Stony Brook University

Cis-regulatory contributions to the regulation of *sloppy-paired 1* transcription initiation and elongation

A Dissertation Presented

by

Saiyu Hang

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

August 2012

Stony Brook University

The Graduate School

Saiyu Hang

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Dr. J. Peter Gergen – Dissertation Advisor
Professor, Department of Biochemistry and Cell Biology

Dr. Gerald H. Thomsen - Chairperson of Defense
Professor, Department of Biochemistry and Cell Biology

Dr. Rolf Sternglanz
Distinguished Professor, Department of Biochemistry and Cell Biology

Dr. Robert S. Haltiwanger
Professor and Chair, Department of Biochemistry and Cell Biology

Dr. Maurice Kernan
Associate Professor, Department of Neurobiology and Behavior
Stony Brook University

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

Cis-regulatory contributions to the regulation of *sloppy-paired 1* transcription initiation and elongation

by

Saiyu Hang

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2012

The expression of the *sloppy-paired 1 (slp1)* gene in the gastrula stage *Drosophila* embryo is controlled by the interplay of four transcription factors Runt, Even-skipped (Eve), Fushi-tarazu (Ftz), and Odd-paired (Opa) with two distinct cis-regulatory enhancer elements, the distal early stripe element (DESE) and the proximal early stripe element (PESE). The stripe pattern of *slp1* is the result of non-additive interactions between these two enhancers and is context-dependent. Using chromatin immunoprecipitation (ChIP) to examine a number of reporter constructs, I found DESE mediates Runt dependent activation by facilitating pre-initiation complex formation on the *slp1* promoter. This DESE-dependent activation is influenced by the extent of promoter-proximal DNA upstream of the transcription start site and involves a mechanism that induces nucleosome depletion around the promoter. This effect is specifically important for DESE activation but not PESE activation. ChIP experiments comparing wild-type versus repressed states of DESE-lacZ and PESE-lacZ reporter genes indicate that Eve represses PESE-lacZ expression by blocking the elongation step of the transcription cycle. This repression involves the regulated association of the elongation factor P-TEFb and phosphorylation of Ser2 in the C-terminal domain of RNA polymerase II. Runt and Ftz repress both DESE and PESE. Interestingly, Runt and Ftz repress DESE-lacZ by the same mechanism as Eve dependent repression of PESE, that is by inhibition of transcription elongation. However, Runt and Ftz repress PESE-lacZ by blocking transcription initiation. Analysis of the conserved domains of Runt revealed that C-terminal domain of Runt (region VIII) is involved in the repression of both DESE and PESE, and that this region is not required for activation of DESE.

Table of Contents

List of Figures

Acknowledgements

A doctoral thesis project like this is never the work of anyone alone. The contributions of many different people around me have made this dissertation possible, to only some of whom it is possible to extend my appreciation here.

Above all, I would like to thank my parents, despite me being thousands of miles away for five years, they have given me their unequivocal support throughout, as always, for which my mere expression of thanks does not suffice.

This thesis would not have been possible without the guidance, support and patience of my advisor, Dr. Peter Gergen, his enlightening and encouraging advice has been invaluable on both academic and personal level, for which I am extremely grateful. My responsible thesis committee, Dr. Gerald Thomsen, Dr. Robert Haltiwanger, Dr. Rolf Stanglanz and Dr. Maurice Kernan, their insightful comments and unsurpassed knowledge has been essential input to my research.

I would like to acknowledge the financial, academic and technical support of the Stony Brook University, Department of Biochemistry and Cell Biology. I would also like to thank the director of BSB program Dr. Erwin London and program staff Carol Juliano and Beverly Piaza, for their help since my start of graduate school.

I am deeply grateful to my lab colleagues and friends Xiaoling Wang, Lisa Prazak, Kim Bell and Mike Higgins for their help and encouragement throughout. I am most grateful to Lisa Prazak, her work laid out the foundation of my thesis.

Chapter I:

General introduction

The genome, which has been called the "blueprint for life", describes all the hereditary information of an organism. It is comprised of deoxyribonucleic acid (DNA) containing both coding and non-coding regulatory regions. The code of the genome is interpreted through the process of gene expression, which is a pathway of many steps connecting gene sequences to active proteins. A multicellular organism start its life as a single cell-a fertilized egg, which give rise to dozens or even hundreds of cell types with various cellular functions. However, the genome is normally identical in every cell; the cells differ not because they contain different genetic information, but because they express different sets of genes. This begs a fundamental question of how gene expression is regulated in different types of cells.

Transcription, the process by which the information in DNA is copied into ribonucleic acid (RNA) for protein production, is the first and one of the most highly regulated steps in eukaryotic gene expression (Tamkun, 2007). Transcription of protein coding genes is catalyzed by RNA polymerase II (Pol II). The cycle of events that characterize Pol II transcription can be divided into three phases: initiation, when Pol II is recruited to the promoter and begins to synthesize RNA, elongation, during which the polymerase extends the RNA transcript, and termination, when both polymerase and the transcript disengage from the DNA template (Tamkun, 2007). These phases can be further divided into distinct sub-phases and each step of this multi-step process can be targeted by transcription factors and subject to regulation (Core and Lis, 2008; Sims et al., 2004).

In the context of development, transcriptional regulation is achieved by the recognition of sequence specific trans-acting factors with cis-regulatory elements and the interaction of factor bound regulatory DNA with the promoter region of target genes. With the recent advance of genome-wide ChIP-chip analysis and the ENCODE and modENCODE projects, numerous cis-regulatory elements have been discovered, which require more extensive studies at the molecular level (Negre et al., 2011; Roy et al., 2010).

The segmentation pathway of the *Drosophila* embryo has one of the most well-defined transcription regulatory networks for any known developmental process and provides an excellent model system to study transcriptional regulation. Segmentation is controlled by a hierarchical network of transcription factors and the resulting sequential expression of different sets of genes establishes the body plan. Maternal gene products, such as *bicoid* and *nanos*, which are synthesized and deposited in the egg during oogenesis, provide positional information that regulates expression of the zygotic gap genes such as *giant* and *Kruppel*. Gap genes divide the embryo into a series of major domains that define the 7-stripe expression pattern of the pair-rule genes including *runt*, *even-skipped (eve)* and *fushi tarazu (ftz)*. Pair-rule genes then determine the 14-stripe expression of segment-polarity genes, including *sloppy-paired 1*, that specify the pattern within each segment, as reviewed in (Andrezj Nasiadka, 2002).

A textbook example of maternal to gap transition is the gap gene *hunchback* (*hb*) being activated by maternal gene *bicoid* in a concentration dependent manner (Tautz,

1988). Three Bicoid binding sites of different binding affinity within the promoter region 300 to 50 base-pairs upstream of the transcription start site (TSS) are required for the activation of *hb* (Driever and Nusslein-Volhard, 1989). An example of gap to pair-rule gene regulation is the *even-skipped* gene stripe #2. This stripe is activated by Bicoid and Hunchback with its boundary defined by the gap repressors Giant and Kruppel (Small et al., 1992). These combinatorial cues of activators and repressors are integrated in the region from -1070 to -1550 base-pair upstream of *eve* transcription start site (TSS) (Small et al., 1992). One principle that emerged from studies of pair-rule to segment-polarity transition is the distinction between the regulation of the odd and even-numbered stripes of segment-polarity genes by pair-rule transcription factors. Parasegment specific cis-regulatory elements have been identified for a number of the segment-polarity genes including *engrailed* and *wingless* (Kassis, 1990; Lessing and Nusse, 1998).

As one of the key pair-rule genes, *runt* is the founding member of an evolutionarily conserved gene family known as the Runx family that encodes a group of DNA-binding transcription factors that play a role in regulating lineage-specific gene expression during a variety of developmental processes. The structural signature of this family is the Runt domain, a conserved 128-amino acid motif, which is responsible for binding to sequence-specific DNA (consensus recognition sequence: A/GCCA/GCA) in a complex with partner proteins. The partner proteins are Drosophila Brother (Bro), Big-brother (Bgb) and mammalian core binding factor β (CBFβ). Runx proteins interact with a variety of nuclear proteins including transcriptional co-activators and co-repressors that enables them to function as both activators and repressors of transcription in a context-dependent manner in response to various developmental cues (Adya et al., 2000; Stifani and Ma, 2009; Wheeler et al., 2000).

Almost a decade of work on the *sloppy-paired 1 (slp1)* gene has made it an ideal model to study the mechanism of transcriptional regulation by Runt and other pair-rule transcription factors in the cellular blastoderm stage embryo (Prazak et al., 2010; Swantek and Gergen, 2004; Wang et al., 2007). This segment-polarity gene *slp1* was first discovered as one of Runt target genes by an ectopic expression strategy of NGT (nano-Gal4-tubulin) maternal Gal4/UAS system (Tracey et al., 2000), which led to the identification of its relatively simple combinatorial code in response to the pair-rule transcription factors (Swantek and Gergen, 2004). This gene has a relatively compact cis-regulatory region. The *slp* locus consists of two structurally related genes transcribed in the same direction, with *slp1* located 10 kb upstream of *slp2* (Grossniklaus et al., 1992). Although the two genes have similar expression patterns, *slp1* is expressed more strongly in the early embryo and makes the major quantitative contribution in the early segmentation pathway (Grossniklaus et al., 1992).

The fourteen-stripe pattern of *slp1* in the body of the embryo consists of seven repetitive units (Fig. 1-1). Each unit consists of one odd numbered parasegment and one even numbered parasegment. *slp1* is expressed in the posterior half of each parasegment with a higher level of expression in the even numbered parasegments. In four different transcription contexts, Runt and three other pair-rule transcription factors regulate *slp1* expression in a combinatorial fashion. Eve represses *slp1* transcription in the two cells in

the anterior half of odd numbered parasegments, Runt in combination with Odd-paired (Opa) activates *slp1* in the two cells in the posterior half of odd numbered parasegments to make the odd numbered stripes of *slp1*. Runt and Ftz repress *slp1* in the two cells that define the anterior half of even numbered parasegments and Opa activates *slp1* in the two cells comprising the posterior half of even number parasegments to make the even numbered stripes of *slp1* (Swantek and Gergen, 2004).

Previous work revealed that the fourteen-stripe pattern of *slp1* is governed by two cis-regulatory elements, DESE and PESE, that are located from 8.1 kb to 7.1 kb and 3.1 to 2.5 kb upstream of the *slp1* TSS respectively (Prazak et al., 2010). DESE drives the expression of a *lacZ* reporter gene in cells expressing both the odd and even numbered *slp1* stripes and shows inappropriate expression in the cells anterior to odd stripes. PESE only drives expression of the even numbered stripes with very little expansion into the anterior cells. Interestingly, however, composite *lacZ* reporter genes containing both DESE and PESE shows a pattern that precisely recapitulates the endogenous *slp1* pattern in the body of the embryo (Fig. 1-2). Accordingly, this 'non-additive' effect of DESE and PESE is proposed to involve the regulation of interactions between these two enhancers and the *slp1* promoter (Prazak et al., 2010).

To understand the molecular mechanisms behind this 'non-additive' regulation of *slp1* enhancers, my thesis concerns three major questions: (1) which enhancer is responsible for the *slp1* regulation in each of the four transcription contexts; (2) the mechanism of DESE and PESE mediated *slp1* activation and (3) the mechanism of *slp1* repression in each transcriptional context.

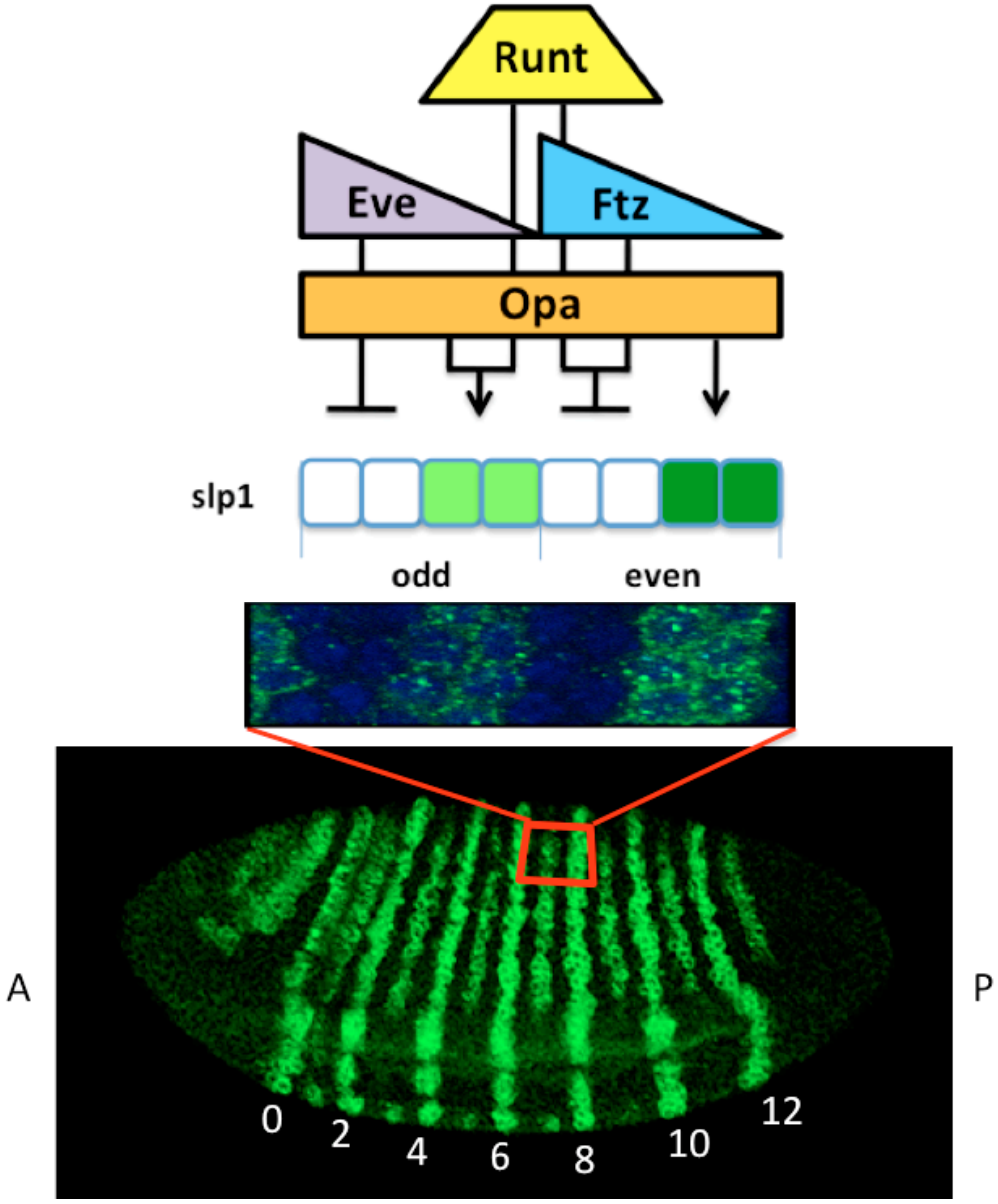

Figure 1-1: Regulation of *slp1* early stripe expression.
From bottom to top: A confocal image of blastoderm stage embryo is oriented anterior head to the left side and dorsal side up. Green staining detects the RNA expression of *slp1* with the even stripes numbered below. The zoomed-in area from the embryo depicts one of the seven repetitive odd and even stripe containing parasegments, illustrated above by a row of eight squares, which represent cells. The green shading indicates the expression of *slp1* in the two most posterior cells in each parasegment, with darker shading indicating the stronger expression in even-numbered parasegments. The schematic diagram above the eight cells shows the pair-rule regulatory inputs that generate the periodic expression pattern of *slp1*. Arrows indicate activation and horizontal bar with a vertical line represents repression.

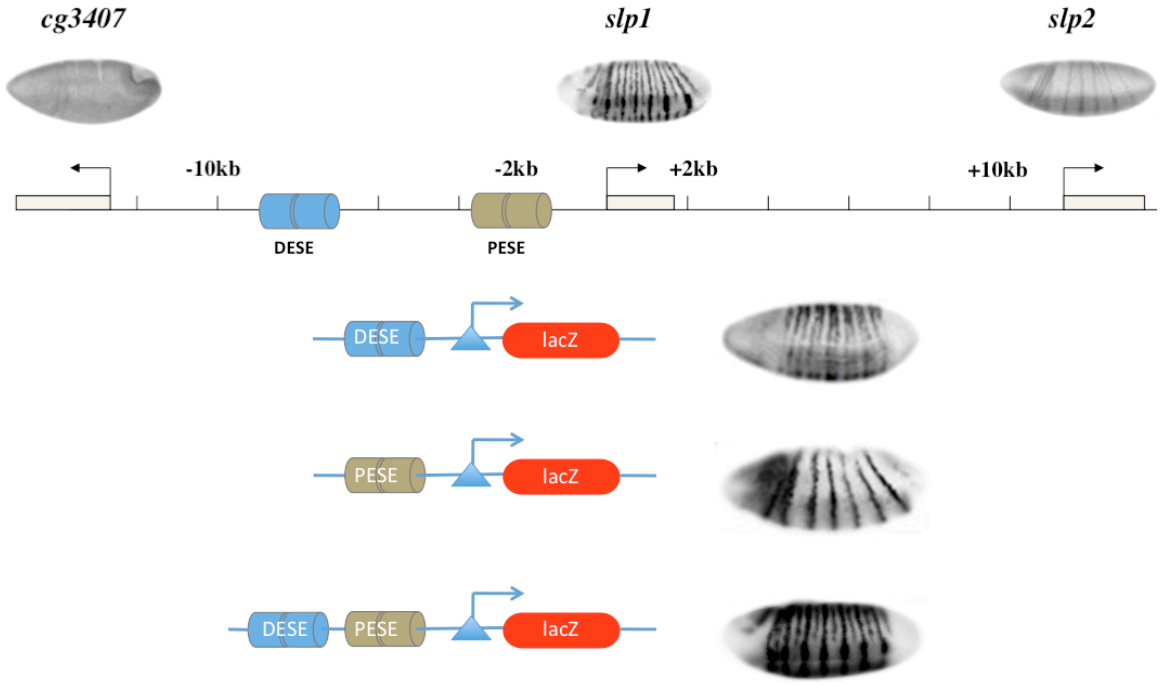

Figure 1-2: The *slp1* locus contains two distinct early stripe elements.
The horizontal line depicts the chromosome region containing the *slp1* locus, extending to the flanking *cg3407* and *slp2* transcription units. Embryos above each of the three genes show their respective RNA expression patterns as visualized by immuohistochemical detection of in situ hybridization. The position of DESE and PESE are indicated on the line. The DESE, PESE and composite *lacZ* reporter constructs are schematically represented below, with embryos on the right showing *lacZ* mRNA expression pattern generated by the different reporter constructs.

Chapter II:

Role of the *slp1* cis-regulatory elements in different transcriptional contexts

Abstract

In the gastrula stage Drosophila embryo, there are four types of cellular contexts for the regulation of *slp1* transcription that is mediated by the DESE and PESE enhancer elements. To determine which enhancer is responsible for *slp1* regulation in each of the four cellular contexts, genetic experiments using both gain of function and loss of function approaches to investigate the roles of the pair-rule factors Eve, Runt, Ftz, Opa and Upd were performed. The results provide descriptive evidence of the enhancer element responsible for mediating *slp1* regulation in each particular context and further demonstrated that DESE is responsible for *slp1* regulation in Runt-expressing cells whereas PESE is responsible for *slp1* regulation in cells that lack Runt. The results also identified the activated dSTAT transcription factor as a second activator of the DESE enhancer.

Introduction

As described in chapter I, and also as shown in figure 2-1, each of the seven repetitive units that compose the fourteen-stripe pattern of *slp1* includes four types of cellular contexts for *slp1* transcription: type I cells correspond to the two cells located in the anterior half of odd numbered parasegments; type II cells correspond to the two cells in the posterior half of odd numbered parasegments; type III cells are the two cells in the anterior half of even numbered parasegments; type IV cells are the two cells in the posterior half of even numbered parasegments. DESE and PESE respond to the pair-rule DNA-binding transcription factors Runt, Eve, Ftz and Opa to regulate *slp1* transcription in the gastrula stage embryo. Each enhancer drives different subsets of the early stripe pattern. However, when in combination, the composite [DESE and PESE] reporter gene generates a pattern beyond what is expected from the additive combination of the two independent patterns and recapitulates the endogenous *slp1* expression pattern. This "non-additive" effect suggests that DESE and PESE compete for the *slp1* promoter resulting in a particular transcription outcome of *slp1* in different cellular contexts. This type of regulation begs the question which enhancer is responsible for the regulation of *slp1* in each of the four contexts. To answer this question, *lacZ* reporter constructs for the two individual enhancers were used in a number of genetic experiments to determine which enhancer-*lacZ* reporter gene behaved similarly to the endogenous *slp1* gene in response to different perturbations.

Results

Context I: anterior half of odd-numbered parasegments

With the presence of Eve, *slp1* transcription is turned off in the two cells located in the anterior half of odd numbered parasegments (type I cells). Expression of the PESE-*lacZ* reporter is also repressed in these two cells, whereas the DESE-*lacZ* reporter showed inappropriate expression in this same cellular context (Fig. 2-1). These observations suggest that PESE is responsible for *slp1* repression in type I cells (Prazak et al., 2010).

Transient elimination of Eve activity, using a temperature sensitive allele, resulted in de-repression of both endogenous *slp1* and the PESE-*lacZ* reporter in type I cells (Fig. 2-2A). Ectopic-expression of Eve throughout the embryo did not significantly affect DESE-*lacZ* expression, but completely eliminated PESE-*lacZ* expression in all cells of embryo (Fig. 2-2B). Therefore, both loss of function and gain of function experiments indicate that PESE but not DESE is sensitive to Eve manipulation, which strongly suggests that PESE mediates Eve-dependent repression of *slp1* (Prazak et al., 2010).

Context II: posterior half of odd numbered parasegments

Activation of *slp1* transcription in the two cells located in the posterior half of the odd numbered parasegments (type II cells) is due to the presence of both Runt and Opa (Swantek and Gergen, 2004). The DESE-*lacZ* reporter is also activated in type II cells, but PESE-*lacZ* is not (Fig. 2-1), which suggests that DESE mediates the Runt- and Opa-dependent activation of *slp1* in these two cells.

Elimination of Runt obliterated the *slp1* expression in type II cells, but DESE-*lacZ* still showed residual expression that is due to its insensitivity to Eve and partial activation by Opa alone (Fig. 2-3A) (Swantek and Gergen, 2004). More importantly, removal of Opa from the embryo abolished the expression of endogenous *slp1* and DESE-*lacZ* in type II cells that generate the *slp1* odd-numbered stripes (Fig. 2-5A). Ectopic-expression of Runt and Opa together activates *slp1* expression in the head region of embryo. The same pattern was observed for DESE-*lacZ* but not the PESE-*lacZ* reporter gene (Fig. 2-3B). Taken together, it is clear that DESE is responsible for mediating Runt- and Opa-dependent activation of *slp1*.

Context III: anterior half of even numbered parasegments

In the two cells located in the anterior half of even number parasegments (type III cells), *slp1* transcription is repressed by Runt and Ftz (Swantek and Gergen, 2004). In fact, both DESE-*lacZ* and PESE-*lacZ* reporter genes are mostly repressed in type III cells with PESE displaying a low level of expression anterior to even number stripes (Fig. 2-1).

It is hard to decide which enhancer is mediating *slp1* repression in these two cells simply based on their individual pattern in wild-type embryos, or by their pattern in *runt*

mutant embryos because both enhancer reporters as well as endogenous *slp1* showed de-repression in response to the transient elimination of *runt* (Fig. 2-3A). However, *ftz* mutant embryos showed clear de-repression in the expression of endogenous *slp1* and DESE-*lacZ* (Fig. 2-4A). PESE-*lacZ* expression is still largely repressed in these two cells (Fig. 2-4A). This suggests DESE is driving activation of *slp1* in the absence of Ftz. An experiment that demonstrates DESE is responsible for the de-repression of *slp1* in type III cells is the ectopic-expression of Eve in the *ftz* mutant background (Fig. 2-4B). As described above, DESE-*lacZ* is insensitive to Eve ectopic-expression, whereas PESE-*lacZ* can be completely repressed by ectopic-expression of Eve (Fig. 2-2B). In situ results showed that ectopic-expression of Eve in the *ftz* mutant background abolished PESE-*lacZ* expression in the body of the embryo but did not eliminate endogenous *slp1* expression in type III cells (Fig. 2-4B), which in turn supports the idea that DESE is responsible for the de-repression of *slp1* observed in cells that lack Ftz and that contain Runt. These observations are consistent with the current hypothesis that Runt plays a role in regulating the interactions of the two *slp1* early stripe enhancers with the promoter and that the presence of Runt is sufficient to block the interaction with PESE. Stated alternatively, DESE is responsible for regulating the transcriptional output of *slp1* in Runt-expressing cells, either activation in response to Runt and Opa in type II cells or the repression in type III cells that also express Ftz.

Context IV: posterior half of even numbered parasegments

Expression of *slp1* is activated in the two cells located in the posterior half of the even-numbered parasegments (type IV cells). Both DESE-*lacZ* and PESE-*lacZ* are expressed in most type IV cells except that DESE-*lacZ* does not activate stripe 0 expression, whereas PESE-lacZ does (Fig. 2-1C). As shown above, ectopic expression of Eve represses the even-numbered stripes of *slp1* and a PESE-*lacZ* reporter, but has no effect on a DESE-*lacZ* reporter (Fig. 2-2B). This result strongly suggests that PESE is primarily responsible for mediating *slp1* activation in type IV cells and raises a question as to whether DESE ever contributes to the activation of *slp1* in these type IV cells.

Prior work predicted the existence of a factor X, in addition to Opa, which is involved in the activation of *slp1* expression in type IV cells, as a low level of *slp1* even-stripe expression is retained in *opa* mutant embryos (Fig. 2-5A) (Prazak et al., 2010). The work, by undergraduate student Liujing Xing identified the factor X as the activated dSTAT, a transcription factor activated by Janus kinase pathway (Aaronson and Horvath, 2002). Similar to Opa, dSTAT can activate *slp1* through DESE in type IV cells.

As shown in figure 2-5, with the exception of stripe 0, knocking out of *opa* completely eliminated PESE-*lacZ* expression in type IV cells, but endogenous *slp1* and DESE-*lacZ* still showed reduced level of expression in these two cells (Fig. 2-5A). This indicates Opa can activate both DESE and PESE. Reduced expression of *slp1* in these two cells could be due to the loss of PESE activation and/or reduction of DESE activation. Double knockout of both *opa* and *unpaired* (*upd*), a ligand of the JAK/STAT pathway (Aaronson and Horvath, 2002), abolished the residual expression of both *slp1* and the DESE-*lacZ* reporter gene. This experiment demonstrates that both Opa and

dSTAT participate in *slp1* activation that is contributed by both PESE and DESE in type IV cells.

In addition to the above genetic experiments, the expression pattern of two composite reporter genes also provided useful insights into the question of which enhancer mediates *slp1* activation in type IV cells. As described later in Chapter 3, reporter genes containing the small p126 (-79 to +47) *slp1* promoter do not allow DESE-*lacZ* expression but are permissive for PESE-*lacZ* expression. A composite reporter containing both the DESE and PESE enhancers with the p126 promoter showed *lacZ* expression in cells corresponding to the *slp1* even number stripes (type IV cells) but not in cells expressing the *slp1* odd-numbered, DESE-dependent stripes (see chapter III, Fig. 3-2B, 3-2D, 3-2H in detail). This indicates that when both DESE and PESE are present and DESE activity is not allowed by p126, PESE is able to activate *slp1* even number stripes, which supports that PESE is responsible for the activation of *slp1* in type IV cells.

As previously described by Prazak et al, the central region of PESE, referred to as the C1 region is required for PESE-dependent expression of the even-numbered stripes (Prazak et al., 2010). Deletion of the PESE:C1 region from full length PESE eliminates PESE activation in type IV cells (Fig. 2-6A, B). Interestingly, a composite construct *slp[DESE/PESE:ΔC1/p381]lacZ* that contains DESE and a PESE that lacks the C1 interval and that also includes a larger *slp1* promoter that is permissive for mediating the response to DESE retained expression in type IV cells (Fig. 2-6C). Consistent with the idea that these even-numbered stripes are due to DESE, their expression in type IV cells was insensitive to Eve repression (Fig. 2-6D). Thus, when PESE activity is removed, either by elimination of Opa or by removal of the essential PESE:C1 interval, DESE is able to activate *slp1* expression in type IV cells.

In conclusion, except for stripe 0, in type IV cells, the *slp1* promoter can receive inputs from both DESE and PESE, and the seemingly redundant activity of both enhancers contributes to the robust transcription of *slp1* in these two cells.

Discussion:

As described above, the fourteen-stripe pattern of *slp1* in the gastrula stage Drosophila embryo is the outcome of context-dependent regulation, involving a combinatorial interaction between a number of DNA binding transcription factors and two cis-regulatory enhancer elements. Interestingly, both DESE and PESE can mediate activation and repression, but depending on the particular cellular context, only one of the two enhancers appears to be primarily responsible for regulating the transcriptional output at the *slp1* promoter.

In type I cells, PESE mediates Eve-dependent *slp1* repression. DESE, by itself, is insensitive to Eve repression and able to activate transcription (Fig. 2-1). Its interaction with *slp1* promoter must be prevented by the presence of PESE and Eve in these two cells. Thus, ectopically introducing Eve into type IV cells, which normally do not contain Eve, resulted a repression of *slp1* and PESE-*lacZ* but not DESE-*lacZ* (Fig. 2-2B). These results indicate Eve represses PESE in a way that also does not allow DESE to activate the *slp1* promoter. How DESE-dependent activation is prevented in type I cells is still an open question. The mechanism by which Eve represses PESE is further investigated in chapter IV.

In type II cells, DESE mediates Runt- and Opa-dependent *slp1* activation. Both Runt and Eve can prevent PESE activation. Interestingly, transient elimination of Runt from type II cells completely removed the expression of endogenous *slp1*, which indicates *slp1* regulation switched back to PESE when Runt is eliminated because a reporter that contains only DESE still drove inappropriate expression in this context (Fig. 2-3A). Ectopic expression of Runt and Opa in all cells showed a perfect match of the endogenous *slp1* expression pattern with the DESE but not the PESE pattern (Fig. 2-3B). Based on these results it has been proposed that when Runt is present in the cell, *slp1* relies on DESE for transcription. The results from loss of function and gain of function experiments with Runt strongly implicate DESE in Runt-dependent *slp1* regulation. When Runt is present in the cell, DESE is responsible for controlling *slp1* transcription regulation. When Runt is not present, PESE mediates *slp1* regulation.

In type III cells, DESE mediates Runt- and Ftz-dependent *slp1* repression. For endogenous *slp1* expression to be off, both DESE and PESE need to be repressed in these two cells. Elimination of Ftz resulted in the de-repression of both endogenous *slp1* and DESE-*lacZ* but not PESE-*lacZ* (Fig. 2-4A). This indicates repression of DESE requires both Runt and Ftz, whereas Runt alone is a potent repressor for PESE. The question of whether Runt represses DESE and PESE by different mechanisms is discussed in chapter IV.

In type IV cells, *slp1* is normally activated by PESE. PESE mediates *slp1* activation by Opa. However, when the input from PESE is compromised by either a mutation within PESE or elimination of Opa, DESE can compensate for this loss of activation. The transcription rule for stripe 0 is distinct and still unclear. This stripe is activated through PESE but unlike the other even stripes is not fully dependent on the

activator Opa as this stripe is still present in *opa* mutant embryos (Fig. 2-5A). Prior work showed Opa is present in cells posterior of PESE stripe 0 (Lisa Prazak unpublished data). Further investigation is required to test whether dSTAT also directly contributes to stripe 0 expression. A more detailed description of the involvement of JAK/STAT pathway in *slp1* regulation is provided in Liujing Xing's undergraduate Honors thesis. The mechanism of DESE and PESE activation is described in chapter III.

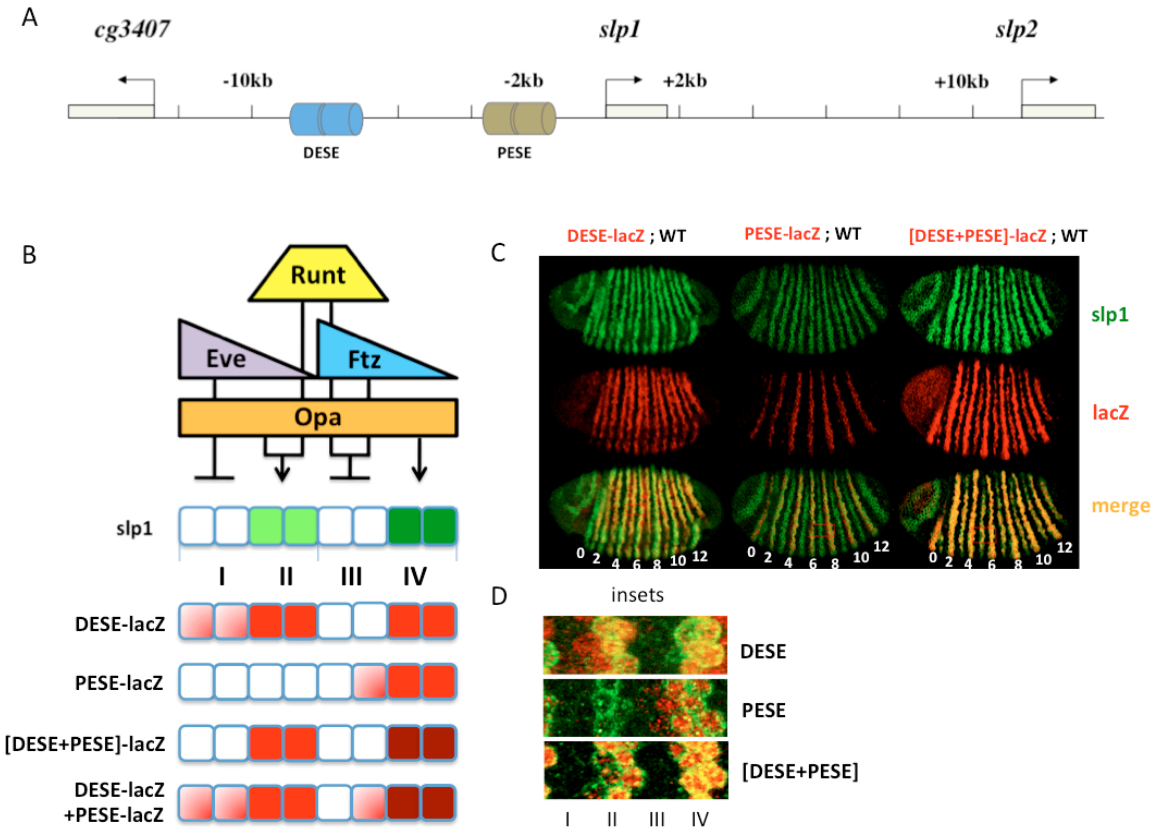

Figure 2-1: Non-additive effect of the two *slp1* cis-regulatory elements.
(A) The *slp1* locus, flanked by *cg3407* and *slp2*, contains two distinct early stripe elements DESE and PESE. (B) Schematic diagram represents each of the seven repetitive units of the fourteen number stripe pattern of *slp1*, indicating pair-rule inputs Opa, Runt, Eve and Ftz regulating *slp1* in the four transcription contexts. Shading corresponds to the expression pattern showing the insets in panel D. (C) Confocal images of the fluorescent in situ hybridization of endogenous *slp1* (green) and *lacZ* reporter gene (red) in wild-type embryo. (D) Insets from the merged confocal images in C.

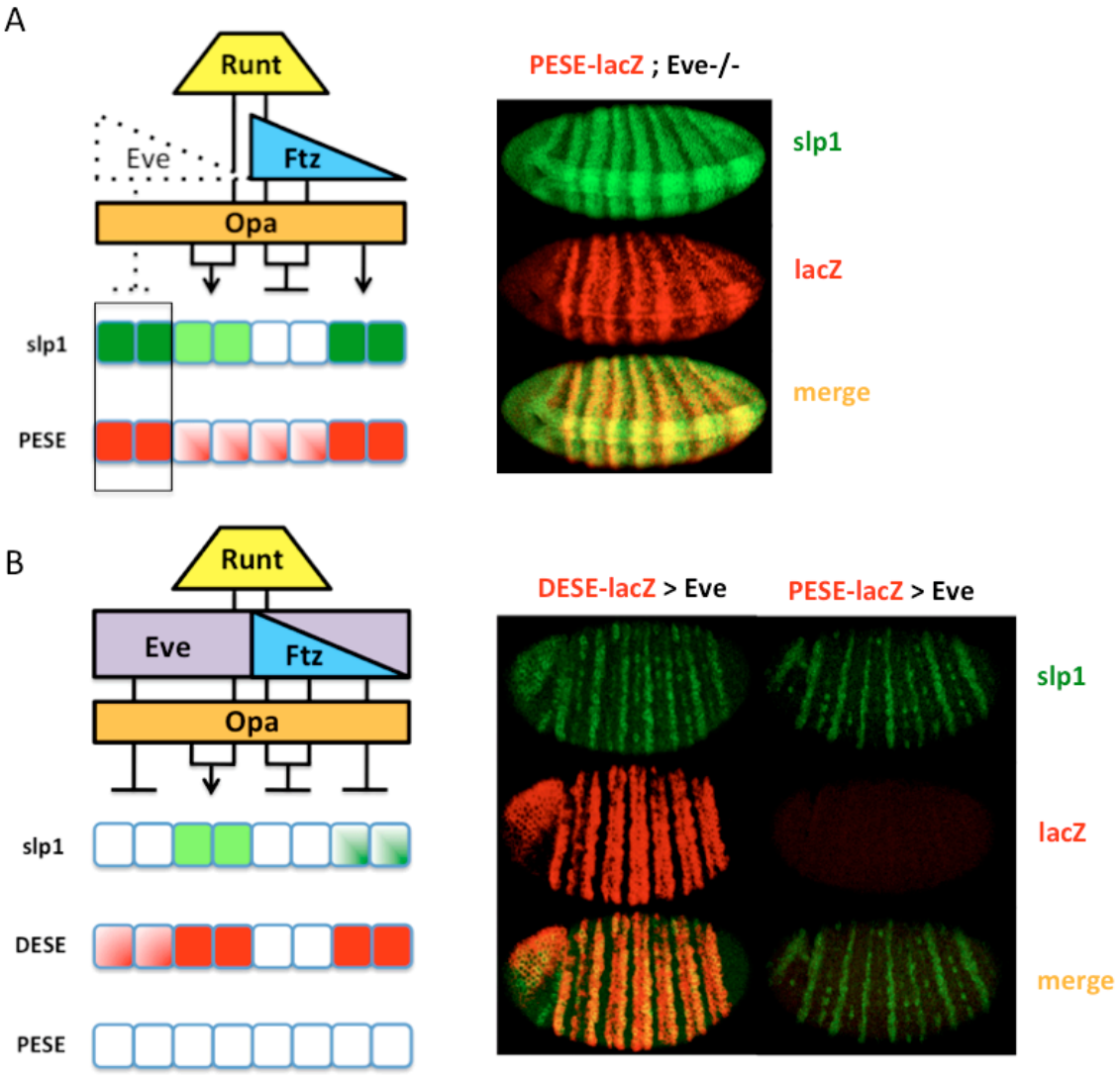

Figure 2-2. PESE mediates Eve dependent *slp1* repression in type I cells.
Fluorescent in situ hybridization of *slp1* (green), *DESE-lacZ* (red) and *PESE-lacZ* (red). (A). Loss of function of *eve*. Embryo was progeny of flies doubly heterozygous of eve^{1} and PESE[3125p381]-*lacZ*. (B). Gain of function of *eve*. Embryo were generated from the cross *NGT40; DESE[8771p126]-lacZ* × *UAS-eve*^{12}, and *NGT40; PESE[3125p381]-lacZ* × *UAS-eve*^{12}.

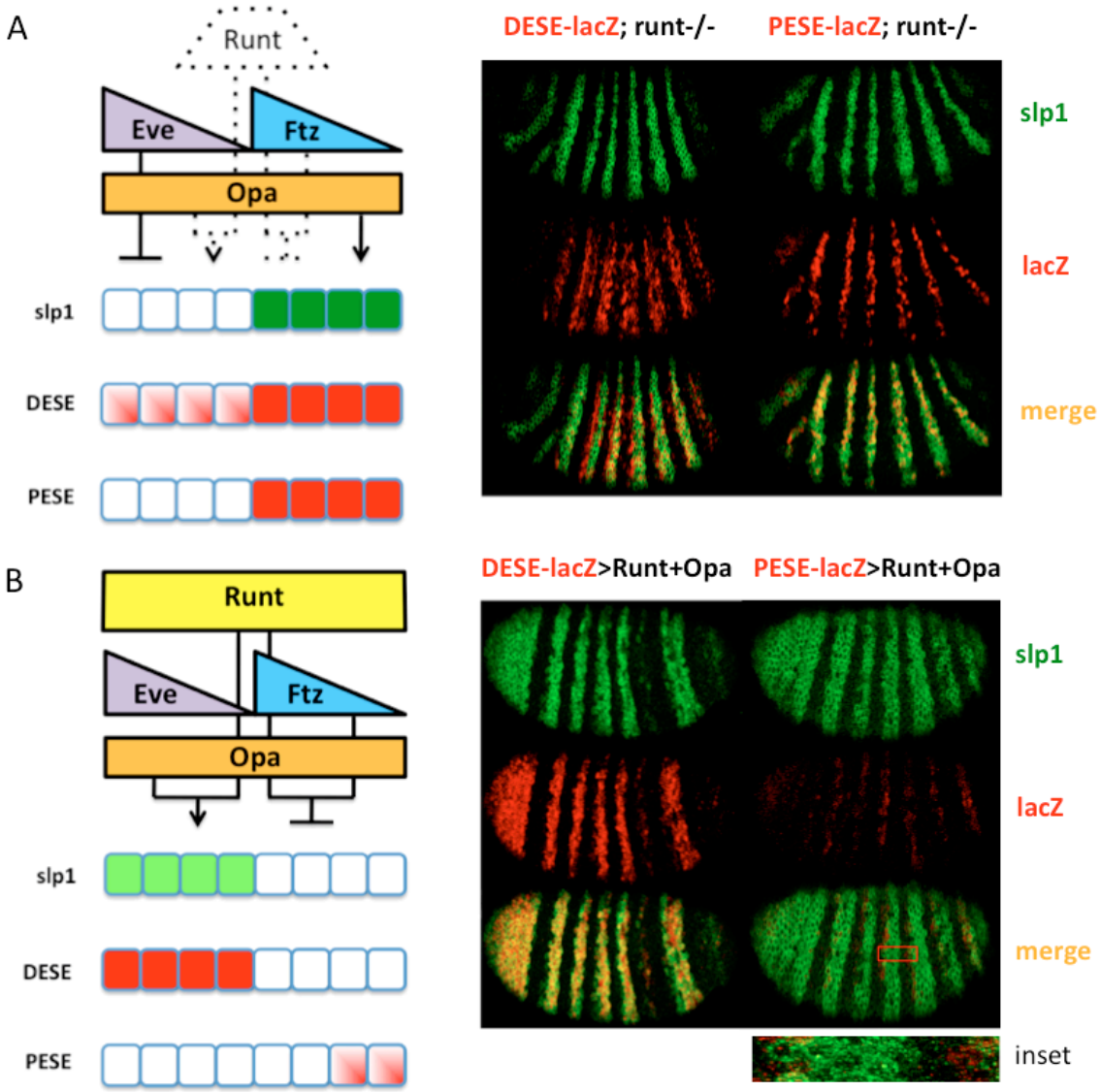

Figure 2-3: DESE mediates *slp1* activation in type II cells.
Fluorescent in situ hybridization of *slp1* (green), *DESE-lacZ* (red) and *PESE-lacZ* (red). (A). Loss of function of runt. Embryos were generated from the cross between run^{29}/*y w* females and males that are homozygous for *DESE[8771p126]-lacZ* or *PESE[3918p126]-lacZ* (Prazak et al. 2010). (B). Gain of function of runt and opa. Embryos were generated from the cross between *NGT40; DESE[8771p126]-lacZ* or *PESE[3918p126]-lacZ* females and *UAS-runt*^{15}; *UAS-opa*^{14} males.

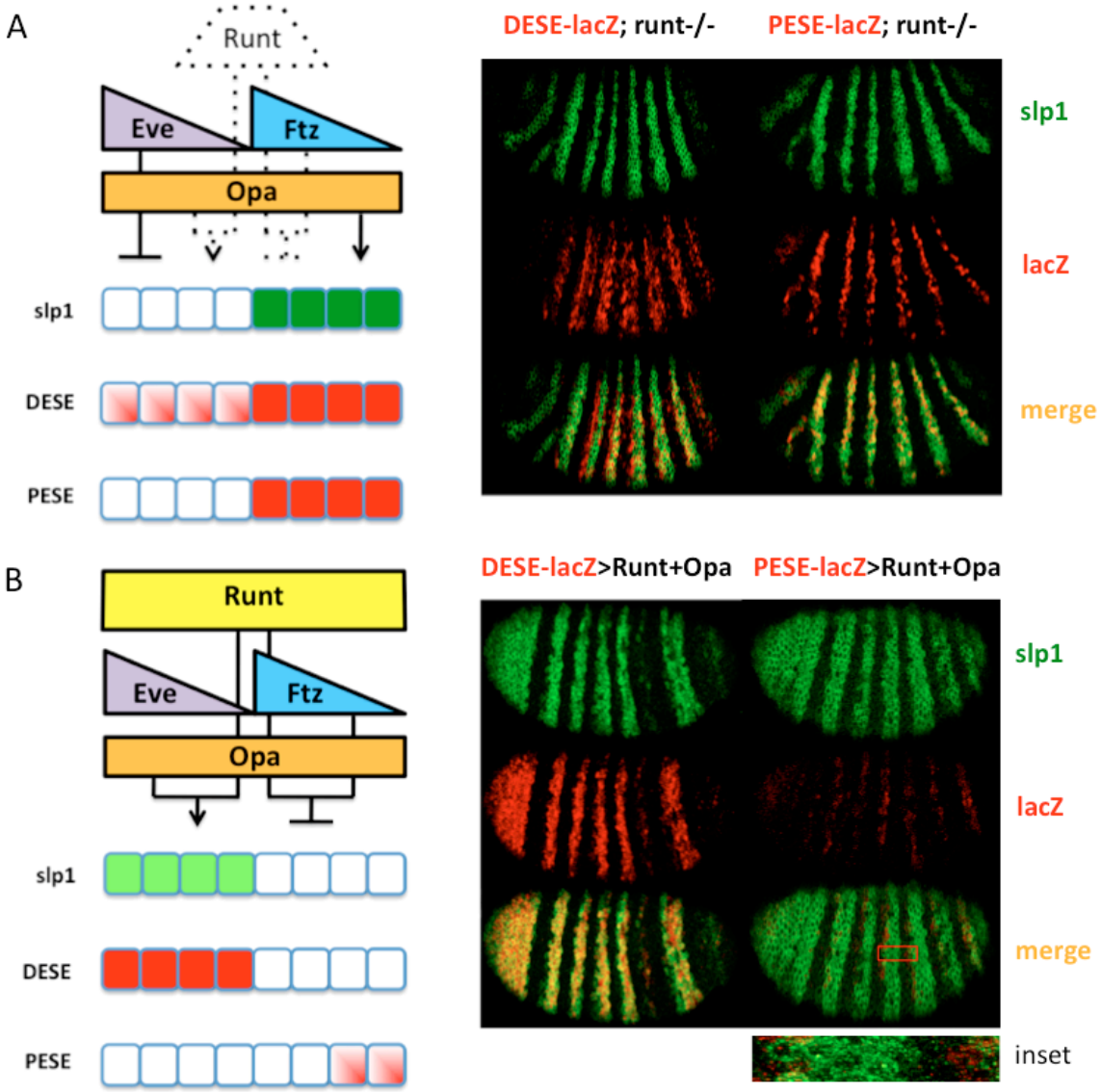

Figure 2-4: DESE mediates *slp1* repression in type III cells.
Fluorescent in situ hybridization of *slp1* (green), *DESE-lacZ* (red) and *PESE-lacZ* (red). (A). Loss of function of *ftz*. Embryos were the progeny of flies with a recombinant third chromosome containing the ftz^{11} mutation and *DESE[8771p126]-lacZ* or *PESE[3918p126]-lacZ*. (B) ectopic-expression of Eve in *ftz* mutant background. Embryo was generated from the cross between *NGT40; PESE[3918p126]-lacZ*, ftz^{11}*/TM3* females and *UAS-eve*^{12}*; PESE[3918p126]-lacZ*, ftz^{11}*/TM3* males.

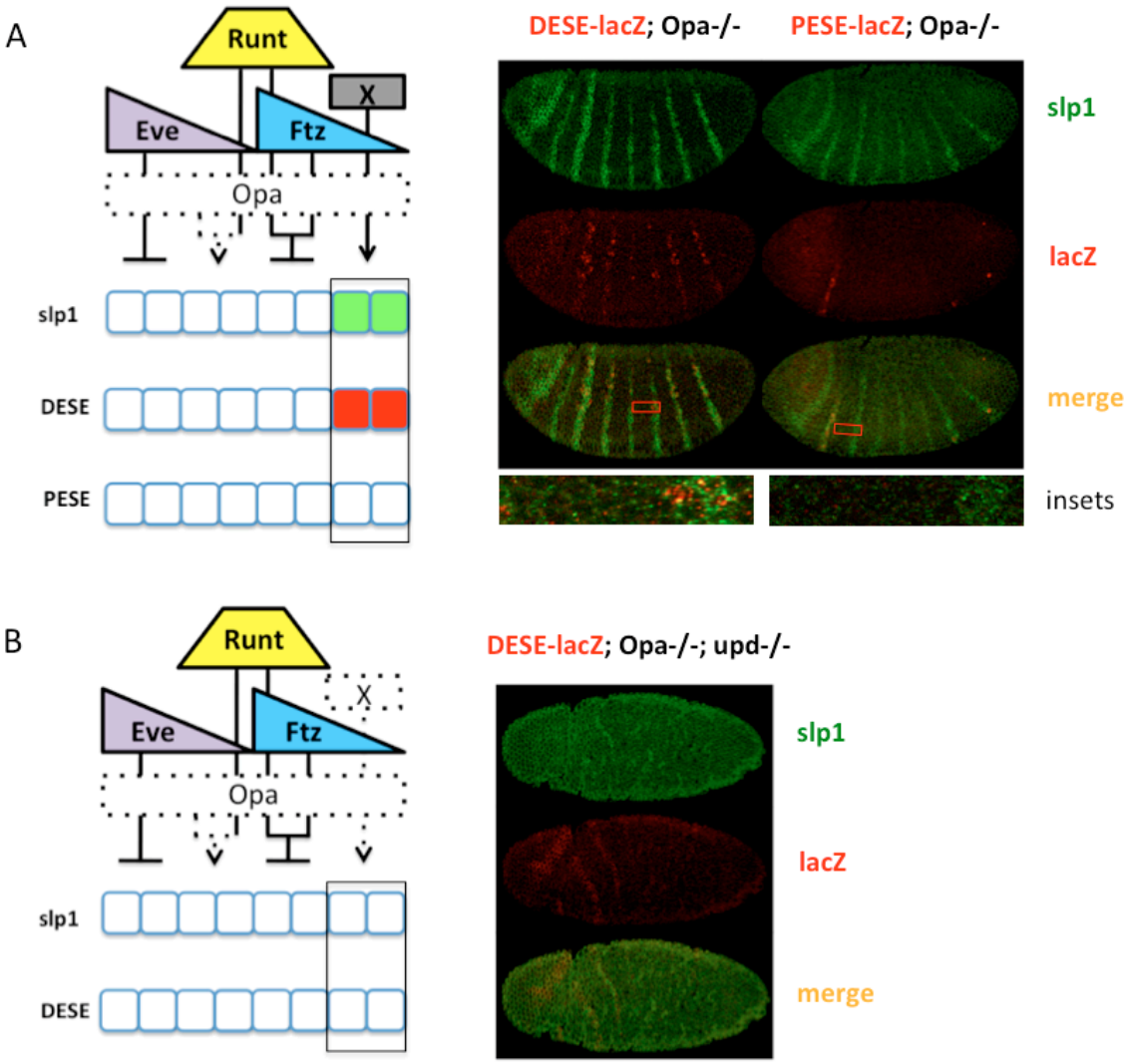

Figure 2-5: DESE and PESE both contribute to *slp1* activation in type IV cells. Fluorescent in situ hybridization of *slp1* (green), *DESE-lacZ* (red) and *PESE-lacZ* (red). (A). Loss of function of *opa*. Embryos were the progeny if flies with a recombinant third chromosome containing the opa^{1} mutation and *DESE[8771p126]-lacZ* or *PESE[3918p126]-lacZ*. (B). Loss of function of *opa* and *upd*. Embryo was the progeny of the cross between $Df(1)os^{UE69}$*/y w; DESE[8771p126]-lacZ,* opa^{1}*/TM3* females and *DESE[8771p126]-lacZ,* opa^{1}*/TM3* males.

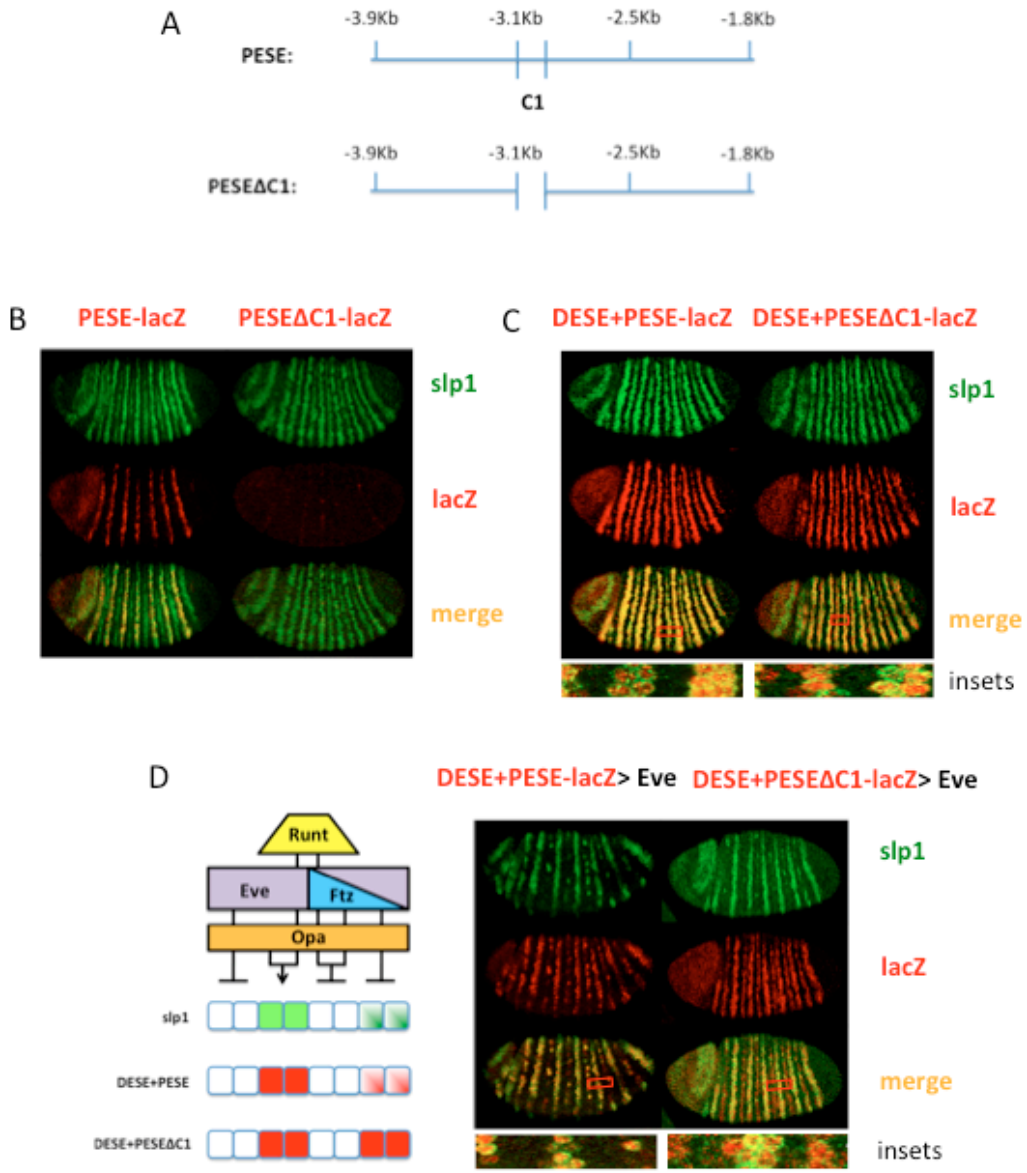

Figure 2-6: DESE replenishes *slp1* activation in type IV cells when PESE activity is eliminated.

(A). Schematic diagram of PESE and PESEΔC1. Numbers on top the lines indicate the sequence position relative to slp1 transcription start site.

From B to D showing fluorescent in situ of *slp1* (green) and *lacZ* reporter (red). (B). Expression of *PESE[3918p126]-lacZ* and *PESEΔC1[3918ΔC1p126]-lacZ* in wild-type embryo. (C). Expression of *DESE+PESE[8765:3918p381]-lacZ* and *DESE+PESEΔC1[8765:3918ΔC1p381]-lacZ* in wild-type embryo. (D). Expression of *DESE+PESE-lacZ* and *DESE+PESEΔC1-lacZ* in embryos with ectopic-expression of eve. Embryos were generated from the cross between *NGT40; DESE+PESE[8765:3918p381]-lacZ* or *NGT40; DESE+PESEΔC1[8765:3918ΔC1p381]-lacZ* and *UAS-eve*[12] males.

Acknowledgement:

The work listed in this chapter consists of some previously published data and some recent unpublished data, and is a result of the combined efforts from many people involved in this project. Peter Gergen generated recombinant fly lines. Lisa Prazak made most of the reporter constructs and did the experiments for figure 2-2B and 2-3A. Liujing Xing did the experiments for figure 2-5 and Shabbir Alam did the *in situ* for figure 2-6.

Chapter III:

Cis-regulatory requirements for *slp1* transcription activation

Abstract

Both DESE and PESE can mediate *slp1* activation. In an attempt to understand the mechanisms involved in DESE and PESE-driven activation, a number of reporter genes carrying different *slp1* promoters with or without enhancers were investigated. Results from chromatin immune-precipitation (ChIP) experiments indicate the DESE and PESE enhancers promote recruitment of general transcription factors and Pol II to the promoter for transcription initiation. Moreover, DESE-driven activation and transcription initiation is influenced by the extent of promoter proximal DNA sequence upstream of the *slp1* core promoter. This sequence induces a nucleosome-depleted region that is specifically important for DESE- but not for PESE-dependent activation.

Introduction

The multi-step process of transcription starts with initiation by RNA polymerase II (Pol II) at specific DNA sequences termed promoters that comprise the gene-specific transcription start site (TSS) and other sequence motifs such as TATA box, Initiator (Inr) and downstream promoter element (DPE) (Butler and Kadonaga, 2002; Smale and Kadonaga, 2003). Interestingly, purified Pol II incubated in a test tube with purified DNA carrying promoter sequences cannot initiate transcription specifically. Additional factors have been found to be necessary for the recruitment of Pol II and confer to this enzyme the full competence for initiating and elongating transcription. These are so-called general transcription factors (GTFs), which are recruited to promoters in an ordered fashion in vitro. So far, this class of proteins comprises the TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. These proteins contribute different activities necessary for the transcription process. For example, TBP, the core component of TFIID complex, recognizes the TATA box and bends the DNA. The TFIIH complex contains a helicase and kinase activity that unwinds DNA and phosphorylates the C-terminal domain (CTD) of Pol II at Ser5 residues (Lee and Young, 2000).

In eukaryotic cells, DNA is packaged into nucleosomes, a structure consisting of a segment of DNA wound in sequence around eight histone protein cores, two H2A/H2B and H3/H4 heterodimers (Luger et al., 1997). As a consequence, accessibility to DNA of the Pol II and GTFs is an essential issue of transcription. Clearly, a number of factors, ranging from information in the DNA sequence to protein complexes that affect nucleosome assembly are found to affect gene regulation (Radman-Livaja and Rando, 2010). An interesting case of nucleosome positioning affecting the gene expression comes from the human β-interferon "slide" promoter, where a nucleosome positioned over the TATA box slides downstream upon activation by three different transcription factors (Lomvardas and Thanos, 2002).

Enhancers, by classic definition, are cis-regulatory DNA sequences that can activate gene transcription irrespective of their orientation or distance relative to the promoters of target genes (Muller et al., 1988). They are responsible for the spatial-temporal regulation of gene expression. While promoters serve as the binding platform for the general transcription apparatus, enhancers are bound by sequence specific DNA-binding transcription factors. The mechanisms by which enhancers can activate transcription from distant promoters remain elusive. Several models have been proposed, including a looping model that involves a physical contact of enhancer-bound transcription factors with the promoter region, with concomitant looping out of the intervening DNA (Bulger and Groudine, 1999; Li et al., 2006). Recent advances in numerous labs using high through put technologies indicate chromatin looping is a wide spread organizing principle of the chromatin fiber (Gondor and Ohlsson, 2009). One well established example is the β-globin Locus Control Region (LCR) whose looping to the promoter by GATA1 and Ldb1 is essential for transcription initiation of this gene (Deng et al., 2012).

DESE and PESE both activate *slp1* transcription. This chapter investigates the mechanisms involved in the activation of *slp1* and a differential requirement for promoter proximal sequences outside of the core promoter in mediating the regulation by the DESE and PESE enhancers.

Results

DESE and PESE facilitate pre-initiation complex formation on *slp1* promoter

In situ hybridization of gastrula stage embryos carrying *lacZ* reporter constructs showed that, without DESE or PESE, a DNA segment spanning from 260 basepairs upstream to 121 basepairs downstream of the *slp1* transcription start site (TSS) does not activate *lacZ* transcription in cells located in the body of the embryo except for a patch of cells in the posterior head (Fig. 3-1A). This result indicates that this *slp1*-promoter containing segment, hereafter referred to as the p381 promoter, does not by itself directly respond to the segmentation regulatory circuitry. When either the DESE or PESE enhancer is included with the *slp1* p381 promoter, the *lacZ* reporter gene showed a characteristic DESE and PESE-driven pattern as described in chapter II and as shown here in figure 3-1B and C.

For transcription to occur, protein-coding genes need to assemble a pre-initiation complex (PIC) on their promoter, which consists of RNA polymerase II (Pol II) and a number of general transcription factors such as TFIID, TFIIB, TFIIF et cetera. Enhancers can activate gene transcription by facilitating the formation of PIC complex on their cognate promoters (Maston et al., 2006). Indeed, Chromatin IP experiments comparing PIC formation on the promoter regions of the p381-*lacZ*, DESE-p381-*lacZ* and PESE-p381-*lacZ* reporters in the embryo showed that TBP, but not TFIIB or TFIIF, was associated with the promoter region of p381-*lacZ* in the absence of either enhancer in the reporter construct. The specificity of the TBP signal with the promoter regions is confirmed by the lack of any significant signal with primers from the middle of the *lacZ* gene (Fig. 3-1D, E, F). Pol II association with p381-*lacZ* was found barely above background on the promoter region, which corresponded to the low expression level only in the posterior head region (Fig. 3-1G, A). Inclusion of DESE or PESE in the reporter constructs slightly increased TBP association with the promoter (Fig. 3-1D), but also resulted in significant increases in the association of TFIIB and TFIIF with the promoter (Fig. 3-1E, F). A similar significant increase in the association of Pol II with the promoter regions of DESE-p381-*lacZ* and PESE-p381-*lacZ* suggests efficient transcription initiation is taking place on the promoter of these two reporter genes (Fig. 3-1G). These results indicate that both DESE and PESE can activate the *slp1* p381 promoter by stimulating the recruitment and/or stabilizing the association of general transcription factors with the promoter, which is essential for Pol II to initiate transcription.

Upstream sequence of core promoter is important for DESE-driven expression

Enhancers can exhibit specificity to certain promoter sequences (Butler and Kadonaga, 2002; Juven-Gershon et al., 2008; Smale, 2001). To assess the extent of promoter sequence requirement for DESE and PESE, I examined the expression of various DESE-*lacZ*, PESE-*lacZ* and composite [DESE+PESE]-*lacZ* reporter constructs containing different size of promoter sequence around the *slp1* transcription start site (TSS) (Fig. 2A). A small basal promoter, p126, extending from 79 base pair upstream to 47 base pair downstream of TSS (-79 to +47), contains a TATA box and initiator (Inr). A large basal promoter, p381, contains the sequence from -260 to +121, including the entire 5' untranslated region of *slp1* (Fig. 3-2A). Interestingly, I found these two promoters respond differently to DESE and PESE. PESE-dependent activation was essentially not affected by the size of promoter in stage-6 embryo (Fig. 3-2B, C), as PESE-p126 and PESE-p381 both showed *lacZ* expression in type IV cells that make *slp1* even number stripes. In contrast, p126 did not allow DESE-driven activation at levels comparable to that obtained with the p381 (Fig. 3-2D and E). In a composite reporter construct, DESE is responsible for the expression of odd number stripes. As shown in Figure 3-2H, *lacZ* expression in the odd number stripes (type II cells) was greatly reduced when the composite reporter contains p126, whereas the reporter construct containing p381 showed clear expression of odd number stripes (Fig. 3-2I). The expression patterns of composite reporter containing these two promoters agreed with the results of individual enhancer reporters, and suggests a specific requirement of an additional region that is outside of p126 but within p381 for DESE-driven activation.

To uncover whether the upstream extension, -260 to -79, or the downstream extension, +47 to +121, is required for DESE-driven activation, I tested two other promoter p200 (-79 to +121) and p307 (-260 to +47) in DESE and composite constructs. Expression of the DESE-*lacZ* reporter was completely eliminated in the construct that has p200 (Fig. 3-2F), but fully recovered in the construct carrying p307 (Fig. 3-2G). Consistent with the results obtained with the DESE reporters, the composite construct with p200 was deficient in the expression of odd number stripes and had some defects in activating even number stripes (Fig. 3-2J). However, when linked to p307, the composite reporter impeccably drove *lacZ* expression of both odd and even number stripes (Fig. 3-2K). Further extension of the promoter upstream region to -940 bp or to -1.8 kb did not subsequently affect the DESE expression pattern within the segmented region of the embryo, although there is an increase in anterior head expression due to head stripe elements in these upstream extensions. These results indicate that the region from -260 to -79 is sufficient for allowing DESE-driven activation and further demonstrates that this effect is not simply due to a close juxtaposition of DESE with the -260 region (Fig. 3-3). Therefore, reporter assays with different size promoter regions revealed that the sequence immediately upstream of core promoter, from -260 to -79, is necessary and sufficient for DESE dependent activation, but is not necessary for PESE activity in gastrula stage embryos.

Upstream sequence affects initiation of transcription

As shown above, DESE activates the *slp1* p381 promoter by enhancing the association of general transcription factors such as TFIIB, TFIIF and Pol II with the promoter. To see if the inability of the p126 promoter to respond to DESE dependent activation is due to insufficient transcription initiation, I used chromatin IP experiments to compare TBP and Pol II association with four different promoters in different reporter gene constructs. Very low TBP and Pol II signals were found on DESE-p126 and DESE-p200, whereas abundant TBP and Pol II association were detected on the DESE-p307 and DESE-p381 promoters along with a detectable increase in Pol II (but as expected, not TBP) association in the middle of the *lacZ* gene (Fig. 3-4A, C). Similar TBP and Pol II association patterns were detected on PESE-*lacZ* constructs (Fig. 3-4B, D). It was surprising that, with little to no detectable TBP binding on promoter region of PESE-p126 (Fig. 3-4B), this reporter was still able to activate gene transcription (Fig. 3-2B, 3-4D, 3-5). One explanation is that PESE activation of transcription is TBP independent. Additional experiments indicate that PESE has little promoter specificity and can activate expression on constructs containing basal promoter regions from the *engrailed* (*en*), *fushi tarazu* (*ftz*) and *nanos* (*nos*) genes (Fig. 3-6). The promoter region of *ftz* is similar to that of *slp1* as it contains a TATA box and Inr element, but lacks a DPE. In contrast, the *en* and *nos* promoters do not contain a TATA box. Indeed, the *nos* promoter region also lacks both the Inr and DPE promoter elements. Taken together, these results provide strong evidence that DESE requires sequences upstream of the core promoter region to initiate transcription and also strongly suggest that PESE is more promiscuous than DESE with respect to the promoter sequences needed to support enhancer-driven transcription activation.

Different promoters affect Runt-dependent regulation

DESE activation requires Runt and Opa (Chapter II, Fig. 2-3) (Swantek and Gergen, 2004). Simultaneous ectopic co-expression of both Runt and Opa induced a strong activation of DESE-p381-*lacZ* expression in the anterior head region (Fig. 3-7B and chapter II, 2-3B). However, this Runt- and Opa-dependent activation of DESE was significantly reduced when DESE was linked to the small p126 *slp1* promoter (Fig. 3-7A), which not only provides additional evidence of DESE's requirement for the promoter-proximal upstream extension, but also indicates that the region between -260 and -79 is specifically required to allow for Runt dependent activation.

Runt is also responsible for PESE repression, as elimination of Runt resulted in de-repression of PESE-*lacZ* expression in type III cells (Chapter II, Fig 2-3A) (Prazak et al., 2010). While ectopic-expression of Runt through out the embryo entirely abolished PESE-p126-*lacZ* expression (Fig. 3-7C), ectopic expression of Runt alone failed to fully repress PESE-p381-*lacZ* expression. Simultaneous ectopic co-expression of Runt and Ftz repressed expression of PESE-p381-*lacZ*, but still was not as potent as Runt repression of PESE-p126-*lacZ*. In contrast, the two promoters p126 and p1381 did not affect Eve-dependent repression of PESE-*lacZ* (Fig 3-7F, G).

These functional assays of DESE and PESE reporter genes with different size promoter regions revealed that Runt dependent *slp1* activation by DESE and repression by PESE are influenced by the promoter. The larger promoter region is required for Runt-dependent activation via DESE, but inclusion of this promoter-proximal upstream extension also reduces the sensitivity of PESE-driven activation to repression by Runt.

Upstream sequence creates a nucleosome-depleted region around promoter

As shown in figure 3-4, the smaller p126 and p200 *slp1* promoter regions do not allow for association of TBP with the promoter, which significantly affects DESE-*lacZ* expression. It has been suggested that nucleosome positioning can regulate gene expression in many organisms (Bai and Morozov 2010). Positioning of a nucleosome on the promoter could create a barrier for TBP binding and thus block transcription initiation. I performed Chromatin IP analysis of Histone H3 to test the nucleosome occupancy on the promoters of different *lacZ* reporters. As expected, on both the DESE and PESE reporters, the small p126 promoter and the p200 promoter lacking the promoter-proximal upstream extension exhibited higher level of H3 association than the large promoter p381 and p307 (Fig. 3-8A and B). All the reporter constructs showed similar levels of H3 association ~300 bp further downstream into the *lacZ* gene.

A more detailed picture of nucleosome positioning around the promoter regions of the DESE-p126 and DESE-p381 reporter genes was obtained using the NOMe-Seq approach, a single molecule technique taking advantage of in vitro methylation of CpG sites with isolated nuclei followed by bisulfite sequencing. DNA sequences assembled into a nucleosome are protected from the methylation and thus revealed by the bisulfite conversion of CpG sites in over a continuous span of around 150bp (Liang et al., 2008).

This experiment revealed that 30% of the DESE-p126 molecules have a nucleosome covering the promoter region, whereas no nucleosomes were found on DESE-p381 (Fig. 3-8C, D and E). These results strongly suggest the involvement of nucleosome positioning in *slp1* regulation, and provides an explanation of the reduced association of TBP with the smaller promoter. The failure of the smaller promoter region to support DESE-*lacZ* expression, and mediate activation in response to Runt and Opa strongly suggests that, in contrast to PESE, DESE-driven activation is TBP dependent and relies on a nucleosome depleted region (NDR).

Discussion:

In this chapter, I presented evidence that transcription initiation on the *slp1* promoter is stimulated by the presence of either the DESE or PESE enhancers. In addition, the activity of DESE and PESE is in turn influenced by the component of the promoter. DESE and PESE distinctively activate the *slp1* promoter and drive proper patterning of gene expression. DESE and PESE contain similar sets of transcription factor binding sites (see chapter IV fig. 4-2). The distinct expression pattern of their *lacZ* reporter genes and their differential responses to the small and large promoter are likely owing to the differential affinity, arrangement, spacing and stoichiometry of the activator and repressor binding sites in these two enhancers.

The involvement of DNA sequences close to gene promoters in transcription regulation are found in many organisms (Bulger and Groudine, 2010). I found here, the sequence immediate upstream of *slp1* core promoter, -260 to -79, affects enhancer activity. A model to explain this phenomenon is this sequence could induce a nucleosome-depleted region (NDR) that allows TBP recognition of TATA box, which is essential for DESE activation (Fig. 3-10A). Without this region nucleosomes assemble around the promoter region, which covers the TATA box and TSS, keeping the TATA box inaccessible to TBP association and thereby not allowing for DESE-driven activation. However, PESE does not rely on TBP to activate transcription, so it is still functional (Fig. 3-10B).

It is interesting that substantial amounts of TBP were detected on p381 alone (Fig. 3-1D). This result suggests that the p381 region alone is sufficient for creating a nucleosome-depleted region that allows for TBP association. It is very likely that TBP is present on the *slp1* promoter in all cells regardless of the transcription status of *slp1*. This is consistent with the previous discovery that in embryos where *slp1* transcription is repressed by Runt and Ftz, TBP was associated with the *slp1* promoter (Wang et al., 2007).

The reporter genes investigated here provide more information to understand the importance of the upstream promoter sequence and suggest a link of Runt function with TBP. Figure 3-9 illustrates models that explain the differential behavior of the reporters. The upstream promoter sequence (-260 to -79bp) contributes to the "attractiveness" of promoter by making a nucleosome free region that allowing TBP to bind so that enhancers are more readily able to "find" the promoter. DESE activation requires Runt (Fig. 2-3). Compared to DESE-p381, the inability of DESE to activate p126 indicates Runt-dependent activation of transcription depends on the presence of TBP (Fig. 3-7A, B). PESE-p126 is able to activate, but no TBP is associated with this reporter (Fig. 3-2B, 3-4B), because PESE activation does not require Runt. Opa is responsible for PESE activation (Fig. 2-5). Thus, Opa activation is very likely independent of TBP. While on the other hand, Runt alone completely represses PESE-p126 (Fig. 3-7C), but not PESE-p381 (Fig. 3-7D). Previously, an extended *PESE[3918-p126]* reporter was used (Prazak et al., 2010). Runt alone was able to repress this reporter, which is consistent with this

PESE[3125-p126] used in this chapter. These results indicate that, with the presence of TBP, Runt functions better as an activator than a repressor. Without TBP, Runt functions better as a repressor than an activator.

Three aspects have been suggested to govern nucleosome positioning: DNA sequence preference, sequence-specific factors and chromatin remodelers (Bai and Morozov, 2010; Iyer, 2012). The 5' UTR of *slp1* contains an A-T rich region that was previously found to exclude nucleosome and induce Pol II pausing (Nechaev and Adelman, 2010). However, the A-T rich region, which is included in p200, showed highest level of H3 association around the promoter, and the p307 construct that did not include this poly-AT region showed the lower level of H3 association consistent with a nucleosome depleted region. This contradicts the involvement of the poly-AT in nucleosome positioning, at least for the *slp1* gene. It is possible that non-histone factors specifically bind to the -260 to -79 region and play a role in nucleosome distribution. Analysis of the sequence of this region identified potential binding sites for GAGA factor, Button-head (Btd) and Zelda (Zld). GAGA factor, encoded by the essential *Trithorax-like* (*Trl*) gene of *Drosophila melanogaster*, has been shown to facilitate transcription of distantly located enhancers by linking them to their cognate promoter through oligomerization when bound to DNA (Katsani et al., 1999; Mahmoudi et al., 2002). Btd is a transcriptional activator related to vertebrate Sp1 and functionally interacts with TBP associated factors (TAFs) to provide interfaces for enhancer-bound transcription factors that contact the basal transcriptional apparatus to activate transcription (Hoey et al., 1993; Smale et al., 1990). Zld is uniformly present in the early embryo and participates in the activation of zygotic genome (Liang et al., 2008). Although initial attempt of mutating the putative binding sites for these three strong candidates individually failed to result in any significant expression defect of DESE (data not shown), it is still likely that combination of these three or other factors might contribute to DESE-driven activation.

In addition to nucleosome occupancy around the promoter, other possibilities including promoter tethering and looping could also explain the effects of upstream promoter region on DESE activity. Nevertheless, the NDR formation mechanism for *slp1* promoter and its requirement for DESE activation requires further elucidation. In metazoans, developmental genes are more likely to be associated with multiple enhancers, and thus a proper promoter architecture is crucial for these genes to respond to multiple regulatory inputs (Lenhard et al., 2012).The mechanisms of enhancer promoter interaction and ability of a promoter to integrate complex regulatory inputs in different cellular contexts requires more attention in future studies. The *slp1* gene certainly provides a unique opportunity for this purpose.

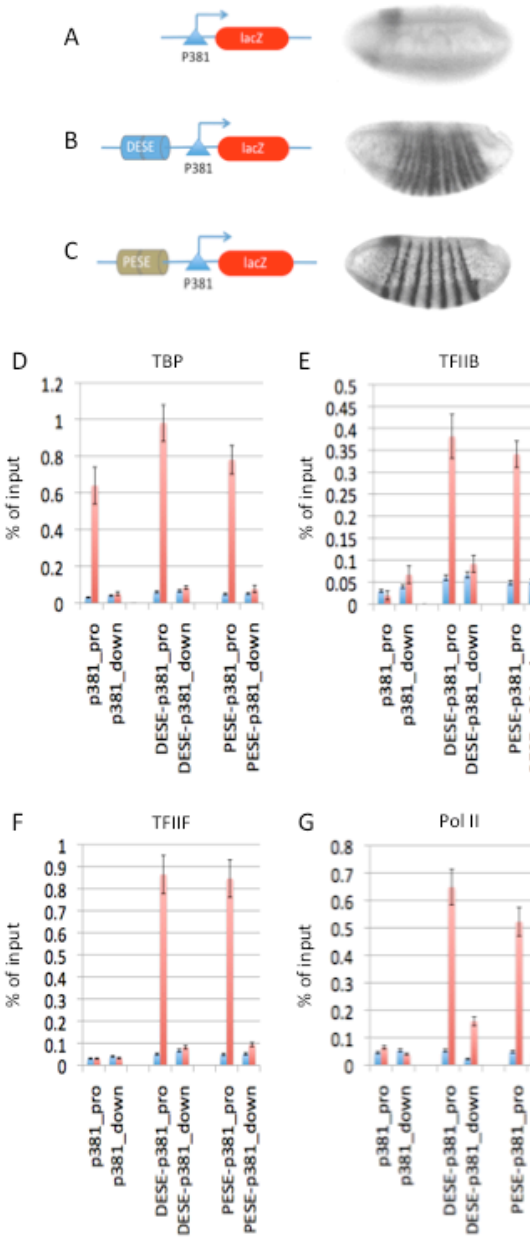

Figure 3-1: DESE and PESE stimulate pre-initiation complex formation on the *slp1* promoter.

Immunohistochemical in situ hybridization of stage 6 embryo carrying reporter construct *p381-lacZ* (A); *DESE[8765]-p381-lacZ* (B) or *PESE[3125]-p381-lacZ* (C). From D to G, ChIP using specific antibodies detecting the association of TBP (D), TFIIB (E), TFIIF (F) and Pol II (G) with the three *lacZ* reporter genes shown in A, B and C. Y-axis is the percentage of the input chromatin, X-axis comparing two regions of the three reporter genes, promoter region (pro) and a downstream region that is in the middle of the *lacZ* gene (down). Rabbit serum was used as negative control (blue). ChIP signals of the three genotypes were normalized with the ChIP signal on the promoter of *hsp70a* gene.

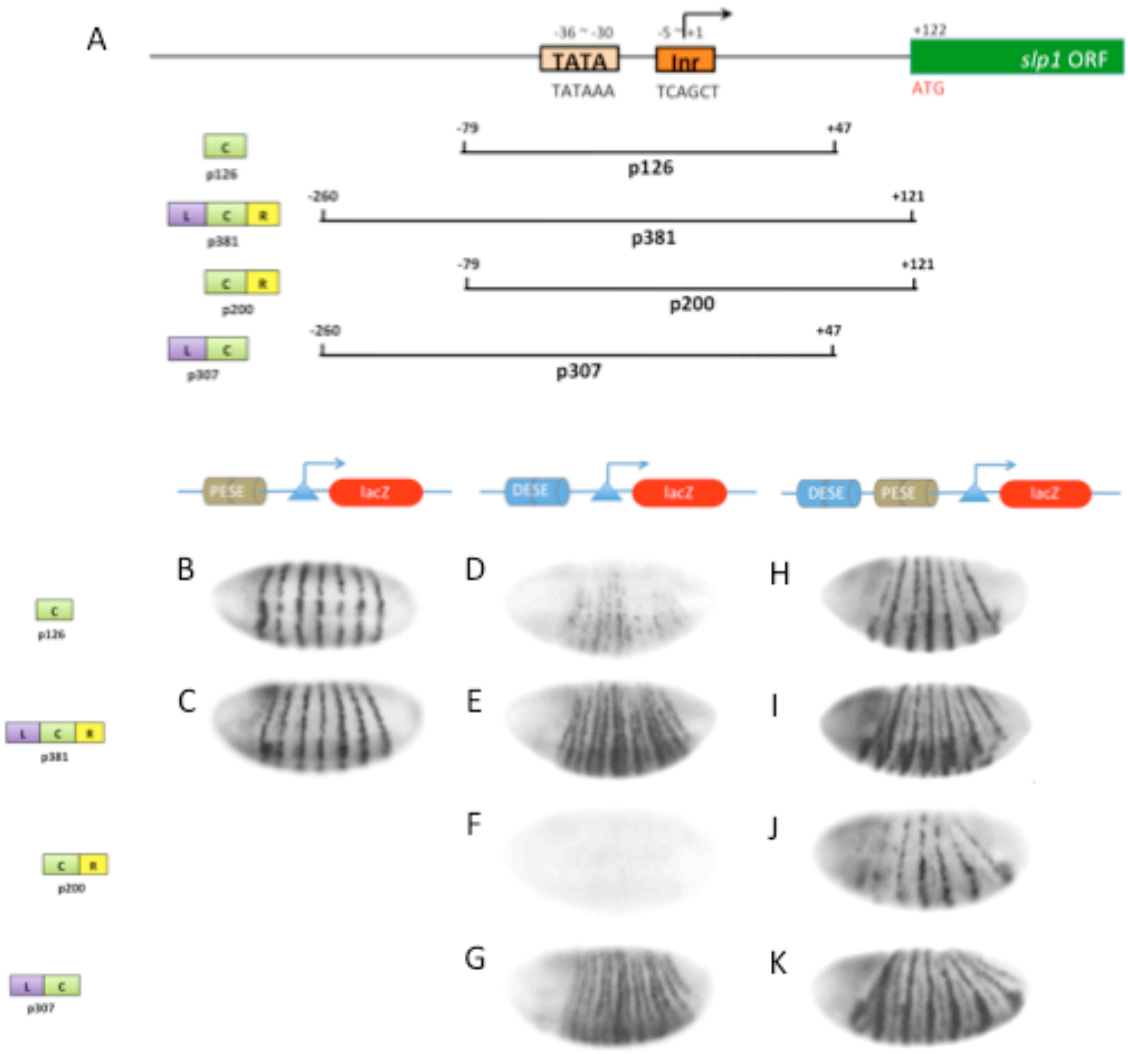

Figure 3-2: The sequence from -260 to -79 base pairs upstream of the *slp1* TSS is important for DESE-, but not for PESE-mediated activation.
(A) schematic representation of *slp1* core promoter. The lines underneath indicate the region included in the four different promoters p126, p200, p307 and p381, spanning the *slp1* transcription start site. From B to K, is the in situ of *lacZ* reporter gene of proximal enhancer *PESE[3125]-p126* (B), *PESE[3125]-p381* (C), distal enhancer *DESE[8765]-p126* (D), *DESE[8765]-p381* (E), *DESE[8765]-p200* (F), *DESE[8765]-p307* (G), composite *DESE+PESE[8765:3918]-p126* (H), *DESE+PESE[8765:3918]-p381* (I), *DESE+PESE[8765:3918]-p200* (J) and *DESE+PESE[8765:3918]-p307* (K).

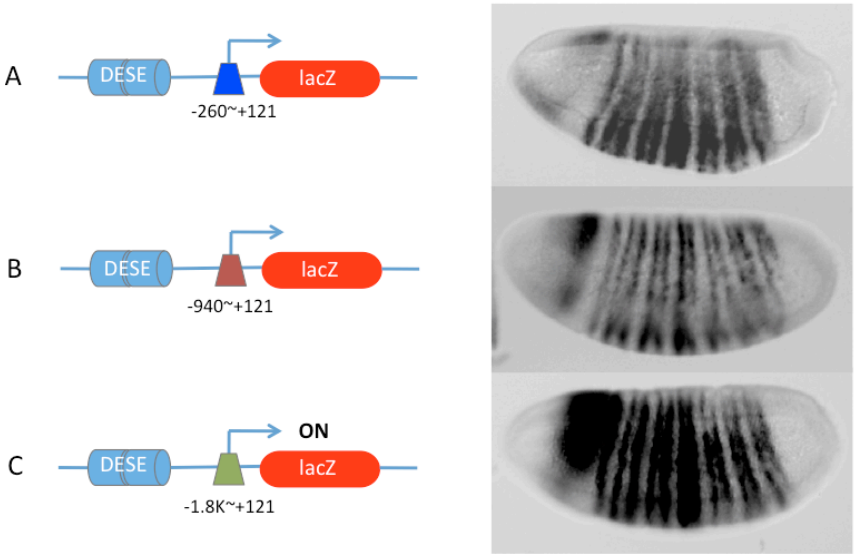

Figure 3-3. Function of DESE with *slp1* large promoters.
in situ of three *DESE[8765]-lacZ* reporters with promoter region spanning from -260 to +121 (A), -940 to +121 (B) and -1.8k to +121 bp (C). These three reporters are P-element constructs randomly inserted in the genome.

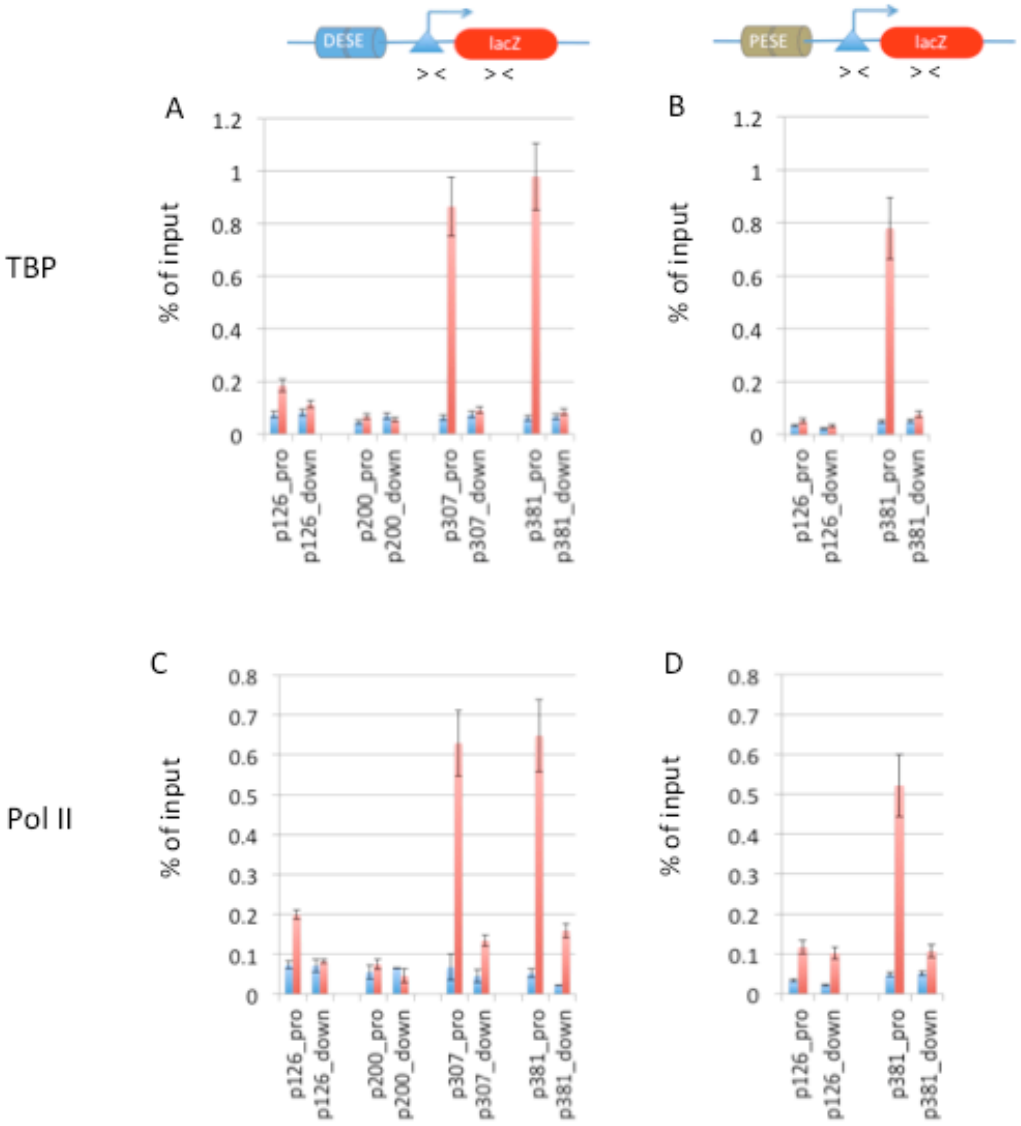

Figure 3-4: The sequence from -260 to -79 base pairs upstream of the *slp1* TSS affects recruitment of TBP and Pol II to the promoter.
ChIP detection of TBP association with four DESE reporter genes (A): *DESE[8765]-p126-lacZ*, *DESE[8765]-p200-lacZ*, *DESE[8765]-p307-lacZ*, *DESE[8765]-p381-lacZ*, and two PESE reporter genes (B): *PESE[3125]-p126-lacZ*, *PESE[3125]-p381-lacZ*. ChIP detection of Pol II association with four DESE reporter genes (C): *DESE[8765]-p126-lacZ*, *DESE[8765]-p200-lacZ*, *DESE[8765]-p30-lacZ* 7, *DESE[8765]-p381-lacZ*, and two PESE reporter genes (D): *PESE[3125]-p126-lacZ*, *PESE[3125]-p381-lacZ*. Y-axis is the percentage of input Chromatin; X-axis shows promoter region (pro) and a downstream region that is in the middle of *lacZ* gene (down) of the individual reporter gene. Arrow heads indicate the position of the primers. Rabbit serum was used as negative control (blue). ChIP signals of the embryos carrying different reporter constructs were normalized with the ChIP signal on the promoter of *hsp70a* gene.

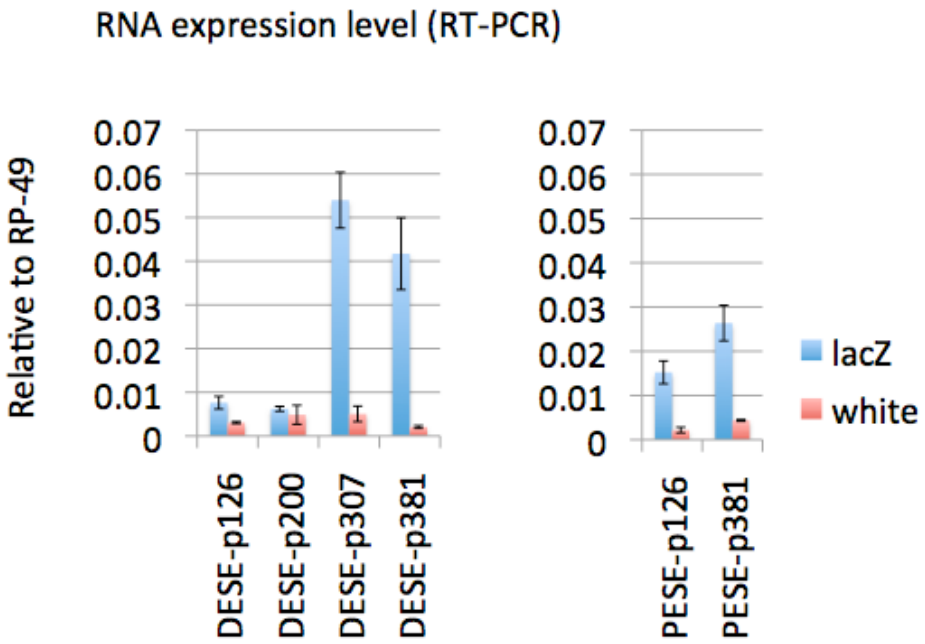

Figure 3-5: Quantitative RT-PCR detection of *lacZ* reporter gene expression.
3-4 hours old embryo were collected and total RNA isolated subjected to reverse transcription assay. Primers corresponding to 1kb downstream of the *lacZ*, *white* and *rp49* genes were used for q-PCR. *white* gene is in the same *lacZ* construct in the opposite direction, used as internal control. Q-PCR signal is relative to *rp49*.

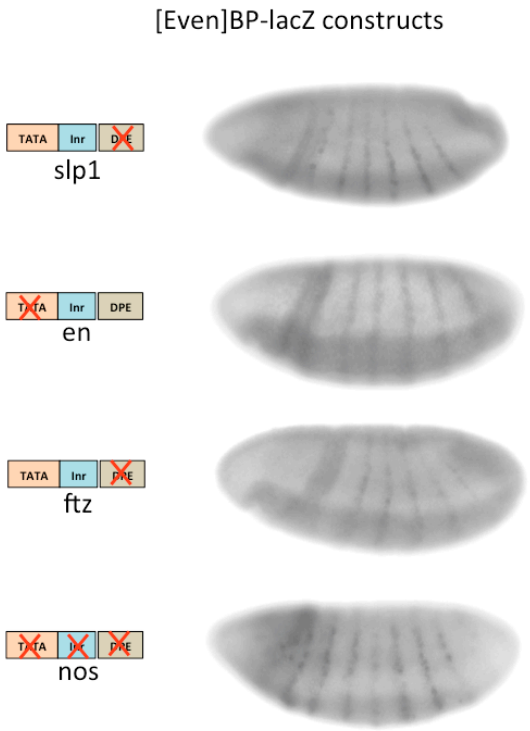

Figure 3-6: Function of PESE with different promoters.
In situ of *PESE-lacZ* constructs with *slp1* promoter (*slp1*), *engrailed* promoter (*en*) that lacks TATA box, *ftz* promoter (*ftz*) that lacks DPE, *nanos* promoter (*nos*) that lacks TATA box, initiator and DPE elements.

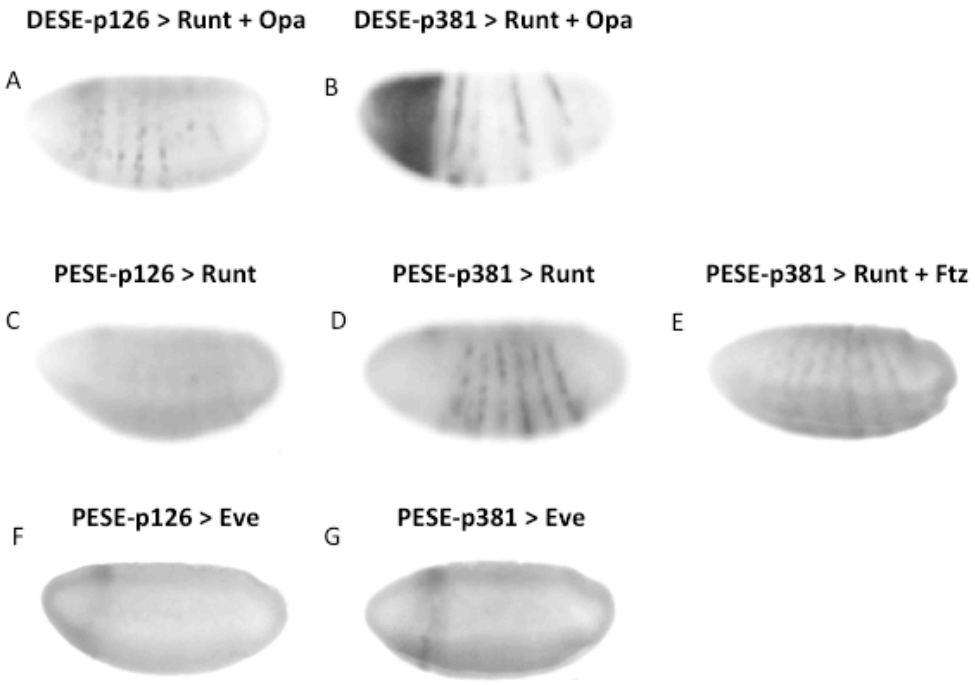

Figure 3-7: Sequence around the *slp1* promoter affects Runt-dependent activation of DESE and repression of PESE.
In situ hybridization detecting the expression of *lacZ* reporter genes *DESE[8765]-p126-lacZ* (A) and *DESE[8765]-p381-lacZ* (B) in Runt and Opa over-expressing embryos; expression of *PESE[3125]-p126-lacZ* (C) and *PESE[3125]-p381-lacZ* in embryos over-expressing Runt; expression of *PESE[3125]-p381-lacZ* in an embryo over-expressing Runt and Ftz (E). Expression of *PESE[3125]-p126-lacZ* and *PESE[3125]-p381-lacZ* in embryos over-expressing Eve. Embryos used in these experiments were obtained from the crosses of between homozygous females with *NGT40* on the second chromosomes and *lacZ* reporter gene on the third chromosomes with the males carrying *UAS-runt15; UAS-opa14* (A and B) or *UAS-runt232* (C and D) or *UAS-runt232*; *UAS-ftz263* (E) or *UAS-eve12* (F and G)

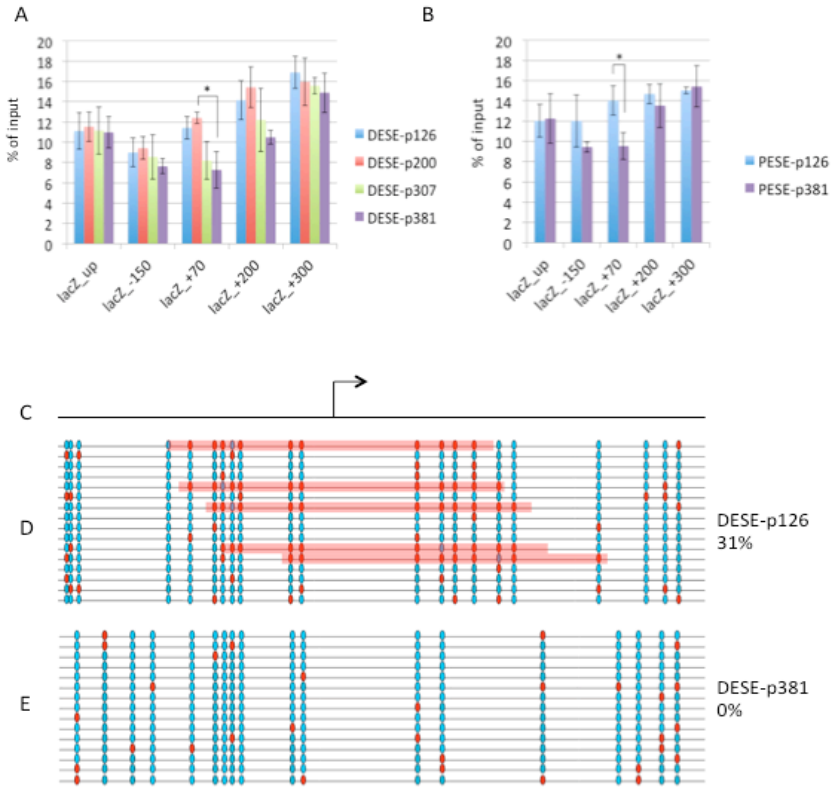

Figure 3-8: The sequence from -260 to -79 base pairs upstream of the *slp1* TSS affects nucleosome assembly around the promoter.

A-B: ChIP using antibody against histone H3 detecting its association with four *DESE[8765]-lacZ* reporter genes (A) and two *PESE[3125]-lacZ* reporter genes (B) in the embryo. Y-axis is percentage of input chromatin, X-axis shows four regions of *lacZ* reporter genes: *lacZ*_up corresponds to the region immediately upstream of the DESE or PESE enhancer in the *lacZ* reporter constructs; *lacZ*_-150 is centered at ~150bp upstream of the TSS of the reporter construct using primers flanking the enhancer:promoter junction; *lacZ*_+70 is centered at ~70 bp downstream of the TSS of the reporter gene using primers flanking the junction of the promoter region and the *lacZ* gene; *lacZ*_+200 and *lacZ*_+300 corresponds to primer pairs centered at +200bp and +300bp to the TSS of the *lacZ* reporter constructs. * $p<0.05$

C-E: NOMe-Seq analysis of nucleosome occupancy for DESE reporter genes. C: representation of 300bp around the transcription start site of the DESE reporter gene. Nuclei were extracted from embryos carrying *DESE[8765]-p126-lacZ* (D) and *DESE[8765]-p381-lacZ* (E) and treated with *M.SssI* CpG methyltransferase and subjected to bisulfite conversion and cloning. Horizontal lines represent single molecular of the region indicated in C. Circles represent CpG dinucleotides (red, unmethylated and inaccessible to *M.SssI;* blue, methylated and accessible to *M.SssI).* Pink bars are 150 bp, regions with several contiguous unmethylated CpG nucleotides representing sites associated with nucleosomes. Numbers on the right indicates the percentage of molecules that are occupied by a nucleosome for each genotype.

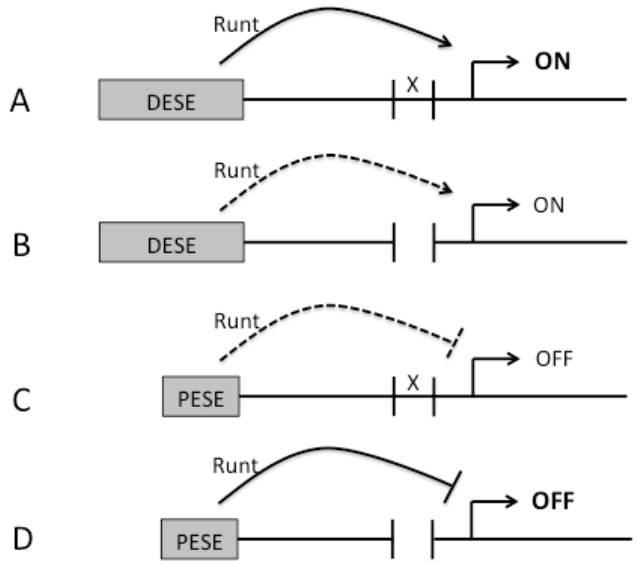

Figure 3-9: Model for the effects of *slp1* upstream promoter sequence on Runt-dependent regulation.
Upstream promoter sequence from -260 to -79bp (area within the two vertical bars) presumably associated with some unknown factor X that can induce a nucleosome free region that allows for TBP binding. DESE mediated activation by Runt is supported by the presence of this region (A), without which Runt-dependent activation is comprised (B). PESE mediates repression by Runt is disfavored in the presence of upstream promoter sequence (C). When this sequence is absent, Runt-dependent repression of PESE is enhanced (D).

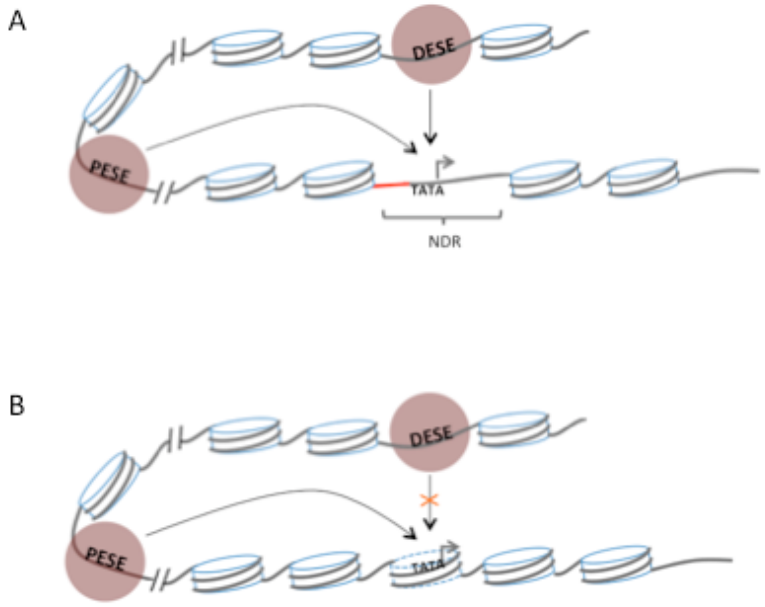

Figure 3-10: A model of enhancer promoter interaction of the *slp1* locus.
DESE and PESE activate *slp1* transcription by stimulation of pre-initiation complex formation on the promoter. This process is affected by the sequence immediately upstream of *slp1* core promoter, from -260 bp to -79 bp upstream of TSS, which induces a nucleosome depleted region (NDR) around the promoter that facilitates enhancer promoter interaction (A). Without this sequence, nucleosome occupancy on the promoter region prevents TBP binding to TATA box, which in turn inhibits DESE dependent activation at the *slp1* promoter. PESE can activate the *slp1* promoter in a TBP-independent way (B).

Chapter IV:

Repression of DESE- and PESE-mediated *slp1* transcription

Abstract

The metameric expression of *slp1* involves repression of both the PESE and DESE enhancers by Runt and other pair-rule transcription factors. PESE is responsible for the Eve-dependent repression of *slp1* in type I cells, and DESE is responsible for the repression by Runt and Ftz in type III cells. By investigating a number of enhancer-*lacZ* reporter genes in wild-type embryos versus genetically manipulated embryos using Chromatin IP, I found Eve represses PESE-dependent expression by regulating transcription elongation. In type III cells, Runt and Ftz repress both DESE and PESE. Interestingly, Runt and Ftz repress DESE-*lacZ* by the same mechanism as Eve-dependent repression of PESE, that is by inhibition of transcription elongation. However, Runt and Ftz repress PESE-*lacZ* by blocking transcription initiation. Finally, I present evidence that Eve-dependent repression of PESE in type I cells occurs in a manner that prevents DESE-dependent activation irrespective of the relative location of the PESE and DESE enhancers and the promoter.

Introduction

There are two types of transcriptional repression: general repression, when repressors sequester or modify a central component of the transcription apparatus so that it is unavailable for transcription, and gene-specific repression, when a particular gene or set of genes is controlled by the activity of a gene-specific repressor or co-repressor (Gaston and Jayaraman, 2003). Gene-specific transcriptional repression plays a central role in gene regulation, especially for developmental genes, whose activity and specificity are strictly controlled (Mannervik et al., 1999).

In contrast to general repressors that usually affect nucleosomes, gene-specific repressors often bind to DNA and they can regulate transcription from binding sites proximal to, or at a distance from, the target promoter (Mannervik et al., 1999). Gene-specific repressor proteins are a large group of diverse proteins that negatively regulate transcription. Repressor proteins have been categorized in a number of ways such as long- or short-range repressors, based on the range of their effects (Courey and Jia, 2001), and as repressors or co-repressors based on whether or not they bind to DNA directly or are recruited by other sequence-specific DNA-binding proteins. However, many repressors are not easily defined due to their ability to act via multiple mechanisms and often with effects that depend on the specific target. For example, the tumour supressor protein Rb uses several mechanisms to repress transcription and represses different promoters using different combinations of mechanisms (Zhang and Dean, 2001). In addition, many developmental regulators, such as the Runx family members, are context-dependent transcription factors that bind DNA and act both as an activator and a repressor depending on the structure of their binding sites, the interacting proteins and other environmental cues (Aronson et al., 1997; Dubnicoff et al., 1997; Kramer et al., 1999; Wheeler et al., 2000). The dual regulatory properties of Runt were previously exemplified by the parasegment-specific effects of Runt on the segment-polarity genes *en*, *wg* and *slp1* (Manoukian and Krause, 1993; Swantek and Gergen, 2004; Tracey et al., 2000).

A great number of studies have been performed to assess the mechanisms of transcriptional activation and repression by these context-dependent regulators in many animal systems (Javed et al., 2000; Peng and Jahroudi, 2002; Sakabe et al., 2012; Seufert et al., 2005). However, it is difficult to define or control the particular context in many of these studies. This study of the Drosophila segmentation pathway with the rich framework of information available provides a major advantage to unravel the mechanisms that are responsible for activation and repression by Runt and other transcription factors.

As described in chapter II, the pair-rule transcription factors Runt, Eve and Ftz regulate *slp1* transcription. Eve represses *slp1* in type I cells through PESE; Runt and Ftz repress *slp1* in type III cells through DESE. The work presented in this chapter investigates the mechanisms by which transcription repression is achieved in each context.

Results

Runt, Eve, Ftz associate with both DESE and PESE

ChIP experiments examining the distribution pattern of pair-rule transcription factors on the *slp1* locus revealed that Runt, Eve and Ftz are all associated with both the DESE and PESE regions. Shown in figure 4-1, primer pairs with ~500 base pair interval were used to tile along the region of DESE, from -8.7 kb to -6.5kb, and PESE, from -3.9 kb to -1.8 kb. All three factors Runt, Eve and Ftz showed similar binding pattern, with broad association across DESE and more restricted association with PESE, indicating different arrangement of binding sites within these two enhancers.

The location of putative binding sites for Opa, Runt, Eve and Ftz within DESE (from -8.1 to -7.1Kb) and PESE (from -3.1 to -2.5 Kb) were analyzed using the JASPAR online database, an open-access database of annotated, high-quality, matrix-based transcription factor binding site profiles for multicellular eukaryotes. The profiles were derived exclusively from sets of nucleotide sequences experimentally demonstrated to bind transcription factors (Sandelin et al., 2004). As shown in figure 4-2, binding sites for the activator Opa are spread over a range of 500 bp in DESE, whereas in PESE they are constricted within a 200 bp region that corresponds well with the PESE:C1 interval required for PESE-driven activation. It is further notable that the putative Opa sites in PESE are tightly surrounded by or overlapping with Eve and Ftz binding sites and removed from most of the putative Runt sites, whereas in DESE the Opa and Runt sites are more interdigitated and frequently removed from putative sites for Eve and Ftz. These observations strongly suggest that differences in enhancer architecture account for the differential responses of these two enhancers to this set of transcription factors.

Differential mechanisms of Runt and Ftz repression of DESE and PESE

Previous studies discovered that repression of the endogenous *slp1* gene by Runt and Ftz occurs at the step of transcription elongation (Wang et al., 2007). Since we believe DESE mediates this repression, the DESE-*lacZ* reporter gene should recapitulate the same response to the elongation control by Runt and Ftz. Indeed, similar to *slp1*, ectopic co-expression of Runt and Ftz repressed DESE-*lacZ* expression, with some residual expression in some cells that also showed residual expression of *slp1* in the body of the embryo (Fig. 4-3A). By comparing wild-type embryos (WT) with Runt and Ftz expressing embryos (RF), ChIP experiments confirmed that DESE-*lacZ* transcription is also repressed by Runt and Ftz at the elongation step of transcription. In RF embryo, DESE-*lacZ* showed abundant association of TBP and PolII with promoter comparable to the signal in WT embryo (Fig. 4-3B, C).

In eukaryotes, the Pol II transcription cycle is coupled with a phosphorylation cycle on the C-terminal domain of the largest subunit of Pol II, which consists of tandem heptapeptide repeats with the consensus sequence $Y_1S_2P_3T_4S_5P_6S_7$ (Meinhart et al., 2005). Phosphorylation on Serine 5 (pSer5) and Serine 2 residue (pSer2) are hallmarks of initiation and elongation, respectively (Hirose and Ohkuma, 2007). Here, an increased

pSer5 signal in RF embryo might reflect a less dynamic turnover of Pol II complexes that are paused on the repressed promoter (Fig. 4-3D). The reduced association of pSer2 modified Pol II both with the *slp1* promoter region as well as with the downstream region of the *lacZ* gene in repressed RF embryos compared to WT embryos provides evidence that repression involves blocking conversion to the elongating form of Pol II (Fig. 4-3E).

Three protein complexes control the early Pol II pausing step of elongation: DSIF and NELF that induce pausing and inhibit elongation, and P-TEFb that releases the pausing and promotes elongation (Li and Gilmour, 2011; Marshall and Price, 1995; Wada et al., 1998a; Yamaguchi et al., 1999). The kinase activity of P-TEFb, which consists of Cyclin T (CycT) and Cdk9, is responsible for the Ser2 phosphorylation and the release of the negative effect from the DSIF and NELF complexes (Fujinaga et al., 2004; Wada et al., 1998b; Yamada et al., 2006). The recruitment of P-TEFb does appear to be a rate-limiting step for transcription. Previous studies showed that the tethering of Cdk9 or CycT to promoters activates gene expression (Bieniasz et al., 1999; Lis et al., 2000; Majello et al., 1999), and the inhibition of P-TEFb reduces total cellular gene expression in human cells (Chao and Price, 2001).

In accordance with the reduced levels of pSer2 modified Pol II obtained in RF repressed embryos, CycT association with the *slp1* promoter was also reduced in RF embryos (Fig. 4-3H). The association of Spt5, a subunit of DSIF complex, was not affected by ectopic co-expression of Runt and Ftz (Fig. 4-3F). However, NELF-E, a component of the NELF complex, showed increased association (Fig. 4-3G), which could be due to the reduced level of P-TEFb as dissociation of NELF complex from the paused complex of Pol II is due to P-TEFb activity (Lis, 2007).

Although genetic experiments demonstrated that DESE mediates *slp1* repression by Runt and Ftz in type III cells (Chapter II Fig. 2-4), inhibition of *slp1* expression in these two cells also requires the prevention of PESE activation. Similar to DESE-*lacZ*, ectopic expression of Runt and Ftz throughout the embryo also significantly repressed PESE-*lacZ* expression (Fig. 4-4A). Interestingly, ChIP detection of TBP and Pol II on PESE-*lacZ* revealed a different repression mechanism. Ectopic co-expression of Runt and Ftz resulted in reduced levels of TBP and Pol II association with the promoter region of PESE-*lacZ* (Fig. 4-4B and C). There is also a reduced pSer2 signal (Fig. 4-4E), which could be due to less Pol II association. The signal of pSer5 is weak compared to the serum control and therefore is inconclusive (Fig. 4-4D), but further examination of other general transcription factors involved in initiation, TFIIB and TFIIF, confirmed that transcription initiation is inhibited when Runt and Ftz are repressing PESE (Fig. 4-4F and G). GAGA factor could function in initiation by facilitating the binding of TFIID and modifying chromatin structure to establish a nucleosome-free region (Giot et al., 2003; Nakayama et al., 2007). Interestingly, there is reduced association of GAGA with the promoter region of the PESE-*lacZ* in response to Runt and Ftz (Fig. 4-4H), indicating that the inhibition of transcription extends beyond blocking recruitment of general transcription factors and may occur at a step before the recruitment of TBP.

P-TEFb is involved in Runt-dependent *slp1* regulation

ChIP experiments with the DESE-*lacZ* reporter gene showed that Runt and Ftz affect the recruitment of CycT and Serine 2 phosphorylation, which suggests the involvement of P-TEFb in the Runt-dependent regulation of endogenous *slp1* transcription. This initial experiment involved comparing WT embryos with RF embryos. If P-TEFb recruitment were central to the Runt-dependent regulation of *slp1* expression, it would show an increased association in embryos that have increased levels of Runt-dependent *slp1* expression. In order to test this hypothesis, I generated embryos that use the NGT expression system to simultaneously co-express Runt, Opa and Hairy (ROH). A central function of *hairy* in segmentation is to repress expression of *ftz*, which in turn should allow for activation by Runt and Opa in cells where Ftz has now been eliminated, i.e. in nearly all cells of the ROH embryos (Fig. 4-5A). Chromatin IP experiments comparing the *slp1*-expressing ROH and *slp1*-repressed RF embryos show Pol II association with *slp1* promoter was slightly increased in activated embryos, whereas the Pol II signal was as abundant in repressed embryos as in wild-type embryos. Compared to wild-type embryo, the pSer2 signal was slightly reduced in repressed embryo but increased more than two fold in the activated ROH embryos (Fig. 4-5B). Consistent with the pSer2 signal, CycT association was also increased on the activated promoter and reduced on the repressed promoter. These results provide strong evidence for the involvement of P-TEFb in Runt dependent regulation of the *slp1* DESE enhancer.

Mechanism of Eve repression of PESE

Previous work demonstrated that Eve represses *slp1* in type I cells through PESE. To test what mechanisms are involved in PESE mediated *slp1* repression by Eve, I took advantage of the PESE:C1+ reporter construct, which is insensitive to repression by Runt and Ftz (Fig. 4-6A) (Prazak et al., 2010). As shown in chapter II, the PESE:C1 region is responsible for PESE activation in type IV cells and removal of this region abolished PESE-driven expression (chapter II, Figure 2-6B). The PESE:C1+-*lacZ* reporter is expressed in type III and IV cells that do not contain Eve and was repressed in type I and II cells that have Eve (Fig 4-6A). Ectopic expression of Eve throughout the embryo eliminated expression of the PESE:C1+-*lacZ* reporter in all cells (Fig. 4-6A). Importantly, and very similar to DESE-*lacZ* in RF embryos, the PESE:C1+-*lacZ* promoter was also associated with TBP and Pol II in response to repression by Eve (Fig. 4-6B, C). Further, the significant reduction of pSer2 and CycT association with the reporter gene promoter indicates an inhibition of transcription elongation (Fig. 4-6E, H). Increased pSer5 association with promoter in these Eve expressing embryos confirmed that transcription initiation was not prevented (Fig. 4-6D). A slight reduction of Spt5 association and a significant enhancement of NELF-E association were also observed (Fig. 4-6F, G) as found for repression of DESE by Runt and Ftz. Thus, PESE mediated *slp1* repression by Eve in type I cells also occurs at the step of transcription elongation, very likely involving inhibition of P-TEFb association with the paused complex. Transcription initiation of PESE-*lacZ* is not affected by Eve. This observation indicates that the interactions with the promoter that account for the PESE-driven increase in initiation on the p381-*lacZ* are not affected by the presence of Eve.

Interestingly, DESE alone is able to activate transcription in type I cells, but the composite reporter gene DESE+PESE did not show expression in these two cells (Fig. 4-7). One possibility is that PESE bound by Eve can block DESE-dependent activation of the *slp1* promoter due to its chromosome location, where it is positioned in between DESE and the *slp1* promoter and can potentially prevent signal tracking from DESE to promoter. One test of this tracking model is to flip the order of these two enhancers. As shown in figure 4-7, placing PESE upstream of both DESE and the promoter did not affect the inhibition of DESE-driven activation in type I cells. Thus suppression of DESE by PESE is independent of their positions relative to promoter. This experiment together with the ChIP experiments presented above strongly suggests that in the absence of Runt, PESE can loop and physically interact to promote initiation of transcription on the *slp1* promoter. In the absence of Eve this results in expression, whereas in the presence of Eve elongation is blocked resulting in Pol II that is paused in a complex with NELF.

Discussion

In the Drosophila segmentation network, transcription repressors are traditionally classified into two distinct groups: short-range and long-range, based on their local or global range of action. Short-range repressors such as Knirps quench the nearby activator within 100 bp and permit other enhancer function, whereas a long-range repressor such as Hairy functions over a distance of 1kb (Gray and Levine, 1996a, b). This distinction between the two has been suggested to stem from the co-repressors involved: short-range repressors recruit CtBP and long-range repressors recruit Groucho (Courey and Jia, 2001; Martinez and Arnosti, 2008; Nibu et al., 1998). However, other studies found the differential recruitment of these two distinct co-repressors by short- and long-range repressors are not absolute and many repressors can recruit both CtBP and Groucho (Payankaulam and Arnosti, 2009; Poortinga et al., 1998). Moreover, a recent study, showing that Knirps does not affect TBP and Pol II recruitment to the promoter but Hairy inhibits the recruitment of basal machinery, provides a new insight into the mechanistic distinction between short- and long-range repressors (Li and Arnosti, 2011). Our observations on Eve repression of PESE and Runt+Ftz repression of DESE both involve inhibition of transcription without affecting TBP and Pol II recruitment and are very similar to Knirps, which suggests a short-range effect. Thus whether or not short- and long-range repressors share a common co-repressor or target the same step of transcription cycle within their own groups is still indistinct and very likely depends on the specific target gene. Any classification of repressor is complicated by the complexity and diversity of protein-protein interactions involved in regulation of transcription.

Control of transcription elongation is gaining more and more recognition as a predominant mechanism of gene regulation beyond initiation, and many developmental control genes are associated with paused Pol II (Chiba et al., 2010; Core and Lis, 2008; Gilmour, 2009; Muse et al., 2007; Sims et al., 2004; Wang et al., 2007). Three complexes are involved in the control of the pausing: DSIF, NELF and P-TEFb. Although it is possible that transcription could be regulated by differential recruitment of the negative factor DSIF and NELF, there is no strong evidence that this occurs under normal cellular conditions (Price, 2008). Knockdown of NELF-B by siRNA reduced Pol II association on promoters of 60% of the genes with poised polymerase; however, there was no concomitant increase in the polymerase density in the coding regions of these genes (Muse et al., 2007). Also, knockout of NELF-A did not affect *slp1* expression in the early embryo examined by in situ hybridization (Wang et al., 2010). P-TEFb is the only known factor that can rescue poised polymerases by generating productive elongation complexes (Price, 2008). Here we found *slp1* repression by Eve and by Runt+Ftz both occur at the step of early elongation and inhibition of P-TEFb recruitment could be the common mechanism. In fact, one of the mammalian homologs of Runt, Runx1, was found to inhibit transcription elongation of CD4 gene during T-cell development by directly binding to P-TEFb (Jiang et al., 2005). Eve has been shown to directly interact with TBP, which previously was interpreted as direct evidence of Eve preventing pre-initiation complex formation (Han and Manley, 1993; McKay et al., 1999). Alternatively, this could be explained as a strategy to target promoter by Eve bound enhancer element

thereby inhibit the transcription process close to the promoter such as recruitment of P-TEFb.

How do transcription factors bound to their cis-regulatory elements several kilo-bases away regulate the processes occuring close to the promoter? Substantial evidence indicates remote enhancers can loop to the promoter region of target genes (Bulger and Groudine, 1999; Engel and Tanimoto, 2000; Li et al., 2006). More importantly, a recent study on the β-globin gene demonstrated that looping of the enhancer to the promoter is crucial for transcription initiation but productive elongation requires additional positive input from activators to recruit P-TEFb complex (Deng et al., 2012). This provides an attractive model to explain the Eve and Runt+Ftz repression of *slp1*. In type I cells, PESE directs Eve to the promoter to initiate transcription but it also prevents P-TEFb recruitment to *slp1* promoter while simultaneously also preventing DESE looping to promoter to activate expression (Fig. 4-8A). In type III cells, Runt and Ftz interact with PESE in a manner that blocks the ability of this enhancer to recruit Pol II and stimulate transcription initiation while also interacting with DESE in a manner that allows this enhancer to promote initiation, but not the productive elongation of transcription by Pol II (Fig. 4-8B).

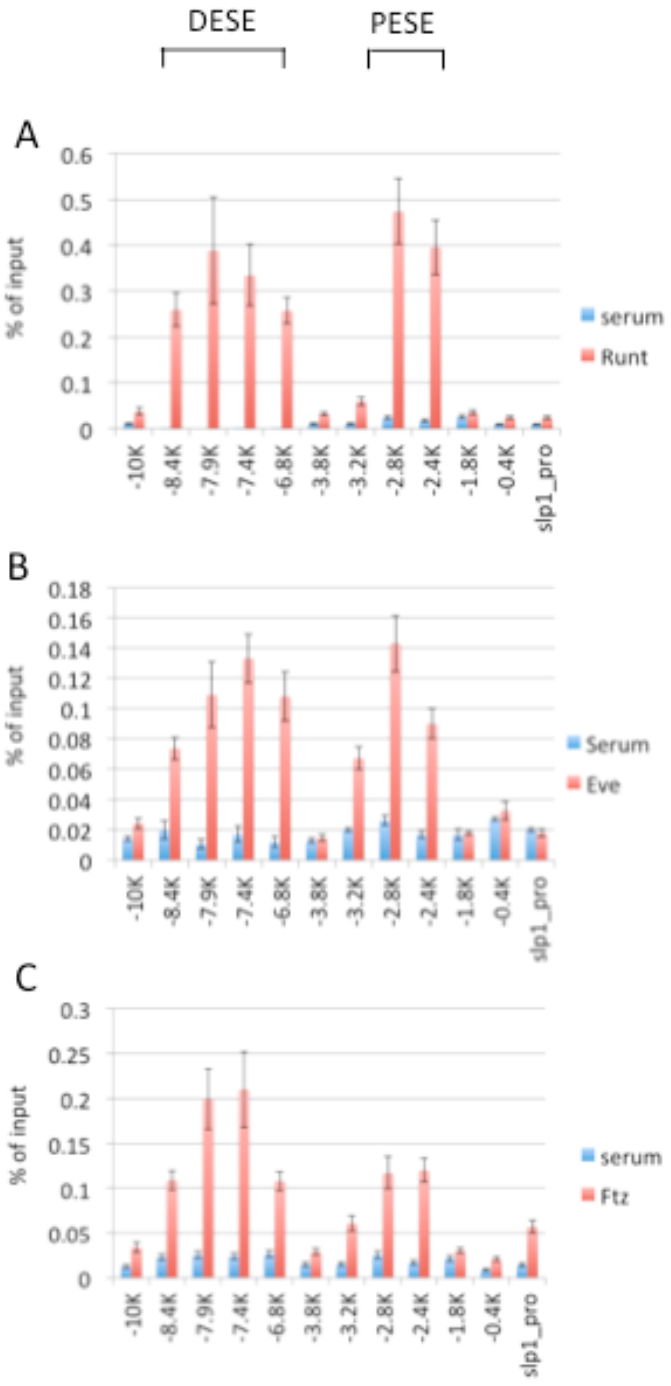

Figure 4-1: Association of Runt, Eve and Ftz with the *slp1* locus.
ChIP using antibodies against Runt (A), Eve (B), and Ftz (C). Rabbit serum is used as negative control. Y-axis is the percentage of input chromatin; X-axis represents the central position of the primer pair used for Q-PCR relative to the *slp1* transcription start site.

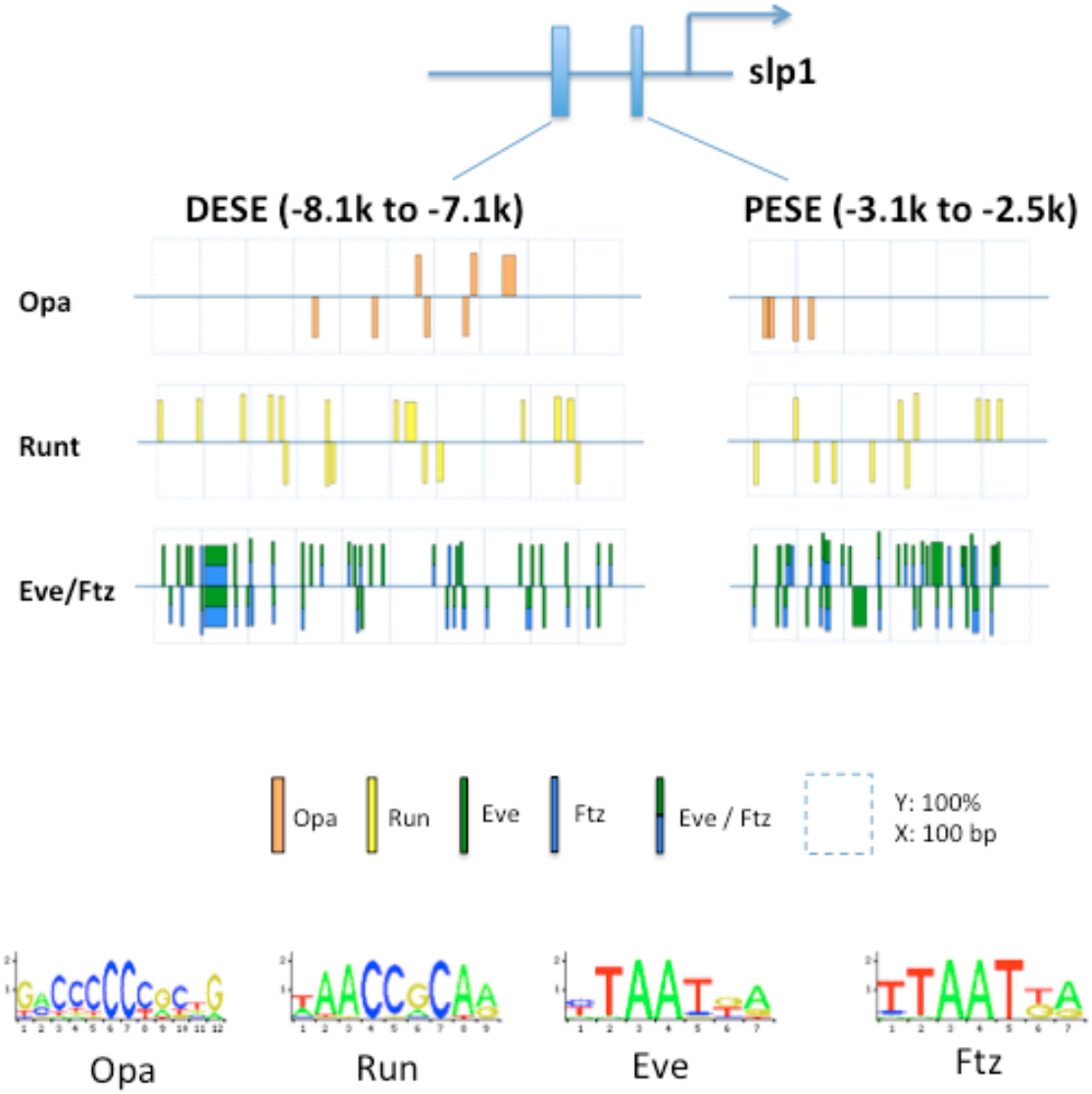

Figure 4-2: Schematic mapping of DNA binding sites of Opa, Runt, Eve and Ftz within DESE and PESE based on JASPAR database predictions.
(http://jaspar.genereg.net/cgibin/jaspar_db.pl?rm=browse&db=core&tax_group=insects). The relative profile score threshold was set as 70%. The output scores were plotted in a square with its Y-axis represents relative score of 100% and X-axis represents 100 bp interval. Orange bar is Opa binding site; yellow is Runt; green is Eve; blue is Ftz; half green and half blue represents the overlapping binding sites for Eve and Ftz. The bars on top of the line represent the plus strand, and the ones below the line represents the minus strand.

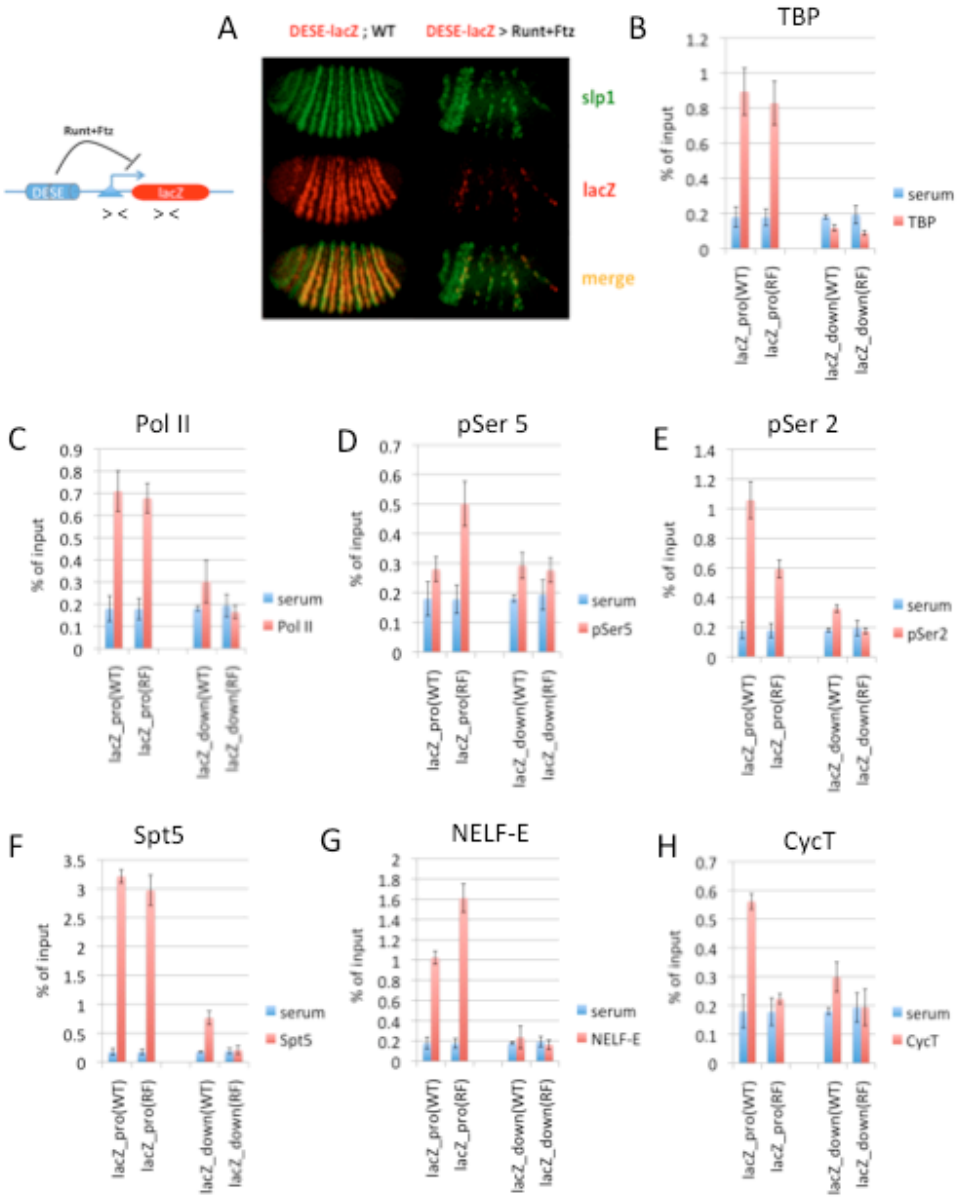

Figure 4-3: Runt and Ftz inhibit *DESE-lacZ* transcription elongation.
(A) Expression of *slp1* (green) and *DESE[*8765p381]-*lacZ* (red) in wild-type embryo versus Runt and Ftz over-expressed embryo. From B to H, ChIP of TBP (B), Pol II (C), pSer5 (D), pSer2 (E), Spt5 (F), NELF-E (G) and CycT (H). Y-axis is percentage of the input Chromatin; X-axis comparing two regions of DESE reporter gene in wild-type (WT) embryo and Runt and Ftz over-expressed embryo (RF): the promoter region (*lacZ*_pro) and the middle of *lacZ* gene (*lacZ*_down). ChIP Signals from two genotypes are normalized to the signal on hsp70a promoter. Arrow-heads underneath the schematic DESE reporter construct indicate position of the primer pairs used to detect promoter region and downstream region of this reporter gene.

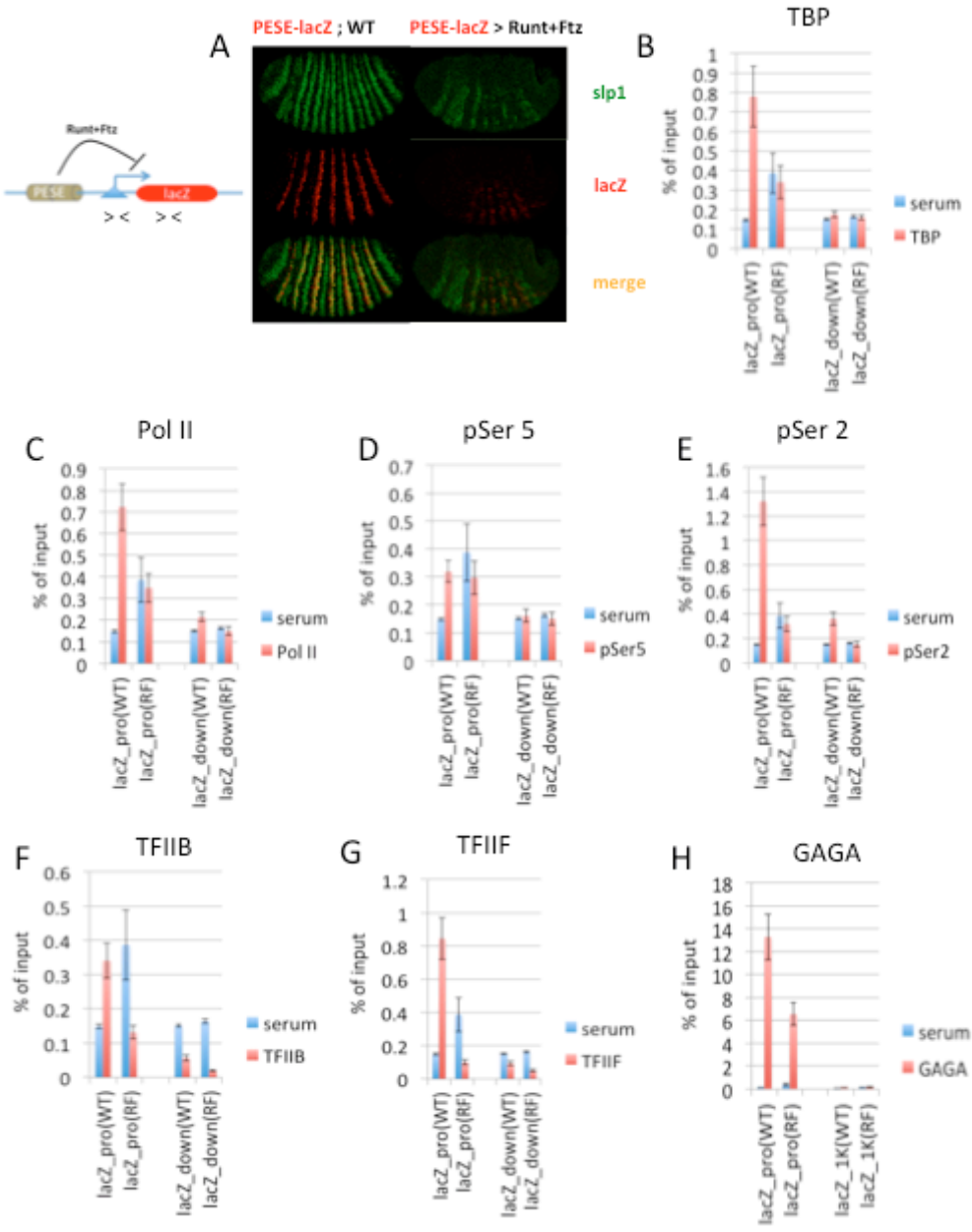

Figure 4-4: Runt and Ftz inhibit *PESE-lacZ* transcription initiation.
(A) Expression of *slp1* (green) and *PESE[3125p381]-lacZ* (red) in wild-type embryo versus Runt and Ftz over-expressed embryo. From B to H, ChIP of TBP (B), Pol II (C), pSer5 (D), pSer2 (E), TFIIB (F), TFIIF(G) and GAGA factor (H). Y-axis is percentage of the input Chromatin; X-axis comparing two regions of PESE reporter gene in wild-type (WT) embryo and Runt and Ftz over-expressed embryo (RF): the promoter region (*lacZ*_pro) and the middle of *lacZ* gene (*lacZ*_down). ChIP Signals from two genotypes are normalized to the signal on hsp70a promoter. Arrow-heads underneath the schematic PESE reporter construct indicate position of the primer pairs used to detect promoter region and downstream region of this reporter gene.

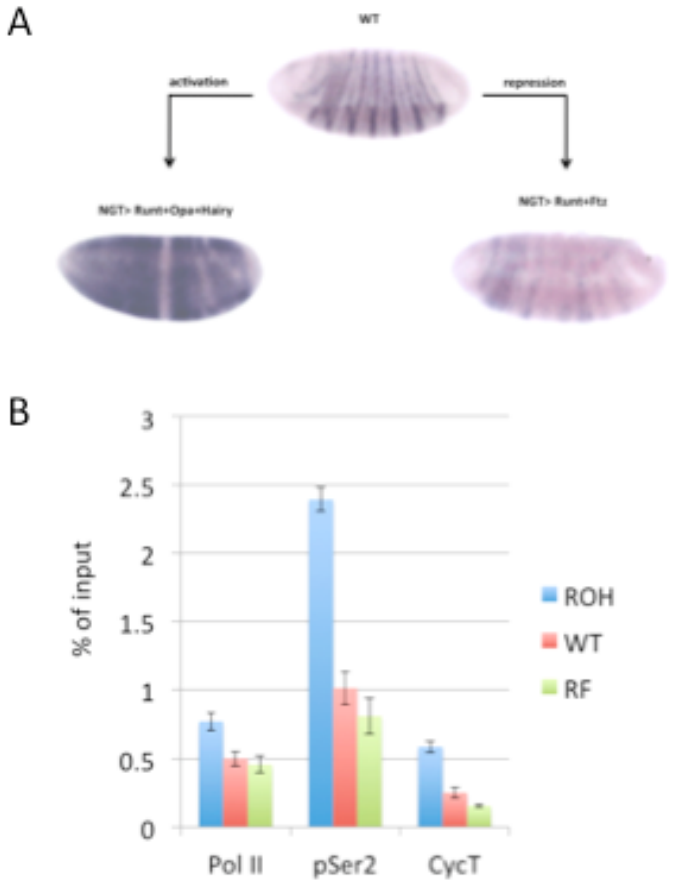

Figure 4-5: Runt dependent *slp1* activation versus repression.
(A) Immunohistochemical in situ of *slp1* gene of wild-type embryo (WT); embryo that simultaneously over-expressed Runt, Opa and Hairy (ROH); embryo that over-expressed Runt and Ftz (RF). (B) Chromatin IP detecting the association of Pol II, pSer2 and CycT with the promoter region of slp1 in Runt, Opa and Hairy over-expressed embryo (ROH); wild-type embryo (WT); and Runt and Ftz over-expressed embryo (RF). Signals were normalized with signals on the *hsp70a* promoter of respective genotypes.

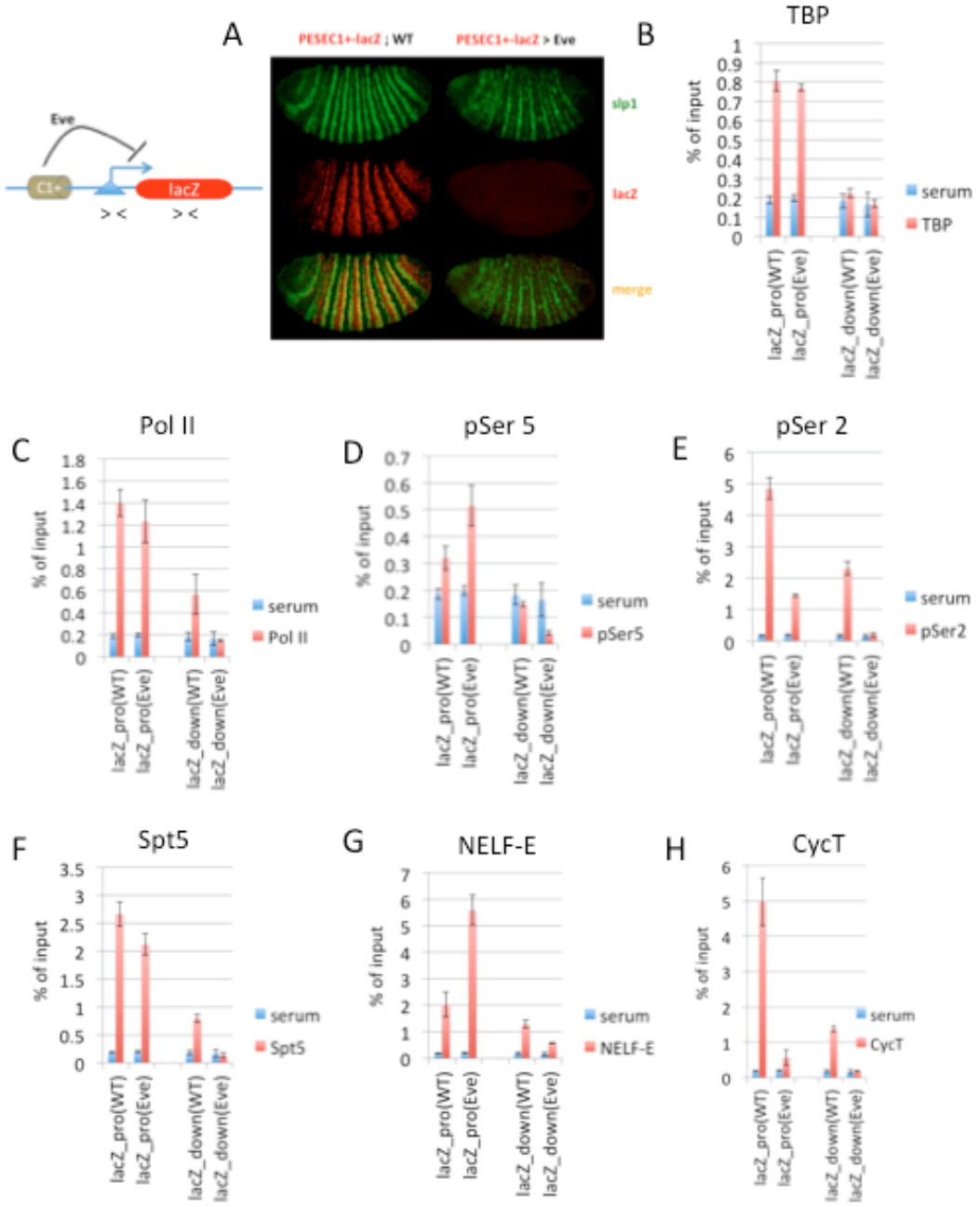

Figure 4-6: Eve inhibits *PESE:C1+-lacZ* transcription elongation.
(A) Expression of *slp1* (green) and *PESE:C1+-lacZ* (red) in wild-type embryo versus Eve over-expressed embryo. From B to H, Chromatin IP of TBP (B), Pol II (C), pSer5 (D), pSer2 (E), Spt5 (F), NELF-E (G) and CycT (H). Y-axis is percentage of the input Chromatin; X-axis compare two regions of PESEc1+ reporter gene in wild-type (WT) embryo and Eve over-expressed embryo (Eve): the promoter region (*lacZ*_pro) and the middle of *lacZ* gene (*lacZ*_down). ChIP Signals from two genotypes are normalized to the signal on *hsp70a* promoter. Arrow-heads underneath the schematic C1+ reporter construct indicate position of the primer pairs used to detect promoter region and downstream region of this reporter gene.

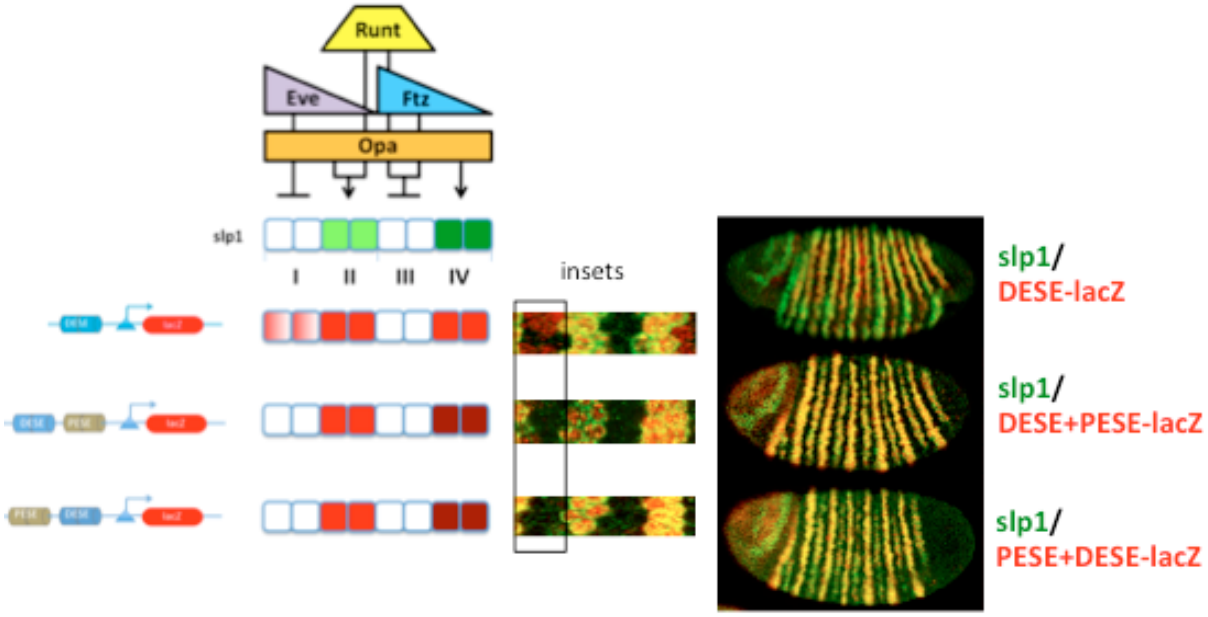

Figure 4-7: PESE mediates *slp1* repression by Eve independent of its position relative to DESE.
On the left is schematic representation of the pattern of *slp1*, *DESE[8765p381]-lacZ*, *DESE+PESE[8765:3918p381]-lacZ* and *PESE+DESE[3918:8765p381]-lacZ* according to respective inset in the middle, which come from the merged confocal images of fluorescent double in situ of *slp1* (green), and *lacZ* (red) on the right.

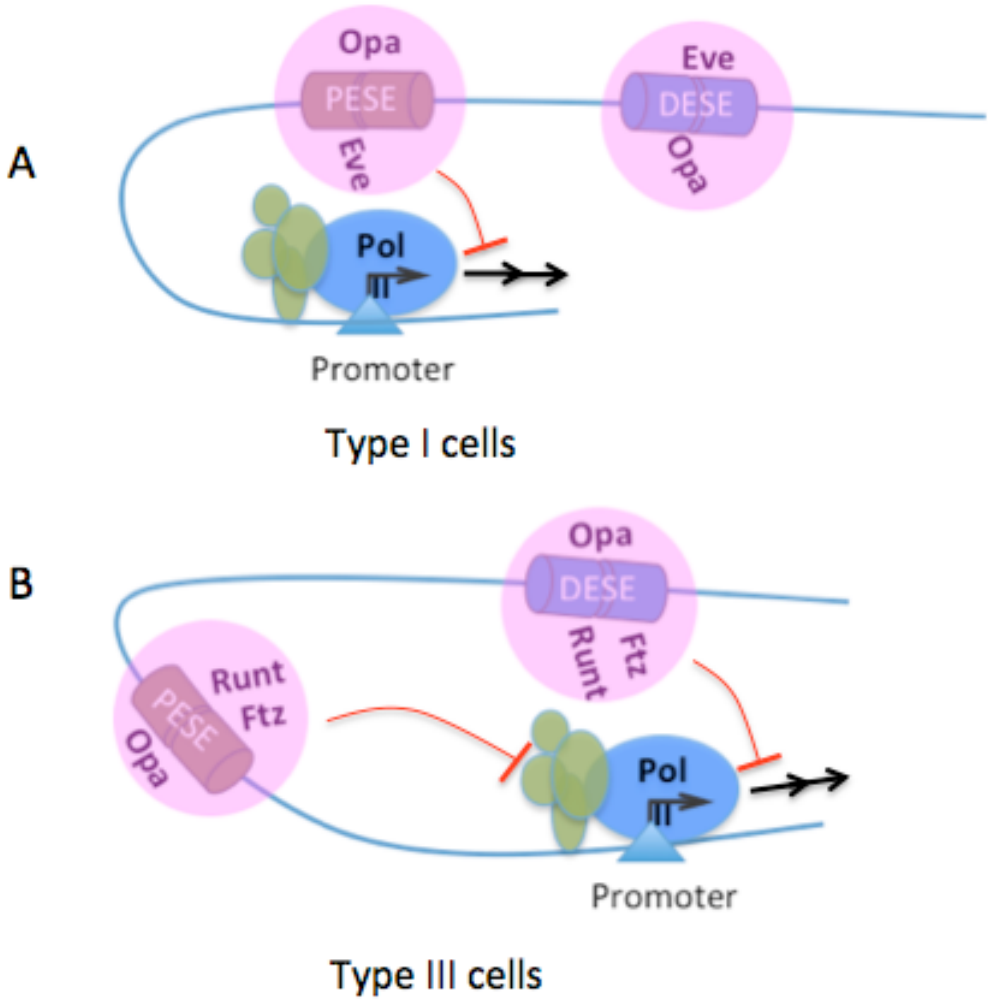

Figure 4-8: A model for transcription repression of *slp1* in type I and III cells.
(A) In type I cells, PESE mediates Eve dependent *slp1* repression. Eve inhibits the process of early transcription elongation by directly target the paused Pol II complex close to the promoter likely via preventing the association of P-TEFb. The interaction of DESE and promoter is prevented by the presence of PESE. (B) In type III cells, DESE mediates slp1 repression by Runt and Ftz. Similar to Eve, Runt+Ftz association with DESE target slp1 paused Pol II complex and thereby block the P-TEFb recruitment to inhibit transcription elongation progress. Runt and Ftz also prevents PESE mediated transcription initiation on slp1 promoter.

Chapter V:

Contributions of Runt conserved domains in Runt-dependent *slp1* regulation

Preface

Runt is responsible for *slp1* activation in type II cells and repression in type III cells. As the founding member of the Runx family regulators, Runt shares two characteristic features with mammalian Runx proteins: the 126-amino acid Runt domain, which is responsible for the DNA binding to the consensus sequence ACC(A/G)CA and the five amino acid VWRPY motif at the C-terminal end that has been shown to act as a platform for the recruitment of Groucho/TLE transcriptional co-repressors (Ito, 2004). Previous work published by Walrad et al. shows that different conserved regions of Runt have distinct functions (Fig. 5-1) (Walrad et al., 2010). Here I extended the work by looking at two of the regions from the C-terminal of Runt, region VII and VIII, to see if constructs lacking those have different effects on enhancer reporter genes. Over-expression of these two RuntΔ7 and RuntΔ8 derivatives using NGT/UAS system provides dominant negative phenotypes, which reveal the attributes of these two regions in the DESE and PESE regulation.

Results and Discussion

Runt region VII and VIII is not required for DESE activation

DESE activation requires Runt and Opa (chapter II, Fig. 2-3; Fig. 2-4A). Over-expression of both full-length Runt and Opa activated *slp1* in the anterior head region of the embryo (Fig. 5-2B). Over-expression of RuntΔ7 and Opa or RuntΔ8 and Opa both resulted a strong head activation (Fig. 5-2C, 5-2D), which indicates that both region VII and region VIII are not required in DESE mediated activation of *slp1* by Runt and Opa. Interestingly, although over-expression of RuntΔ7 and Opa showed very similar *slp1* and DESE-*lacZ* pattern to the over-expression of full length Runt and Opa pattern (Fig. 5-2C), over-expression of RunΔ8 and Opa showed differential pattern for both endogenous *slp1* and DESE-*lacZ* (Fig. 5-2D). DESE-*lacZ* and *slp1* expression remained in the type IV cells suggest RuntΔ8 fails to repress DESE-*lacZ* and *slp1* contrast to full length Runt and RuntΔ7. Also, distinct from full length Runt and RuntΔ7, RuntΔ8 did not fully activate endogenous *slp1* expression in type I cells indicates, like in WT embryos, PESE is still mediating Eve dependent repression in these two cell in the presence of RuntΔ8, which cannot induce DESE-promoter interaction as full length Runt and RuntΔ7. In contrast, DESE-*lacZ* expression was enhanced in type I cells by RuntΔ8 further confirms that RuntΔ8 retains the ability to activate DESE. Similar results were observed for the composite construct DESE+PESEΔC1 (Fig. 5-4B), however, this construct remained its sensitivity to Eve due to the retained sequenced in PESE, thus it faithfully recapitulated the endogenous *slp1* pattern in the Runt Δ8 and Opa over-expressing embryos. All in all, these experiments demonstrate that RuntΔ8 is able to activate DESE together with Opa but also suggest RuntΔ8 fails to repress DESE in combination with Ftz.

Runt region VIII is required for repression of both DESE and PESE

To further investigate the ability of RuntΔ8 to repress DESE, we examined DESE-*lacZ* reporter gene expression in embryos simultaneously over-expressing RuntΔ8 and Ftz. Over expression of full length Runt or RuntΔ7 together with Ftz almost eliminated all *slp1* and DESE-*lacZ* expression with some residual expression in type I cells (Fig. 5-3B, 5-3C). In contrast, RuntΔ8 and Ftz did not repress DESE or *slp1* expression in type IV cells (Fig. 5-3D), which were confirmed by the expression of composite construct DESE+PESEΔC1 in the same genetic background (Fig. 5-4C, 5-4D). Thus, RuntΔ8 loses the ability to repress DESE and *slp1*, so region VIII is required for DESE repression, suggesting the involvement of Groucho in DESE-mediated repression of *slp1* by Runt.

Runt and Ftz can also inhibit PESE activation in type III cells. Similar to DESE, over-expression of RuntΔ8 and Ftz failed to repress PESE-*lacZ* expression, in contrast to Runt+Ftz and Runt Δ7+Ftz (Fig. 5-5). Therefore Runt region VIII is also involved in the repression of PESE dependent activation.

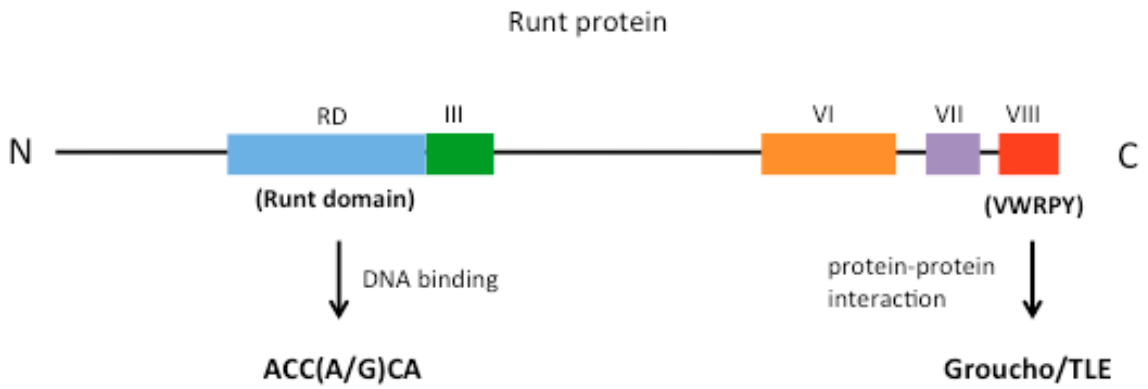

Figure 5-1: conserved domains of Runt.
Runt contains four conserved domains: III, VI, VII and VIII among 12 *Drosophila* species and other insects. It shares a DNA binding Runt domain (RD) and a C-terminal VWRPY with other mammalian Runx family members. Runt domain binds to the consensus sequence ACC(A/G)CA. VWRPY, at the C-terminal end of regions VIII, is involved in the interaction with cofactor Groucho or mammalian TLE.

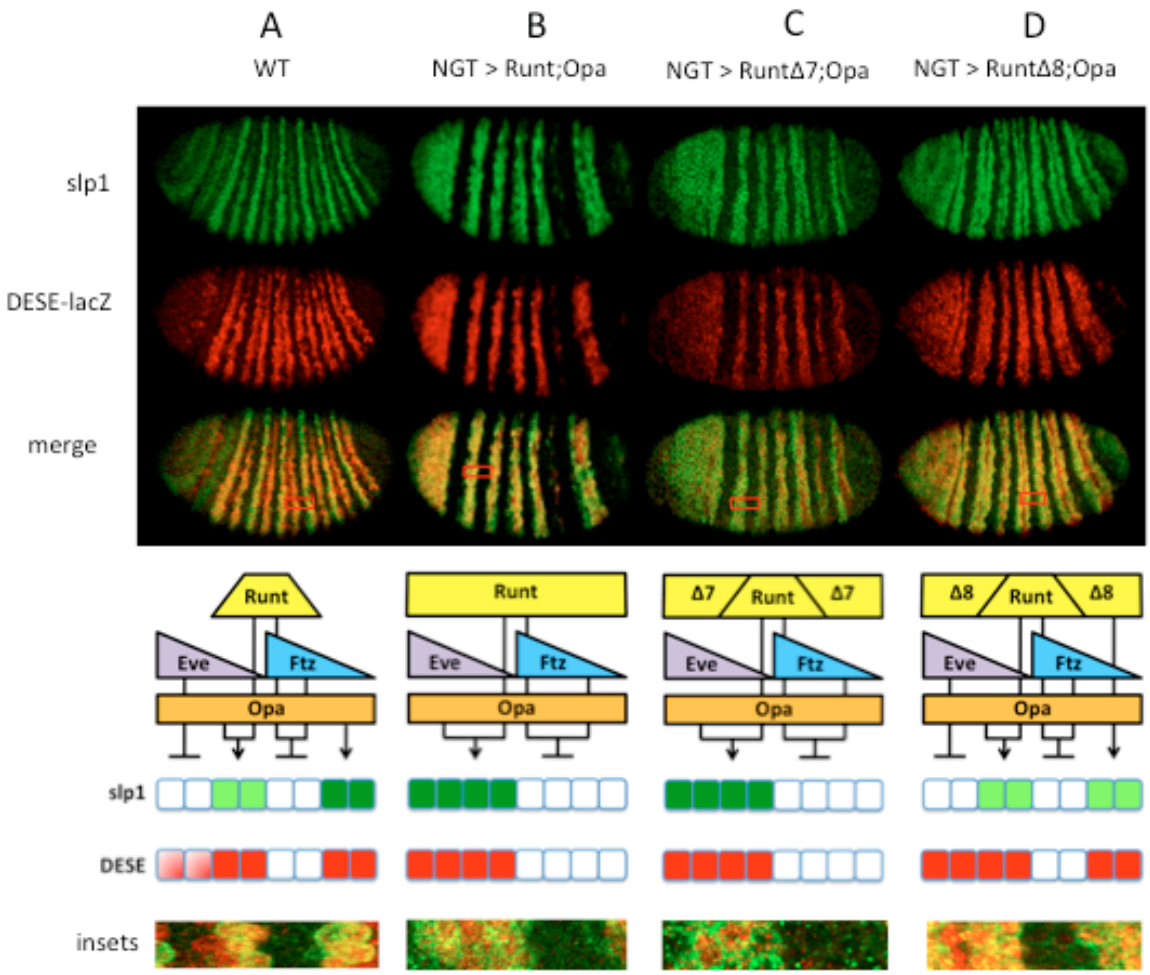

Figure 5-2: Region VII and region VIII is not involved in DESE-mediated *slp1* activation by Runt and Opa.
Fluorescent in situ hybridization of *slp1* (green) and *DESE[8765p381]-lacZ* (red) in wild-type embryo (A), Runt and Opa over-expressed embryo (B), RuntΔ7 and Opa over-expressed embryo (C), and RuntΔ8 and Opa over-expressed embryo (D). Gain of function of *runt* and *opa* embryos were obtained from the crosses between *NGT40; 8765p381-lacZ* females and *UAS-runt15; UAS-opa14* males or *UAS-runtΔ7; UAS-opa14* males or *UAS-runtΔ8; UAS-opa14* males.

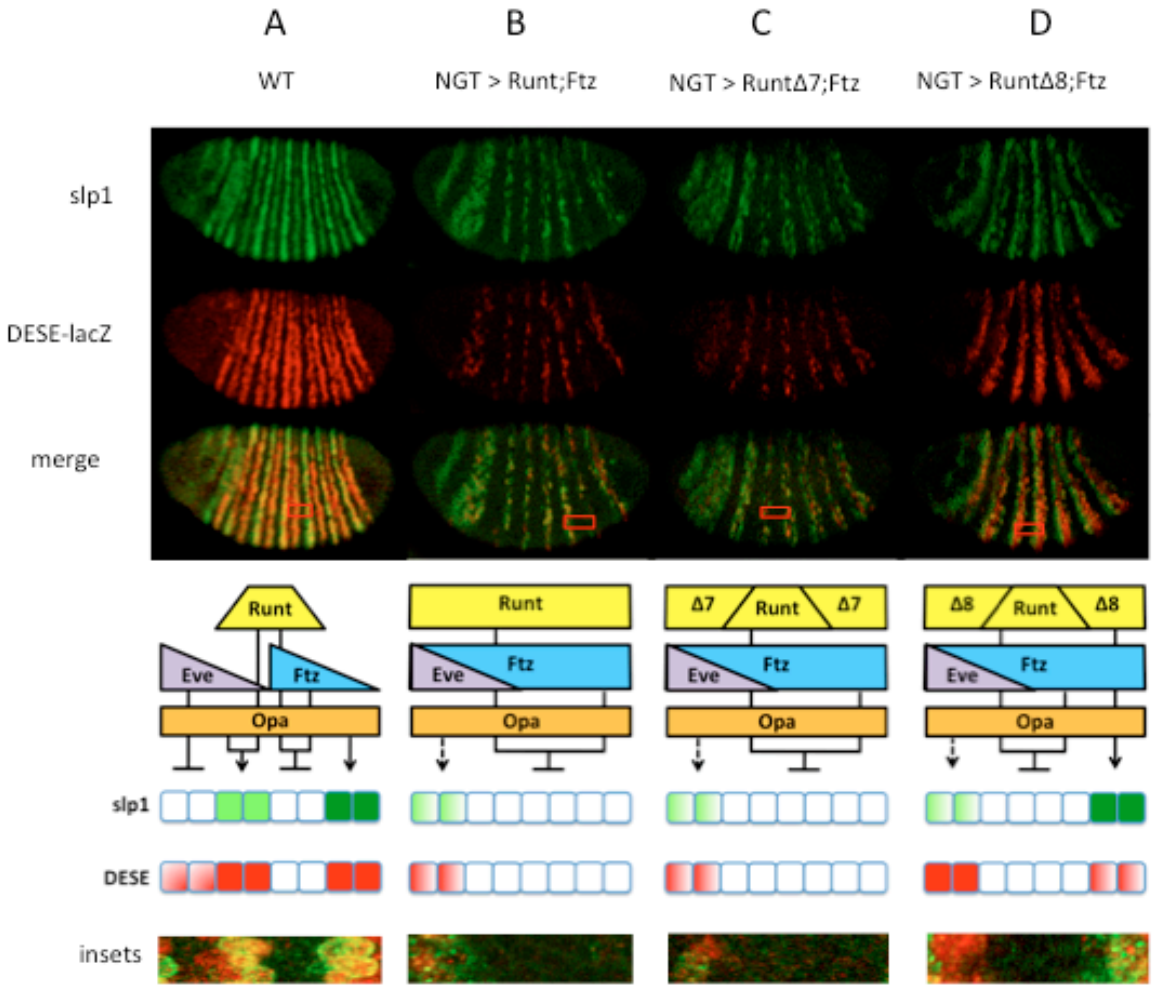

Figure 5-3: Region VIII is involved in DESE-mediated *slp1* repression by Runt and Ftz.

Fluorescent in situ hybridization of *slp1* (green) and *DESE[8765p381]-lacZ* (red) in wild-type embryo (A), Runt and Ftz over-expressed embryo (B), RuntΔ7 and Ftz over-expressed embryo (C), and RuntΔ8 and Ftz over-expressed embryo (D). Gain of function of *runt* and *ftz* embryos were obtained from the crosses between *NGT40; 8765p381-lacZ* females and *UAS-runt232; UAS-ftz263* males or *UAS-runtΔ7; UAS-ftz263* males or *UAS-runtΔ8; UAS-ftz263* males.

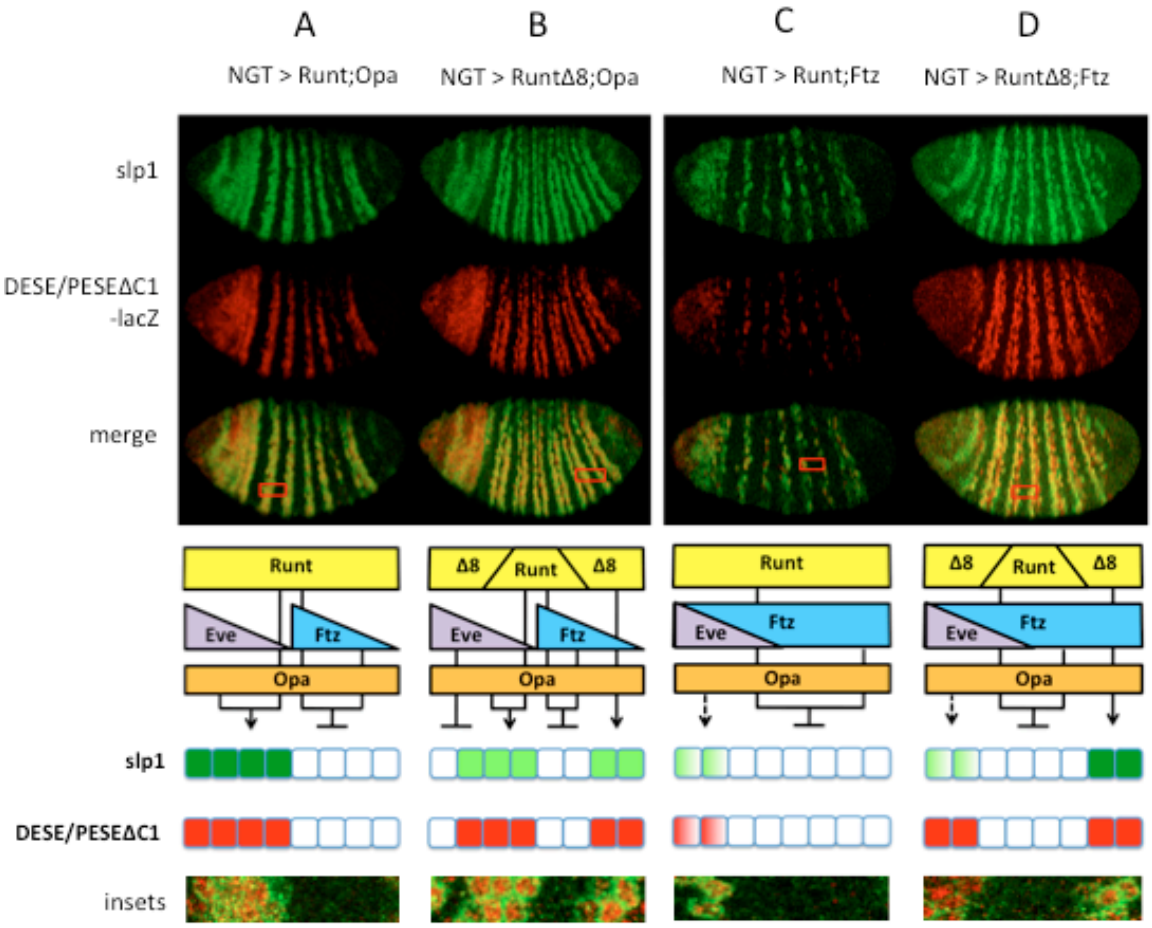

Figure 5-4: Region VIII is involved Runt-dependent *slp1* repression but not activation.
Fluorescent in situ hybridization of *slp1* (green) and the composite[*8765:3918ΔC1p381]-lacZ* (red) in Runt and Opa over-expressed embryo (A), RuntΔ8 and Opa over-expressed embryo (B), Runt and Ftz over-expressed embryo (C), RuntΔ8 and Ftz over-expressed embryo (D). Gain of function embryos were obtained from the crosses between *NGT40; [8765:3918ΔC1p381]-lacZ* females and *UAS-runt15; UAS-opa14* males or *UAS-runtΔ8; UAS-opa14* males or *UAS-runt232; UAS-ftz263* males or *UAS-runtΔ8; UAS-ftz263* males.

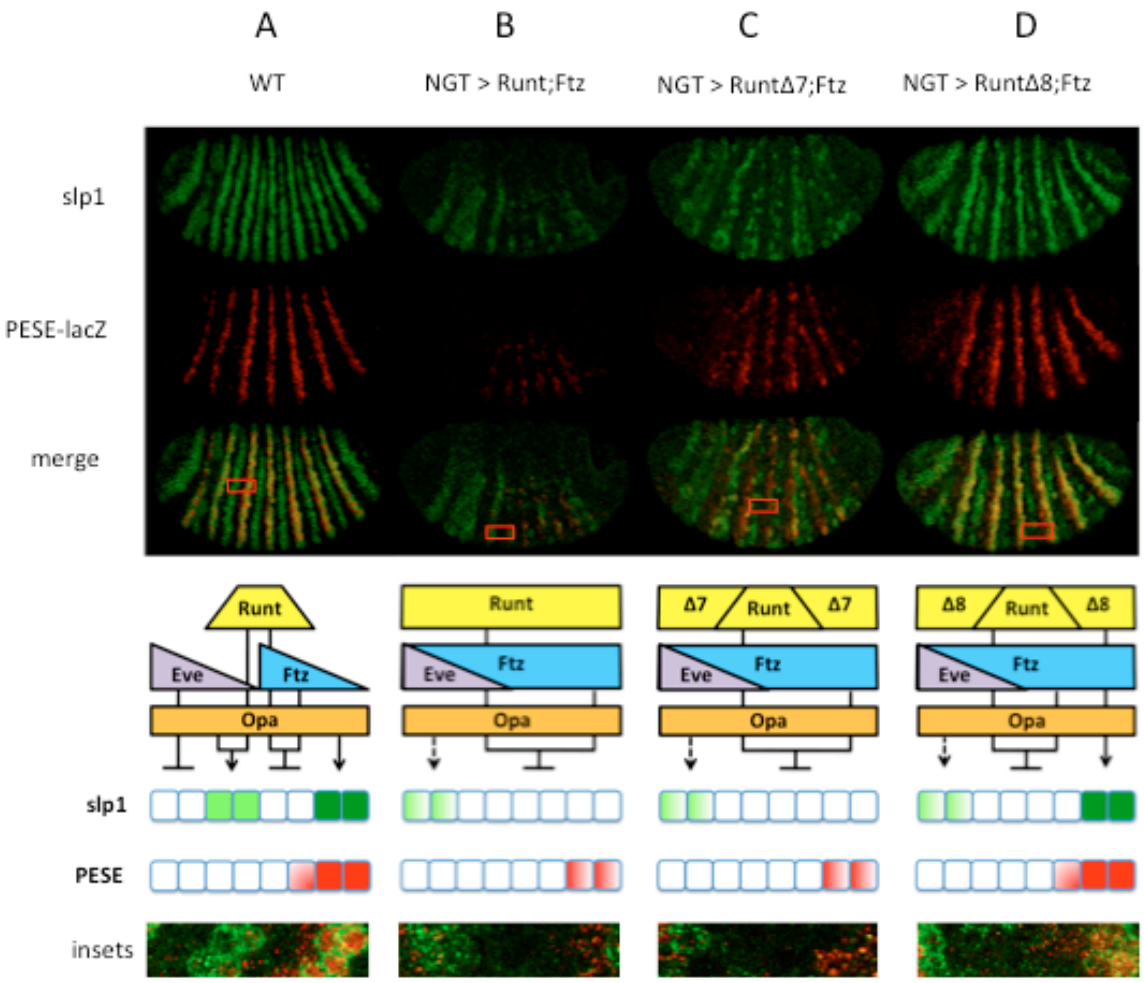

Figure 5-5: Region VIII is involved in repression of PESE by Runt and Ftz.
Fluorescent in situ hybridization of *slp1* (green) and *PESE[3125p381]-lacZ* (red) in wild-type embryo (A), Runt and Ftz over-expressed embryo (B), RuntΔ7 and Ftz over-expressed embryo (C), and RuntΔ8 and Ftz over-expressed embryo (D). Gain of function of *runt* and *ftz* embryos were obtained from the crosses between *NGT40; PESE[3125p381]-lacZ* females and *UAS-runt232; UAS-ftz263* males or *UAS-runtΔ7; UAS-ftz263* males or *UAS-runtΔ8; UAS-ftz263* males.

Chapter VI:

In vivo role of NELF in gene transcription

(my contributions for the paper published in Wang et al., 2010)

Results

Previous finding that Negative elongation factor (NELF) is associated with promoter region of *slp1* and some other segmentation genes leads us to investigate its role in transcription regulation (Wang et al., 2007) (Fig. 6-1). NELF is a complex composed of four subunits: NELF-A, B, D and E. Elimination of either one of the four subunits leads to disassembly of the NELF complex and disruption of NELF induced pol II pausing (Narita et al., 2003; Wu et al., 2005; Wu et al., 2003; Yamaguchi et al., 2002). We took advantage of P-element insertion mutations in NELF-A and NELF-E gene. Q-RT-PCR and Western blot indicated elimination of NELF-A transcripts in NELF-A[KG] germline clone embryos and reduction of NELF-E protein in NELF-E[PB] germline clone embryos respectively (data not shown).

To our surprise, the in situ results revealed no overt changes in the blastoderm stage expression of endogenous *slp1* in embryos that lack maternal NELF-A, whereas we find expression of the DESE-*lacZ* reporter gene is nearly eliminated in these same embryos (Fig. 6-2A). Consistent with this result, we also find a reduction, but not total elimination of DESE-*lacZ* expression in embryos with reduced levels of maternally provided NELF-E (Fig. 6-2A). A partial loss of expression in NELF-E germline clone embryos is similarly observed with the PESE-*lacZ* reporter (Fig. 6-2B). We conclude that NELF contributes to the expression of these reporter genes in a manner that is sensitive to NELF dosage. The expression of a composite reporter DESE+PESE-*lacZ*, which faithfully recapitulates endogenous slp1 expression pattern, is also greatly reduced in NELF-A[KG] germline clone embryo, but is not as significantly affected in embryos with reduced NELF-E level (Fig. 6-2C). The differential response of these reporters to NELF-E depletion suggests that NELF makes a quantitative contribution to transcription that can be superseded by flanking cis-regulatory information. As a further test of this idea we examined expression of slp1[05965], an enhancer trap P-transposon inserted 44 basepairs upstream of the slp1 promoter. Transcription of *lacZ* mRNA from this enhancer trap initiates at the P-element promoter in response to endogenous *slp1* cis-regulatory DNA and faithfully recapitulates the full *slp1* expression pattern in gastrula stage embryos. In contrast to reporter genes containing only subsets of flanking DNA, the enhancer trap is expressed in NELF-A deficient embryos (Fig. 6-2D). This result not only provides additional evidence that the requirement for NELF is dependent on the flanking cis-regulatory DNA, but also rules out explanations based on differences in the processing or stability of the *slp1* and *lacZ* mRNAs

The finding that NELF is dispensable for endogenous *slp1* is surprising. However, perhaps even more interesting is the finding that NELF has a positive role in supporting transcription for the *slp1* reporter genes, which is contradictory to the well-characterized properties of NELF as a transcriptional inhibitor in vitro. Further experiments are required to characterize the differences between the endogenous gene and artificial reporter genes. It is important to know whether or not these reporter genes have the same paused PolII as the endogenous *slp1* gene, which may account for the differential requirement of NELF in expression of the endogenous gene and reporter genes. The

repression of endogenous slp1 by Runt and Ftz clearly does not directly involve NELF. The increased association of NELF could just be due to reduced association of P-TEFb.

In order to determine if NELF-dependence is restricted to the *slp1* reporters we examined the expression of reporters that emulate aspects of the blastoderm stage expression of other genes involved in embryonic pattern formation. NELF-A deficient embryos failed to express reporters containing the minimal element for stripe #2 of the even-skipped gene, the 6.3 kb upstream element in the ftz-lacC reporter, and the NEE element of the dorsal-ventral patterning gene rhomboid and brinker (Fig.6-3). As was found for *slp1*, the expression of each of the endogenous cognate genes was relatively normal in these same embryos. These results indicate that the requirement for NELF is revealing a common functional distinction between the properties of these several different reporter genes and the endogenous chromosomal loci.

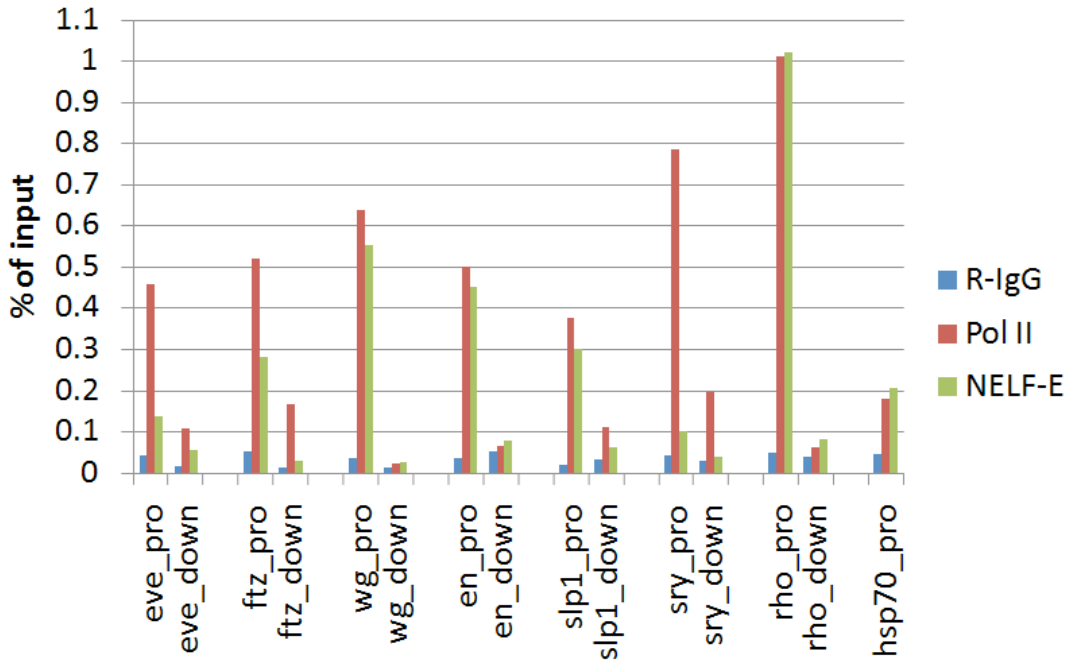

Figure 6-1. NELF specifically associates with promoter regions.
This graph shows a plot of the ChIP signals obtained with antisera against RNA polymerase II (red bars) and NELF-E (green bars) as well as a background control using rabbit IgG (blue bars). Results with two different primer pairs, one pair near to promoter and a second pair centered more than 500 bp downstream of the transcription start site are shown for seven different genes as labeled across the bottom. These experiments were conducted using chromatin from 2:45 – 3:30 AED embryos. For each gene, the NELF-E signal is stronger with the promoter proximal primer pair, and the downstream signal is close to background levels. In contrast, and as expected for genes that are expressed at this developmental stage, Pol II ChIP signals above background are obtained with the downstream primer pair for *eve, ftz, slp1,* and *srya*.

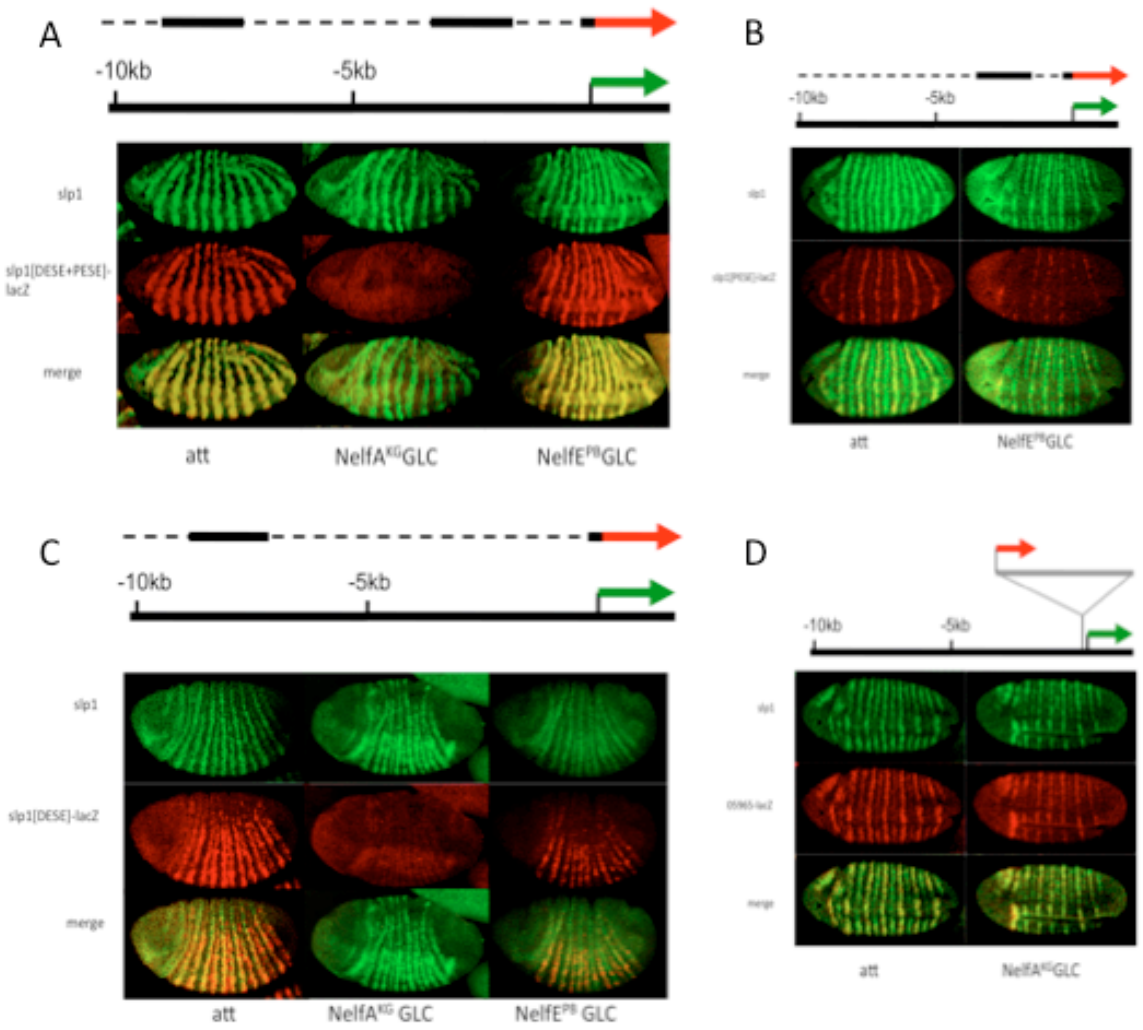

Figure 6-2. NELF is required for expression of *slp1-lacZ* reporter genes: Fluorescent double *in situ* hybridization was used to compare the expression of the endogenous *slp1* (green) and *lacZ* (red) mRNAs in embryos of different genotypes: wild-type control embryo (att), *NELF-A[KG]* germline clone embryo (NelfA^{KG} GLC) and *NELF-E[PB]* germline clone embryo (NelfE^{PB} GLC). Schematic diagrams shown on the top of each panel are different *slp1-lacZ* reporter gene, containing distinct *slp1* enhancer DESE (A), PESE (B) and DESE+PESE (C) fused to *slp1* basal promoter segment followed by *lacZ* structure gene. The solid black line at the bottom of the diagram represent the *slp1* locus with coordination given at positions 5 and 10 kb upstream of the promoter. DNA segments included in the reporter transgene are shown as solid line with the dotted line indicating flanking DNA that is omitted from the transgene. 05965-lacZ is an enhancer trap inserted 44 base-pairs upstream of slp1 transcription start site (D).

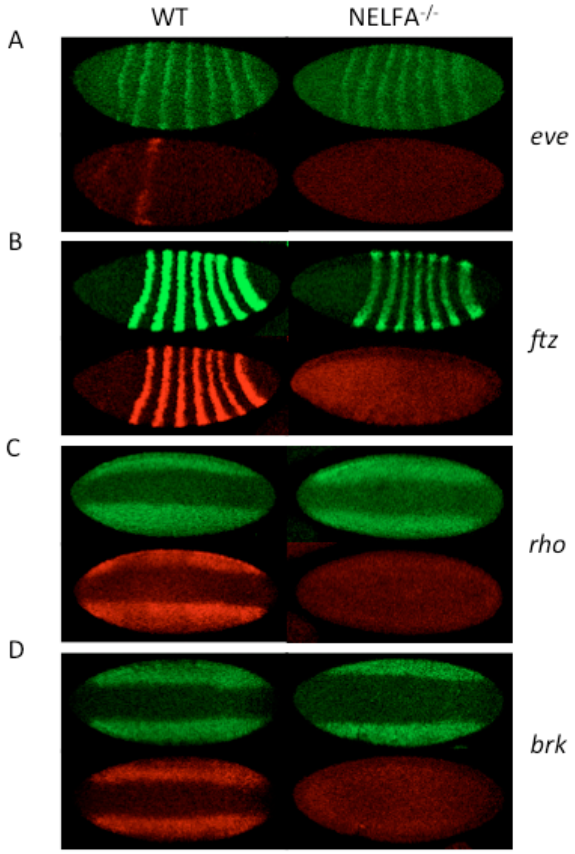

Figure 6-3. NELF-dependent activity of cis-regulatory elements that mediate blastoderm patterning.

in situ hybridization revealed the expression of *eve*, *ftz*, *rho and brk* mRNAs, relative to the expression of *lacZ* reporter genes which emulate aspects of the blastoderm-stage expression of the endogenous genes. In each case, expression of the endogenous gene is shown in green, and *lacZ* expression is shown in red. Expression of anterior-posterior patterning genes *eve* and expression of dorsal-ventral patterning genes such as *rho and brk*. (A). P[MSE-lacZ] was expressed in cells corresponding to stripe #2 of the pair-rule gene *eve* in wild-type embryos (WT), but failed to be expressed in *NELF-A[KG]* GLC embryos that showed pair-rule expression of the endogenous gene (NELFA-/-). (B) P[ftz/lacC] was expressed in a pair-rule pattern similar to *ftz* in wild-type embryos, but failed to be expressed in *ftz*-expressing NELF-A deficient embryos. (C) and (D).The P[Dm rho[NEE]-lacZ] and P[Dm brk[NEE]-lacZ] transgenes faithfully emulated the early activation of *rho* and *brk* in the neurogenic ectoderm in wild-type embryos, but was not expressed in embryos that lacked maternal NELF-A. The intensity of expression of the endogenous loci was somewhat variable in NELF-A deficient embryos, with occasional defects in patterning.

Chapter VII:

General conclusions and future perspectives

As the first and one of the most important mechanisms for controlling gene expression, transcriptional regulation is mediated by the interplay of two complementary components: (1) Trans-acting transcription factors associate with the regulatory DNA, and activate or repress the transcription of target gene; (2) The regulatory DNA that control when and where it is expressed are arranged in units that are termed cis-regulatory modules (CRM), which contains a cluster of different transcription factor binding sites. A cis-regulatory module acts like an information processor: the input that it reads is the regulatory state of the cell and the output is either activation or repression of the gene that it controls (Ben-Tabou de-Leon and Davidson, 2007). Transcriptional regulation of *slp1* gene illustrates this basic principle of how cis-regulatory elements and transcription factors control gene expression in a context dependent manner (Fig. 7-1):

(1) In the anterior half of the odd-numbered parasegments (type I cells), PESE mediates *slp1* repression by Eve that occurs at the step of elongation in Pol II transcription cycle. DESE alone is able to activate *slp1* transcription in type I cells. However, the presence of PESE inhibits DESE-dependent activation regardless of their relative position.
(2) In the posterior half of the odd-numbered parasegment (type II cells), DESE mediates *slp1* activation by Runt and Opa. It stimulates the formation of initiation complex on the *slp1* promoter, and this process is facilitated by the sequence from -260 to -79 base pair upstream of *slp1* TSS. Binding of Runt to DESE is required for this activation.
(3) In the anterior half of the even-numbered parasements (type III cells), DESE mediates *slp1* repression by Runt and Ftz. This repression occurs at the step of transcription elongation and requires Runt binding to DESE and the C-terminal domain of Runt, which suggests the involvement of the co-repressor Groucho.
(4) In the posterior half of the even-numbered parasegment (type IV cells), both PESE and DESE contribute to the activation of *slp1*. PESE mediates the activation by Opa and DESE mediates the activation by both Opa and dSTAT.

One central question raised by our results is how transcription factors that are associated with the specific DNA sequences several kilobases away can regulate the behavior of transcription machinery on the promoter of the target gene. Chromatin looping, which involves a physical association of an enhancer with the target promoter by DNA bending, is by far the most attractive model to explain the enhancer and promoter interaction (Carter et al., 2002; Dekker et al., 2002; Li et al., 2006; Tolhuis et al., 2002). The development of the Chromosomal Conformation Capture (3C) technique and its variants has provided the means to map interactions between distal sequences within the nucleus, including those that may occur between enhancers and promoters (Dekker et al., 2002). This technique is valuable to detect the direct interaction of DESE or PESE with *slp1* promoter. However, there are four distinct cellular contexts. Fluorescent activated cell sorting (FACS) provides a powerful tool for the separation of cellular populations with distinct fluorescent labeling. When a particular cell type is marked using an introduced transgene, that population of cells could be isolated from whole embryo.

ChIP experiments and genetic experiments with Runt deletion mutants strongly suggest the involvement of P-TEFb and Groucho in the Runt dependent *slp1* regulation. While further genetic experiments are required to verify these results, other unbiased

approaches could be adopted to identify cofactors that participate in the *slp1* transcription regulation. An approach for identifying proteins associated with a specific genomic locus known as PICh (Proteomics of isolated Chromatin segments) has recently been described by Kingston lab (Dejardin and Kingston, 2009). The basis of this approach is the use of nucleic acid hybridization to isolate cross-linked protein-DNA complexes from cells. Capture probes are made using LNA-(locked nucleic acid)-containing oligonucleotides in order to stabilize probe-chromatin interactions, with a long spacer to minimize steric hindrance and with desthiobiotin as the immobilization tag. Desthiobiotin is a biotin analog with a weaker affinity for avidin, allowing for elution of captured complexes using biotin, and the associated proteins were identified by gel purification and mass spectrometric analysis. Adaption of this protocol for use on *Drosophila* embryos and investigate *slp1* locus will generate a more detailed description of this locus, and allow identification of factors that participate in the activation and repression of *slp1* transcription also might provide information for the factors that associate with -260 to -79 bp region that are critical for DESE activation.

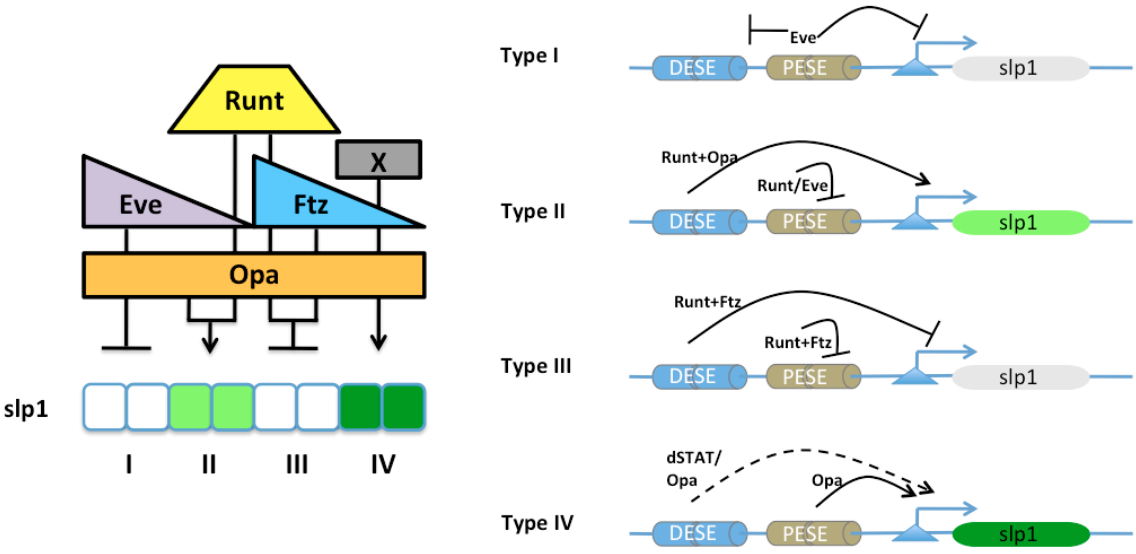

Figure 7-1: A model for the interplay of transcription factors and cis-regulatory enhancer elements in *slp1* transcription regulation.

Chapter VIII:

Materials and Methods

Construction of *slp1-lacZ* reporters and fly transgenics

Unless otherwise indicated, all the constructs use backbone vector p381-*lacZ*, pC:slp1-link-*lacZ*^{att}, that promote transgene integration into the same chromosomal site using ΦC31-mediated recombination as initially described (Prazak et al., 2010). The DESE containing constructs, pC:slp1[8765]*lacZ*^{att}, contain sequences from 8710 to 6506 bp upstream of *slp1*. PESE containing constructs, pC:slp1[3125]*lacZ*^{att}, contain sequences from 3140 to 2519 upstream of *slp1*. The composite constructs DESE+PESE, pC:slp1[8765:3918]*lacZ*^{att}, contains sequences from 8710 to 6506 and 3926 to 1774 bp upstream of *slp1*. Different promoters were obtained by PCR amplification from genomic DNA with the addition of upstream SpeI and downstream KpnI sites and ligated into DESE-, PESE- and composite-*lacZ* constructs that are linearized with the same two restriction enzymes. All of the cloned fragments were verified by sequencing and sent for injection at BestGene Inc. Reporter gene used for specific experiment and the genetic crosses for the experiments are indicated in the figure legend.

Fly stocks

All the mutant stocks: *eve*^{1}, *run*^{29}, *ftz*^{11}, *opa*^{1} and a deficiency allele *Df(1)os*^{UE69} that deletes *upd* and *upd*^{3}, maternal Gal4 line: *NGT40*, *NGT40+A*, and the UAS transgenes: *UAS-runt*^{15}, *UAS-runt*^{232}, *UAS-eve*^{12}, *UAS-ftz*^{263}, *UAS-opa*^{14}, *UAS-opa*^{D10}, *UAS-hairy*^{211} were obtained from the Bloomington Stock Center or maintained in Gergen lab.

In situ hybridization

Immunohistochemical and fluorescent in situ hybridization was carried out as described before (Prazak et al., 2010). Staged embryos were collected and fixed in 20% formaldehyde. Antisense RNA probes were in vitro transcribed using T7 or T3 RNA polymerase and digoxigenin-UTP (Roche) or FITC-UTP (Roche). Antibodies for fluorescent in situ mouse anti-DIG, rabbit anti-FITC, Alexa Fluor555 goat anti-mouse, Alexa Fluor647 donkey anti-rabbit, Alexa Fluor555 donkey anti-goat were purchased from Molecular Probes. After in situ confocal images were obtained on the Leica TCS SP5 Microscope system.

Chromatin IP and qPCR

The ChIP experiments followed by qPCR were performed as described previously (Wang et al., 2007). Rabbit serum was purchased from Sigma; Pol II antibodies 8WG16, pSer5 (H14) and anti-pSer2, anti-H3 were purchased from abcam; Anti-dTBP, anti-TFIIB, anti-TFIIF were gifts from Kadonaga lab in UCSD; dSpt5 and NELF-E antibodies were gifts from Gilmour lab in Penn State University; CycT antibody was a gift from Nakamura lab in Japan. Primers used in qPCR are listed in appendix. For each PCR, DNA standards were included for quantitation. Samples were also immunoprecipitated with rabbit serum to control for background enrichment. Immunoprecipitated DNA was calculated as a percentage of input DNA. Different genotypes were normalized to hsp70 promoter.

RT-PCR

RNA used for reverse transcriptase assay was isolated from homogenates of about 200 3-4 hours old embryos. RNA was extracted using the High Pure RNA isolation Kit (Roche). cDNA was synthesized with the Transcriptor First Strand cDNA Synthesis Kit (Roche). SYBR green-based real-time quantitative PCR was conducted with primer pairs centered 1kb downstream of each gene. RT-PCR signal obtained from different genotypes was normalized using the RT-PCR signal for rp49. Primer sequences are listed in appendix.

Nucleosome Occupancy and Methylome Sequencing (NOMe-Seq)

Nuclei extracted from 3-4 hours old embryos were used for a methylase-based, single-molecule assay as previously described (Taberlay et al., 2011; You et al., 2011). Briefly, 100mg embryos were de-chorionated and transferred into eppendorf tube and wash with ice-cold RSB buffer (10 mM Tris [pH 7.4], 10 mM NaCl, 3 mM $MgCl_2$), resuspended in 1 ml ice-cold nuclei isolation buffer (10 mM Tris [pH 7.4], 10 mM NaCl, 3 mM MgCl2, 0.1 mM EDTA, and 0.5% NP-40, plus protease inhibitors). Followed by Dounce homogenization and incubation on ice for 2 hours, nuclei were recovered by filter through 60um filter followed by centrifugation at 5000rpm for 10 min and washed in nuclei wash buffer (10 mM Tris [pH 7.4], 10 mM NaCl, 3 mM MgCl2, and 0.1 mM EDTA containing protease inhibitors). Freshly prepared nuclei were resuspended in *M.SssI* reaction buffer (NEB), and then treated with 60 U of *M.SssI* (NEB). Reactions were quenched by the addition of an equal volume of Stop Solution (20 nM Tris-HCl [pH 7.9], 600 mM NaCl, 1% SDS, 10 mM EDTA, 400 mg/ml Proteinase K) and incubated at 55 °C overnight. DNA was purified by phenol/chloroform extraction and ethanol precipitation. Bisulfite conversion was performed using the Epitect Bisulfite Kit (QIAGEN). Molecules were cloned using the Topo TA Kit (Invitrogen), both according to the manufacturers' instructions.

References

Aaronson, D.S., and Horvath, C.M. (2002). A road map for those who don't know JAK-STAT. Science *296*, 1653-1655.

Adya, N., Castilla, L.H., and Liu, P.P. (2000). Function of CBFbeta/Bro proteins. Semin Cell Dev Biol *11*, 361-368.

Andrezj Nasiadka, B.H.D.a.H.M.K. (2002). Anterior-posterior patterning in the Drosophila embryo. Advances in Developmental Biology and Biochemistry *12*.

Aronson, B.D., Fisher, A.L., Blechman, K., Caudy, M., and Gergen, J.P. (1997). Groucho-dependent and -independent repression activities of Runt domain proteins. Molecular and cellular biology *17*, 5581-5587.

Bai, L., and Morozov, A.V. (2010). Gene regulation by nucleosome positioning. Trends Genet *26*, 476-483.

Ben-Tabou de-Leon, S., and Davidson, E.H. (2007). Gene regulation: gene control network in development. Annual review of biophysics and biomolecular structure *36*, 191.

Bieniasz, P.D., Grdina, T.A., Bogerd, H.P., and Cullen, B.R. (1999). Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. Proceedings of the National Academy of Sciences of the United States of America *96*, 7791-7796.

Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes & development *13*, 2465-2477.

Bulger, M., and Groudine, M. (2010). Enhancers: the abundance and function of regulatory sequences beyond promoters. Developmental biology *339*, 250-257.

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

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nature genetics *32*, 623-626.

Chao, S.H., and Price, D.H. (2001). Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. The Journal of biological chemistry *276*, 31793-31799.

Chiba, K., Yamamoto, J., Yamaguchi, Y., and Handa, H. (2010). Promoter-proximal pausing and its release: molecular mechanisms and physiological functions. Experimental cell research *316*, 2723-2730.

Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science *319*, 1791-1792.

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

Dejardin, J., and Kingston, R.E. (2009). Purification of proteins associated with specific genomic Loci. Cell *136*, 175-186.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell *149*, 1233-1244.
Driever, W., and Nusslein-Volhard, C. (1989). The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. Nature *337*, 138-143.
Dubnicoff, T., Valentine, S.A., Chen, G., Shi, T., Lengyel, J.A., Paroush, Z., and Courey, A.J. (1997). Conversion of dorsal from an activator to a repressor by the global corepressor Groucho. Genes & development *11*, 2952-2957.
Engel, J.D., and Tanimoto, K. (2000). Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. Cell *100*, 499-502.
Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004). Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Molecular and cellular biology *24*, 787-795.
Gaston, K., and Jayaraman, P.S. (2003). Transcriptional repression in eukaryotes: repressors and repression mechanisms. Cellular and molecular life sciences : CMLS *60*, 721-741.
Gilmour, D.S. (2009). Promoter proximal pausing on genes in metazoans. Chromosoma *118*, 1-10.
Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E.*, et al.* (2003). A protein interaction map of Drosophila melanogaster. Science *302*, 1727-1736.
Gondor, A., and Ohlsson, R. (2009). Chromosome crosstalk in three dimensions. Nature *461*, 212-217.
Gray, S., and Levine, M. (1996a). Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in Drosophila. Genes & development *10*, 700-710.
Gray, S., and Levine, M. (1996b). Transcriptional repression in development. Current opinion in cell biology *8*, 358-364.
Grossniklaus, U., Pearson, R.K., and Gehring, W.J. (1992). The Drosophila sloppy paired locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. Genes & development *6*, 1030-1051.
Han, K., and Manley, J.L. (1993). Transcriptional repression by the Drosophila even-skipped protein: definition of a minimal repression domain. Genes & development *7*, 491-503.
Hirose, Y., and Ohkuma, Y. (2007). Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. Journal of biochemistry *141*, 601-608.
Hoey, T., Weinzierl, R.O., Gill, G., Chen, J.L., Dynlacht, B.D., and Tjian, R. (1993). Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell *72*, 247-260.
Ito, Y. (2004). Oncogenic potential of the RUNX gene family: 'overview'. Oncogene *23*, 4198-4208.
Iyer, V.R. (2012). Nucleosome positioning: bringing order to the eukaryotic genome. Trends in cell biology.

Javed, A., Guo, B., Hiebert, S., Choi, J.Y., Green, J., Zhao, S.C., Osborne, M.A., Stifani, S., Stein, J.L., Lian, J.B.*, et al.* (2000). Groucho/TLE/R-esp proteins associate with the nuclear matrix and repress RUNX (CBF(alpha)/AML/PEBP2(alpha)) dependent activation of tissue-specific gene transcription. Journal of cell science *113 (Pt 12)*, 2221-2231.
Jiang, H., Zhang, F., Kurosu, T., and Peterlin, B.M. (2005). Runx1 binds positive transcription elongation factor b and represses transcriptional elongation by RNA polymerase II: possible mechanism of CD4 silencing. Molecular and cellular biology *25*, 10675-10683.
Juven-Gershon, T., Hsu, J.Y., and Kadonaga, J.T. (2008). Caudal, a key developmental regulator, is a DPE-specific transcriptional factor. Genes & development *22*, 2823-2830.
Kassis, J.A. (1990). Spatial and temporal control elements of the Drosophila engrailed gene. Genes & development *4*, 433-443.
Katsani, K.R., Hajibagheri, M.A., and Verrijzer, C.P. (1999). Co-operative DNA binding by GAGA transcription factor requires the conserved BTB/POZ domain and reorganizes promoter topology. The EMBO journal *18*, 698-708.
Kramer, S.G., Jinks, T.M., Schedl, P., and Gergen, J.P. (1999). Direct activation of Sex-lethal transcription by the Drosophila runt protein. Development *126*, 191-200.
Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. Annual review of genetics *34*, 77-137.
Lenhard, B., Sandelin, A., and Carninci, P. (2012). Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nature reviews Genetics *13*, 233-245.
Lessing, D., and Nusse, R. (1998). Expression of wingless in the Drosophila embryo: a conserved cis-acting element lacking conserved Ci-binding sites is required for patched-mediated repression. Development *125*, 1469-1476.
Li, J., and Gilmour, D.S. (2011). Promoter proximal pausing and the control of gene expression. Current opinion in genetics & development *21*, 231-235.
Li, L.M., and Arnosti, D.N. (2011). Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. Current biology : CB *21*, 406-412.
Li, Q., Barkess, G., and Qian, H. (2006). Chromatin looping and the probability of transcription. Trends in genetics : TIG *22*, 197-202.
Liang, H.L., Nien, C.Y., Liu, H.Y., Metzstein, M.M., Kirov, N., and Rushlow, C. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in Drosophila. Nature *456*, 400-403.
Lis, J.T. (2007). Imaging Drosophila gene activation and polymerase pausing in vivo. Nature *450*, 198-202.
Lis, J.T., Mason, P., Peng, J., Price, D.H., and Werner, J. (2000). P-TEFb kinase recruitment and function at heat shock loci. Genes & development *14*, 792-803.
Lomvardas, S., and Thanos, D. (2002). Modifying gene expression programs by altering core promoter chromatin architecture. Cell *110*, 261-271.
Luger, K., Rechsteiner, T.J., Flaus, A.J., Waye, M.M., and Richmond, T.J. (1997). Characterization of nucleosome core particles containing histone proteins made in bacteria. Journal of molecular biology *272*, 301-311.

Mahmoudi, T., Katsani, K.R., and Verrijzer, C.P. (2002). GAGA can mediate enhancer function in trans by linking two separate DNA molecules. The EMBO journal *21*, 1775-1781.
Majello, B., Napolitano, G., Giordano, A., and Lania, L. (1999). Transcriptional regulation by targeted recruitment of cyclin-dependent CDK9 kinase in vivo. Oncogene *18*, 4598-4605.
Mannervik, M., Nibu, Y., Zhang, H., and Levine, M. (1999). Transcriptional coregulators in development. Science *284*, 606-609.
Manoukian, A.S., and Krause, H.M. (1993). Control of segmental asymmetry in Drosophila embryos. Development *118*, 785-796.
Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. The Journal of biological chemistry *270*, 12335-12338.
Martinez, C.A., and Arnosti, D.N. (2008). Spreading of a corepressor linked to action of long-range repressor hairy. Molecular and cellular biology *28*, 2792-2802.
Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. Annual review of genomics and human genetics *7*, 29-59.
McKay, L.M., Carpenter, B., and Roberts, S.G. (1999). Evolutionary conserved mechanism of transcriptional repression by even-skipped. Nucleic acids research *27*, 3064-3070.
Meinhart, A., Kamenski, T., Hoeppner, S., Baumli, S., and Cramer, P. (2005). A structural perspective of CTD function. Genes & development *19*, 1401-1415.
Muller, M.M., Gerster, T., and Schaffner, W. (1988). Enhancer sequences and the regulation of gene transcription. European journal of biochemistry / FEBS *176*, 485-495.
Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nature genetics *39*, 1507-1511.
Nakayama, T., Nishioka, K., Dong, Y.X., Shimojima, T., and Hirose, S. (2007). Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. Genes & development *21*, 552-561.
Narita, T., Yamaguchi, Y., Yano, K., Sugimoto, S., Chanarat, S., Wada, T., Kim, D.K., Hasegawa, J., Omori, M., Inukai, N*., et al.* (2003). Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Molecular and cellular biology *23*, 1863-1873.
Nechaev, S., and Adelman, K. (2010). Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. Biochim Biophys Acta.
Negre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R*., et al.* (2011). A cis-regulatory map of the Drosophila genome. Nature *471*, 527-531.
Nibu, Y., Zhang, H., and Levine, M. (1998). Interaction of short-range repressors with Drosophila CtBP in the embryo. Science *280*, 101-104.

Payankaulam, S., and Arnosti, D.N. (2009). Groucho corepressor functions as a cofactor for the Knirps short-range transcriptional repressor. Proceedings of the National Academy of Sciences of the United States of America *106*, 17314-17319.
Peng, Y., and Jahroudi, N. (2002). The NFY transcription factor functions as a repressor and activator of the von Willebrand factor promoter. Blood *99*, 2408-2417.
Poortinga, G., Watanabe, M., and Parkhurst, S.M. (1998). Drosophila CtBP: a Hairy-interacting protein required for embryonic segmentation and hairy-mediated transcriptional repression. The EMBO journal *17*, 2067-2078.
Prazak, L., Fujioka, M., and Gergen, J.P. (2010). Non-additive interactions involving two distinct elements mediate sloppy-paired regulation by pair-rule transcription factors. Developmental biology *344*, 1048-1059.
Price, D.H. (2008). Poised polymerases: on your mark...get set...go! Molecular cell *30*, 7-10.
Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome positioning: how is it established, and why does it matter? Developmental biology *339*, 258-266.
Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F.*, et al.* (2010). Identification of functional elements and regulatory circuits by Drosophila modENCODE. Science *330*, 1787-1797.
Sakabe, N.J., Aneas, I., Shen, T., Shokri, L., Park, S.Y., Bulyk, M.L., Evans, S.M., and Nobrega, M.A. (2012). Dual transcriptional activator and repressor roles of TBX20 regulate adult cardiac structure and function. Human molecular genetics *21*, 2194-2204.
Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic acids research *32*, D91-94.
Seufert, D.W., Prescott, N.L., and El-Hodiri, H.M. (2005). Xenopus aristaless-related homeobox (xARX) gene product functions as both a transcriptional activator and repressor in forebrain development. Developmental dynamics : an official publication of the American Association of Anatomists *232*, 313-324.
Sims, R.J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. Genes & development *18*, 2437-2468.
Smale, S.T. (2001). Core promoters: active contributors to combinatorial gene regulation. Genes & development *15*, 2503-2508.
Smale, S.T., and Kadonaga, J.T. (2003). The RNA polymerase II core promoter. Annual review of biochemistry *72*, 449-479.
Smale, S.T., Schmidt, M.C., Berk, A.J., and Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proceedings of the National Academy of Sciences of the United States of America *87*, 4509-4513.
Small, S., Blair, A., and Levine, M. (1992). Regulation of even-skipped stripe 2 in the Drosophila embryo. The EMBO journal *11*, 4047-4057.
Stifani, S., and Ma, Q. (2009). 'Runxs and regulations' of sensory and motor neuron subtype differentiation: implications for hematopoietic development. Blood Cells Mol Dis *43*, 20-26.

Swantek, D., and Gergen, J.P. (2004). Ftz modulates Runt-dependent activation and repression of segment-polarity gene transcription. Development *131*, 2281-2290.
Taberlay, P.C., Kelly, T.K., Liu, C.C., You, J.S., De Carvalho, D.D., Miranda, T.B., Zhou, X.J., Liang, G., and Jones, P.A. (2011). Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell *147*, 1283-1294.
Tamkun, J.W. (2007). Stalled polymerases and transcriptional regulation. Nature genetics *39*, 1421-1422.
Tautz, D. (1988). Regulation of the Drosophila segmentation gene hunchback by two maternal morphogenetic centres. Nature *332*, 281-284.
Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Molecular cell *10*, 1453-1465.
Tracey, W.D., Jr., Ning, X., Klingler, M., Kramer, S.G., and Gergen, J.P. (2000). Quantitative analysis of gene function in the Drosophila embryo. Genetics *154*, 273-284.
Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G.A., Winston, F.*, et al.* (1998a). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes & development *12*, 343-356.
Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D., and Handa, H. (1998b). Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. The EMBO journal *17*, 7395-7403.
Walrad, P.B., Hang, S., Joseph, G.S., Salas, J., and Gergen, J.P. (2010). Distinct contributions of conserved modules to Runt transcription factor activity. Molecular biology of the cell *21*, 2315-2326.
Wang, X., Hang, S., Prazak, L., and Gergen, J.P. (2010). NELF potentiates gene transcription in the Drosophila embryo. PloS one *5*, e11498.
Wang, X., Lee, C., Gilmour, D.S., and Gergen, J.P. (2007). Transcription elongation controls cell fate specification in the Drosophila embryo. Genes & development *21*, 1031-1036.
Wheeler, J.C., Shigesada, K., Gergen, J.P., and Ito, Y. (2000). Mechanisms of transcriptional regulation by Runt domain proteins. Seminars in cell & developmental biology *11*, 369-375.
Wu, C.H., Lee, C., Fan, R., Smith, M.J., Yamaguchi, Y., Handa, H., and Gilmour, D.S. (2005). Molecular characterization of Drosophila NELF. Nucleic acids research *33*, 1269-1279.
Wu, C.H., Yamaguchi, Y., Benjamin, L.R., Horvat-Gordon, M., Washinsky, J., Enerly, E., Larsson, J., Lambertsson, A., Handa, H., and Gilmour, D. (2003). NELF and DSIF cause promoter proximal pausing on the hsp70 promoter in Drosophila. Genes & development *17*, 1402-1414.
Yamada, T., Yamaguchi, Y., Inukai, N., Okamoto, S., Mura, T., and Handa, H. (2006). P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Molecular cell *21*, 227-237.
Yamaguchi, Y., Inukai, N., Narita, T., Wada, T., and Handa, H. (2002). Evidence that negative elongation factor represses transcription elongation through binding to a

DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. Molecular and cellular biology *22*, 2918-2927.
Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., and Handa, H. (1999). NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. Cell *97*, 41-51.
You, J.S., Kelly, T.K., De Carvalho, D.D., Taberlay, P.C., Liang, G., and Jones, P.A. (2011). OCT4 establishes and maintains nucleosome-depleted regions that provide additional layers of epigenetic regulation of its target genes. Proceedings of the National Academy of Sciences of the United States of America *108*, 14497-14502.
Zhang, H.S., and Dean, D.C. (2001). Rb-mediated chromatin structure regulation and transcriptional repression. Oncogene *20*, 3134-3138.

Appendix

Primers for ChIP and RT-PCR

Name	Sequence	Notes
slp1-pro_F(20007)	gggctctcttcgtgtagacttcgt	
slp1-pro_R(20010)	ggagaagttgctcttgaattccatt	product size 224bp, centered at +35bp
slp1+0.5K_F	tcccagcaagaaacagaaaatgac	8592
slp1+0.5K_R	ttggccttgaagtaggggaatcta	8593
slp1_+1.2K_F	agggacaaaagggtcgcttg	
slp1_+1.2K_R	ttccaagtcgattcggagga	3'UTR
lacZ_pro_F	tccggggctgtagagaatca	Centered at +77bp, for all lacZ reporter constructs small and large promoter
lacZ_pro_R	cagaccaatgcctcccagac	
lacZ-link_F	gtctgggaggcattggtctg	for Lisa's constructs, Miki's construct should use 20007/link_R
lacZ-link_R	gcgggcctcttcgctattac	
lacZ-300_F	aaccctggcgttacccaact	
lacZ-300_R	agtttgaggggacgacgaca	
lacZ_+0.5K_F	tttcggcggtgaaattatcg	
lacZ_+0.5K_R	gacatcgcaggcttctgctt	
lacZ_+0.7K_F	ggcgagttgcgtgactacct	
lacZ_+0.7K_R	cgggttttcgacgttcagac	
lacZ_1K_F	tcgcattatccgaaccatcc	
lacZ_1K_R	ccagcgaccagatgatcaca	
lacZ_2K_F	gagctcctgcactggatggt	
lacZ_2K_R	gggagcgtcacactgaggtt	
DESE-Rt34_F	acgtttctcgggcaaactga	
DESE-Rt34_R	aatgcgaatgggaatccaga	-7.4k
PESE-Rt_F	agaaaaagggaagccgcaaa	
PESE-Rt_R	cgaaagagacggcaatcgag	-1.8K
8271	tcgaggcgttcaaatcttcattta	
8272	agttaaaagtgcccgaaagagacg	tail of 3918 (-1.8K)
8275	tatgagactgaatgctcacccaca	
8276	ctgcctcattagctcacaaaaacg	C1+ tail (-2.8k)
8395	atctgggattgacaccggacttat	
8396	atgatcgatctatcgttgcacacc	C1+ head (-3.2K)
slp1_-2.4K_F	gtgtgccctcaaggatctcg	tail of 3125,use this as Forward primer for junction of 3125-BP
slp1_-2.4K_R	gcaaggtgtgcctatgcagtt	
slp1_-8.4K_F	gcttgcacacgccctactct	
slp1_-8.4K_R	tcgctcaccactcgattcaa	
slp1_-7.9K_F	tcaaaatcaacgcccaaacc	
slp1_-7.9K_R	cgcaaaactgtggccataca	

slp1_-6.8K_F	tgcacgttcatctgcgactt	
slp1_-6.8K_R	acatcgcgtacggaaaagga	
slp1_-2.5K_F	ggccacggtgtctacttgga	middle of 3125
slp1_-2.5K_R	gatcaggacgcgatcaggac	
lacZ-8765up_F	gtcacgaaaccgctgacaaa	lacZ reporter construct. immediately upstream of DESE
lacZ-8765up_R	tcaaagtcagcgctgtttgc	
lacZ-8765sBPjn_F	caatcctggctgcactgaaa	lacZ reporter construct, 8765 p126 junction
lacZ-8765sBPjn_R	tcacacacgacttgggatcg	
lacZ-8765lBPjn_F	tccttttccgtacgcgatgt	lacZ reporter construct, 8765 p381 junction
lacZ-8765lBPjn_R	gttccactcgccttccacac	
wht_1K_F	acgcggactattctgcaacg	
wht_1K_R	atcgaaaggcaagggcattc	
wht_2K_F	gcggcttcttcttgaactcg	
wht_2K_R	acgtggtgttcgacgatgtg	