Fine mapping of chromatin structure in *Drosophila melanogaster* embryos using micrococcal nuclease[4]

Abstract

The structure of chromatin in eukaryotes exerts significant influences on many DNA related processes, including transcription, replication, recombination and repair. A useful tool for mapping chromatin structure is micrococcal nuclease (MNase), which induces double-strand breaks within nucleosome linker regions, and with more extensive digestion, single-strand nicks within the nucleosome itself. Many studies, carried out largely with microbes and cell cultures, have used MNase to determine the positions of nucleosomes within a region of DNA to identify dynamic changes induced during gene regulation. To measure similar processes in a developmental context, we turned to a tractable model system, the *Drosophila* embryo. Here we describe a protocol that enables MNase mapping of the enhancer chromatin structure in the embryo, and show how it can be used to identify structural changes on a cis-regulatory element targeted by the Knirps repressor.

Key words

Micrococcal nuclease mapping, *Drosophila* embryo, enhancer

4 This work was published as the following manuscript: **Li, L.M. and Arnosti, D.N.** (2010) Fine mapping of chromatin structure in *Drosophila melanogaster* embryos using micrococcal nuclease. Fly (Austin). 2010 Jul 29;4(3).

Introduction

In eukaryotes, DNA is packaged into chromatin, which imposes a natural barrier for various DNA-related processes including transcription and replication [43]. Eukaryotic cells deploy multiple mechanisms to alter the positions of nucleosomes in order to enhance the accessibility of DNA [44]. In the case of transcription, regulatory factors recruit chromatin remodeling complexes such as SWI/SNF and ISWI to remove, transfer, or slide a nucleosome along the DNA template [45]. Alternatively, histone modifying enzymes can also be recruited by transcription factors to acetylate, methylate, phosphorylate, monoubiquitinate, sumoylate, or ADP-ribosylate histones, or carry out reverse reactions [44]. Such modifications can directly affect inter- and intranucleosomal interactions, or recruit downstream effectors to modulate the structure of nucleosomes [46, 47]. Thus, measurement of chromatin structure and changes associated with transcriptional activation and repression is essential for understanding of gene regulatory mechanisms.

Chromatin immunoprecipitation (ChIP) is used to analyze histone occupancy and modifications, and can pinpoint whether a specific histone protein or modification is associated with a particular piece of DNA; however the method does not provide direct information on overall chromatin structure [48]. Nuclease digestion does provide such structural information; DNase I has been used to map hypersensitive sites created by the loss or remodeling of nucleosomes, providing an indication of an active regulatory function [49]. Alternatively, fine structure mapping can be obtained by the use of micrococcal nuclease (MNase). This nuclease can induce double-stranded breaks within the nucleosome linker region, providing an indication of

whether a region of interest is protected within a nucleosome . In standard MNase experiments, chromatin accessibility is determined in a qualitative manner by Southern blot using radioactive-labeled probes against the DNA of interest . To obtain a more quantitative picture of nucleosome structure, MNase digested chromatin can be analyzed by quantitative PCR using overlapping primers. This technique has been used to determine chromatin structure in yeast, as well as tissue-culture cells [50-52]. Current studies are also analyzing chromatin from multicellular organisms, but no protocols are available for application of MNase mapping in embryos. Here, we describe a protocol we developed and optimized for use with the *Drosophila* blastoderm embryo, providing a method to study chromatin structure in a developmental setting.

Collection of *Drosophila* embryos and formaldehyde crosslinking

To ensure reproducibility, it is critical to standardize the amount of starting materials. For each MNase experiment, we start with approximately 500 mg of 2-4 hr old embryos collected from 10 laying bottles. Embryos are rinsed free of yeast paste and any adults by washing through coarse nylon mesh into a large collection basket, dechorinated by bleaching for 2.5 min, rinsed in stream of tap water for 1 min and blotted dry from below using Kimwipes.

Embryos are then transferred into a 50 ml Corning tube and crosslinked by vigorously shaking for 30 min in 10 ml crosslinking buffer (3% formaldehyde in 50 mM HEPES, pH 7.6, 1 mM EDTA, 0.5 mM EGTA, 100 mM NaCl) and 30 ml heptane. Heptane is added to the crosslinking mixture to increase the permeability of the embryos. Embryos are centrifuged at 1000 rpm (167 g) for 1 min in a clinical centrifuge. The top organic layer and the crosslinking buffer are carefully

removed so as not to lose too many embryos, and the reaction is quenched by addition of 25 ml crosslinking stop buffer (0.125 M glycine in PBS (137mM NaCl, 4.3mM Na_2HPO_4, 1.4mM NaH_2PO_4) -Triton-X-100 0.01%). After gently mixing at room temperature for 30 min, the mixture is centrifuged at 1000 rpm (167 g) for 1 min, and the supernatant is discarded. At this point, the embryos can be flash frozen by immersing the Corning tube in liquid nitrogen for later processing, or the material can be directly processed as described below.

Preparation of nuclei and MNase digestion

Embryos are washed three times in 10 ml PBS-Triton-X-100 (0.01%) with gentle rocking for 5 min. The crosslinked embryos are resuspended in 5 ml homogenization buffer (10mM HEPES, pH 7.6, 0.3M sucrose, 10mM KCl, 1.5mM $MgCl_2$, 0.5mM EGTA, 1mM DTT, 1 pellet of protease inhibitor (*cOmplete* mini, Roche, Catalog No. 04 693 124 001) per 10 ml) per 0.5 gram of embryos and disrupted in a Dounce homogenizer 10 times with a loose pestle, and 15 times with a tight pestle. To collect the nuclei, the lysate is transferred to a 15 ml Corning tube and centrifuged at 1000 rpm (167 g) for 10 min at 4^{o}C in a clinical centrifuge. The supernatant is removed and the pellet is resuspended in 1 ml MNase digestion buffer (10mM Tris-HCl, 15mM NaCl, 60mM KCl, 0.15mM spermine, 0.5mM spermidine). After transferring to a 1.7 ml microcentrifuge tube, the material is centrifuged at 14,000 rpm (16,000 g) for 10 min at 4^{o}C. The supernatant is removed and discarded, and the pellet is flash frozen until ready for use.

Nuclei from 500 mg embryos are suspended in 1ml of MNase digestion buffer in a microcentrifuge tube. Half of the volume is reserved to serve as the undigested control. 100 U MNase (USB, Catalog No. 70196Y) is added to the rest of the sample and incubated at 37^{o}C for

30 min, long enough to digest most of the chromosomal DNA into mononucleosomal DNA. The sample is chilled on ice for 10 min and then EDTA is added to the final concentration of 10 mM to quench the digestion. The digested nuclei are collected by centrifuging at 14,000 rpm (16,000 g) for 10 min at 4^{o}C in a microcentrifuge.

Dissolving the nuclei and reversing crosslinking and DNA purification

Digested nuclei are washed twice in 1 ml sonication buffer (10mM HEPES pH 7.6, 1mM EDTA, 0.5mM EGTA, 0.1% sodium dioxycholate). For each wash, the pellet is completely resuspended, and then recollected by centrifuging at 14,000 rpm (16,000 g) for 10 min at 4^{o}C using a microcentrifuge. Nuclei are then suspended by addition of 500 µl sonication buffer and sonicating three times at output 4, 60% duty cycle and 20 pulses with a Branson sonicator, using a microtip. The material is then centrifuged at 14,000 rpm (16,000 g) for 20 min to pellet the nuclear membrane; the supernatant contains the mononucleosomes.

The crosslinking of both the digested sample and the undigested control is reversed by addition of SDS to a final concentration of 1%, and NaCl is added to a final concentration of 0.2 M. The samples are heated at 65^{o}C overnight to reverse the crosslinking. RNA within the samples is removed by incubating tubes at 37°C for 30 min with 10 µg RNAse (Roche, Catalog No. 10109142001) per reaction. To remove protein, the solution is adjusted to 10 mM EDTA, 40mM Tris-Cl pH 6.5, and 20 µg Proteinase K (Roche, Catalog No. 03115836001) is added to the reaction. Samples are incubated at 42°C for 2 hours. DNA is purified by extraction once with buffered phenol-chloroform. To precipitate DNA, 400 µl of supernatant is transferred into a

microcentrifuge tube (avoid sucking up the interface), and incubated with 1 µl of GlycoBlue (Ambion, AM9515), 44 µl NaOAc 0.3 M, and 444 µl of room temperature isopropanol for at least 30 min. DNA is then collected by centrifugation at 14,000 rpm (16,000 g) for 15 min at room temperature, washed once with 70% EtOH, and dried in a Speedvac for 10 min at 65°C or in air overnight. The pellet is dissolved in 100 µl water; typically 1 µl of this solution is sufficient for real-time PCR analysis.

Analysis of digested nucleosomes

To obtain a detailed map of chromatin structure, the resulting materials (both the digested and undigested) are analyzed by real-time PCR with an Applied Biosystem 7500 thermocycler using Power Syber Master Mix (Catalog No. 4367659). For a detailed mapping of nucleosome structure, primer pairs that generate short-overlapping PCR products are normally used {Sekinger, 2005 #56}. If the purpose of the experiment is to identify changes in nucleosome density, less densely spaced, non-overlapping primer pairs can be used. Primers were designed using Primer Express software (Applied Biosystems) with lengths of 100-150bp and Tms 57-60°C. Nucleosome density of a given region is defined by the ratio of digested DNA versus undigested control DNA.

If a nucleosome is reproducibly positioned on a length of DNA in all cells, the DNA will be protected from MNase digestion and an amplicon from a primer pair located totally within this nucleosome will yield a digested/undigested ratio of 1. On the other hand, if part of the amplicon is not protected by a nucleosome, MNase digestion will reduce the amount of intact genomic DNA for this amplification, yielding a digested/undigested DNA ratio of as little as 0.

Thus, a positioned nucleosome array will generate peaks and valleys, with the valleys corresponding to linker regions between nucleosomes. If only a fraction of cells contain a positioned nucleosome, the ratio of digested/undigested DNA will lie between 0-1.

Analysis of chromatin structure at early *Drosophila* zygotic genes using MNase mapping

During early Drosophila development, zygotic gap, pair-rule and segment polarity genes are expressed in a spatially dynamic manner representing a well-studied transcriptional cascade [1]. A set of transcriptional repressors encoded by *knirps*, *Krüppel* and *giant* gap genes dictate critical positional information in the blastoderm embryos, but little is known of these proteins' biochemical activities. We applied this MNase mapping method to assess the effect of the short-range repressor Knirps on the chromatin structure of a cis-regulatory region located 9.5 kbp downstream of the *hunchback* (*hb*) transcription start site, and compared it to the chromatin structure of *adam*, a non-target locus. Knirps protein was uniformly expressed in 2-4 hr blastoderm embryos to repress target genes using a heatshock driver [25]. We observed a significant increase in the resistance to MNase digestion across the 400 bp *hb* enhancer, whereas the digestion pattern remained unchanged at the nonspecific *adam* gene, as well as an intergenic regions tested on the 3rd chromosome (Fig. 1 and data not shown). We have observed similar increases in nuclease resistance at specific *even-skipped* enhancers targeted by Knirps (Li and Arnosti, submitted). These results suggest that chromatin structure is significantly altered during transcriptional repression by this factor, and that such effects can be observed using the whole *Drosophila* embryos. This protocol is suitable for local analysis of

individual regulatory elements controlled by gap gene repressors, and may provide useful insight into chromatin alterations mapped by more global approaches [53].

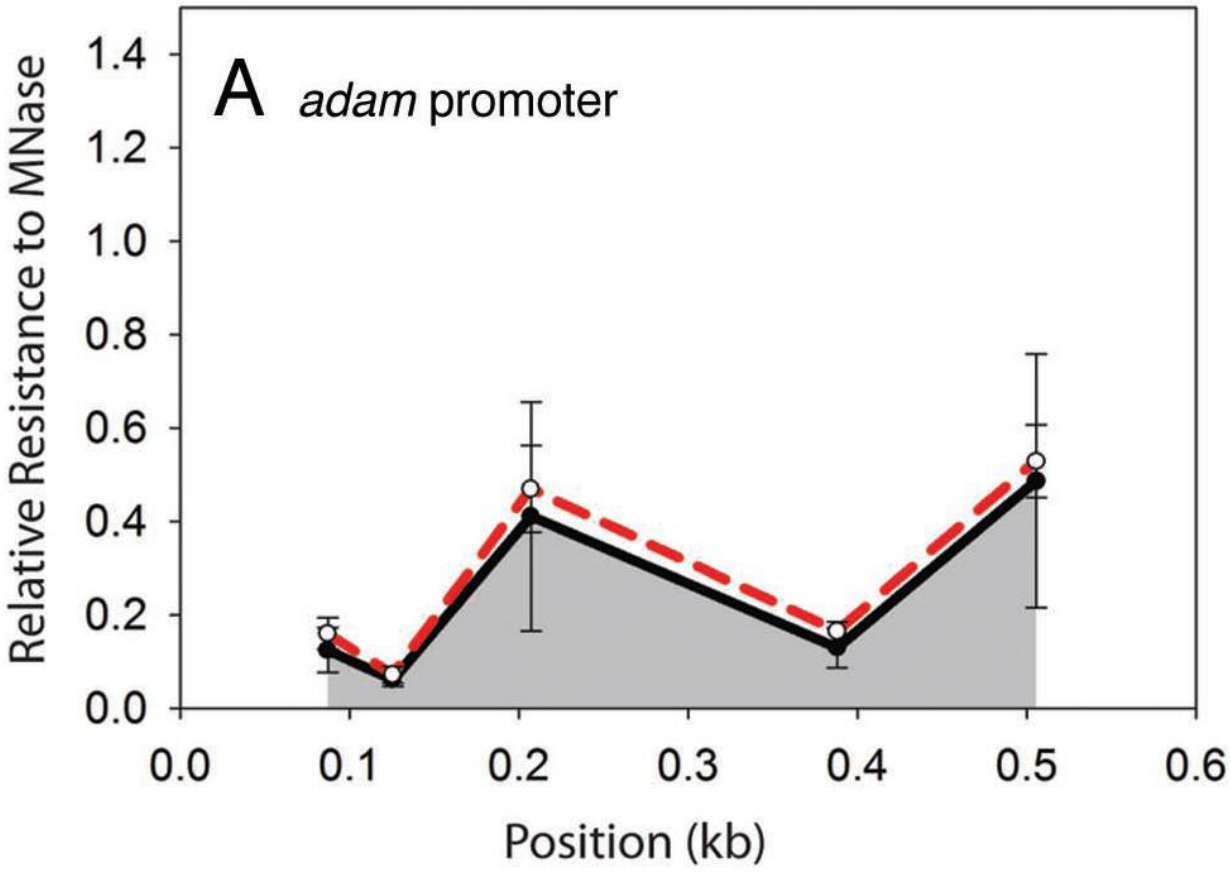


Figure A-1. Chromatin sensitivity to MN ase digestion in the Drosophila embryo, before and after induction of Knirps repressor. (A) Chromatin accessibility of *adam* promoter region before (solid line) and after (dashed line) induction of Knirps repressor. (B) Increase in resistance to MN ase digestion at *hb* enhancer located at +9.5 kbp after repression of the gene by Knirps. Data represent averages of 3 biological replicates; error bars indicate standard errors; points in the graph represent the centers of the amplicons used in real-time PCR analysis.

Figure A-1 continued

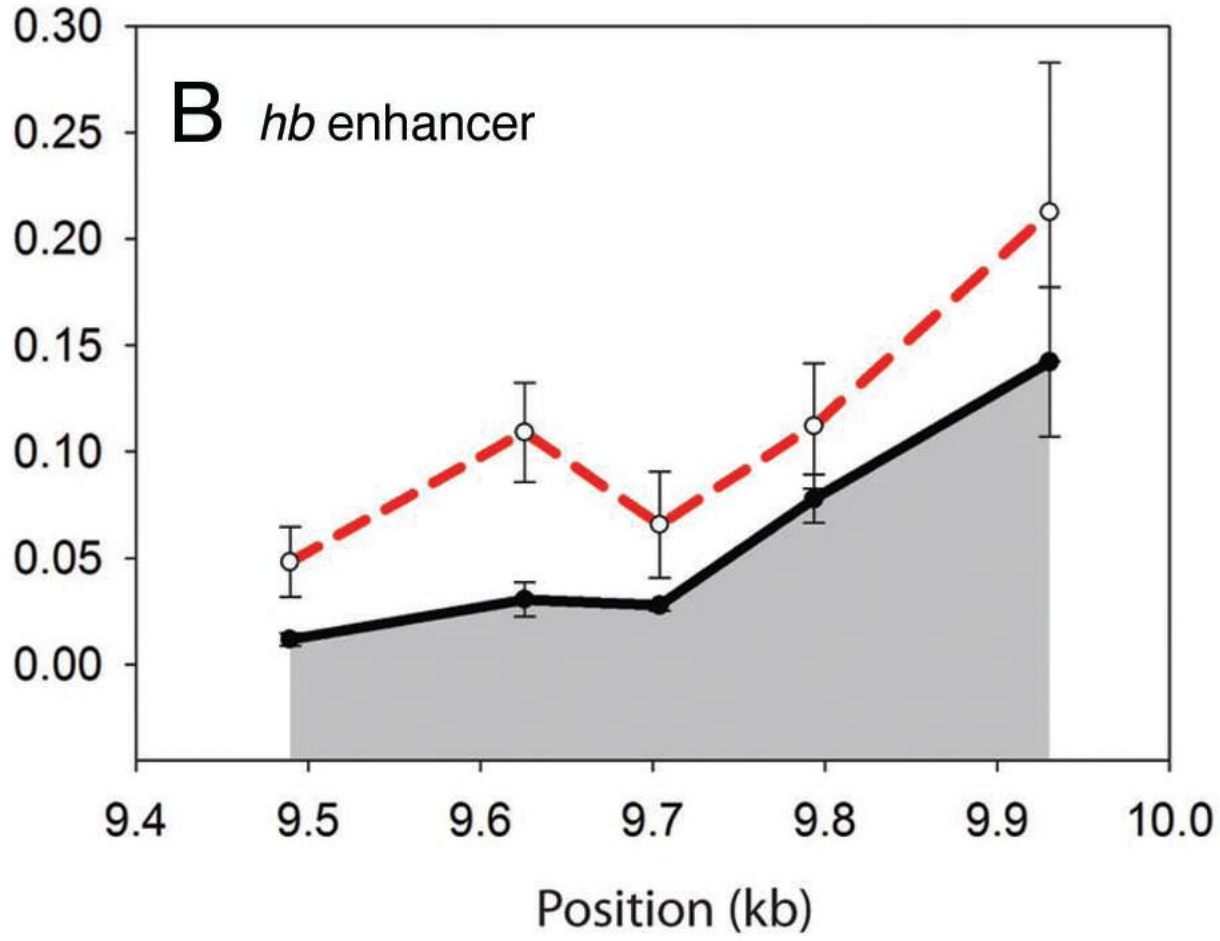


Figure A-1. Chromatin sensitivity to MN ase digestion in the Drosophila embryo, before and after induction of Knirps repressor. (A) Chromatin accessibility of *adam* promoter region before (solid line) and after (dashed line) induction of Knirps repressor. (B) Increase in resistance to MN ase digestion at *hb* enhancer located at +9.5 kbp after repression of the gene by Knirps. Data represent averages of 3 biological replicates; error bars indicate standard errors; points in the graph represent the centers of the amplicons used in real-time PCR analysis.

References

References

1. Rivera-Pomar, R., and Jackle, H. (1996). From gradients to stripes in Drosophila embryogenesis: filling in the gaps. Trends Genet *12*, 478-483.

2. Small, S., and Levine, M. (1991). The initiation of pair-rule stripes in the Drosophila blastoderm. Curr Opin Genet Dev *1*, 255-260.

3. Cai, H.N., Arnosti, D.N., and Levine, M. (1996). Long-range repression in the Drosophila embryo. Proc Natl Acad Sci U S A *93*, 9309-9314.

4. Gray, S., Szymanski, P., and Levine, M. (1994). Short-range repression permits multiple enhancers to function autonomously within a complex promoter. Genes Dev *8*, 1829-1838.

5. Small, S., Arnosti, D.N., and Levine, M. (1993). Spacing ensures autonomous expression of different stripe enhancers in the even-skipped promoter. Development *119*, 762-772.

6. Paroush, Z., Finley, R.L., Jr., Kidd, T., Wainwright, S.M., Ingham, P.W., Brent, R., and Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell *79*, 805-815.

7. Jimenez, G., Paroush, Z., and Ish-Horowicz, D. (1997). Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. Genes Dev *11*, 3072-3082.

8. Song, H., Hasson, P., Paroush, Z., and Courey, A.J. (2004). Groucho oligomerization is required for repression in vivo. Mol Cell Biol *24*, 4341-4350.

9. Chen, G., Fernandez, J., Mische, S., and Courey, A.J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in Drosophila development. Genes Dev *13*, 2218-2230.

10. Flores-Saaib, R.D., and Courey, A.J. (2000). Analysis of Groucho-histone interactions suggests mechanistic similarities between Groucho- and Tup1-mediated repression. Nucleic Acids Res *28*, 4189-4196.

11. Courey, A.J., and Jia, S. (2001). Transcriptional repression: the long and the short of it. Genes Dev *15*, 2786-2796.

12. Martinez, C.A., and Arnosti, D.N. (2008). Spreading of a corepressor linked to action of long-range repressor hairy. Mol Cell Biol *28*, 2792-2802.

13. Nibu, Y., Zhang, H., Bajor, E., Barolo, S., Small, S., and Levine, M. (1998). dCtBP mediates transcriptional repression by Knirps, Kruppel and Snail in the Drosophila embryo. EMBO J *17*, 7009-7020.

14. Subramanian, T., and Chinnadurai, G. (2003). Association of class I histone deacetylases with transcriptional corepressor CtBP. FEBS Lett *540*, 255-258.

15. Shi, Y.J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell *19*, 857-864.

16. Payankaulam, S., and Arnosti, D.N. (2009). Groucho corepressor functions as a cofactor for the Knirps short-range transcriptional repressor. Proc Natl Acad Sci U S A *106*, 17314-17319.

17. Arnosti, D.N., Gray, S., Barolo, S., Zhou, J., and Levine, M. (1996). The gap protein knirps mediates both quenching and direct repression in the Drosophila embryo. EMBO J *15*, 3659-3666.

18. Ish-Horowicz, D., and Pinchin, S.M. (1987). Pattern abnormalities induced by ectopic expression of the Drosophila gene hairy are associated with repression of ftz transcription. Cell *51*, 405-415.

19. Yu, Y., and Pick, L. (1995). Non-periodic cues generate seven ftz stripes in the Drosophila embryo. Mech Dev *50*, 163-175.

20. Calhoun, V.C., and Levine, M. (2003). Long-range enhancer-promoter interactions in the Scr-Antp interval of the Drosophila Antennapedia complex. Proc Natl Acad Sci U S A *100*, 9878-9883.

21. Hiromi, Y., Kuroiwa, A., and Gehring, W.J. (1985). Control elements of the Drosophila segmentation gene fushi tarazu. Cell *43*, 603-613.

22. Macarthur, S., Li, X.Y., Li, J., Brown, J.B., Chu, H.C., Zeng, L., Grondona, B.P., Hechmer, A., Simirenko, L., Keranen, S.V., 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.

23. Fujioka, M., Emi-Sarker, Y., Yusibova, G.L., Goto, T., and Jaynes, J.B. (1999). Analysis of an even-skipped rescue transgene reveals both composite and discrete neuronal and early blastoderm enhancers, and multi-stripe positioning by gap gene repressor gradients. Development *126*, 2527-2538.

24. Clyde, D.E., Corado, M.S., Wu, X., Pare, A., Papatsenko, D., and Small, S. (2003). A self-organizing system of repressor gradients establishes segmental complexity in Drosophila. Nature *426*, 849-853.

25. Struffi, P., Corado, M., Kulkarni, M., and Arnosti, D.N. (2004). Quantitative contributions of CtBP-dependent and -independent repression activities of Knirps. Development *131*, 2419-2429.

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

27. Schubeler, D., MacAlpine, D.M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D.E., O'Neill, L.P., Turner, B.M., Delrow, J., et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes Dev *18*, 1263-1271.

28. Struffi, P., and Arnosti, D.N. (2005). Functional interaction between the Drosophila knirps short range transcriptional repressor and RPD3 histone deacetylase. J Biol Chem *280*, 40757-40765.

29. Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

30. Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell *111*, 185-196.

31. Dearolf, C.R., Topol, J., and Parker, C.S. (1989). The caudal gene product is a direct activator of fushi tarazu transcription during Drosophila embryogenesis. Nature *341*, 340-343.

32. Li, X.Y., MacArthur, S., Bourgon, R., Nix, D., Pollard, D.A., Iyer, V.N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C.L., et al. (2008). Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. PLoS Biol *6*, e27.

33. Janssens, H., Hou, S., Jaeger, J., Kim, A.R., Myasnikova, E., Sharp, D., and Reinitz, J. (2006). Quantitative and predictive model of transcriptional control of the Drosophila melanogaster even skipped gene. Nat Genet *38*, 1159-1165.

34. Pankratz, M.J., Busch, M., Hoch, M., Seifert, E., and Jackle, H. (1992). Spatial control of the gap gene knirps in the Drosophila embryo by posterior morphogen system. Science *255*, 986-989.

35. Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat Genet *39*, 1512-1516.

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

37. Adelman, K., Wei, W., Ardehali, M.B., Werner, J., Zhu, B., Reinberg, D., and Lis, J.T. (2006). Drosophila Paf1 modulates chromatin structure at actively transcribed genes. Mol Cell Biol *26*, 250-260.

38. Strahl-Bolsinger, S., Hecht, A., Luo, K., and Grunstein, M. (1997). SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. Genes Dev *11*, 83-93.

39. Papp, B., and Muller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev *20*, 2041-2054.

40. Fakhouri, W.D., Ay, A., Sayal, R., Dresch, J., Dayringer, E., and Arnosti, D.N. Deciphering a transcriptional regulatory code: modeling short-range repression in the Drosophila embryo. Mol Syst Biol *6*, 341.

41. Payankaulam, S., Li, L.M., and Arnosti, D.N. Transcriptional repression: conserved and evolved features. Curr Biol *20*, R764-771.

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

43. 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.

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

45. Saha, A., Wittmeyer, J., and Cairns, B.R. (2006). Mechanisms for nucleosome movement by ATP-dependent chromatin remodeling complexes. Results Probl Cell Differ *41*, 127-148.

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

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

48. Das, P.M., Ramachandran, K., vanWert, J., and Singal, R. (2004). Chromatin immunoprecipitation assay. Biotechniques *37*, 961-969.

49. Giresi, P.G., and Lieb, J.D. (2006). How to find an opening (or lots of them). Nat Methods *3*, 501-502.

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

51. Sekinger, E.A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol Cell *18*, 735-748.

52. 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.

53. Dorschner, M.O., Hawrylycz, M., Humbert, R., Wallace, J.C., Shafer, A., Kawamoto, J., Mack, J., Hall, R., Goldy, J., Sabo, P.J., et al. (2004). High-throughput localization of functional elements by quantitative chromatin profiling. Nat Methods *1*, 219-225.

Chapter III

Analysis of Hairy-mediated Long-range Transcriptional Repression in a Reporter System

Abstract

Endogenous targets allow for studying repressor mechanisms in the native setting of the repressors. However, the complexity of gene regulatory circuits and the inhomogeneity of cell populations in an intact organism complicate the interpretation of the results from molecular studies. In this chapter, reporter systems with defined positions of activator and repressor binding sites are used to study the repression mechanisms of the long-range repressor Hairy. One set of reporter constructs allows for placement of a constitutively expressed activator on the promoter and induction of a LexA-Hairy fusion protein in a facultative manner, converting the embryo from an “all on” to “all off” state. This system allows chromatin studies in a homogenous cell population. Changes of RNA polymerase II occupancy, histone density, histone acetylation and histone methylation at the reporter upon Hairy repression resemble the observations made with endogenous targets. Another reporter system contains the endogenous *twist* enhancer close to the basal promoter, with Hairy binding sites introduced at various distances, which was used to study the distance limitation in Hairy-mediated repression. This section describes a series of preliminary experiments that led to the development of the system described in Chapter II.

Introduction

Studies on endogenous targets have limitations

To obtain physiologically meaningful information, it is important to study long- and short-range repression in the context of endogenous target genes. However, such studies have several technical limitations. First, endogenous targets for Hairy and Knirps are expressed in a spatially and temporally limited fashion. Chromatin studies performed on embryos involve a mixture of cell populations. Second, endogenous genes are often controlled by complex cis-regulatory modules, which are bound and regulated by multiple transcription factors. Due to the complexity in the gene circuit during early Drosophila development, transcription factors bound to the same regulatory modules might cross-regulate each other. Thus, when Hairy or Knirps is overexpressed, the final effects, either at the chromatin or transcription level, might be also caused by other transcriptional factors whose expression levels are regulated by these two repressors.

A reporter system to study the mechanisms of Hairy-mediated repression

Studies described by Martinez and Arnosti outline a useful approach to bypass these limitations [1]. In their study, they engineered a gene regulatory system in which the activator is constitutively expressed and binds to a *lacZ* reporter promoter, and the LexA-Hairy fusion repressor is expressed in an inducible manner driven by the *hsp70* promoter, converting the embryo from an “all on” to “all off” state. The *lacZ* reporter gene contains five binding sites for the Gal4 activator, flanked by two pairs of LexA binding sties to accommodate the repressors (Figure 3-1).

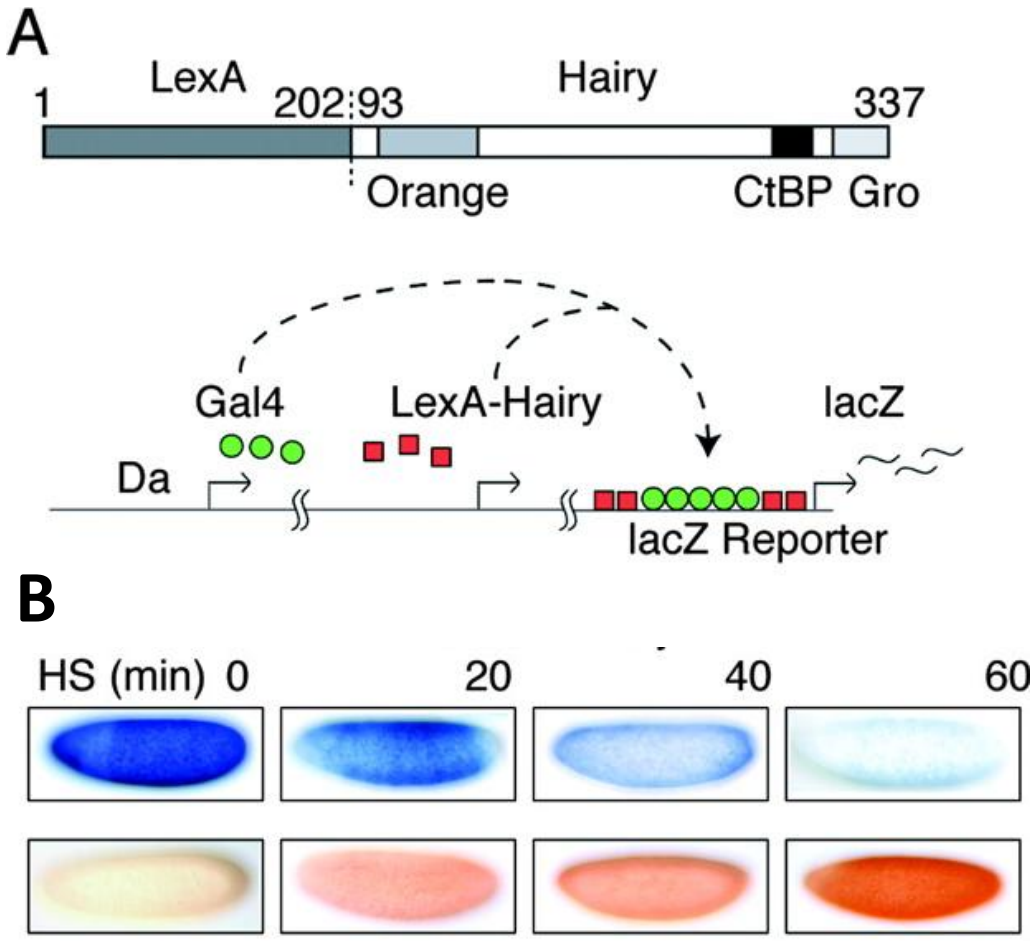


Figure 3-1. Hairy-regulated reporter system (A) The bacterial LexA protein was fused in frame to a Hairy fragment lacking the bHLH DNA-binding domain (93-337). The resulting LexA-Hairy construct contains the Orange, CtBP binding and Groucho binding motifs, but not the dSir2 corepressor binding motif. The Hairy regulated reporter system contains three transgenes, which were combined onto a single chromosome. The lacZ reporter contains Gal4 and LexA sites. The Gal4 activator is driven by the *daughterless* (*da*) enhancer for ubiquitous expression in the early Drosophila embryo. The LexA-Hairy repressor is driven by the *hsp70* promoter to achieve inducible expression. B) Expression of lacZ mRNA and LexA-Hairy protein upon heat shock treatment for varying time as indicated on top of the panels. As LexA-Hairy protein accumulates, the *lacZ* mRNA decreases. (Figures are adopted from Figure1 of reference [1])

As measured by in situ hybridization, *lacZ* transcription levels in the entire embryo drop markedly after even a short (5 min) heat shock. Successively longer heat shocks resulted in the complete loss of *lacZ* staining. In typical experiments, the percentage of embryos showing strong staining dropped from almost half to less than 1% 1 hour after the induction of the repressor, indicating that the repression was effective in the vast majority of nuclei and embryos.

Previous chromatin studies in the reporter system

After establishing the efficacy of the system, chromatin studies were performed to determine the chromatin changes associated with Hairy-mediated repression. Overall levels of H3 decreased in the activated state relative to those of the unactivated gene throughout the open reading frame, and are not much further affected during repression. Strong increases in relative H3 and H4 acetylation were observed at promoter and +1kb of lacZ reporter with the Gal4 activator present, and this acetylation decreases after the induction of the repressor. No change within the transcription unit was observed during activation or repression. Activation of the reporter gene is associated with a decrease in relative H3K27 methylation levels, whereas no changes in relative H3K27 methylation levels were observed during repression [1].

Unanswered questions

RNA polymerase II occupancy upon LexA-Hairy repression

My studies of the *ftz* locus suggest that Pol II is excluded from this gene when it is repressed by Hairy. To determine whether this effect on Pol II is a general property of Hairy repression, I turned to the *lacZ* reporter system. Previous studies suggested that upon LexA-

Hairy repression, activator and coactivator binding is not affected on the *lacZ* reporter, similar to my observations made on endogenous *ftz* locus. This implies that repression affects the cross-talk between activators and the basal transcriptional machinery. Examining Pol II occupancy at the repressed *lacZ* reporter will indicate whether inhibition of Pol II recruitment by Hairy is the general rule or an effect specific to the *ftz* locus.

Changes of histone H3 lysine4 and lysine 36 methylation marks upon Hairy repression

Besides histone acetylation, histone H3 lysine4 and lysine36 methylation has also been linked with active transcription. The common theme from genome-wide studies in various eukaryotes is that high levels of H3K4 trimethylation are associated with the 5' regions of almost all active genes, and there is a strong positive correlation between this modification, transcription rates, active Pol II occupancy, and histone acetylation [2-4]. The function of H3K4 methylation in transcription regulation is still to be elucidated; however, accumulating evidence suggests that this modification permits the recruitment of downstream effector proteins, which in turn affect chromatin modification and remodeling. These effectors include Chd1 in ATP-dependent chromatin remodeling, the ING proteins in histone acetylation and deacetylation, and JMJD2A in histone demethylation [5-7].

Histone H3K36 methylation is associated with transcription elongation. Both di- and trimethylation are enriched at the open reading frame (ORF), however, only trimethylation displays a positive correlation with transcription rates [3]. The main function of histone H3K36 trimethylation is to recruit the Rpd3S histone deacetylase complex through the interaction with the chromodomain Eaf3. The Rpd3S complex then produces a hypoacetylated environment in

the recently transcribed chromatin region to inhibit spurious transcription and ensure genome integrity [8, 9].

While the correlation between histone H3 lysine 4 and lysine 36 methylation and active transcription has been well established, few studies have looked into the changes of these marks as the target locus changes from an active state to a repressed state. The inducible repression reporter system provides a unique opportunity to examine the changes of these active marks during transcription repression.

Histone variants and Hairy-mediated repression

Chromatin structure is not regulated solely by modification of the core histones but also by the incorporation of histone variants. The histone variant H2A.Z is conserved across eukaryotes [10]. Genome-wide studies suggest that in yeast H2A.Z marks both active and inactive promoters, whereas in *Drosophila* and Human H2A.Z deposition is positively correlated to transcription level, and marks the promoters of actively transcribed genes [11-14]. An increased level of H2A.Z is associated with the process of activation, whereas gene silencing is associated with decreased H2A.Z level [12, 15]. In yeast, repression by Tup1, a Groucho homolog, also deposits H2A.Z at the targeted promoter to mark it for future reactivation [16]. It will be interesting to investigate whether Hairy, a repressor that heavily replies on Groucho, also deposits H2A.Z at the promoter upon repression.

Repressing a distal enhancer

One advantage of working with the synthetic enhancers is the ease of manipulation of transcription factor binding sites. The Lex-Hairy repressor binds close to both the Gal4 activator and the transcription start site in the transcription system used in Martinez and Arnosti. This transcription system can be modified to study Hairy-mediated repression in a long-range context, by putting LexA-Hairy repressor binding sites away from the activator binding sites and the transcription start site at the *lacZ* reporter.

The studies in this chapter first summarize the changes of RNA polymerase II occupancy, histone H3 lysine4 and lysine36 methylation marks, and histone variant H2A.Z level upon repression when LexA-Hairy repressor binds close to the activators and the basal promoter. Then, I discuss efforts to study Hairy-mediated repression in a long-range context.

Materials and Methods

Chromatin immunoprecipitation

Embryo collections and chromatin immunoprecipitations were performed according to a previously described protocol [1].

Quantitative PCR analysis

The samples from ChIP were analyzed using real-time PCR (Applied Biosystem 7500). Primer pairs had a Tm in the range of 58-60^{o}C, and amplicons ranged from 50 to 150 bp. Primer sequences are available upon request. A standard curve was generated by serially diluting input samples to quantify IP samples. All values used were collected from the linear range of amplification. The primers used to amplify the different regions of the reporter were promoter (**DA 1780:** 5'-TATACAGAAGCTTGCCTGCAGGTCGGA-3' and **DA 1781:** 5'-TATACAGCCATGCTAGAGTCTCCGCT-3'), +1kb (**DA1764:** 5'-AGCGTGGTGGTTATGCCG-3' and **DA1765:** 5'- GTTCAACCACCGCACGATAGA-3'), +2kb (**DA1751:** 5'-AAATGGCTTTCGCTACCTGGA-3' and **DA1752:** 5'-ATTTAGCGAAACCGCCAAGA-3'), 3rd chromosome intergenic region (**DA1938:** 5'-CTTTCCCGACATCCTGCAAA-3' and **DA1939:** 5'-TCAGCCAACTTTAGCTATTGACAG-3') and *hsp70A* promoter (**DA1705:** 5'-GTCACACAGTAAACGGCGCACT-3' and **DA1706:** 5'-CGCTTGTTTGTTTGCTTAGCTTT-3').

Construction of *lacZ* reporter plasmids

The lacZ reporter plasmids with LexA sites distal from the 2 X PE were constructed by cloning SphI-spacer-2 X PE-SphI fragments from pBluescript into the pC2L5U2L used in Martinez and Arnosti, 2008. The two SphI sites were introduced into pBluescript through KpnI using oligos **DA1214/DA1215** (**DA1214**: 5'-GTTA**GCATGC**TACGGTAC-3' and **DA1215**: 5'-CGTA**GCATGC**TAACGTAC-3', KpnI site underlied and SphI site in bold) and SacII sites using oligos **DA1205/DA1206** (**DA1205**: 5'-GCTTA**GCATGC**TACCCGC-3' and **DA1206**: 5-GGGTA**GCATGC**TAAGCGC-3', SacII site underlined and KpnI site in bold). The 2XPE enhancer was digested out of the Twi 105 plasmid and cloned in pBluescript through NotI site. The 1.5 kbp spacer (in p4L2S2PE) consists of two fragments. SpacerI is from nucleotide 1 to 750 bp of *knirps* cDNA, which was amplified from phskni1-429 plasmid using **DA1212/DA1218** (**DA1212**: 5'-GTCATCGATATGAACCCAGACATGCAAA-3' and **DA1218**: 5'-TATAAGCTTCGGGACTAAAGCGGTTCT-3') and cloned into pBluescript as a ClaI-HindIII fragment. SpacerII is the nucleotides from 1-740 bp of *gfp* cDNA, which was amplified from pPelican plasmid using **DA1209/DA1216** (**DA1209**: 5'-TATAAGCTTCGCCACCATGGTGAGCAA-3' and **DA1216**: 5'-TTATCTAGAGCCGGCCGCTTTACTTGT-3') and cloned into pBluescript through HindIII and XbaI sites. The lacZ reporter with 750 bp (p4LS2PE) only contains the *gfp* spacer. The Hairy sites were cloned into the p4LS2PE and p4L2S2PE using **DA1391** (5'-AATTCGCGG**CACGCG**ACATGACCCGCGG**CACGCG**ACATG-3', EcoRI sites underlined and core Hairy binding motif in bold)**/DA1392** (5'-AATTCATGT**CGCGTG**CCGCGGGTCATGT**CGCGTG**CCGCG-3', EcoRI sites underlined and core Hairy binding motif in bold).

In situ hybridization of Drosophila embryos

Embryos were fixed for hybridization and stained using a digoxigenin-UTP labeled antisense RNA probe to lacZ as described [17].

Results and Discussion:

Gal4 activator and LexA-Hairy repressor occupancy at the lacZ reporter measured by Real-time PCR

Previous studies established that on a *lacZ* reporter integrated into the fly genome, the Gal4 activator occupies the promoter when the reporter is active, and continue to occupy the promoter even when the LexA-Hairy represses the *lacZ* reporter. This previous study used conventional PCR and gel electrophoresis analysis following ChIP to monitor protein occupancy [1]. To obtain a more quantitative picture of changes on the *lacZ* reporter, I analyzed the precipitated materials using real-time PCR. This quantitative analysis yielded similar trends of activator and repressor occupancies at the *lacZ* reporter compared to the observations made using densitometry analysis of gels, and we calculated the fold changes of the activator and repressor occupancy at the reporter (Figure 3-2). Activation of the *lacZ* reporter resulted in at least 10 fold increase of Gal4 activator at the promoter. Upon repression, no significant change of Gal4 occupancy was observed. LexA-Hairy repressor occupancy increases by around 10 fold upon heat shock induction. This level of quantitation is especially useful for determining the changes of histone modifications associated with different states of the *lacZ* reporter as described below.

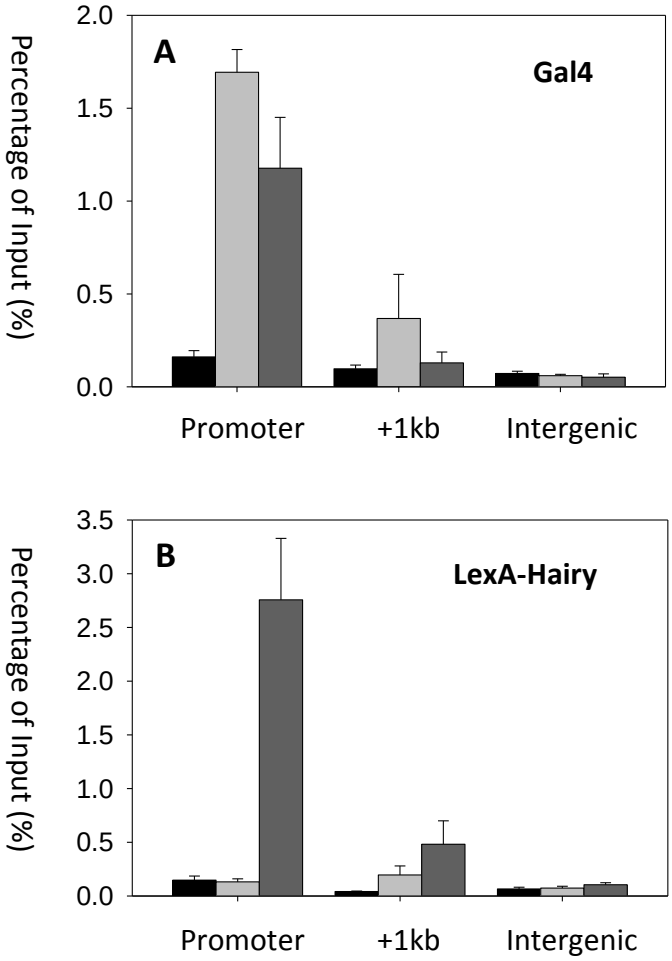


Figure 3-2. Promoter occupancy by Gal4 activator and LexA-Hairy repressor was measured by chromatin immunoprecipitation followed by Real-time PCR analysis. (A) Background level of Gal4 and LexA was observed at the lacZ using materials prepared from embryos containing only the reporter (black bar). Strong Gal4 signal was observed at the promoter from the embryos containing both the activator and the reporter (light grey bar). The activator occupancy was not significantly affected upon binding of the repressor, consistent with previous reporter (grey bar). (B) In the embryos contain all three transgenes, heatshock induction increased LexA-Hairy repressor level at the promoter (grey bar).

Hairy repression inhibits RNA polymerase II recruitment to the targeted reporter

Upon Gal4 mediated activation, RNA polymerase II is recruited to the promoter and body of the lacZ reporter. Upon induction of LexA-Hairy, RNA polymerase II occupancy decreases, similar to the effect observed when Hairy represses the *ftz* locus (Figure 3-3). The loss of polymerase occupancy suggests that Hairy-mediated repression generally inhibits RNA polymerase II recruitment on all of its targets, unless there are specific features shared by the *ftz* and *lacZ* reporter genes. A genome-wide analysis of all Hairy mediated repression events will clarify this question.

Activation marks increase upon activation by Gal4 activator and decrease upon repression by LexA-Hairy

I observed that H3K4 mono-, di- and tri-methylation all increased upon transcription activation by the Gal4 activator, consistent with the earlier report. Upon repression, H3K4 methylation levels decreased throughout the transcribed region of the *lacZ* reporter. Changes of H3K36 tri-methylation resembled the pattern of H3K4 methylation[18-20] (Figure 3-4).

Decreased activation marks upon Hairy-mediated repression may be a result of active histone demethylation, mediated by enzymes in the Hairy corepressor complex. Alternatively, when the reporter gene is switched off, histone methyltransferases can no longer be recruited by activators or the polymerase, baseline demethylation continues at a steady rate, and overall levels of methylation drop.

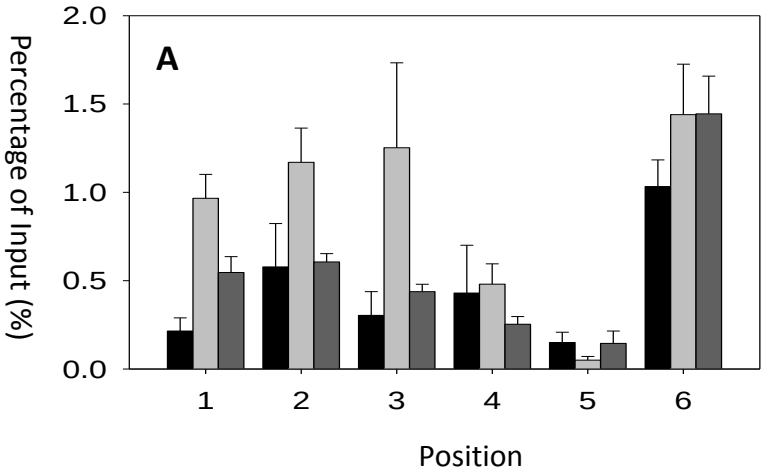


Figure3-3. Repression of lacZ **reporter expression by LexA-Hairy results in inhibition of RNA polymerase II recruitment measured by chromatin immunoprecipiation using antibodies again (A) unphosphorylated, (B) ser-5 phosphorylated, (C) ser-5 and ser-2 phosphorylated forms of RNA polymerase II**. Background levels of Pol II were detected with chromatin prepared from embryos containing only the *lacZ* reporter (black bar). Increased Pol II signals were observed from the embryos containing both the reporter and Gal4 activators (light grey bar). Induction of LexA-Hairy repressor using heat shock decreased Pol II occupancy at loci tested out on the reporter gene (gray bar). Numbers on the Y-axis represent the precipitated materials relative to the percentage of inputs. Numbers on the X-axis represent the positions tested out. Specifically, 1: promoter; 2: +1kb; 3: +2kb; 4: +4kb; 5: one intergenic region on the 3rd chromosome used as negative control; 6: the *hsp70* promoter used as a positive control. All the embryos have been subjected to heat shock treatment in order to eliminate non-specific effects caused by this treatment.

Figure 3-3 continued

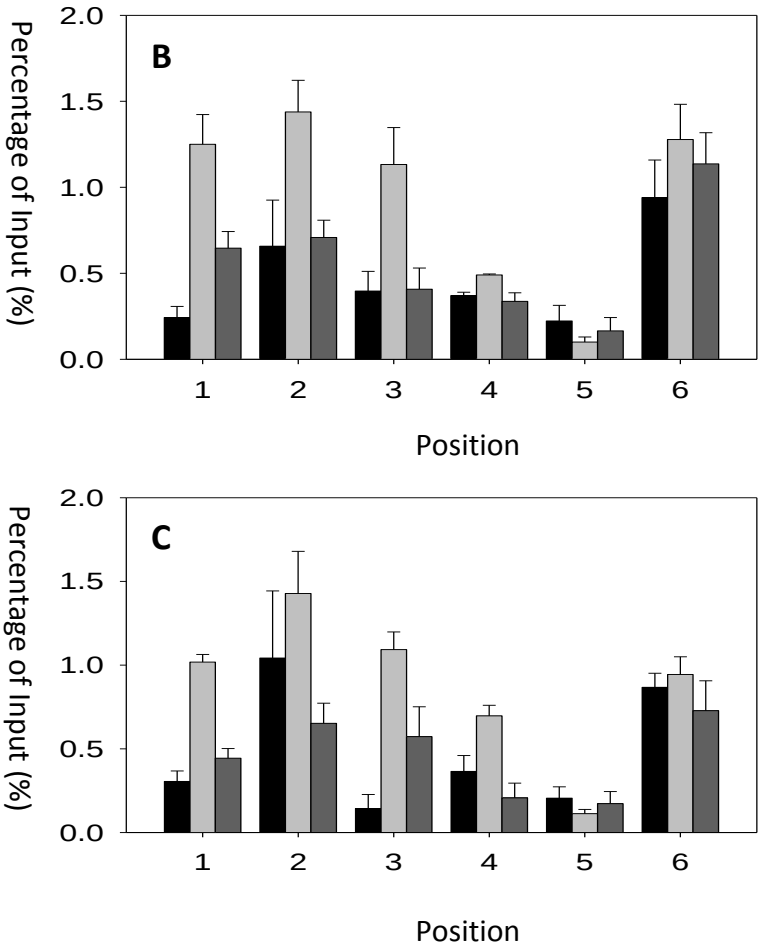


Figure3-3. Repression of lacZ **reporter expression by LexA-Hairy results in inhibition of RNA polymerase II recruitment measured by chromatin immunoprecipiation using antibodies again (A) unphosphorylated, (B) ser-5 phosphorylated, (C) ser-5 and ser-2 phosphorylated forms of RNA polymerase II**. Background levels of Pol II were detected with chromatin prepared from embryos containing only the *lacZ* reporter (black bar). Increased Pol II signals were observed from the embryos containing both the reporter and Gal4 activators (light grey bar). Induction of LexA-Hairy repressor using heat shock decreased Pol II occupancy at loci tested out on the reporter gene (gray bar). Numbers on the Y-axis represent the precipitated materials relative to the percentage of inputs. Numbers on the X-axis represent the positions tested out. Specifically, 1: promoter; 2: +1kb; 3: +2kb; 4: +4kb; 5: one intergenic region on the 3rd chromosome used as negative control; 6: the *hsp70* promoter used as a positive control. All the embryos have been subjected to heat shock treatment in order to eliminate non-specific effects caused by this treatment.

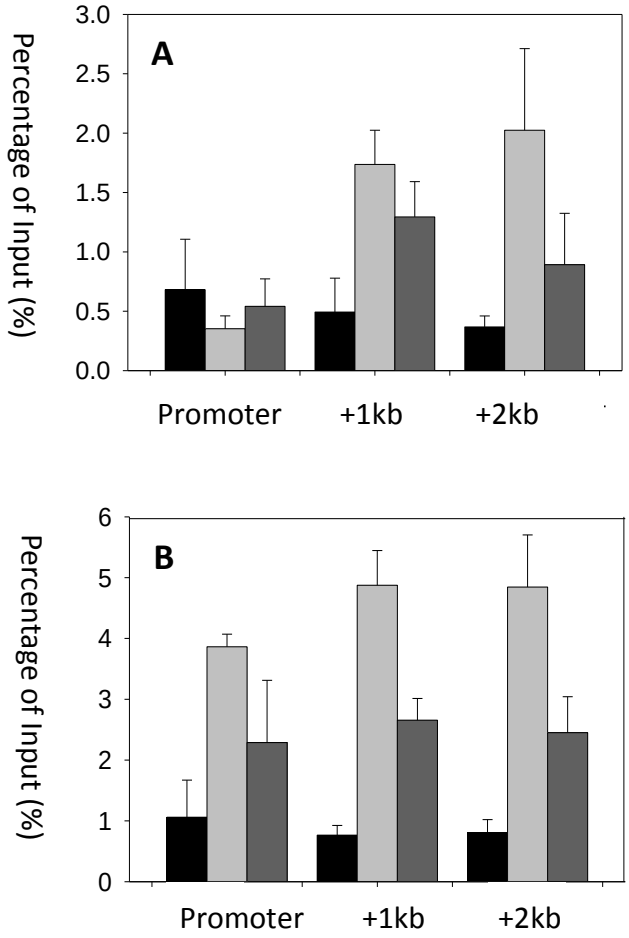


Figure 3-4. Histone H3 lysine 4 and lysine 36 methylation marks track with transcription activity of the *lacZ* reporter measured by chromatin immunoprecipitation using antibodies specific to (B) Histone H3 Lysine 4 and (A) Lysine 36 methylation. A) Gal4 activation of the *lacZ* reporter gene (light grey bar) resulted in a significant increase of H3K36 tri-methylation at +1kb and +2kb of the *lacZ* reporter, but not the promoter compared to the inactivated state (black bar). LexA-Hairy repression decreases H3K36 methylation downstream of the *lacZ* reporter at around +2kb (grey bar). B) Gal4 activation (light grey bar) causes increased H3K4 methylation level at all regions tested for the *lacZ* reporter compared to before activation (black bar). LexA-Hairy repression resulted in decreases of H3K4 methylation level throughout (grey bar).

Density of H2A.Z histone variant tracks with H3 level

Upon both activation and repression of the lacZ reporter, no significant change in the level of H2A.Z was observed at the *lacZ* reporter, although H2A.Z levels at the promoter are enriched compared to an intergenic region (Figure 3-5, and data not shown). H2A.Z may simply mark a promoter for future activation, but Groucho-mediated repression does not deposit more H2A.Z at the promoter, unlike its homolog Tup1 in yeast.

Study of Hairy-mediated long-range repression using the reporter systems

The long-range effect of Hairy was first identified using transgenic embryo assays, in which Hairy was shown to repress the *rhomboid* NEE enhancer and the *twist* 2XPE enhancer when bound 1kb from the nearest activators within these elements [21]. To test whether the LexA-Hairy fusion protein also functions as a long-range repressor, the *twist* 2XPE enhancer was inserted between the binding sites for LexA-Hairy and the *lacZ* reporter. Two spacer DNA sequences derived from the *knirps* and *gfp* ORFs, or a single spacer from *gfp* ORF, were inserted between the LexA sites and the *twist* 2XPE enhancer to separate the repressor binding sites 1.5 kbp or around 750 bp from the enhancer and the *lacZ* reporter (Figure 3-6, A). In the absence of the LexA-Hairy repressor, the 2XPE enhancer drives the expression of lacZ at the ventral region of the blastoderm embryos (Figure 3-6, A).

To test the ability of LexA-Hairy to repress the 2XPE enhancer-activated reporter in transgenic embryos, we expressed LexA-Hairy in a central stripe in embryos using a *Kruppel* driver. LexA-Hairy protein is able to repress a *rhomboid* enhancer-driven lacZ reporter, when bound in proximity of the enhancer (Figure 3-6, A) [1]. LexA-Hairy fusion protein driven by the *Kruppel*

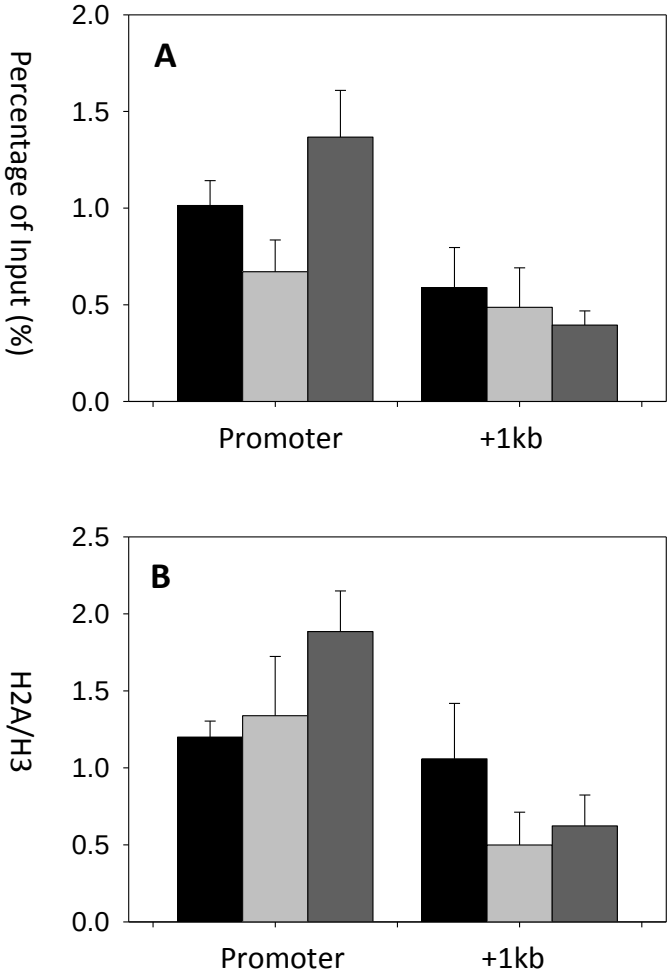


Figure 3-5. Histone variant H2A.Z tracks with histone H3 during activation and repression of the *lacZ* reporter measured by immunoprecipitation using antibody against H2A.Z. (A) Upon activation of the reporter (light grey bar), H2A.Z level decreases compared to the inactivated state of the reporter (black bar). Induction of LexA-Hairy repressor increases the level of H2A.Z (grey bar). This trend is similar to that of histone H3. (B) When normalized to the H3 chromatin IP signals, no significant change was observed upon activation and repression.

enhancer failed to repress *lacZ* expression in the corresponding region of the embryos, whether it was situated from either 1.5 kbp or 750 bp upstream of the 2XPE enhancer (Figure 3-6, A). This result was not expected, as the native Hairy protein can regulate enhancers over similar distances.

Two possibilities may account for the inability of the LexA-Hairy fusion protein to repress the 2XPE. The *Kruppel* driver used to express LexA-Hairy protein expression may not produce sufficient levels of the transgene, or alternatively, the removal of the Hairy DNA binding domain (DBD) abolishes interaction with the dSir2 corepressor, which may be important for Hairy's long-range function.

To increase expression levels of the LexA-Hairy protein, I used the *hsp70A* promoter to drive the ubiquitous expression of LexA-Hairy protein. Upon a 20 minute heat shock, LexA-Hairy protein partially (around 50% of the embryos in the blastoderm stage) repressed the lacZ reporter expression (Figure 3-6, C). This repression is less efficient, compared to the situation when LexA-Hairy binds close to the activators and the transcription start site, in which after 20 minute heat shock *lacZ* expression was repressed in most embryos.

To compare this repression efficacy with the effectiveness of the endogenous Hairy protein, and to test whether dSir2 binding is important for Hairy-mediated long-range repression, a *lacZ* reporter with six Hairy binding sites 1.5kbp upstream of the 2XPE enhancer was used. Only a limited number of blastoderm embryos exhibited repression by endogenous Hairy protein, suggesting that Hairy, for reasons we do not yet understand, is not able to interfere with the 2XPE enhancer when bound 1.5 kbp away 5' of this enhancer (Figure 3-6, B).

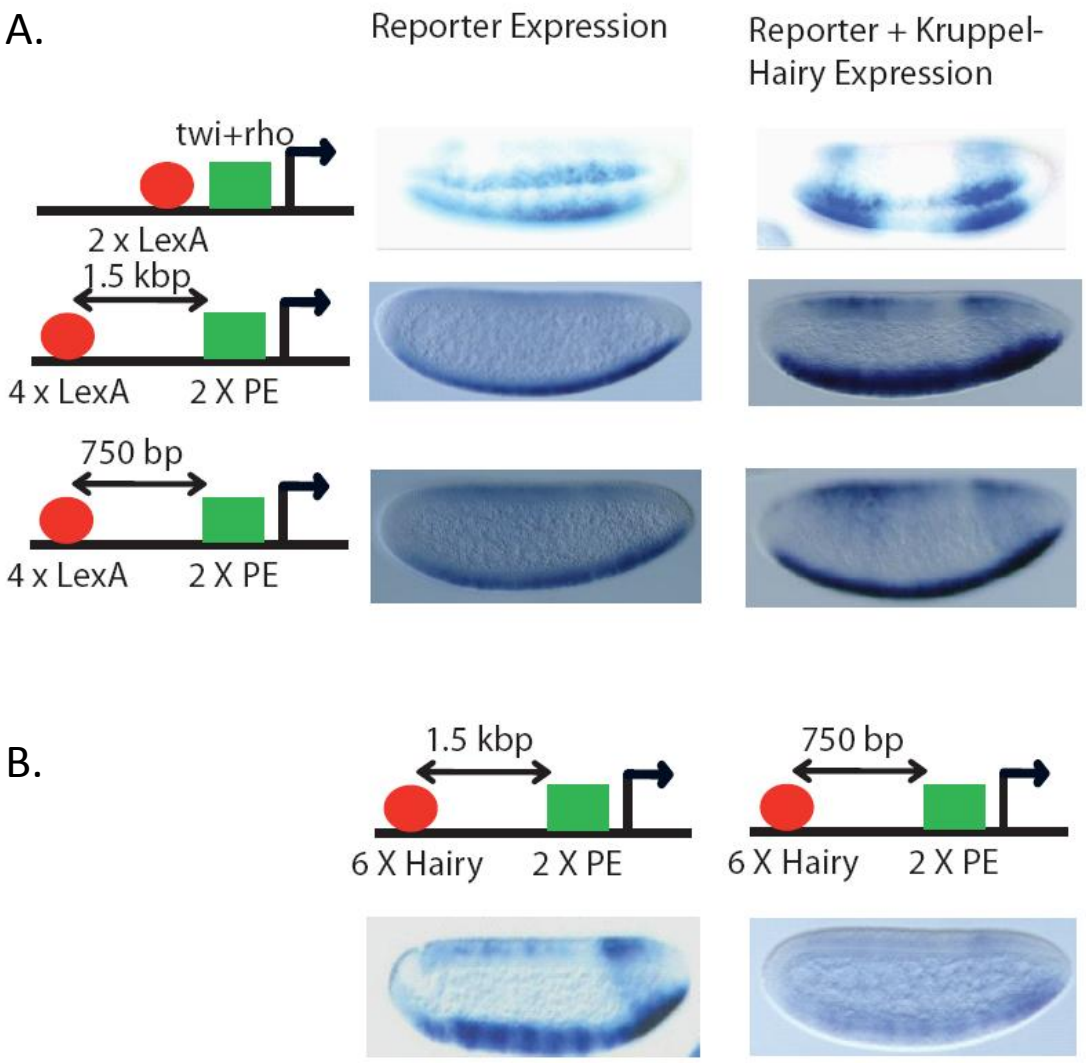


Figure3-6. Repression of twist 2 X PE enhancer by LexA-Hairy or endogenous Hairy protein from a distance. A) LexA-Hairy protein driven by *Kruppel* enhancer fails to repress the 2XPE enhancer from 1.5 kbp or 750 bp. Top panel, lacZ reporter driven by rhomboid (rho) and twist (twi) enhancers with 2 LexA sites nearby is repressed by LexA-Hairy protein driven by a *Kruppel* driver (The figures were adapted from figure 1 of Martinez and Arnosti, 2008). Lower two panels, 2 X PE enhancer drives lacZ reporter express in a single stripe across the ventral region of the embryos. This enhancer was not repressed by LexA-Hairy driven by *Kruppel* enhancer from a distance. (B) Endogenous Hairy protein is able to repress 2XPE driven lacZ reporter from 750bp but not from 1.5 kbp.

Figure 3-6 continued

C.

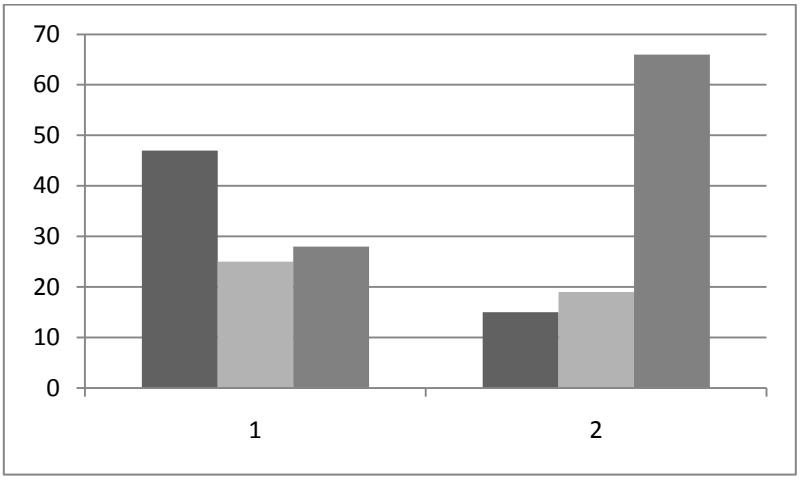


Figure 3-6. Repression of twist 2 X PE enhancer by LexA-Hairy or endogenous Hairy protein from a distance. (C) Quantification of *lacZ* staining pattern in embryos expressing both the lacZ reporter and LexA-Hairy protein under a heat shock driver with no heat shock (1) or 20 minute heat shock (2). After 20 minute heat shock, the number of unstained embryos increased by three-fold. Black: strong staining; light grey: light staining; grey: no staining; Y-axis: numbers of embryos in each category; n=100.

In light of previous reports demonstrating Hairy's activity over long distances, it is puzzling that even the endogenous Hairy protein was unable to function effectively in this transgenic reporter assay. One explanation involves a "hot chromatin" model for Hairy-mediated long-range repression, in which Hairy can only bind active or 'open' enhancers. This means that in order for Hairy to mediate long-range repression, its binding sites need to localize within an active enhancer.

Recruitment of LexA-Hairy fusion protein to the LexA sites 1.5 kbp away from the 2XPE enhancer

In order to test whether the LexA-Hairy fusion protein indeed fails to access the binding sites when there is no activator at the vicinity, we performed chromatin immunoprecipitation to determine the occupancy of LexA-Hairy binding. There is indeed a lower level of LexA-Hairy occupancy at the LexA sites, compared to when there is Gal4 activator binding close by. This result implies an open chromatin environment is more favorable for Hairy binding (Figure 3-7).

I attempted to measure changes in chromatin structure and the RNA polymerase II machinery upon LexA-Hairy repression of the reporter. However, little change was observed (data not shown). This is possibly due to two effects; first, the 2XPE enhancer only drives the expression of *lacZ* in the ventral region of the embryo; second, the repression observed is only less than 50%. To set up a reporter system suitable for study of Hairy mediated repression in a long-range context, enhancers that drive ubiquitous expression of the *lacZ* reporter should be used. Also, activator binding sites might need to be added to the vicinity of the LexA sites in order to open up the chromatin.

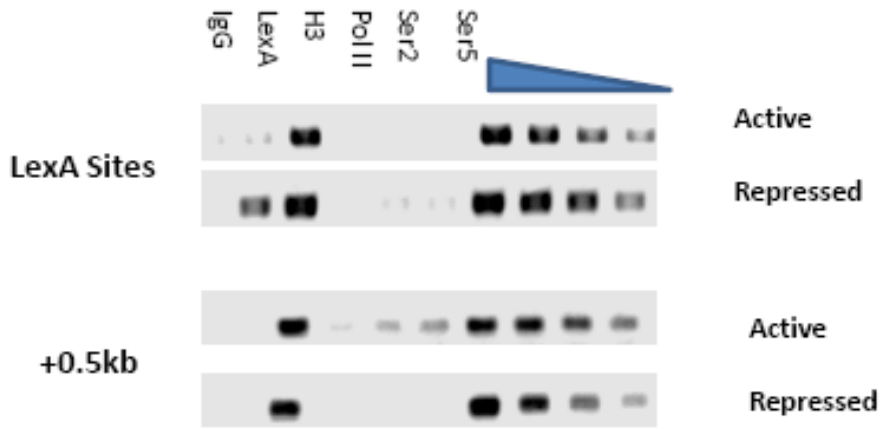


Figure3-7. Closed chromatin environment inhibits LexA-Hairy recruitment. LexA-Hairy fusion protein can still be recruited to the LexA sites in the absence of closeby activators measured by chromatin immunoprecipitation, although the occupancy is at a lower level. Input titration 2%, 1%, 0.5%, 0.25%.

Conclusion

Studies presented here use reporter systems with well-defined activator and repressor binding sites to study the effects of Hairy-mediated repression on the transcription machinery and chromatin structure. When binding to the vicinity of activators and the basal promoter, repression of the reporter by LexA-Hairy inhibits the recruitment of RNA polymerase II and decreases the levels of histone H3 lysine 4 and lysine 36 methylation marks. Chromatin studies in the reporter system not only set the stage for investigation of the endogenous locus described in Chapter II, but also provide complementary information for Hairy repression mechanisms. Both the LexA-Hairy and the endogenous Hairy protein are less potent in repressing a distal *twist* enhancer, suggesting the recruitment of this repressor and possibly its corepressor complex requires an open chromatin environment.

APPENDICES

Appendix B

Cofactor contributions to Hairy Repression Activities

Introduction

Hairy and its multiple corepressors

Given the important roles of Hairy protein during various biological processes, many studies have looked into its repression mechanism. Groucho was the first purified as a Hairy corepressor recruited by the WRPW motif, and was shown to be important for Hairy function [22-24]. Hairy also recruits short-range corepressor CtBP through the PLSLV motif. CtBP maternal depletion suppresses h mutant phenotype, and CtBP is required for Hairy repression in a cell-type specific manner [25, 26]. In ectopic expression assays, CtBP antagonize Groucho-dependent Hairy repression of *slx*, *fkh*, *hkb* and *tll* [27]. dSir2 protein also interacts physically and genetically with Hairy. This interaction maps to a highly conserved "RRAR" motif in the basic region of the Hairy DNA binding domain [28]. Another critical motif for Hairy function is the Orange domain, which is required for rescue of *hb-scute* male lethality [29]. DamID mapping studies suggest that dSir2, CtBP and Groucho corepressors differentially overlap with Hairy binding, suggesting that Hairy might utilize these different corepressors in a context-dependent manner [30].

Requirement of multiple corepressors is a common theme in repression

It is frequently observed that a single repressor uses multiple corepressors for repression in eukaryotes. Drosophila repressor Brinker requires both CtBP and Groucho for maximal and versatile repression during the Dpp signaling. The conclusion is based on analysis of transcriptional outcomes arising from the genetic removal of these corepressors, and from

ectopically expressing Brk variants in the embryo [31]. A new corepressor dNAB for Brinker was recently indentified, which is required for regulation of dMyc expression [32]. Besides Groucho binding motif, repressor Runt repressor also has other repression domains [33].

Although previous studies suggest that Hairy is likely to deploy different corepressors in a context-dependent manner, the specific contexts where particular repressors are used are still not clear. Also, most studies of Hairy repression mechanism focus on Drosophila embryonic stage and cell culture system; it is not clear how these corepressors contribute to Hairy repression during other developmental stages? In order to understand the general rules of corepressor usage, it is necessary to analyze Hairy repression at a genome-wide level.

We plan to use previous established Hairy ectopic expression assays to investigate the contribution of different corepressors in Hairy repression, by overexpressing mutant versions of Hairy that does bind to different corepressors. The purpose is to set up a platform that can be used for future analysis at a genome-wide level. We focus the assays on the role of Groucho corepressor in Hairy mediated repression, as previous studies suggested Groucho accounts for most of Hairy repression activity.

Materials and Methods:

Constructs for protein overexpression

Heat shock inducible wild-type Hairy and mutant Hairy that does not bind to Groucho constructs were generated as described in Chapter II. Heat shock inducible Orange domain mutant Hairy (NEVSR/AAAAA) was generated through site-directed mutagenesis using primer DA2248 (5' -GCCGGATTCGCCGACTGTGTGGCCGCGGCTGCCGCCTTTCCCGGCATCGAGCCCG-3')/DA2249 (5'- CGGGCTCGATGCCGGGAAAGGCGGCAGCCGCGGCCACACAGTCGGCGAATCCGGC-3') using wild-type Hairy plasmid as a template. Heat shock inducible Groucho plasmid was generated using modified pCasper plasmid used for Knirps overexpression in previous studies [34]. Full length *groucho* cDNA was amplified using primer DA1580 (5'-CGTCTAGAATAAATAACTTCGTAGAC-3')/DA1579 (5'-ATATGGTACCATGTATCCCTCACCGGTG-3') from Pet17b-Gro used in previous studies [22].

Crude embryo lysate preparation

Approximately 50 mg of dechrionated embryos were resuspended in 1.2 ml of lysis buffer (25 mM HEPES pH 7.9, 150 mM NaCl, 1 mM DDT, 1 mM PMSF, 1 mM Na-metabisulfite, 1 mM benzamidine, 10 uM pepstatin A) and disrupted by sonication using a Branson Sonifier 250 (2 cycles of 12 pulses each, output 3, duty cycle 60%). After sonication, lysates were centrifuged for 15 minutes at 14,000 rpm using an Eppendorf centrifuge, and the protein centrifuge of the supernatant was determined using the Bradford assay, with BSA as the standard.

Western blot analysis

Immunoblotting was performed according to standard protocols using a tank transfer system (Mini Trans-Blot Cell, BioRad). Sequi-BlotTM PVDF membranes (BioRad) were used and antibody incubation was in TBST(20 mM Tris-HCl, pH 7.5, 120 mM NaCl, 0.1% Tween-20) supplemented with 5% (w/v) nonfat dry milk as blocking agent. The primary anti Hairy monoclonal antibody (Abcam, catalog #: ab20165) was used at 1:5,000 dilution incubated at 4 °C for overnight. The primary anti Tubulin antibody (Hybridoma Bank) was used at 1:10,000 dilution. The secondary ImmunoPure Goat Anti-Mouse HRP-conjugated antibody (Pierce) was used at 1:10,000 dilution incubated at room temperature for 3 hour. The blot is detected using SuperSignal Western Blotting Kit (Pierce), and visualized using Fuji LAS-1000 imager.

Quantification of mRNA expression levels

hairy, ftz, ac, prd, stg mRNA expression levels were quantified through real time PCR analysis. Actin5C mRNA expression level, whose expression level is not affected by Hairy overexpression, was used as an internal control. Primers used for quantification of mRNA expression levels are as followed: DA2013(5'-TCCAAAACCGACGCATGAA-3')/DA2014(5'-GTGTAGCCGGCACCACAGT-3') for *ftz*; DA2005 (5'-ACGCCTGGCACCAAGAGTT-3')/DA2006 (5'-CGGAATCGTCGATGTTGCT-3') for *ac*; DA2049 (5'-GGATCCAGCCACTGTGTACCA-3')/DA2050 (5'-GGCGATAGATTACTACTGGCCATT-3') for *prd*; DA2019 (5'-ACCAACGCAACACATAACACAAC-3')/DA2020 (5'-CGGCATAATTACCTGATATTGCAAT-3') for *stg*.

***In situ* staining of mRNA expression pattern**

Embryos were fixed for *in situ* hybridization and stained using anti-digoxigenin-UTP-label RNA probes for targeted genes as previously described [17]. Template from making ac probe was

purchase from Drosophila Genome Resource Center (Stock #: IP01413). The template plasmid was digested using BamH1, and the digoxigenine-labeled RNA was synthesis using SP6 polymerase (NEW ENGLAND BioLabs, Catalog# M0207S) following standard protocol.

Results and Discussions:

Groucho cofactor and Hairy-mediated repression

Mutant Hairy protein that does not bind to Groucho accumulates to a lower level than wild-type Hairy upon heat shock induction

In order to make direct comparison of the repression activity between wild-type and Groucho-binding deficient Hairy, we first compared the mRNA and protein induction levels of Hairy protein. Upon 20 minute induction, wild-type and mutant version of Hairy expressed at compared levels, suggesting positions of the inserted heat shock inducible transgenes do not affect the transcription levels of these two versions of Hairy proteins (Figure B-1, A). However, Hairy protein that does not bind to Groucho reproducibly expresses at a lower level when compared to the wild-type protein. This might suggest that binding to the Groucho cofactor is important for the stability of the Hairy repressor (Figure B-1, B). Alternatively, the Hairy protein is measured using an antibody that is raised using a GST-tag full-length Hairy protein. It is possible that this antibody is less efficient in detecting the mutant Hairy protein.

Hairy repression of *fushi tarazu* is dependent on Groucho corepressor

As described in Chapter II, Hairy repression of *ftz* is dependent on Groucho, and that overexpression of Hairy mutant protein that does not bind to Groucho failed to ectopically

repress the seven-stripe expression pattern of *ftz*. We quantified the mRNA expression of *ftz* upon induction of these two versions of Hairy, and found that only overexpression of wild-type Hairy protein resulted decreased *ftz* mRNA level, consistent with the results from the *in situ* hybridization analysis (Figure B-2).

Other Hairy targets in the blastoderm embryos

To investigate Hairy targets other than *fushi tarazu* in early embryos, we analyzed expression patterns of several important early Drosophila genes. DamID studies suggested that Hairy binds close to *acheate*, *egghead* and *paired* genes in early embryos. Indeed, ectopic expression of wild-type Hairy protein resulted in complete repression of *acheate* and *egghead*, as well as repression of parasegments of *paired*. Hairy overexpression also disrupts *even skipped* expression pattern. This regulation might be indirect, as endogenous Hairy does not occupy this locus according to a recent genome-wide ChIP-chip study [35]. On the other hand, Hairy expression does not affect the expression levels of maternal and gap genes, such as *bicoid* , *caudal* (Caudal protein level before and after induction of Hairy repression was compared by western blotting, data not shown) and *Kruppel* (Figure B-3).

Hairy might use Groucho repressor in a context-dependent manner

We further analyzed the role of Groucho corepressor during Hairy repression of other targets. Overexpression of Hairy decreased *paired* expression level by around two-fold. Overexpression of the mutant version of Hairy also represses *paired* expression with similar efficiency (Figure B-4, A). This indicates that there is Groucho-independent Hairy repression activity.

Analysis of *acheate* expression upon induction of Hairy resulted in similar observation, that both versions of Hairy repress *acheate* with similar efficiency, although even overexpression of wild-type Hairy resulted in a really minor ($p = 0.15$) change in *acheate* mRNA expression level (Figure B-4, B). This might be caused by the fact that *acheate* is only expressed in a few nuclei in the blastoderm embryos, and the assay used is not sensitive enough to detect the difference (Figure B-3, 3rd panel, right).

Role of Orange domain in Hairy repression of *fushi tarazu*

Other than the WRPW motif that recruits Groucho, Orange domain is another motif that is highly conserved among Hairy and Enhancer of Split family members. A five-amino acid deletion in the orange domain (ΔNEVSR) rendered Hairy unable to rescue *hb-scute* lethality [29]. In order to test the role of orange domain in Hairy-mediated transcriptional repression, we generated flies expressing mutant Hairy with NEVSR motif mutated to Alanines under a heat shock driver. The Orange domain mutant Hairy can be induced to similar levels as the wild-type Hairy protein in adult flies after 20 minute heat shock induction (Figure B-5, A). Different from the mutant Hairy protein that does not bind to Groucho, this Orange domain mutant Hairy can still disrupt *fushi tarazu* expression pattern, although less efficient than the wild type. After 20 minute heat shock induction, 39% of the embryos showed normal fushi tarazu overexpression pattern. Among the embryos that were repressed by Hairy overexpression, most of them have only stripe 4 and 7 repressed (Figure B-5, B). This result suggests that although Groucho is indispensable in repression of fushi tarazu expression by Hairy, Orange domain also plays an important role during this process. Stripe 4 and 7 are the stripes that are most sensitive to Hairy repression, with Hairy binding likely to be closest to their enhancers. Orange domain

mutant Hairy is most efficient in regulating these stripes, suggesting that this mutant is only deficient in repressing distal enhancers. Although it is quite speculative, it is possible that Orange domain is required for Groucho to adapt a conformation that is suitable for long-range repression.

Characterization of heat shock inducible Histidine and Flag double-tagged Groucho protein in embryos

To provide direct evidence for the involvement of Groucho spreading in Hairy-mediated long-range repression, it is necessary to directly chromatin immunoprecipitate Groucho onto Hairy endogenous targets. However, attempts to ChIP Groucho using available antibodies have not been successful (data not shown). One possibility is that the qualities of the antibodies are not good enough for chromatin immunoprecipitation. We overexpressed a Histidine and Flag double-taqqed Groucho protein under a heat shock inducible *hsp70A* promoter, in order to provide epitopes that can be precipitated by other antibodies.

We first examined the effects of Groucho overexpression on *fushi tarazu*. To our surprise, Groucho overexpression did not result in obvious disruption of *ftz* expression pattern (data not shown). It is possible that there is not enough endogenous Hairy protein to recruit ectopically expressed Groucho to the *ftz* locus. We then overexpressed the Groucho protein together with wild type Hairy protein, and observed that not only *ftz* expression was repressed; the development of the embryo was also disrupted (data not shown). This suggested that once recruited to the chromatin, the double-tagged Groucho is functional. However, attempts to chromatin immunoprecipitates Groucho using antibodies against either the Histidine or Flag epitopes have not been successful (data not shown). The N-terminal of Groucho protein is

important for its oligomerization, and the C-terminal WD40 motif is important for Groucho interaction with the repressors. It is possible that the epitopes at these two ends are masked by protein-protein interaction.

We also used this overexpression lines to examine the size of Groucho corepressor complex. Overnight embryos were collected and subjected to 20 minute heat shock. Groucho protein was first purified over a Ni^{2+}-NTA matrix and then analyzed by gel filtration chromatography. When fractions were analyzed by immunoblotting, Groucho was observed to migrate in fractions with size larger than 158 KDa, with more intense signals at fractions bigger than 670 kDa. Previous studies suggested Groucho presents in the form of tetramers, our result implied that Groucho can also present in smaller size complex in *Drosphila* embryos.

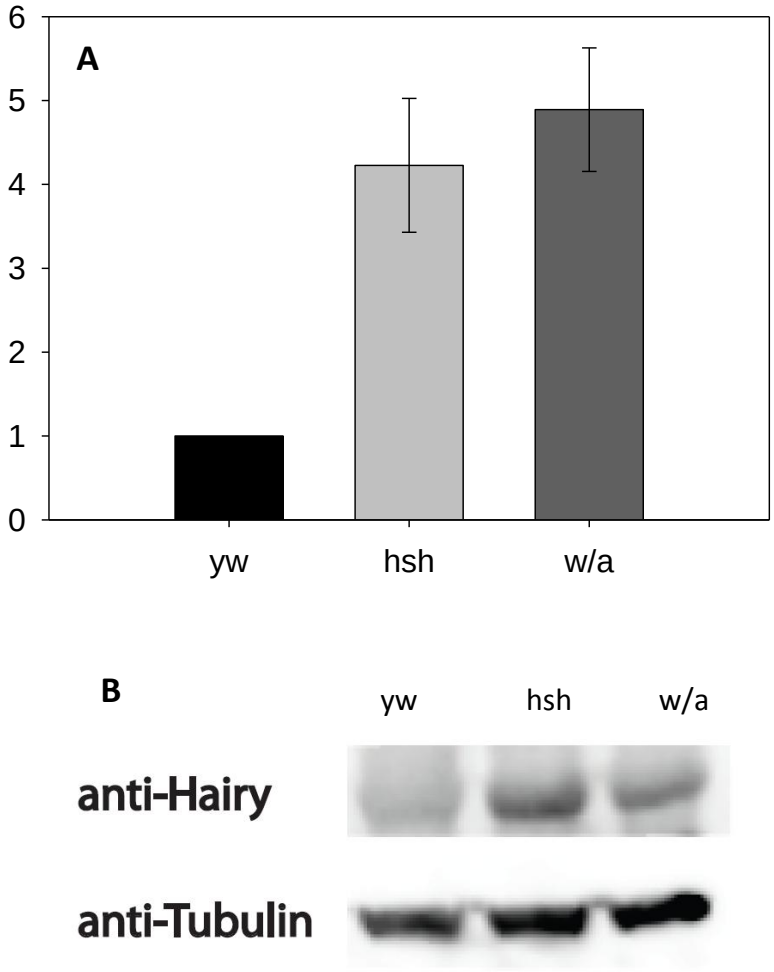


Figure B-1. Mutant Hairy accumulates at a lower level than the wild-type Hairy protein, despite that the mRNA levels are comparable. A) mRNA expression level of Hairy in embryos without transgene (yw), with heat shock inducible wild type hairy transgene (hsh), or mutant hairy transgene that does not bind to Groucho cofactor (w/a). 20 minute heatshock increases the hairy mRNA levels by around 4 folds in the embryos containing heat shock inducible transgenes. Relative mRNA levels were calculated using ΔCt method. The hairy mRNA levels from different samples were normalized to hairy mRNA levels in the embryos containing no transgene. B) Hairy protein level in the embryos containing mutant hairy transgene (w/a) is higher than that in the embryos without transgene (yw), and lower than that in the embryos containing wild type hairy transgene (hsh). Hairy protein is measured by an antibody that is raised using a GST-tag full-length Hairy protein. Tubulin level is as an internal control to make sure equal protein loading.

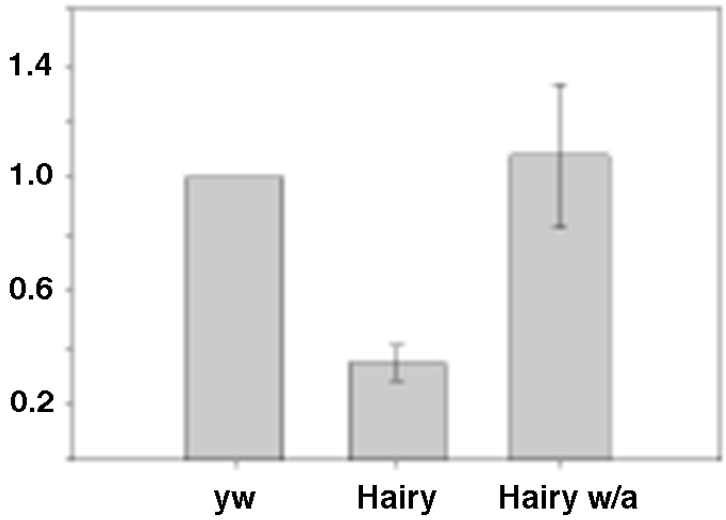


Figure B-2. Abundance of *ftz* mRNA in embryos with overexpressed wild-type Hairy or mutant Hairy lacking the Groucho-binding WRPW motif (Hairy w/a) relative to embryos lacking the Hairy transgne (yw). mRNA was extracted from 2-3 hour embryos that were exposed to 20 minute heat shock with no recovery period. The y axis represents relative levels of *ftz* mRNA, which were calculated by ΔCt method.

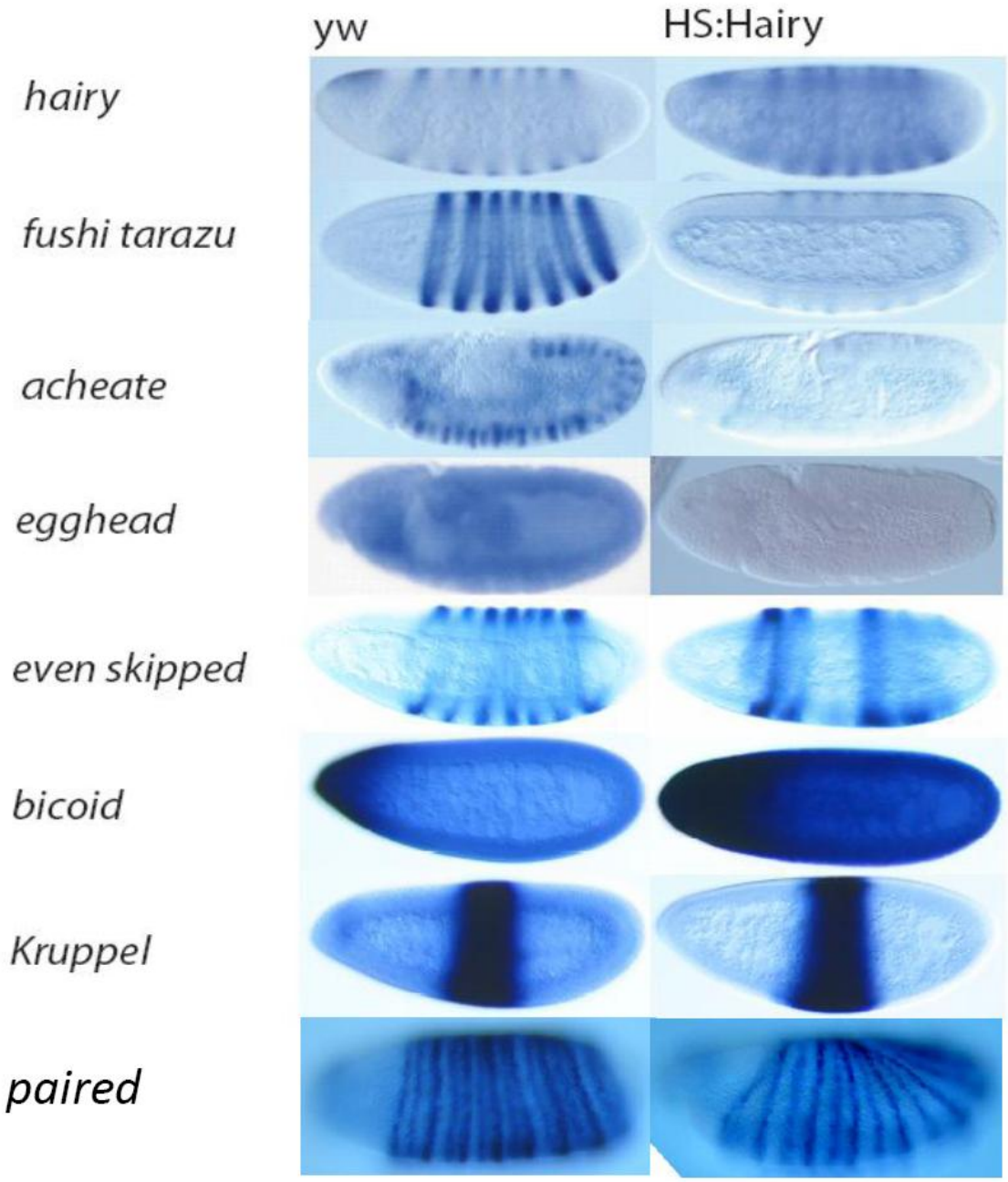
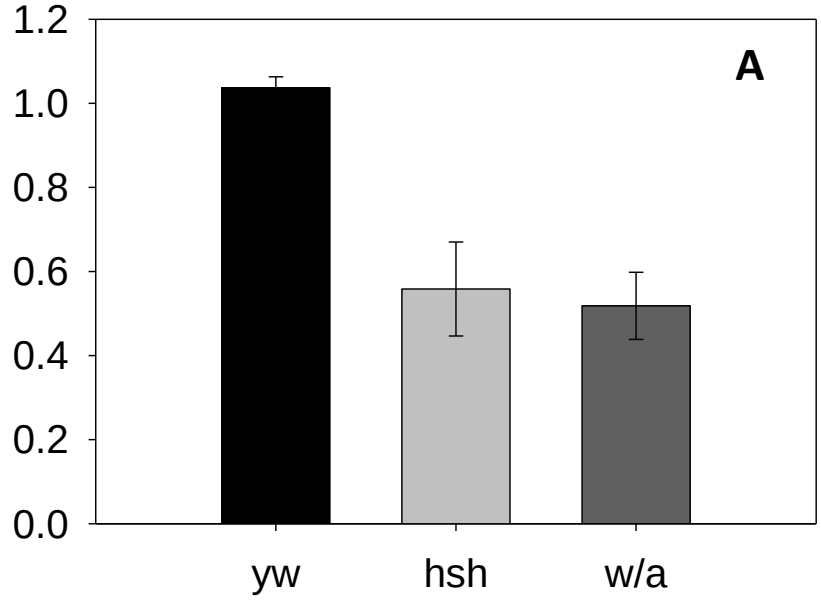


Figure B-3. Expression patterns of early Drosophila genes before and after 20 minute induction of Hairy protein analyzed by *in situ* hybridization.



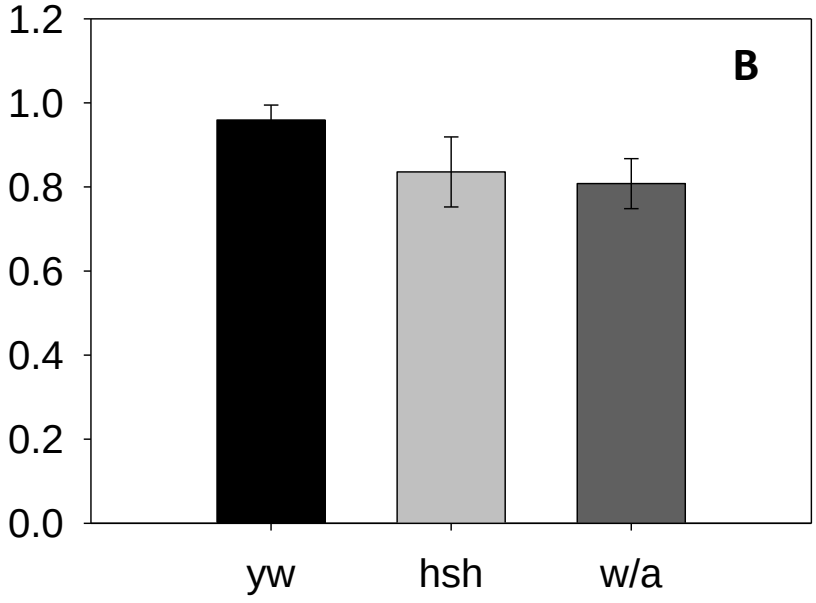


Figure B-4. Abundance of *paird and achaete* mRNA in embryos with overexpressed wild-type Hairy or mutant Hairy lacking the Groucho-binding WRPW motif (Hairy w/a) relative to embryos lacking the Hairy transgne (yw). mRNA was extracted from 2-3 hour embryos that were exposed to 20 minute heat shock with no recovery period. The y axis represents relative levels of mRNA, which were calculated by ΔCt method. (A) *paired* mRNA level; (B) *achaete* mRNA level.

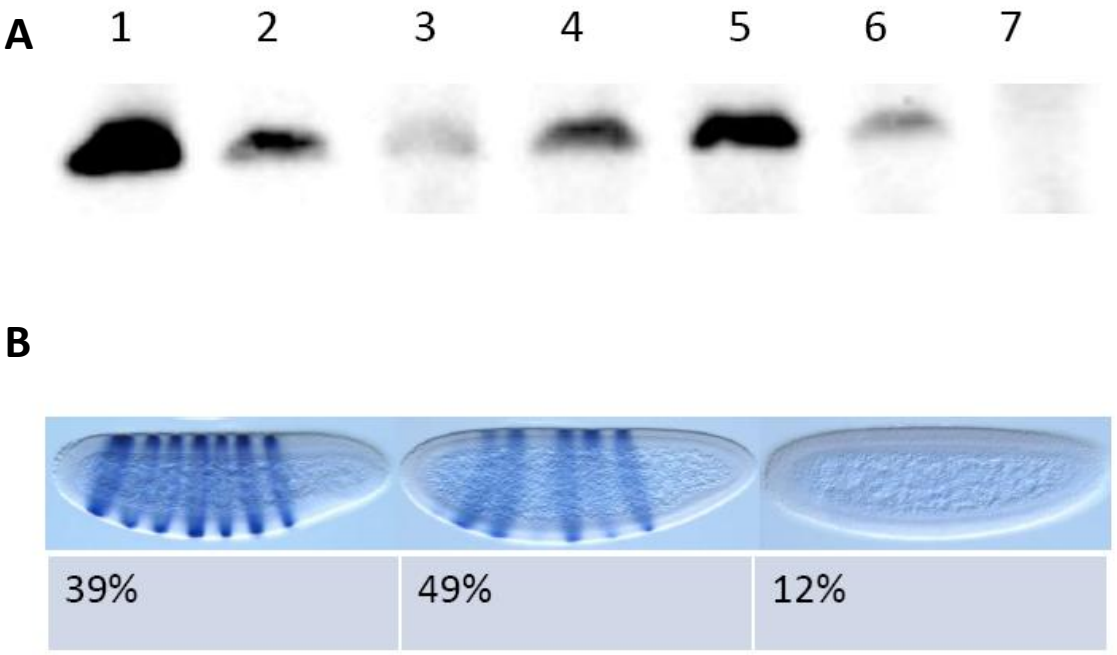


Figure B-5.Mutation in the Orange domain rendered Hairy less efficient in repression of *ftz*.
(A)Protein levels of Hairy in adult flies containing no transgene (lane 7), heat shock inducible wild type Hairy (lanes 5,6), mutant Hairy that does not bind to Grouch (lanes 3,4) and Hairy bearing Orange domain mutations (lanes 1,2) measured by western blotting using anti-Hairy monoclonal antibody. All the transgenic lines expressed Hairy protein to a higher level. (B) *fushi tarazu* mRNA expression pattern upon overexpression of Hairy bearing Orange domain mutations in blastoderm embryos analyzed by *in situ* hybridization. Number below the images showed the percentage of embryos showing according expression patterns (n=100).

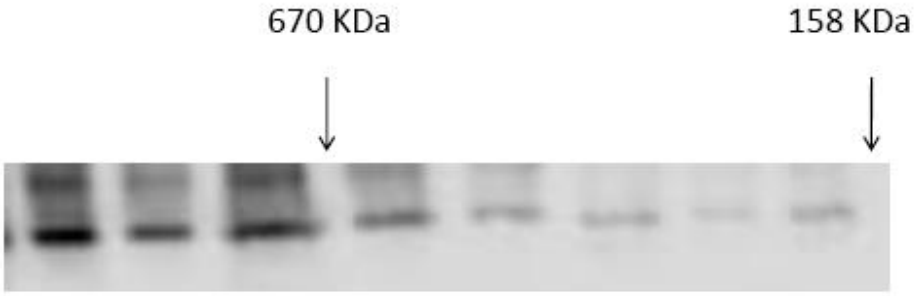


Figure B-6. Groucho migrates in molecular masses from around 158 KDa to above 670 KDa. Peak fractions from Superdex 200 size exclusion experiments with purified Groucho protein were resolved on SDS-PAGE gels and western blotted with anti-M2 antibody (to detect FLAG-tagged purified proteins). Groucho was found in molecular mass that is around 158 KDa, suggesting that other than tetramers, Groucho can also present in monomeric in a complex with other proteins.

References

References

1. Martinez, C.A., and Arnosti, D.N. (2008). Spreading of a corepressor linked to action of long-range repressor hairy. Mol Cell Biol *28*, 2792-2802.

2. Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. Cell *120*, 169-181.

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

4. Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

5. Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434-438.

6. Shi, Y., and Whetstine, J.R. (2007). Dynamic regulation of histone lysine methylation by demethylases. Mol Cell *25*, 1-14.

7. Shi, X., and Gozani, O. (2005). The fellowships of the INGs. J Cell Biochem *96*, 1127-1136.

8. Li, B., Gogol, M., Carey, M., Lee, D., Seidel, C., and Workman, J.L. (2007). Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. Science *316*, 1050-1054.

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

10. Jackson, J.D., and Gorovsky, M.A. (2000). Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. Nucleic Acids Res *28*, 3811-3816.

11. 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.

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

13. Zhang, H., Roberts, D.N., and Cairns, B.R. (2005). Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. Cell *123*, 219-231.

14. Guillemette, B., Bataille, A.R., Gevry, N., Adam, M., Blanchette, M., Robert, F., and Gaudreau, L. (2005). Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. PLoS Biol *3*, e384.

15. Leach, T.J., Mazzeo, M., Chotkowski, H.L., Madigan, J.P., Wotring, M.G., and Glaser, R.L. (2000). Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. J Biol Chem *275*, 23267-23272.

16. Gligoris, T., Thireos, G., and Tzamarias, D. (2007). The Tup1 corepressor directs Htz1 deposition at a specific promoter nucleosome marking the GAL1 gene for rapid activation. Mol Cell Biol *27*, 4198-4205.

17. Small, S., Blair, A., and Levine, M. (1992). Regulation of even-skipped stripe 2 in the Drosophila embryo. EMBO J *11*, 4047-4057.

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

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

20. Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell *123*, 593-605.

21. Barolo, S., and Levine, M. (1997). hairy mediates dominant repression in the Drosophila embryo. EMBO J *16*, 2883-2891.

22. Paroush, Z., Finley, R.L., Jr., Kidd, T., Wainwright, S.M., Ingham, P.W., Brent, R., and Ish-Horowicz, D. (1994). Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. Cell *79*, 805-815.

23. Fisher, A.L., Ohsako, S., and Caudy, M. (1996). The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein-protein interaction domain. Mol Cell Biol *16*, 2670-2677.

24. Jimenez, G., Paroush, Z., and Ish-Horowicz, D. (1997). Groucho acts as a corepressor for a subset of negative regulators, including Hairy and Engrailed. Genes Dev *11*, 3072-3082.

25. Poortinga, G., Watanabe, M., and Parkhurst, S.M. (1998). Drosophila CtBP: a Hairy-interacting protein required for embryonic segmentation and hairy-mediated transcriptional repression. EMBO J *17*, 2067-2078.

26. Phippen, T.M., Sweigart, A.L., Moniwa, M., Krumm, A., Davie, J.R., and Parkhurst, S.M. (2000). Drosophila C-terminal binding protein functions as a context-dependent transcriptional co-factor and interferes with both mad and groucho transcriptional repression. J Biol Chem *275*, 37628-37637.

27. Zhang, H., and Levine, M. (1999). Groucho and dCtBP mediate separate pathways of transcriptional repression in the Drosophila embryo. Proc Natl Acad Sci U S A *96*, 535-540.

28. Rosenberg, M.I., and Parkhurst, S.M. (2002). Drosophila Sir2 is required for heterochromatic silencing and by euchromatic Hairy/E(Spl) bHLH repressors in segmentation and sex determination. Cell *109*, 447-458.

29. Dawson, S.R., Turner, D.L., Weintraub, H., and Parkhurst, S.M. (1995). Specificity for the hairy/enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. Mol Cell Biol *15*, 6923-6931.

30. Bianchi-Frias, D., Orian, A., Delrow, J.J., Vazquez, J., Rosales-Nieves, A.E., and Parkhurst, S.M. (2004). Hairy transcriptional repression targets and cofactor recruitment in Drosophila. PLoS Biol *2*, E178.

31. Hasson, P., Muller, B., Basler, K., and Paroush, Z. (2001). Brinker requires two corepressors for maximal and versatile repression in Dpp signalling. EMBO J *20*, 5725-5736.

32. Ziv, O., Suissa, Y., Neuman, H., Dinur, T., Geuking, P., Rhiner, C., Portela, M., Lolo, F., Moreno, E., and Gerlitz, O. (2009). The co-regulator dNAB interacts with Brinker to eliminate cells with reduced Dpp signaling. Development *136*, 1137-1145.

33. Walrad, P.B., Hang, S., Joseph, G.S., Salas, J., and Gergen, J.P. Distinct contributions of conserved modules to runt transcription factor activity. Mol Biol Cell *21*, 2315-2326.

34. Struffi, P., Corado, M., Kulkarni, M., and Arnosti, D.N. (2004). Quantitative contributions of CtBP-dependent and -independent repression activities of Knirps. Development *131*, 2419-2429.

35. MacArthur, S., Li, X.Y., Li, J., Brown, J.B., Chu, H.C., Zeng, L., Grondona, B.P., Hechmer, A., Simirenko, L., Keranen, S.V., 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.

Chapter IV

Conclusions and Future Directions

Conclusions

Studies presented here provide insights into mechanisms mediated by two major classes of repressors that are conserved throughout the metazoans. Repression mechanisms at the chromatin level not only confirm that there are indeed two distinct modes of repression during early *Drosophila* development, but also suggest two different ways in which repressors bound to distal enhancers can effect transcription at the basal promoter. Key discoveries of this thesis are summarized as follows.

Secondary pair-rule genes are driven by stripe-specific enhancers

My studies suggest that Hairy mediates progressive repression on *fushi tarazu*. This has important implications for the Hairy repression mechanism and the regulation of secondary pair-rule gene expression. Secondary pair-rule genes like *fushi tarazu* have long been considered to be driven by a single enhancer [1]. The fact that different stripes of *ftz* can be repressed in a step-wise manner suggests that there are distinct modular enhancers underlying the seven-stripe expression pattern, similar to the case of primary pair-rule genes like *even skipped*.

Mechanisms of Hairy-mediated long-range repression

The progressive repression of *ftz* mediated by Hairy suggests that Hairy is not directly targeting the basal promoter, in which case all the modular enhancers would be switched off at the same

time. A short induction of the repressor protein followed by recovery time resulted in complete repression of the *ftz* locus. This phenomenon might reflect the kinetics by which Groucho corepressor spreads to repress the whole *ftz* locus.

Hairy-mediated long-range repression involves a "spreading" mechanism. Chromatin analysis suggests that extensive repression of the *ftz* locus by Hairy resulted in wide-spread histone deacetylation, and a mild induction of Hairy repression resulted in localized histone deacetylation around the enhancer region. This implicates that Hairy action first nucleates around its binding sites, and then extends to the whole locus, probably through Groucho corepressor spreading. Repression of *ftz* by Hairy resulted in the inhibition of RNA polymerase II recruitment, unlike in the case of many other developmental regulated genes having a paused RNA polymerase around the basal promoter region (Figure 4-1).

On a reporter gene repressed by Hairy, similar changes of RNA polymerase II occupancy and histone modifications were observed, suggesting these observations are likely to be the general rules of Hairy repression other than gene-specific effects [2].

Analysis of the reporter system also suggests that the levels of histone H3 lysine 4 and lysine 36 marks track with the transcription levels during repression, whose mechanism is not clear. The decrease may result from binding of histone demethylase to the Hairy corepressor complex; alternatively, reduction in activation marks could be just a passive effect of decreased transcription activity.

When the Hairy sites are moved 1.5 kbp away from the closest enhancer. Hairy is a less potent repressor. This observation supports a “hot chromatin” model in which recruitment of Hairy and its corepressors is favored by an open chromatin environment.

Mechanisms of Knirps-mediated short-range repression

Knirps-mediated repression resulted in local changes of chromatin structure. Micrococcal nuclease mapping and chromatin immunoprecipitation reveals that Knirps mediated repression causes a local increase of histone density. The increased histone density might result from Knirps inhibiting the disassembly of the nucleosome or Knirps recruiting factors that deposit nucleosomes around its binding region. These results suggest that in addition to previously proposed models, short-range repression can happen through chromatin-related mechanisms. Knirps repression inhibits the recruitment of Caudal but not Bicoid activator, suggesting Knirps’ effects on activators might be target-dependent. Knirps does not affect the recruitment of RNA polymerase II, though it is still not clear at which downstream step Knirps represses RNA polymerase II (Figure 4-1).

Future Directions

There are several possible future directions that will increase our knowledge of the mechanisms of long- and short-range repression.

Changes in chromatin structure upon repression at a genome-wide level

As more and more transcription factor binding patterns have been profiled, one emerging question is how to define a functional binding event. Elucidation of chromatin features

associated with repression would help to define functional repressor binding sites. The chromatin studies for long- and short-range repression have been performed only on the canonical targets *fushi tarazu* and *even skipped*. To understand whether the observations are

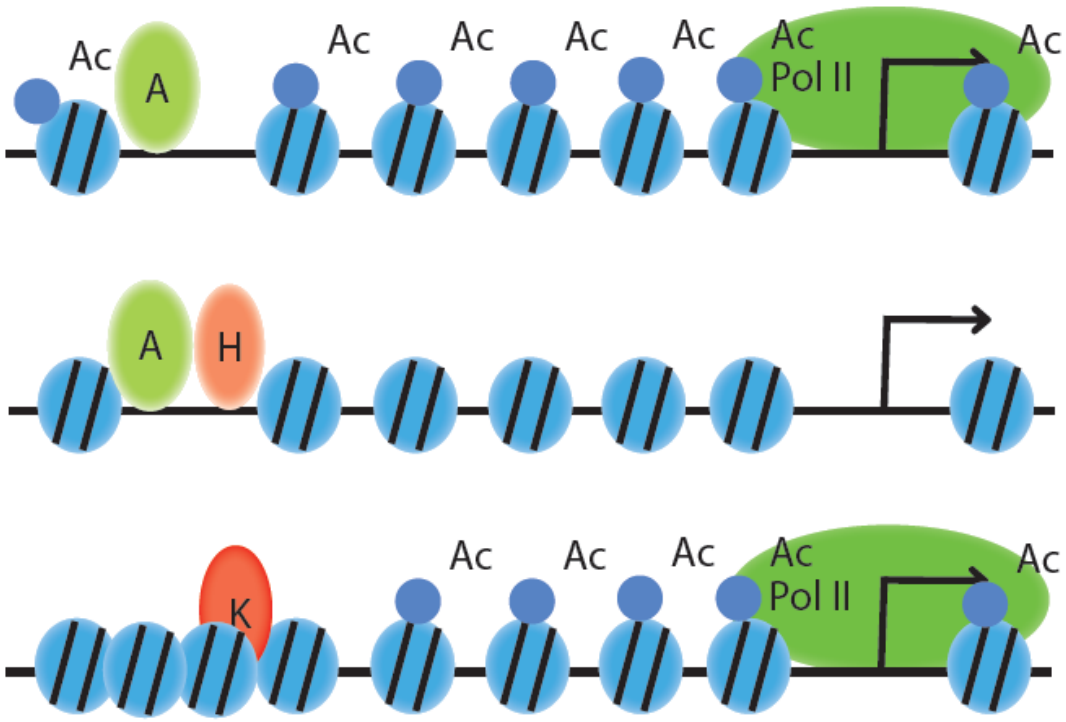


Figure 4-1. Mechanisms of Hairy and Knirps-mediated transcriptional repression. The top panel is a schematic representation of an active gene. Middle panel: Hairy (H) mediated repression causes wide-spread histone deacetylation and inhibition of RNA polymerase II (Pol II) recruitment, without affecting the recruitment of the activator (A). Lower panel: Knirps (K) locally increases histone density and causes histone deacetylation. Knirps inhibits activator recruitment in a context-dependent manner, and affects RNA polymerase II at the steps after recruitment.

target-specific or generally representative, it is necessary to perform similar chromatin studies at a genome-wide scale using techniques like ChIP-chip or ChIP-seq.

How do corepressors dictate context-dependent effects?

Hairy and Knirps recruit similar corepressor complexes including both CtBP and Groucho corepressors, but mediate distinct repression output and chromatin effects [3, 4]. It would be interesting to investigate the mechanisms underlying context-dependent effects of these corepressors. Hairy and Knirps recruit Groucho through different motifs. Because of differences in interaction surfaces, Groucho may adopt alternative configurations when recruited by different transcription factors, so that oligomerization and spreading is only possible when associated with a long-range repressor. This hypothesis would suggest that chimeric repressor proteins carrying suitable interaction surfaces would adopt short- or long-range activities, regardless of the rest of the protein's structure. Alternatively, as noted in previous chapters that Groucho is phosphorylated in response to signaling pathways, so differently modified forms of the protein may exhibit different abilities to oligomerize, affecting short- or long-range activity. Or, as with other corepressor proteins, Groucho may assemble into distinct complexes with different abilities to mediate spreading, and these complexes may be separately recruited by short- or long-range repressors. This hypothesis can be tested by determining the forms of Groucho recruited to different corepressor complexes by chromatin immunoprecipitation (Figure 4-2) [5].

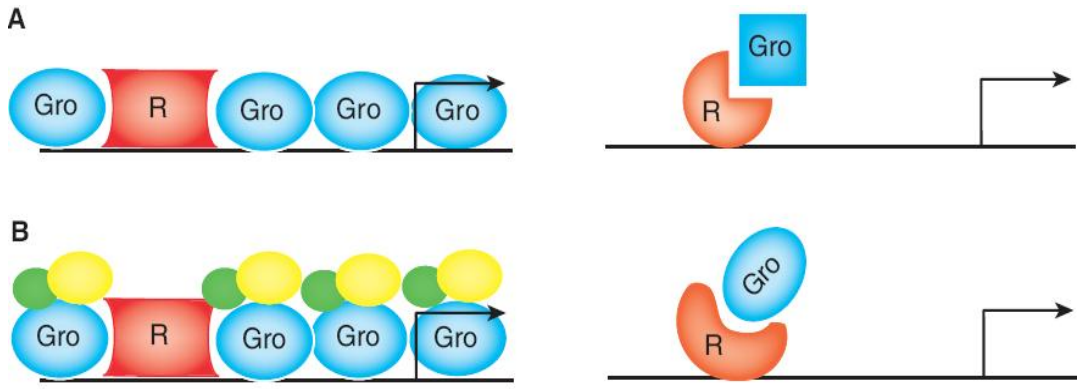


Figure 4-2. Possible mechanisms underlying Groucho-mediated context-dependent repression. (A) Groucho adopts alternative conformations when associated with different repressors, and only spreads when recruited by long-range repressors. (B) Groucho is associated with distinct corepressor complexes that are recruited by different repressors; a form that is capable of spreading is recruited by long-range repressors.

Contributions of distinct corepressors to repression mediated by a single repressor

Similar to many other transcription factors, Hairy and Knirps recruit several different corepressors. It would be interesting to determine the roles of various distinct corepressors during the repression process. It is possible that these corepressors could act together to achieve maximal repression, function redundantly, mediate target-specific or cell-specific repression, or antagonize the functions of the other corepressors. This question can be answered by overexpressing mutant versions of the repressors that do not bind to specific corepressors, and measuring the target expression by quantitative PCR or in situ hybridization. The wild-type and mutant versions of the repressors can be induced at different developmental stages to determine the roles of different corepressors at specific developmental time windows. Transcriptome analysis upon induction of various versions of the repressors can be used to determine the roles of various corepressors at the genome-wide scale.

References

References

1. Hiromi, Y., Kuroiwa, A., and Gehring, W.J. (1985). Control elements of the Drosophila segmentation gene fushi tarazu. Cell *43*, 603-613.

2. Martinez, C.A., and Arnosti, D.N. (2008). Spreading of a corepressor linked to action of long-range repressor hairy. Mol Cell Biol *28*, 2792-2802.

3. Barolo, S., and Levine, M. (1997). hairy mediates dominant repression in the Drosophila embryo. EMBO J *16*, 2883-2891.

4. Arnosti, D.N., Gray, S., Barolo, S., Zhou, J., and Levine, M. (1996). The gap protein knirps mediates both quenching and direct repression in the Drosophila embryo. EMBO J *15*, 3659-3666.

5. Payankaulam, S., Li, L.M., and Arnosti, D.N. (2010). Transcriptional repression: conserved and evolved features. Curr Biol *20*, R764-771.