Dorsal-Ventral Patterning and Gene Regulation in the Early Embryo of *Drosophila melanogaster*

Thesis by

Mayra Garcia

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2012

(Defended May 17, 2012)

To my parents, Pedro and Aureliana, for always being supportive

and

my sisters, Analu and Ybette, for being great rolemodels

Acknowledgments

First I would like to thank my advisor, Angela Stathopoulos, who has been very supportive throughout my graduate career. She is a wonderful advisor and was always available for guidance when I needed help. She also challenged and pushed me to do my best. By leading by example; she has shown me that with a lot of hard work and sometimes a little bit of luck you can accomplish a lot.

I would, also, like to thank all of the past and current members of the Stathopoulos lab who have provided excellent advice and discussion throughout the years. Especially, Snehalata Kadam who has been a wonderful bay-mate and was always available when I had genetics questions. Also, Desirea Mecenas who was very helpful when I first joined the lab and did not know my way around.

I give many thanks to my advisory committee, Marianne Bronner, Carl Parker, and Ellen Rothenberg. I enjoyed meeting with them and talking about my work because they always had great feedback. I am especially indebted to Carl Parker who helped me by sharing large numbers of embryos with me and showing me how to make nuclear extracts.

My undergraduate advisor, Jennifer Armstrong, and the other professors at Scripps College, who did an excellent job preparing me for graduate school, also deserve my sincere gratitude.

Also, I must thank my family who has always been very supportive and understanding. I know they do not really understand what I have been doing for the past seven years, but nonetheless, they are still proud of me. I do not think I would have made it to graduate school without having wonderful parents who taught me to work hard and

not give up. They have, also, allowed me to disappear for long stretches at time when I was overly involved in my work and only gave me a tiny bit of grief for it.

Lastly, I would like to thank Tony for allowing me to get away from my work for a while especially when things became stressful.

Abstract

In order for an embryo to develop and form properly, the anterior-posterior and dorsal-ventral axes must be specified. This is accomplished by controlled regulation of gene expression that allows for the activation and repression of tissue specific genes. Patterning of the *Drosophila* dorsal-ventral axis is an excellent model for understanding how axis specification is controlled. The dorsal-ventral axis is patterned by a nuclear gradient of Dorsal that is highest in ventral regions of the embryo. Dorsal activates genes in a concentration dependent manner to establish early patterning of the embryo. The patterns are refined by interactions between Dorsal and other activators as well as repressors in both dorsal and ventral regions of the embryo. Until recently there was only evidence for repressors acting in ventral regions of the embryo but our studies and other recent studies have provided evidence to suggest that several repressors act in dorsal regions of the embryo to refine Dorsal target genes. Here we show that an element, the A-box, previously identified in the cis-regulatory module (CRM) of the gene *intermediate neuroblast defective (ind)*, is necessary and sufficient to mediate dorsal repression and is also involved in activation of *ind*. We conducted an affinity chromatography assay and identified factors that bound the A-box element. One of the factors that bound to this element, Grh, activates *ind*. We also identified several chromatin-remodeling factors that may function to silence *ind* in dorsal regions of the embryo. Our results also indicate that a second tier of repression that is independent of the A-box element, mediates repression of *ind* via Dpp-signaling. We extended our studies to the CRMs of *ventral neuroblast defective (vnd)* and *short gastrulation (sog)*. Using a chimeric CRM repression assay, we found that strong and weak dorsal repression

are also mediated by the *vnd* and *sog* CRMs, respectively. This suggests that limiting amounts of Dorsal are not sufficient to establish the dorsal borders of dorsal-ventral patterning genes as was previously believed, and rather, repressors are used to establish these borders.

Contents

Chapter 1

Introduction

Embryonic development is a very complex and elegant endeavor. Starting from a single cell, an embryo must become a complex organism with axis polarity and distinctive organ structures. Amazingly the embryo does this with a limited tool kit that it adapts and reorganizes to fit its particular needs. The same signaling pathways are used over and over again throughout embryonic development to activate or repress sets of tissue specific genes that will provide information to the embryo and aid in its development.

After the single cell divides into multiple cells, it forms a hollow ball of cells called a blastula. Or in the case of *Drosophila melanogaster*, a syncytial blastoderm is formed in which one cell holds several nuclei, which are located along the periphery of the cell. The first step in distinguishing these cells is axis specification. The dorsal-ventral as well as the anterior-posterior axis of the developing embryo must be determined in order for proper development to proceed. If this does not occur properly, the embryo's development will not progress and the embryo will not survive. To accomplish this task the embryo uses morphogen gradients; in fact, throughout its development, the embryo uses morphogen gradients to pattern the axis of body parts and organ structures. Often the same morphogens are reused to pattern different body parts. For example, BMP/Dpp signaling is used to pattern the dorsal-ventral axis of the embryo as well as the anterior-posterior axis of the wing disc (Holley et al., 1995; Ingham and Fietz, 1995). Since morphogens play such a large role in development, they are well studied, and it is important to further understand their strengths and limits in patterning fields of cells.

It was once believed that morphogens are sufficient to provide the positional information necessary to pattern fields of cells (Wolpert, 1996). While there is evidence that morphogens, when ectopically expressed, can trigger expression of certain genes in a set pattern (figure 1.1), it is becoming increasingly apparent that other factors, some of which are triggered by the morphogen gradient, are also important in patterning (Balaskas et al., 2012; Liberman et al., 2009; Stathopoulos and Levine, 2005). Our data suggests morphogen gradients are only necessary for initiating gene expression and often in a broader domain than the final pattern. Once the expression is initiated, other factors (some dependent and some independent of the morphogen) refine the expression patterns. In the following chapters, we provide evidence to support this idea. Here we give a brief summary of morphogen gradients. Then we use the Dorsal morphogen gradient as an example of how a morphogen is used to initiate expression patterns, while other factors (including repressors) and signaling pathways are used to refine these patterns.

Morphogen Gradients

A morphogen is a signaling molecule that imparts positional information and organizes a field of cells into a pattern by activating different sets of genes in a concentration-dependent manner. In the simplest example, morphogen concentrations are established by diffusion of a ligand from a localized region. The receptor for the ligand is generally ubiquitously expressed. At the source, the ligand concentration is higher and more receptors are bound leading to higher levels of signal (Rogers and Schier, 2011; Teleman et al., 2001).

The effectors of morphogen gradients are generally transcription factors that are activated, phosphorylated, or translocated into the nucleus by morphogen initiated signaling transduction. The effector then activates the transcription of cell specific genes that leads to axis specification and later differentiation of cells (Ashe and Briscoe, 2006; Rogers and Schier, 2011; Teleman et al., 2001).

The idea of morphogen gradients was first introduced over a century ago. Thomas Hunt Morgan first presented the idea of gradients being used for pattern formation with his studies on regeneration. The idea continued for 40 years and was revived with further works advancing the idea in the 1960s (Wolpert, 1996). In 1968, Wolpert published his theory for solving the French flag problem, or rather how to get three sequential rows of different cell types, using a gradient (figure 1.1 A). Without a means of visualizing morphogen gradients, early studies were focused on using tissue culture studies in which different concentrations of a signaling molecule were added to see if a concentration-dependent response was elicited. Studies in *Xenopus* showed that treating presumptive ectoderm with increasing amounts of the morphogen Activin created different types of mesodermal tissue (Green et al., 1992; Green and Smith, 1990) (figure 1.1 B).

Studies were also focused on proving that morphogens could act at a distance. One such study showed that a Nodal-related TGF-β protein was able to pattern the mesoderm when injected into a single cell of an early stage embryo. Cells close to the site of injection expressed genes that were typical to high levels of the signal, while cells that were further away expressed genes that were typical to low levels of the signal. When lower amounts of the morphogen were injected only low level signaling genes

were expressed. This study was important because it showed that the signal was capable of acting at a distance from the injection site (Chen and Schier, 2001) (figure 1.1 C).

Many studies have explored how morphogen gradients are formed. Evidence supporting active transport by secretion and endocytosis has been shown in many systems by mutation to the endocytosis machinery, which affects how far the gradient can spread (Teleman et al., 2001).

Comprehending how cells interpret the information provided by the gradient is another key question in understanding morphogens. The answer must lie, at least in part, in a concentration-dependent mechanism. In fact, binding studies in *Drosophila* in both anterior-posterior and dorsal-ventral patterning have shown that genes that respond to lower amounts of signal contain higher affinity binding sites while those that respond to higher amounts of a signal contain lower affinity binding sites (Driever et al., 1989; Stathopoulos and Levine, 2004). By this mechanism genes that have low affinity binding sites can only be activated by the highest concentrations of the effector. Now that gradients can be visualized and quantified, it has become increasingly clear that gradient information alone does not provide the positional information necessary for patterning (Balaskas et al., 2012; Liberman et al., 2009; Reeves et al., 2012). Instead, interactions between the target genes and combinatorial interactions between the morphogen effector and other transcription factors come together to pattern the embryo. These interactions include but are not limited to feed-forward loops, cross repression, positive feedback, autoregulation, reciprocal inhibitor gradients, and temporal integration (Ashe and Briscoe, 2006).

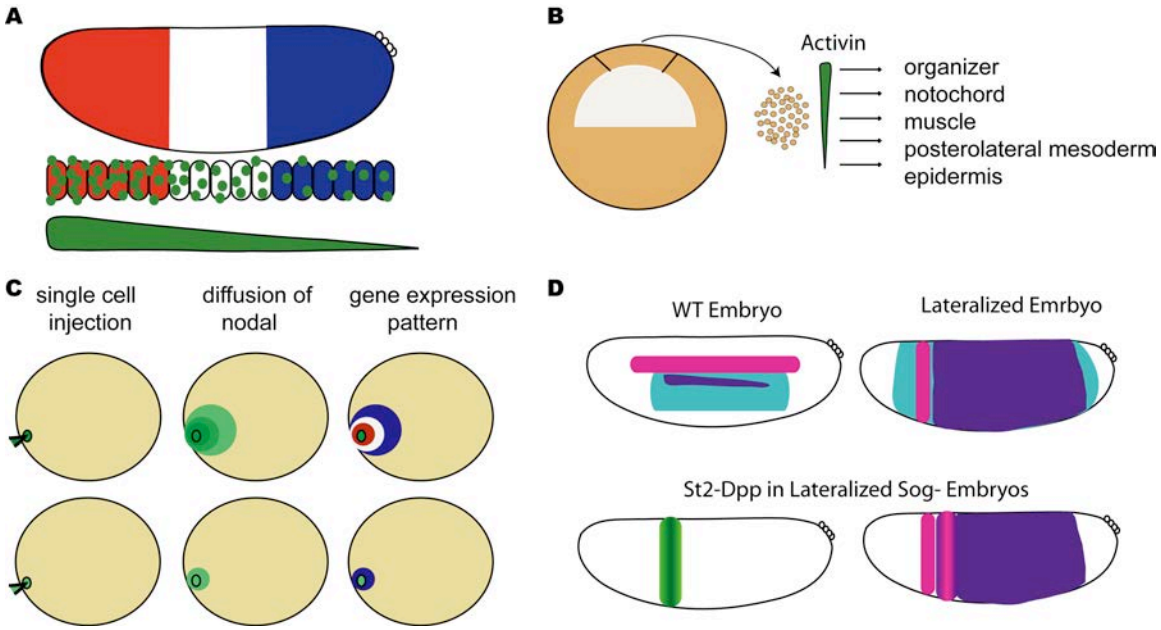


Figure 1.1. Ectopic expression of morphogens can pattern fields of cells.

Several experiments have been conducted in an attempt to show that morphogen gradients are sufficient to pattern the embryo.

(A) The French Flag Model proposed by Wolpert states that a morphogen gradient can provide positional information to a field of cells. The morphogen (green circles) diffuses from its source to form thresholds of signal. In the cartoon, high levels of the morphogen activate the red expression pattern. While, intermediate and low levels of the morphogen activate the white and blue expression patterns, respectively. In this model there is no need for other inputs as the gradient thresholds are sufficient for patterning.

(B) A study in *Xenopus* was conducted in which the animal cap was removed from early blastulas. The cells were dissociated and treated with increasing concentrations of the morphogen, Activin. The cells were then assayed for the presence of tissue specific gene markers. It was determined that increasing levels of Activin alone were able to specify five different tissue types.

(C) In a study to see if a morphogen can act at a distance, Nodal was injected into a single cell of a zebra fish embryo. The morphogen shown in green was able to diffuse and formed a gradient. The embryo was assayed for gene expression and it was found that a pattern of gene expression was formed with genes that respond to high levels of Nodal being expressed at the injection site and genes that respond to lower levels being expressed further from the injection site. When lower concentrations of Nodal were injected a pattern was not formed and only the genes that respond to low levels of Nodal were expressed.

(D) In *Drosophila* it was shown that Dpp signaling was capable of patterning the neurognic ectoderm. The neurogenic ectoderm genes depicted in the cartoon are *msh* (pink), *ind* (purple), and *sog* (teal). The lateralized embryo was created using an activated Toll receptor. In the lateralized embryo *sog* is expressed almost throughout the entire embryo. When Dpp was ectopically expressed in the *eve.stripe2* domain it was necessary to use a *sog*- background, because Sog inhibits Dpp. In this background the Dpp signal was able to diffuse and in areas where the signal was the highest *ind* was repressed allowing *msh* to be expressed, thus creating a pattern in the lateralized embryo.

Dorsal-Ventral Patterning and the Dorsal Nuclear Gradient

Dorsal-ventral pattering in the *Drosophila* embryo begins in the oocyte. The oocyte is surrounded by a group of cells called follicle cells, except at the anterior region where it is bordered by a group of cells called nurse cells. The nurse cells deposit Gurken and other factors into the ooctye; in late stages of oogensis, the nucleus and Gurken migrate to the dorsal-anterior region of the oocyte (Neuman-Silberberg and Schupbach,

1993). Gurken is the ligand for the Egf receptor Torpedo, which is located in the follicle cells. Gurken activates Torpedo imparting a dorsal fate to these cells. Both Gurken and Torpedo mutants result in ventralized embryos (Neuman-Silberberg and Schupbach, 1993). Torpedo signaling results in the limited expression of Pipe to the ventral follicle cells, where it activates a proteolytic cascade through an unknown mechanism (Sen et al., 1998). The cascade ends with the processing of Spätzle, which is the ligand of the Toll receptor (Morisato, 2001). Activation of the Toll receptor results in degradation of Cactus that binds to the transcription factor Dorsal to sequester it in the cytoplasm. Once Dorsal is released it can be transported into the nucleus, where it activates gene expression (Belvin et al., 1995). This establishes a nuclear gradient of Dorsal in the early syncytial blastodem, which will persist up to gastrulation (figure 1.2).

The Dorsal gradient does not follow the classical statues for formation of a morphogen gradient. In this case the ligand, Spätzle, is secreted into the perivitelline space by the follicle cells, which are not part of the embryo proper; but nonetheless, it follows many of the same themes used in interpretation of gradients for patterning. Ultimately, the effector of the signaling is Dorsal, which forms a nuclear gradient and activates target genes in a concentration dependent manner.

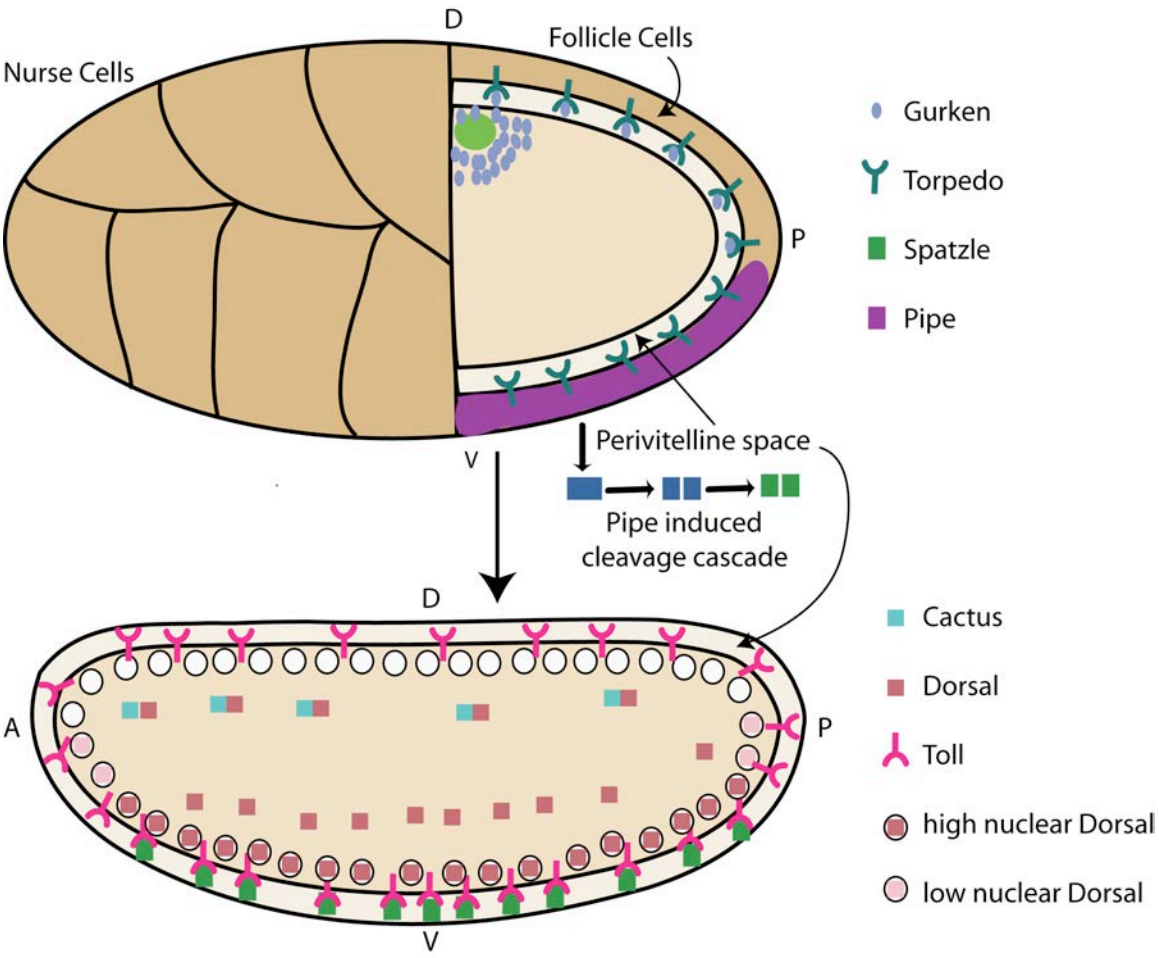


Figure 1.2. Patterning of the eggshell and embryo.

First the eggshell of the drosophila embryo is patterned, and this leads to the patterning of the embryo. The developing oocyte is shown in the upper cartoon and the embryo is shown in the lower one. Gurken (blue circles) is secreted into the oocycte by the nurse cells, and the nucleus and Gurken both migrate to the anterior-dorsal region of the embryo. Gurken is secreted into the perivitelline space where it binds the Egf receptor Torpedo. Activation of Torpedo signaling results in limited expression of Pipe (purple) to ventral follicle cells. Pipe initiates a cleavage cascade that results in the cleavage of Spätzle. Spätzle is secreted into the perivitelline space where it binds to the Toll receptor located in the embryo membrane. Activation of the Toll pathway leads to the

degradation of Cactus (blue square), resulting in the translocation of Dorsal into the nucleus. A gradient is formed with high levels of Dorsal in the ventral nuclei and lower levels in intermediate nuclei. There is little-to-no Toll signaling in the Dorsal region of the embryo resulting in little-to-no nuclear Dorsal.

The nuclear gradient of Dorsal divides the embryo into three tissue types; mesoderm is specified by high levels of nuclear Dorsal, where as, the neurogenic ectoderm and non-neurogenic ectoderm are specified by intermediate and low/no levels of nuclear Dorsal, respectively (Stathopoulos and Levine, 2004) (figure 1.3 B). The genes that specify the non-neurogenic ectoderm are expressed in the dorsal part of the embryo and are repressed by Dorsal, while genes in ventral and lateral regions of the embryo are activated by Dorsal. Over 30 genes have been identified as Dorsal target genes (Stathopoulos et al., 2002). For several of these genes the cis-regulatory modules (CRM) were studied and proved to be useful in deciphering the mechanisms used in interpretation of the Dorsal gradient (Ip et al., 1992b; Liberman and Stathopoulos, 2009; Markstein et al., 2004; Stathopoulos and Levine, 2005).

Analysis of Dorsal binding sites within CRMs revealed binding sites of both high and low affinity. The low affinity binding sites were located in the CRMs of genes such as *snail (sna), twist (twi), and heartless (hrt),* which are expressed in the ventral most part of the embryo and specify the mesoderm. Presumably, these genes can only be expressed in regions with high levels of nuclear Dorsal. The Dorsal binding sites for these genes were also closely associated with Twist binding sites, which suggest the use of a feed-forward loop in the regulation of Dorsal target genes (Ip et al., 1992b; Jiang et al., 1991).

The genes that are expressed in the neurogenic ectoderm contain high-affinity Dorsal binding sites and while some also contain Twist binding sites, they are not as closely associated to the Dorsal binding sites (Stathopoulos and Levine, 2004). Some of the genes expressed in the presumptive neurogenic ectoderm are expressed in broad patterns [*short gastrulation* (*sog*) and *rhomboid* (*rho*)], while others appear to be carved out by repressors and are restricted to fewer cells [*single minded (sim)* and *intermediate neuroblast defective (ind)*]*(Morel and Schweisguth, 2000; Stathopoulos and Levine, 2005)*. In fact cross repression, which is seen as a theme in the regulation of other morphogen gradients, is seen in interpretation of the Dorsal gradient (Cowden and Levine, 2003). To this effect, the genes that respond to high levels of the gradient, which specify ventral regions of the embryo, repress genes that respond to lower levels of the gradient in a ventral-dominant fashion. Thus *sna* represses the genes that are expressed in the neurogenic ectoderm, and within the neurogenic ectoderm genes expressed in more ventral locations repress genes that are expressed more dorsally (figure 1.3 A).

Dynamics of the Dorsal Nuclear Gradient

A recent study has shown that the early expression patterns of dorsal-ventral target genes are closely related to the dorsal gradient and changes in the patterns are seen with changes in the Dorsal nuclear gradient (Reeves et al., 2012). As cell divisions occur, Dorsal is shuttled in and out of the nucleus; the amplitude of the gradient also increases with each successive nuclear cycle. The target genes respond to Dorsal dynamics; at the end of nuclear cycle 13 the target genes are repressed in ventral regions but then in early nuclear cycle 14, when the gradient is being reestablished after nuclear

division, the genes are derepressed in ventral regions (Reeves et al., 2012) (figure 1.3 A). Presumably at these early stages, the broad expression patterns are solely reliant on activation by Dorsal; and at later stages, interactions between the target genes and inputs from other signaling molecules and transcription factors refine and maintain the patterns.

The Dorsal Gradient is Sufficient to Activate Genes that are Expressed along the Dorsal-Ventral Axis

A study found that if an activated Toll receptor was placed at the anterior pole of the early embryo, it was sufficient to create a Dorsal gradient along the anterior-posterior axis (Huang et al., 1997). This ectopic gradient of nuclear Dorsal was able to activate the expression of its target genes. The patterning of the genes was similar to the patterning observed in the dorsal-ventral axis. This would suggest that Dorsal alone is sufficient to pattern the dorsal-ventral axis, although applying a gradient response interpretation model, such as the French flag model (figure 1.1 A), would be an oversimplification because the patterns are not solely established by different concentrations of Dorsal.

We know that there are other inputs such as combinatorial interactions between Dorsal and Twist and cross-repression between the Dorsal target genes that provide positional information for pattern formation. These interactions that are dependent on Dorsal still occurred in the ectopic gradient. Also, combinatorial interactions between Dorsal and Zelda have been shown to be important in the expression of *sog*. Zelda is ubiquitously expressed and thus would be available to interact with ectopic Dorsal to activate dorsal-ventral patterning genes along the anterior-posterior axis (Liberman and Stathopoulos, 2009).

Many of the interactions and signaling pathways that pattern the dorsal-ventral axis are either directly or indirectly dependent on Dorsal. Another example is Egfr signaling, which is dependent on Dorsal activation of the Egfr ligand, *vein (vn)* and the Egfr ligand (*spitz*) processor *rho* (Ip et al., 1992a; Schnepp et al., 1996) (figure 1.3 C). There is no denying that a gradient is necessary for patterning. If only one level of Dorsal is present, then only one tissue type is formed and the axis is not patterned properly (figure 1.1 D), but patterning is more complicated than genes being activated by different threshold levels of a gradient. A more complete model is that the Dorsal nuclear gradient establishes the early expression patterns and also a base for signaling and combinatorial interactions, this along with other maternally deposited factors present in the embryo patterns the dorsal-ventral axis (figure 1.3). As development proceeds, the gene regulatory network established by Dorsal is self-sufficient, allowing patterning and development to continue without the need for the gradient to be maintained.

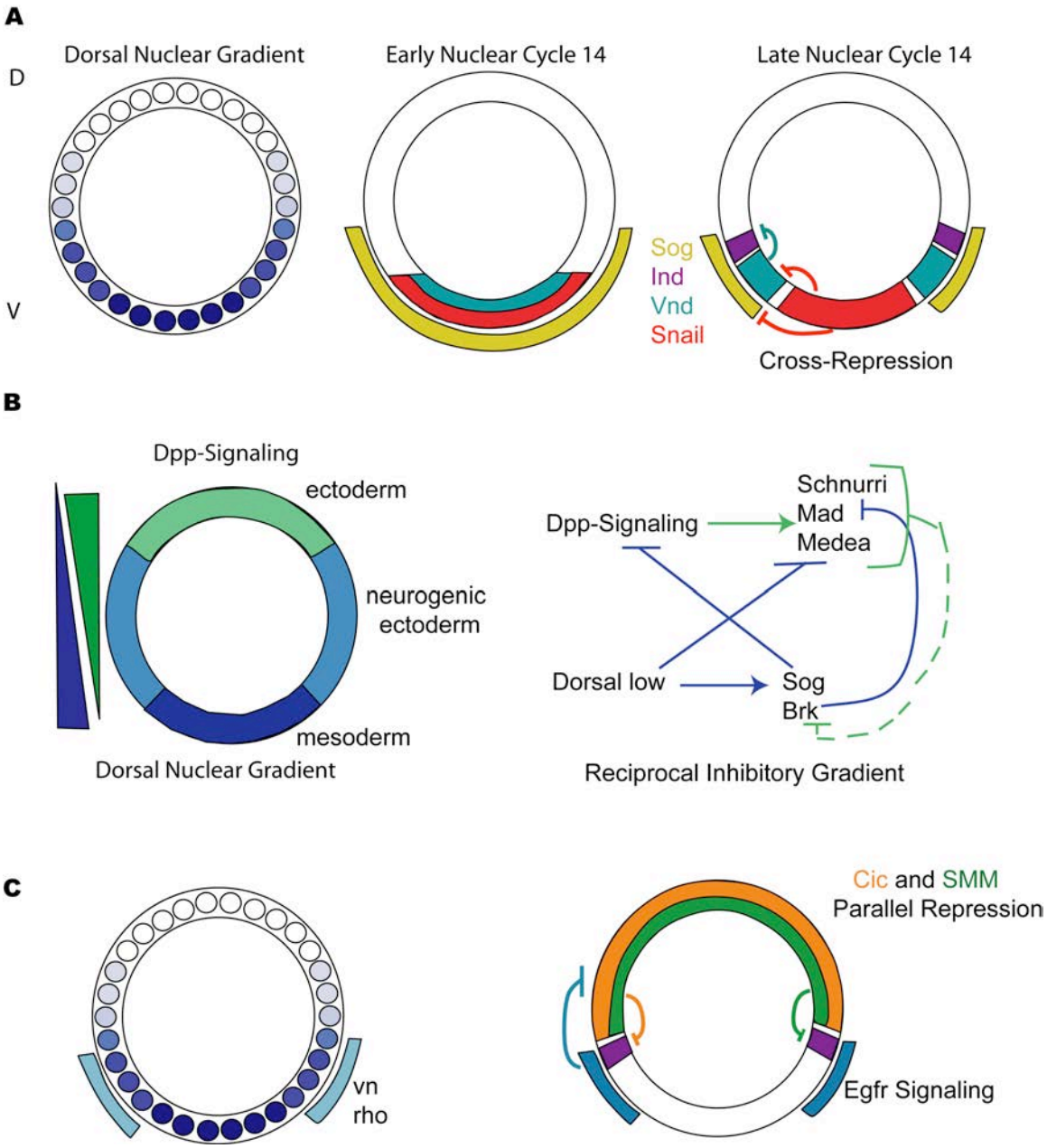


Figure 1.3. Interpretation of the Dorsal gradient and patterning of the dorsal-ventral axis of *Drosophila* embryos.

The Dorsal nuclear gradient establishes gene expression, which is then refined by interactions such as cross-repression and reciprocal inhibitory gradients.

(A) The schematic depicts a cross section through the center of the embryo. The small circles represent nuclei and the different shades of blue represent increasing levels of the

Dorsal nuclear gradient, with the highest levels of Dorsal found in the most ventral nuclei. In early nuclear cycle 14 the genes expressed along the dorsal-ventral axis, *sog* (yellow), *sna* (red), and *vnd* (teal) are expressed in overlapping regions in the ventral part of the embryo. At this point they are only reliant on the forming Dorsal gradient that is being reestablished after nuclear division. In late nuclear cycle 14 Snail represses *vnd* and *sog* restricting their expression to ventral-lateral and lateral regions of the embryo, respectively. At this stage, *ind* (purple) is expressed in lateral regions of the embryo; it is repressed by Vnd and Sna in ventral-lateral and ventral regions of the embryo.

(B) The embryo is divided into three tissue layers; the ectoderm (green) in dorsal regions of the embryo, the neurogenic ectoderm (light blue) in lateral regions of the embryo, and the mesoderm (dark blue) in ventral regions of the embryo. A reciprocal inhibitory gradient of Dpp-signaling (BMP-signaling) opposes the Dorsal nuclear gradient. This gradient is formed by inhibition of Dpp directly by Sog and indirectly by Brk. Dorsal is also believed to bind to the CRMs of Dpp target genes to repress their expression. Due to the presence of Schnurri-Mad-Medea (SMM) binding sites in neurogenic ectoderm CRMs it is believed that Dpp signaling represses them. Thus, the Dpp and Dorsal gradients act antagonistically to each other.

(C) Egfr signaling plays a role in patterning the neurogenic ectoderm. Dorsal activates expression of *vn* and *rho* in ventral-lateral regions of the embryo. They function to turn on Egfr signaling in this region of the embryo. Egfr signaling is believed to inhibit Cic (pink); this allows for the activation of *ind* (purple), as it frees it from repression by Cic. There is a second tier of repression that limits the dorsal border of *ind*, this is mediated by Dpp-signaling acting through the SMM binding site (green).

The Role of Dpp-Signaling in Dorsal-Ventral Patterning of the Early Embryo

Opposing the Dorsal nuclear gradient there is a BMP signaling step gradient with high levels in dorsal-most regions and lower levels in dorsal-lateral regions of the embryo. This gradient is created by graded localization of the ligand, Dpp; while the receptors Thickveins and Punt are ubiquitously expressed (Ashe, 2005; Mizutani et al., 2006). Dpp binds to the receptors either as a homodimer or as a heterodimer with Screw (Scw). Graded Dpp signaling functions to pattern the non-neurogenic ectoderm and might also function to repress genes that are expressed in the neurogenic ectoderm (Ashe, 2005; Mizutani et al., 2006).

Dpp is localized to dorsal and dorsal-lateral regions of the early embryo. Although the expression pattern of *dpp* is generally uniform, a step gradient of Dpp signaling is formed with high levels of signaling in dorsal most regions of the embryo and lower levels of signaling in dorsal-lateral regions of the embryo (Irish and Gelbart, 1987; Ray et al., 1991; Shimmi et al., 2005). There is evidence to suggest that Sog binds to Dpp and inhibits it from binding the receptors (Yu et al., 1996). In lateral regions of the embryo a Dpp-Scw-Sog-Tsg complex is formed which is transported to dorsal regions of the embryo (Shimmi et al., 2005). The protease Tolliod cleaves this complex resulting in higher levels of signaling in dorsal regions of the embryo (Canty et al., 2006; Serpe et al., 2005). In lateral regions of the embryo, where Sog is present, Dpp rebinds Sog keeping signaling levels low. In contrast, Sog is absent in dorsal regions, which allows for high levels of Dpp-Scw heterodimers to accumulate resulting in high levels of signaling. *Sog* is activated by Dorsal, thus the Dorsal gradient indirectly interacts with and helps to shape the Dpp morphogen gradient (figure 1.3 B). It is believed that Dpp

effectors are repressed by Brinker, adding yet another level of interaction between the Dorsal and Dpp gradients (Jazwinska et al., 1999; Rushlow et al., 2001).

Dpp signaling is also thought to act on Dorsal target genes and is believed to aid in patterning the neurogenic ectoderm. This mechanism is highly conserved between vertebrates and invertebrates, although in vertebrates the opposing gradient in the neural tube is Sonic Hedgehog and not Dorsal. It has been shown that, in a lateralized background, activation of Dpp signaling in an anterior-posterior stripe can create a pattern of neurogenic ectoderm genes (Mizutani et al., 2006) (figure 1.1 D). At the ectopic source of Dpp, the most dorsal neurogenic gene, *msh*, is expressed with the second most dorsal neruogenic gene, *ind*, being expressed in the rest of the embryo (figure 1.1 D). In contrast, when Dpp signaling is lost or reduced there is not a dramatic change to the neurogenic ectoderm genes, suggesting that while Dpp signaling can and likely does contribute to patterning of the neurogenic ectoderm, it is not absolutely necessary (Von Ohlen and Doe, 2000). In the following chapters, we provide evidence for the presence of other regulatory factors that are independent of Dpp signaling, that act dorsally to pattern genes in the neurogenic ectoderm. The presence of these factors clarifies some of the results that have been seen when investigating the role of Dpp signaling in regulation of the neurogenic ectoderm. Unlike ectopic activation of Dpp signaling, loss of Dpp signaling has little to no effect on gene expression in the neurogenic ectoderm. The lack of an observable phenotype in Dpp mutants is likely due to the presence of factors that are independent of Dpp signaling that compensate for the loss of Dpp signaling.

Egfr Signaling also Plays a Role in Patterning the Neurogenic Ectoderm

Egfr signaling is present in the presumptive neurogenic ectoderm, and it is essential for proper patterning. The Egfr receptor is ubiquitously expressed at this stage and the ligand *vein* is activated by Dorsal in lateral regions of the embryo (Schnepp et al., 1996). *rho,* which activates Egfr signaling by cleaving the Egfr ligand Spitz, is also activated by Dorsal in a domain similar to *vein (Bier et al., 1990; Ip et al., 1992a).* During later stages of development neuroblasts develop a unique identity based on where they are located and differentiate into specialized neuroblast accordingly. Their location is determined based on which proneural gene they express and whether they receive Egfr signaling. In Egfr mutants expression of *ind* is lost and consequently no intermediate neuroblast are formed (Skeath, 1998). The medial neuroblast, are still specified and express *vnd*, but the lack of Egfr signaling causes them to display some traits specific to lateral neuroblast. The lateral neuroblasts, which are specified by *msh*, expand into regions where intermediate neuroblasts would normally form. Thus it is clear that Egfr signaling plays and important role in patterning the dorsal-ventral axis of the neurogenic ectoderm. Even though Egfr signaling is only necessary for the activation of *ind*, it has a dramatic affect on the patterning of the neurogenic ectoderm; as loss of just one of the neurogenic ectoderm genes results in patterning defects of the entire tissue.

The Egfr receptor is a receptor tyrosine kinase (RTK) with many downstream effectors including transcriptional activators and repressors. Initially it was unclear whether Egfr signaling activated *ind* directly or whether it inhibited a repressor allowing *ind* to be expressed. A recent publication and our work featured in the following chapter support the latter case, suggesting Egfr signaling is responsible for inhibiting a repressor

that binds an 16 bp repeated sequence (the "A-box" element) present in the *ind* CRM (Ajuria et al., 2011; Stathopoulos and Levine, 2005) (figure 1.3 C).

Insights into Dorsal-Ventral Patterning via CRM Analysis

Analysis of CRMs has proven useful in understanding how genes are regulated and how axis specification is determined. In chapter 2, we analyze the *ind* CRM, which drives expression of a sharp dorsal-ventral stripe in the presumptive neurogenic ectoderm. We show that an element, we call the "A-box" is both necessary and sufficient to mediate repression in dorsal regions of the embryo and thus maintains proper patterning of the neurogenic ectoderm by refining the dorsal border of *ind.* We conducted affinity chromatography to identify factors that bind the A-box element; our analysis resulted in the identification of Grh, which we believe acts as an activator rather than a repressor. We also show that a second tier of repression, acting to define the dorsal boundary of *ind*, is mediated by Dpp signaling and acts on the *ind* CRM via a Schnurri-Mad-Media complex (SMM) binding site (figure 1.3 C). Thus, the pattern of *ind* is initiated by Dorsal and Grh activation and is refined by tiers of repressors in both ventral and dorsal regions of the embryo.

Our analysis of the A-box element reveled that as well as the activator Grh several chromatin factors also bound the A-box element. Chapter 3 summarizes the results from the affinity chromatography analysis and discusses a possible role for chromatin remodeling factors in regulating *ind* expression and patterning of the dorsal-ventral axis of the early embryo.

In chapter 4 we extend our CRM analysis to *vnd* and *sog*. We show that dorsal repressors may play a role in regulating the expression pattern of *vnd* and, to a lesser extent, *sog*. We also provide a more detailed discussion of the combinatorial interactions known to pattern the dorsal-ventral axis.

In chapter 5 we discuss the implications of our studies and discuss future directions to determine how genes are regulated to establish axis specification.

Chapter 2

Lateral Gene Expression in *Drosophila* Early Embryos is Supported by Grainyhead-mediated Activation and Tiers of Dorsally-Localized Repression*

*This chapter, first published in *Plos One* in 2011, was written by Mayra Garcia and Angelike Stathopoulos.

Abstract

The general consensus in the field is that limiting amounts of the transcription factor Dorsal establish dorsal boundaries of genes expressed along the dorsal-ventral (DV) axis of early *Drosophila* embryos, while repressors establish ventral boundaries. Yet recent studies have provided evidence that repressors act to specify the dorsal boundary of *intermediate neuroblasts defective* (*ind*), a gene expressed in a stripe along the DV axis in lateral regions of the embryo. Here we show that a short 12 base pair sequence ("the A-box") present twice within the *ind* CRM is both necessary and sufficient to support transcriptional repression in dorsal regions of embryos. To identify binding factors, we conducted affinity chromatography using the A-box element and found a number of DNA-binding proteins and chromatin-associated factors using mass spectroscopy. Only Grainyhead (Grh), a CP2 transcription factor with a unique DNA-binding domain was found to bind the A-box sequence. Our results suggest that Grh acts as an activator to support expression of *ind*, which was surprising as we identified this factor using an element that mediates dorsally-localized repression. Grh and Dorsal both contribute to *ind* transcriptional activation. However, another recent study found that the repressor Capicua (Cic) also binds to the A-box sequence. While Cic was not identified through our A-box affinity chromatography, utilization of the same site, the A-box, by both factors Grh (activator) and Cic (repressor) may also support a "switch-like" response that helps to sharpen the *ind* dorsal boundary. Furthermore, our results also demonstrate that TGF-β signaling acts to refine *ind* CRM expression in an A-box independent manner in dorsal-most regions, suggesting that tiers of repression act in dorsal regions of the embryo.

Introduction

During development the embryo is patterned by the localized expression of genes to discrete parts of the embryo. Such tight spatial regulation of gene expression is necessary to set the boundaries that distinguish different cell types required for proper development. One mechanism to impart spatial information is to regulate gene expression through transcription factors that are spatially localized. Alternately, localized activation of signaling pathways in particular domains can also influence the boundaries of gene expression.

In *Drosophila melanogaster,* the dorsal-ventral (DV) axis of the pre-gastrula embryo is patterned by a nuclear gradient of the NF-κB homologous transcription factor Dorsal (Reeves and Stathopoulos, 2009). High levels of nuclear Dorsal are present in ventral regions of the *Drosophila* embryo and nuclear levels decrease progressively toward more dorsal regions. Due in part to these differing nuclear Dorsal levels, different domains of gene expression are established along the DV axis to specify different cell types (Stathopoulos et al., 2002). In the ventral most regions of the embryo, high concentrations of nuclear Dorsal drive expression of genes such as *twist* and *snail (sna)* to specify the presumptive mesoderm. In ventral lateral regions of the embryo, intermediate levels of Dorsal activate genes such as *rhomboid (rho)* and *ventral neuroblast defective (vnd*) and low levels of Dorsal support expression of genes such as *short gastrulation (sog)* in broad lateral domains of the embryo (that encompass both ventral-lateral and dorsal-lateral regions) to specify distinct domains within the presumptive neurogenic ectoderm (Bier et al., 1990; Ip et al., 1992a; Jimenez et al., 1995). Lastly, as Dorsal can also function as a repressor, the expression of some genes such as *zerknüllt* (*zen*) are

limited to dorsal regions of the embryo, leading cells in this domain to adopt amnioserosa and non-neurogenic dorsal ectoderm cell fates (Jiang et al., 1993; Jiang et al., 1992; Stathopoulos et al., 2002). Even though Dorsal provides positional information through its dorsal-ventrally modulated nuclear gradient, combinatorial interactions of transcription factors are very influential towards DV patterning. Specifically, Dorsal regulates gene expression together with other transcription factors, such as the bHLH factor Twist and the early ubiquitous activator Zelda (e.g. Ip et al., 1992b; Liang et al., 2008; Liberman and Stathopoulos, 2009).

More and more evidence suggests that signaling pathways also help to define gene expression patterns in the early embryo. For example, the expression domains of several Dorsal target genes cannot be explained by changing Dorsal levels (and/or the localization of any other previously characterized transcription factors). Additionally, it is well understood that signaling molecules provide positional information to help define the very specific expression domain encompassed by the gene *single-minded (sim)*. *sim* is expressed as a stripe of a single cell width present in ventrolateral regions of the embryo, within cells located between the presumptive mesoderm and neurogenic ectoderm boundary. *sim* expression is supported by combinatorial interactions of Dorsal and Twist transcription factors and also through Notch-dependent signaling (e.g. Morel and Schweisguth, 2000).

Along similar lines, the gene *intermediate neuroblast defective (ind)* is expressed in dorsal-lateral regions of the embryo in a stripe of 5–7 cells in width, which is narrower than the broad domain encompassed by *sog*. Genetic studies support the view that refined *ind* expression is supported by inputs from both Dorsal and Epidermal growth

factor receptor (Egfr) signaling, suggesting that limiting amounts of both of these inputs help delineate *ind* expression boundaries (Von Ohlen and Doe, 2000). The Egfr gene is ubiquitously expressed in embryos but the receptor is activated locally in ventrolateral regions by the ligands Vein and Spitz (Rutledge et al., 1992; Schnepp et al., 1996). Several binding sites for the ETS transcription factor, which mediates Egfr signaling, are also found in the *ind* cis-regulatory module CRM, but it has not been shown if they are required for activation or whether an indirect mechanism is used for activation of *ind* expression via Egfr signaling (Stathopoulos and Levine, 2005).

No other gene in the *Drosophila* embryo described to date shares the same expression domain as *ind*, yet understanding how the *ind* expression domain is regulated may have far-reaching implications. Interestingly, the genes that pattern the ventral nerve cord of *Drosophila* and the neural tube of higher vertebrates share a conserved organization and function (Cornell and Von Ohlen, 2000; Mizutani et al., 2006). Specifically, the gene *ventral neuroblast defective* (*vnd*)/*Nkx2.2* is expressed ventral to *ind/Gsh*, and the gene *muscle specific homeobox* (*msh*)/*Msx1/2* is expressed dorsally to *ind* (Chu et al., 1998; Isshiki et al., 1997; Weiss et al., 1998). Experiments conducted in the *Drosophila* embryo have suggested that the ventral boundaries of these genes are set following a "ventral dominance rule," in which the more ventral genes repress expression of the more dorsal genes (Cowden and Levine, 2003). In contrast, it had been proposed that the dorsal boundaries of these genes result from limiting amounts of the activator, Dorsal, present in distinct domains along the DV axis (reviewed in Stathopoulos et al., 2002). However, recently it was discovered that the *ind* gene is expressed in a domain along the DV axis where the Dorsal gradient appears uniform without a clear transition

that would be capable of setting a dorsal border (Liberman et al., 2009). A previous analysis of the *ind* CRM suggested evidence for a dorsally-acting repressor which could explain how the dorsal boundary of *ind* is specified (Stathopoulos and Levine, 2005).

Direct evidence for repressor action within dorsal regions of the early embryo was found through analysis of the *cis*-regulatory region of *ind* (Stathopoulos and Levine, 2005). A 1.4 kb (kilobase) DNA fragment located ~2 kb downstream of the *ind* coding sequence was found to support expression in a refined stripe within lateral regions of the embryo, in a pattern comparable to the endogenous gene. However, the promoter proximal half of the *ind* CRM drove expression of a reporter gene within a broad pattern, one that extends into ventral-lateral as well as dorsal-lateral regions, suggesting that the distal half contains repressor binding sites. Using a chimeric CRM assay designed to detect repression along the dorsal-ventral axis by silencing of an associated *even-skipped stripe 3/7* CRM (*eve.stripe3/7*), this previous study found that the 1.4 kb *ind* CRM mediates repression of *eve.stripe3/7* in dorsal and ventral regions of the embryo. A specific search for an element supporting dorsal repression was conducted and identified a 111 base pair (bp) region of the *ind* CRM, which supported dorsal-lateral and dorsal repression of *eve.stripe3/7*. A 12 bp sequence was highlighted, as it repeats twice within these 111 bp, and was called the A-box (WTTCATTCATRA). Importantly, in this previous study, when the A-box was mutated in the context of a minimal element supporting repression in dorsal regions (i.e., 267 bp fragment), repression of the *eve.stripe3/7* CRM was lost. Presumably transcription factors bind to the A-box element to help establish the dorsal boundary of the *ind* gene, but their identities remained unknown.

Additional evidence also suggests that TGF-ß signaling may also regulate the *ind* expression domains, but whether or not this signaling pathway functions through the A-box element was not known. Decapentaplegic (Dpp) is a TGFß/BMP homolog that is limited in its expression to dorsal regions of the embryo and functions as a morphogen to support patterning of the amnioserosa, at higher levels in dorsal-most regions of the embryo, and the non-neurogenic ectoderm, at lower levels in dorsal-lateral regions of the embryo (Ferguson and Anderson, 1992). A previous study found that in mutants in which Dpp signaling is expanded into lateral regions of the embryo, *ind* expression is lost (Von Ohlen and Doe, 2000). Likewise, ectopic expression of *dpp* in lateralized embryos that exhibit expanded *ind* expression throughout the embryo was able to repress *ind* in the domain where Dpp signaling was presented (Mizutani et al., 2006). Also, the *ind* CRM contains a 15 bp DNA sequence implicated in TGF-β signaling-mediated repression (Stathopoulos and Levine, 2005). Similar sites have been shown to mediate repression by recruiting a Dpp-dependent Schnurri/Mad/Medea (SMM) protein complex, but SMM dependent repression of *ind* has never been shown and in fact this mechanism of repression has only been shown to act at later stages of development (Dai et al., 2000; Pyrowolakis et al., 2004).

Therefore, to gain further insight into how patterning is controlled along the dorsal-ventral axis of *Drosophila* embryos, we tracked the repression activity supported by different DNA elements associated with the *ind* CRM. We found that the A-box element facilitates both activation and repression of *ind* and propose that this helps to mediate a sharp border. In addition, we found that TGF-β signaling supports *ind*

repression in dorsal-most regions of the embryo through the SMM site located within the *ind* CRM that is distinct from the A-box.

Results

Chimeric CRM Assays Can Help Identify and Track Repression Activity Associated with CRM Sequences

In order to gain insights into how the boundaries of dorsal-ventral patterning genes are set, we deconstructed the *cis*-regulatory element of *ind* to find direct evidence for dorsal repression activity. We utilized a chimeric cis-regulatory module (CRM) assay, using *eve.stripe3/7* and *ind* CRMs in order to determine whether repressors are present within either of these sequences to help refine the domains of expression (Stathopoulos and Levine, 2005). The *ind* CRM supports expression along the DV axis in a lateral stripe, comparable to the endogenous gene ((Figure 1A; Stathopoulos and Levine, 2005). In turn, the *eve.stripe3/7* sequences supports expression of two stripes located along the anterior-posterior (AP) axis of embryos (figure 2.1 B) (Small et al., 1996). When two CRMs are placed in tandem upstream of a reporter gene (i.e. *lacZ*), if additive expression is observed this result indicates that either repressors are not present or they are not located in range to act on the adjacent CRM; conversely, if non-additive expression is observed this indicates repressors are present and function to silence activators associated with both CRMs. Previously, using a chimeric CRM assay, it was shown that the 1.4 kb *ind* CRM drives repression of *eve.stripe3/7* (Figure 1C; Stathopoulos and Levine, 2005). In this case non-additive expression is observed; the *eve.stripe3/7* CRM is repressed in ventral regions by *snail* and *vnd* repressor sites located

in the *ind* CRM and by unknown transcription repressors in dorsal regions. Concurrently, the *ind* CRM is repressed by Knirps, through sites in the *eve.stripe3/7* CRM, forming a gap in the *ind* expression pattern. It was suggested the unknown transcription repressors located in dorsal regions act through a pair of 12 bp A-box sequences located within the 1.4 kb *ind* CRM. Here we examined the function of the A-box sequence more closely.

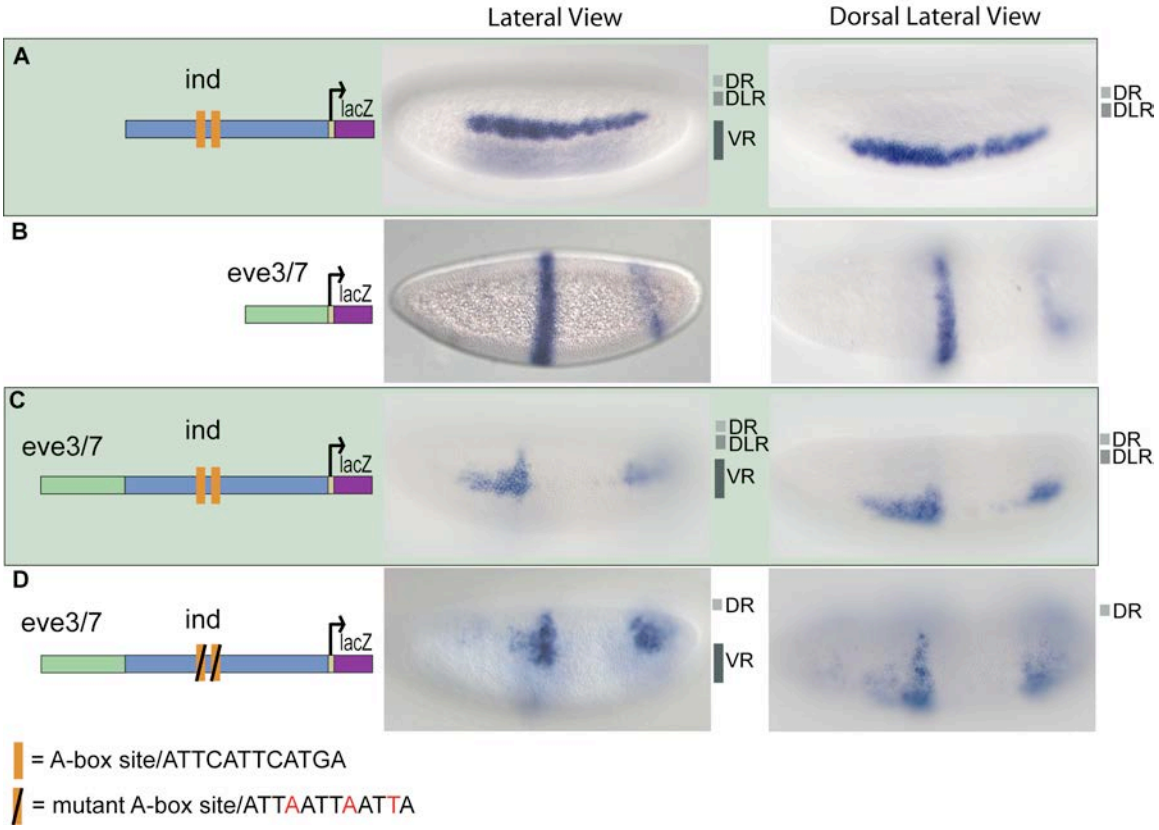


Figure 2.1. The *ind* CRM contains binding sites that mediate repression in dorsal regions.

lacZ reporter expression was visualized within cellularized embryos (late stage 5) by in situ hybridization using a digoxigenin-labeled antisense *lacZ* riboprobe. In this and all subsequent figures, embryos are oriented with anterior to the left. In addition, embryos are oriented to show views of lateral, dorsal on top, (left image) and dorsal (right image) domains. The repression domains are outlined to the right of each image: DR = dorsal

repression, DLR = dorsal lateral repression, and VR = ventral repression. The schematic depicts the chimeric CRM combinations used: **(A)** The 1.4 kb *ind* CRM drives expression of *lacZ* in a 5–7 cell lateral stripe representative of *ind* expression; **(B)** The 0.5 kb *eve.stripe3/7* CRM drives expression of *lacZ* in two anterior-posterior stripes representative of *eve.stripe3/7* expression; **(C)** The *eve.stripe3/7-ind* chimeric CRM drives expression of *lacZ* in a non-additive fashion showing repression of *eve.stripe3/7* in dorsal, dorsal lateral, and ventral regions; **(D)** The *eve.stripe3/7-mut-A-box-ind* chimeric CRM supports non-additive expression with repression of *eve.stripe3/7* in dorsal and ventral regions but not dorsal lateral regions.

The A-box Element Mediates Repression of ind in Dorsal-Lateral Regions of the Embryo, while Other Sequences Support Repression in Ventral and Dorsal-Most Regions of the Embryo

When we mutated both of the A-box sites in the context of the full-length *ind* CRM and assayed the fragment's ability to repress expression of the associated *eve.stripe3/7* CRM, repression of *eve.stripe3/7* was lost in dorsal lateral regions (figure 2.1 D, compare with 2.1 C). This result demonstrated that these two A-box sequences are necessary to mediate dorsal-lateral repression of *eve.stripe3/7* by the *ind* CRM. Next, we assayed the full-length *ind* CRM with two mutant A-boxes alone and found that *lacZ* reporter expression was expanded into dorsal-lateral regions; giving a broad, patchy, and diffuse pattern not a sharp stripe of 5–7 cells in width representative of *ind* (figure 2.2 B, compare with 2.2 A).

However, even in the absence of the A-box sites, repression was retained in dorsal-most and ventral regions of the embryo when the A-box was mutated in the context of the full-length CRM (figure 2.2 B), as well as in the chimeric CRM assay of *ind* and *eve.stripe3/7* CRMs (figure 2.1 D). These results suggest that the A-box sequences mediate dorsal-lateral repression, but that there might be other repressor binding sites in the *ind* CRM which mediate repression in dorsal-most and ventral regions of the embryo. Vnd and Snail binding sites within the *ind* CRM most likely mediate the repression observed in ventral regions (Cowden and Levine, 2003). In contrast, while we were able to track repression in dorsal-most regions, the identity of the responsible transcription factors was unknown.

A-box Elements Limit Expression in Dorsal-Lateral and Dorsal Regions of Embryos

Another important question is whether the A-box elements are sufficient to cause repression of the *eve.stripe3/7* CRM, as perhaps multiple sequences within the *ind* CRM are necessary to support repression. To investigate this, we flanked the *eve.stripe3/7* CRM with the A-box element (i.e., A-box.*eve.stripe3/7*.A-box) and observed clear repression in dorsal-lateral regions, as expected, and also within dorsal regions of the embryo (figure 2.2 D). Weak repression was also observed in ventrolateral regions at lower frequency (data not shown). This result suggests that A-box sequences are sufficient to support repression in dorsal-lateral regions, but also contribute to repression in dorsal-most and ventrolateral regions of the embryo.

The expression supported within the *eve.stripe3/7* domain did extend a few cells above the endogenous *ind* dorsal boundary in the context of the A-box.*eve.stripe3/7*.A-

box reporter. This may indicate the chimeric CRM assay is limited in its ability to track repression activity as the stripe of expression also extended a few cells above *ind* when the full length *ind* CRM was assayed in tandem to *eve.stripe3/7*. Alternatively, sharp definition of the *ind* dorsal boundary may require more input than localized repressor activity.

The Dorsal Transcription Factor Only Partially Supports Activation of ind

We investigated the activation of the *ind* expression pattern by mutagenizing the sole match to the Dorsal binding site consensus present within the *ind* 1.4 kb CRM (figure 2.2 C). *ind* is not expressed in *dorsal* mutants (Von Ohlen and Doe, 2000), thus, we expected loss of the sole Dorsal binding site would severely impair reporter expression. Instead, we found that the expression pattern driven by the mutated CRM is very similar to that driven by the wild-type CRM, except for a gap in the expression pattern (figure 2.2 C).

Early *ind* expression, at the start of cellularization, exhibits a smaller gap in expression at 40% egg length (Stathopoulos and Levine, 2005) which is likely mediated by anterior-posterior patterning factors. In reporter constructs, repression within this domain is more apparent with the 1.4 kb *ind* CRM sequence is oriented in the opposite direction relative to the promoter in reporter constructs (data not shown). The function of activators, including Dorsal and others that act through the A-box sequence, are likely required to counterbalance this repression.

Our results suggest that Dorsal binding contributes to *ind* activation but that other activators also influence *ind* expression. Furthermore, chromatin immunoprecipitation

(i.e., ChIP-seq) experiments did not detect Dorsal binding in the genome at the *ind* CRM (Ozdemir et al., 2011), which indicates Dorsal may not bind to the *ind* CRM (or that it is a very transient interaction). Collectively, these results suggest that additional transcriptional activators likely function to support *ind* expression.

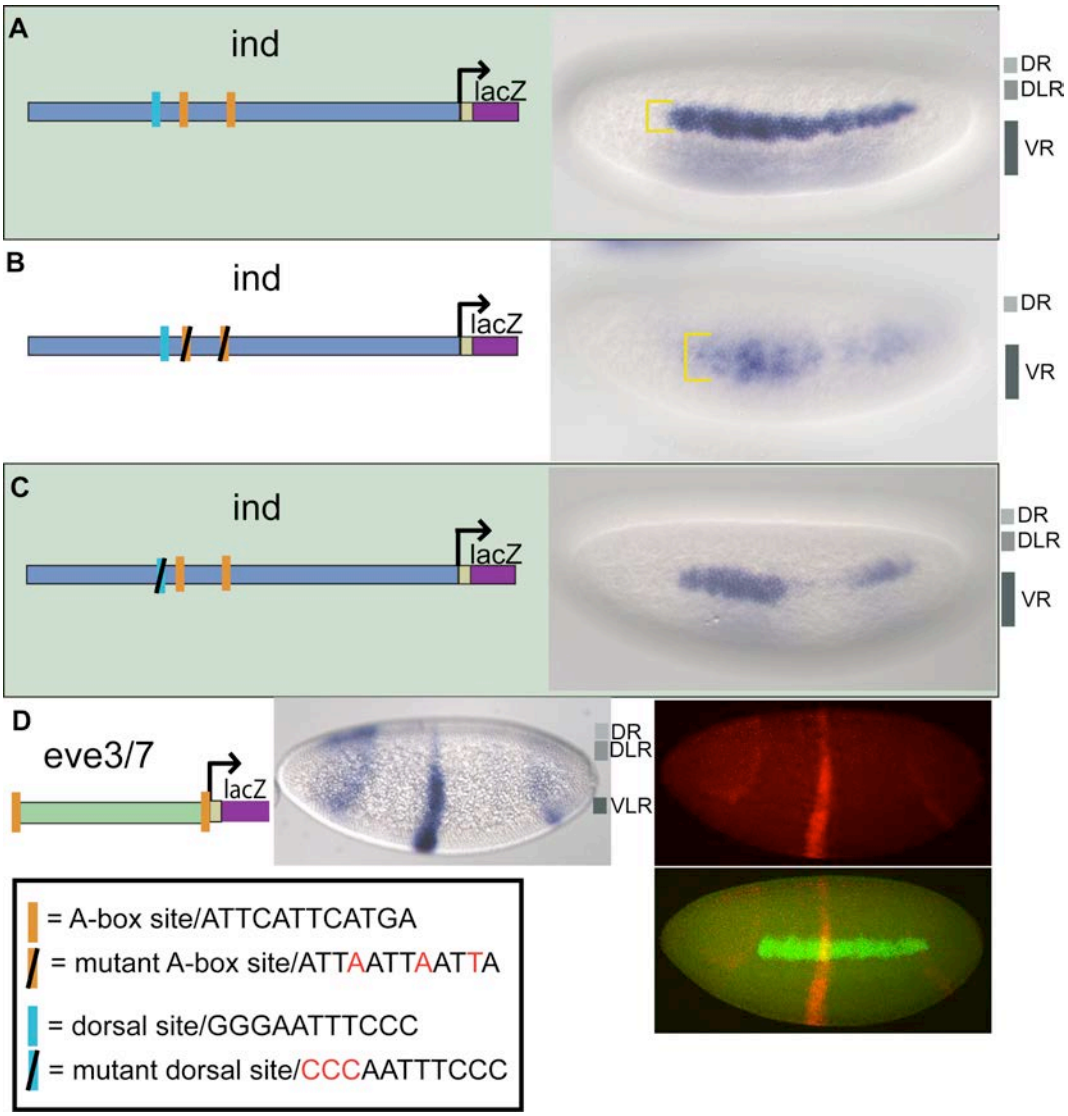


Figure 2.2. The A-box sites are necessary for dorsal lateral repression and sufficient for dorsal and dorsal-lateral repression.

Site-directed mutagenesis was used to mutate regulatory sites in the *ind* CRM. The CRMs depicted in the schematic were used to drive expression of *lacZ* in embryos that were analyzed by in situ hybridization using a *lacZ* anti-sense riboprobe. Cellularized embryos of stage 5 are oriented to show a lateral view, with anterior to the left and dorsal on the top. The yellow brackets mark the height of the expression pattern. The repression domains are outlined to the right of the image: DR = dorsal repression, DLR = dorsal lateral repression, VR = ventral repression, VLR = ventral lateral repression. **(A)** The 1.4 kb *ind* CRM drives expression of a lateral stripe of 5–7 cells in width comparable to *ind* expression. **(B)** The 1.4 kb *mut-A-box-ind* CRM drives expression of 7–10 cell width lateral stripe that is diffuse, weak, and expanded compared to the *ind* CRM. **(C)** The 1.4 kb *mut-dorsal-ind* CRM drives expression that has a gap and is weak in posterior regions compared to the *ind* CRM. **(D)** The *eve.stripe3/7* CRM flanked by A-box sites (*A-box-eve.stripe3/7-A-box)* shows repression in dorsal, dorsal-lateral, and ventral-lateral regions. In the fluorescent image *lacZ* expression is shown in red and endogenous *ind* expression is shown in green, as detected by multiplex fluorescent in situ hybridization (Kosman et al., 2004).

Dorsalized and Lateralized Embryos Provide Insights into the Localization of the A-box Repressor Activity

Next we introduced the *lacZ* reporter gene containing the *eve.stripe3/7* CRM flanked by A-box sequences (i.e., A-*box-eve.stripe3/7-A-box*) into different mutant backgrounds to test whether the repressor activity associated with the A-box sequence is influenced by altered DV positional information. Maternal mutant backgrounds exist that

affect the levels of nuclear Dorsal (i.e., low or absent) to create lateralized or dorsalized embryos, respectively. Expression of Dorsal target genes are affected such that certain genes expressed by a particular level of Dorsal, normally refined in expression to distinct domains along the DV axis, are instead expressed ubiquitously or absent in either of these mutant backgrounds. In sum, our aim was to determine whether the repressor activity was responsive to changes in Dorsal levels, providing additional evidence that the repressor activity we had tracked was indeed functioning in a DV localized manner.

In *pipe* mutants, Dorsal is not able to enter the nucleus thus Dorsal target genes are not activated, resulting in dorsalized embryos (Stathopoulos et al., 2002; Stein et al., 1991). In this mutant background, endogenous *ind* is not expressed. We assayed the *A-box-evestripe3/7-A-box lacZ* reporter construct in the *pipe* mutant background and found that expression of *lacZ* was retained but severely dampened (figure 2.3 B compare with 2.3 A). This result suggests that some repressor activity is present ubiquitously in dorsalized embryos but most likely it is less active, because only partial repression of the reporter is observed.

We also examined reporter expression in $Toll^{rm9/10}$ embryos, which have a partially active form of the Toll receptor allowing low levels of Dorsal to enter the nucleus throughout the embryo (Stathopoulos et al., 2002). In this background *ind* is expressed throughout the embryo, suggesting that repressors are unable to refine the *ind* pattern in this background. We also observed strong uniform expression of the *lacZ* reporter in the *eve.stripe3/7* domain indicating that in this background the repressor activity is gone (figure 2.3 C).

The A-box element clearly supports repression in dorsal regions of the embryo and is responsive to mutations altering DV pattern (figure 2.3). These results suggest the A-box associated repressor exhibits localized expression in dorsal regions of the embryo and/or that its activity is modulated by signaling pathways that exhibit differential activation along the DV axis.

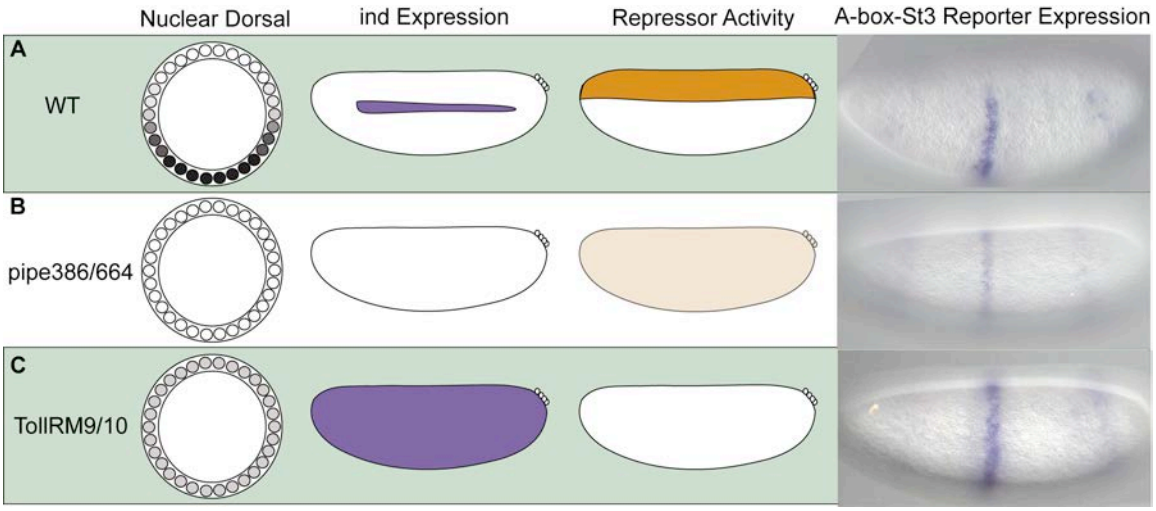


Figure 2.3. Dorsalized and lateralized embryos provide insights into the A-box repressor domain of activity.

The depictions show the Dorsal nuclear gradient within embryo cross-section schematics, whereas *ind* expression and the putative repressor activity are schematized within lateral views. Expression of the *A-box-eve.stripe3/7-A-box* reporter gene was examined by in situ hybridization in (A) wild-type, (B) $pipe^{384}/pipe^{664}$ mutants, and (C) $Toll^{RM9}/Toll^{RM10}$ mutants. The in situ images show *lacZ* expression as such: (A) Repression of *lacZ* is shown in dorsal regions of the embryo in WT embryos. (B) Weak repression of *lacZ* is shown throughout the embryos from *pipe* mutant females (i.e., dorsalized embryos). (C) A lack of repression of *lacZ* is shown in embryos from $Toll^{RM9/10}$ mutant females (i.e., lateralized embryos).

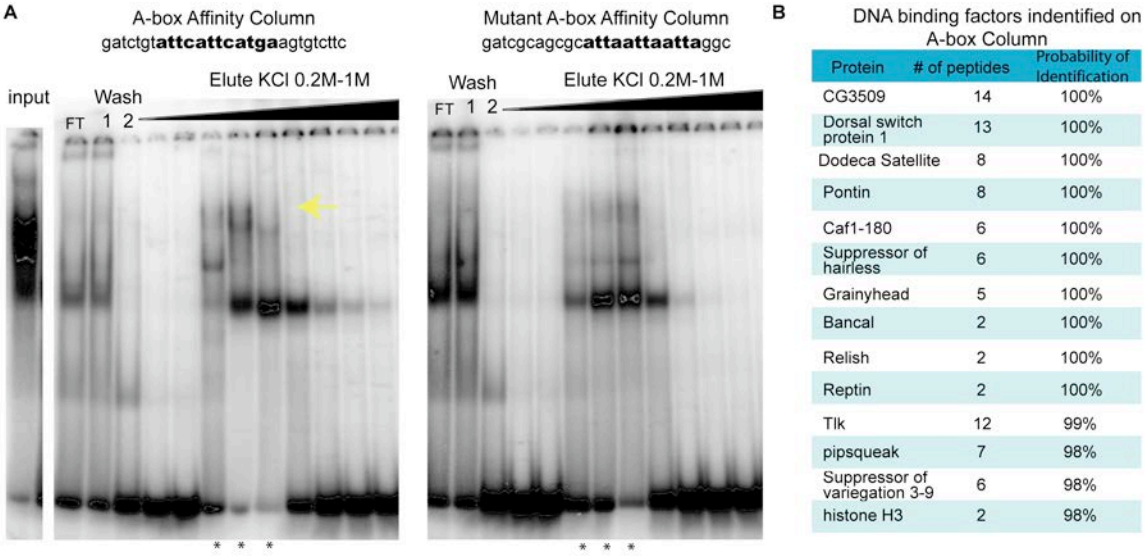
Affinity Chromatography and Mass Spectrometry Identifies Putative A-box Binding Factors

In order to provide molecular insight into the mechanism by which *ind* expression in dorsal regions is limited, we set out to identify the factor that binds the A-box element choosing affinity chromatography using a 22 bp oligonucleotide containing the A-box sequence (12 bp) and endogenous flanking regions (5 bp on either side). As a control, we also compared binding with that obtained with a mutant A-box sequence modified in 3 of 12 bp, which we showed does not support dorsal repression when assayed in the context of a chimeric CRM assay *in vivo* (see figures 1.1 D, 1.2 B) and containing different flanking region].

We used affinity chromatography to purify proteins that recognize the A-box or mutant A-box sequence from early embryonic nuclear extracts age 0–6 hours. The A-box binding activity was tracked throughout a number of biochemical separations (see figure 2.S1 and materials and methods). There were several factors that bound to both columns but some of the binding was specific to the A-box (figure 2.4 A). Cold competition with the A-box versus the mutant A-box confirmed the binding observed was specific to the A-box (data not shown). With advances in mass spectroscopy, we could analyze a complex sample containing a number of proteins. Therefore, at this step, we analyzed samples isolated from either the A-box column or the mutant A-box column by mass spectrometry.

Focusing on factors that only bound the A-box column (figure 2.4 B), we selected targets for future analysis. Several transcription factors were found specifically associated with the A-box, and not the mutant A-box column. Furthermore, several

chromatin-related factors bound to the A-box column but failed to bind the mutant A-box column (figure 2.S2). This suggested to us that the repressor activity associated with the intact A-box sequence may be composed of a large complex of proteins including chromatin components; a role for chromatin in supporting expression in the early *Drosophila* embryo is unclear (see discussion).



Protein	# of peptides	Probability of identification
CG3509	14	100%
Dorsal switch protein 1	13	100%
Dodeca Satellite	8	100%
Pontin	8	100%
Caf1-180	6	100%
Suppressor of hairless	6	100%
Grainyhead	5	100%
Bancal	2	100%
Relish	2	100%
Reptin	2	100%
Tlk	12	99%
pipsqueak	7	98%
Suppressor of variegation 3-9	6	98%
histone H3	2	98%

Figure 2.4. Affinity chromatography and mass spectrometry was used to identify factors that bind the A-box element.

(A) shows the EMSAs preformed using γ^{32}P-labeled A-box oligonucleotides on nuclear extract fractions after they were affinity purified with the A-box column and the mutant A-box column. FT denotes the flow through which did not bind to the column. The yellow arrow marks the area where the A-box specific binding was found. The stars mark the samples used for mass spectrometry identification. **(B)** The table lists the DNA binding factors that bound to the A-box element column but not the mutant A-box column. The "# of peptides" corresponds to the number of unique peptides that contributed to the protein identification. The probability of identification was calculated

by the program Scaffold used to identify the proteins by mass spectrometry analysis and corresponded to the likelihood a correct match was made.

The Grainyhead Transcription Factor Binds to the A-box sequence and is Required to Support ind Expression

In order to narrow down a list of factors to examine in this preliminary analysis, we focused on identifying factors that bind specifically to the A-box DNA sequence. We conducted EMSAs on the following factors, which contain a predicted DNA-binding domain, and for which cDNAs were available: ATP-dependent chromatin assembly factor large subunit (Acf1), Structure specific recognition protein (Ssrp), CG3509, Grainyhead (Grh), Dorsal switch protein 1 (Dsp1) and Pipsqueak (Psq) (data not shown). Of these factors, only Grh exhibited binding to the 22 bp oligonucleotide, containing the 12 bp A-box and endogenous sequences.

Using in vitro translated proteins in EMSAs, we further analyzed Grh and found that while it bound the A-box element it did not bind to the mutant A-box element (figure 2.5 B, full gel figure 2.S3). We, therefore, conducted additional analysis on Grh as it seemed a likely candidate to support the A-box repression activity. The *grh* gene is maternally and zygotically expressed (Huang et al., 1995; Weber et al., 1995), and by in situ hybridization we confirmed that it is ubiquitously expressed in the early embryo (figure 2.5 A). While some evidence exists that *grh* transcripts are localized to dorsal and lateral regions of the embryo (Huang et al., 1995), we could not detect such a localized expression domain by in situ hybridization even though a number of different riboprobes were designed to detect *grh* transcripts.

We generated *grh* germline clone females in order to deplete both maternal and zygotic *grh* expression from embryos. The conventional method of creating germline clones (Perrimon, 1998), which relies on flipase catalyzed mitotic recombination in the context of transheterozygous FRT *ovoD* (dominant female sterile mutation) and FRT *grh* chromosomes, for example, could not be used because *ovoD* within the commonly used FRT *ovoD* chromosome is most likely inserted at the *grh* locus. FRT *ovoD* in combination with all *grh* alleles tested are zygotically lethal, but no lethality was observed with *ovoD* insertions located on other chromosomes. Thus, it was necessary to make germline clones in females of the genetic background FRT *grh*/FRT GFP. Embryos obtained from these females were manually screened for absence of GFP (Luschnig et al., 2004), thus allowing isolation of embryos containing the mutant form of *grh*. To ensure that *grh* zygotic transcripts were absent, females containing germline clones were mated to males containing appropriate balancer chromosomes to allow detection in the early embryo (i.e., FRT *grh*/ Cyo *ftz-lacZ*; see Materials and Methods).

Because manual hand sorting of embryos was required, only a small number of embryos could be examined, but multiplex in situ hybridization allowed us to examine the expression of multiple genes simultaneously. Therefore, in addition to examining the effect of loss of *grh* on *ind* expression, we also assayed whether this mutation affected expression of two other genes, *tailless* (*tll*) and *zen*. In a previous study, embryos devoid of *grh* maternal message were produced X-ray irradiation induced mitotic recombination; *tll* was found to be expanded in *grh* mutant embryos obtained in this manner (Liaw et al., 1995). However, we failed to see expansion of *tll* in embryos lacking both maternal and/or zygotic *grh*; a similar negative result was recently reported (Harrison et al., 2010).

Our results concur with those of Harrison et al. and we agree that the expansion of *tll* observed previously (Liaw et al., 1995) was most likely an artifact induced by X-ray irradiation. We also examined *zen* expression in order to determine if there was any effect on Dpp target genes due to loss of *grh*; a previous study had suggested that *grh* may be involved in repression of *dpp* (Huang et al., 1995). During early stages, *zen* expression is broad, present in dorsal-lateral regions as well as dorsal regions, but by cellularization (late stage 5) its pattern has refined to a dorsal stripe present in dorsal-most regions of the embryo (Doyle et al., 1989). This later pattern is regulated by Dpp-mediated TGF-β signaling (Ray et al., 1991; Rushlow et al., 2001). However, no effect on *zen* expression was identified in embryos lacking maternal and zygotic *grh* (figure 2.5 F, compare 2.5 D; and data not shown).

In contrast to the "normal" expression patterns of the genes *tll* and *zen* within *grh* mutant embryos, we found that *ind* expression was severely dampened in these mutants (figures 2.5 C, E, G compare with 2.5 D, F, H); the data for grh^{IM} is shown. In wild-type embryos, *ind* comes on weakly at first during early stage 5 (precellularization), but by the end of stage 5 upon complete cellularization of embryos *ind* expression becomes sharp and clearly apparent. In the absence of maternal *grh*, the *ind* pattern was severely to weakly affected (figure 2.5 compare F to H), with some embryos showing an almost complete loss of *ind* in late stage 5 and others showing a weak thin uniform stripe compared to the wild-type tapered stripe. It is possible that the *grh* zygotic contribution relates to the variability. Furthermore, only a weak phenotype was observed with the grh^{B37} allele, which is expected because grh^{IM} is the stronger amorphic allele. To confirm that the phenotype observed was due to the grh^{IM} mutation and not a secondary mutation,

we mated the females containing the germline clones to males in which the *grh* gene is absent, *Df(2R)Pcl7B/Cyoftzlacz*. We did not observe a rescue suggesting the phenotype is associated with loss of *grh*.

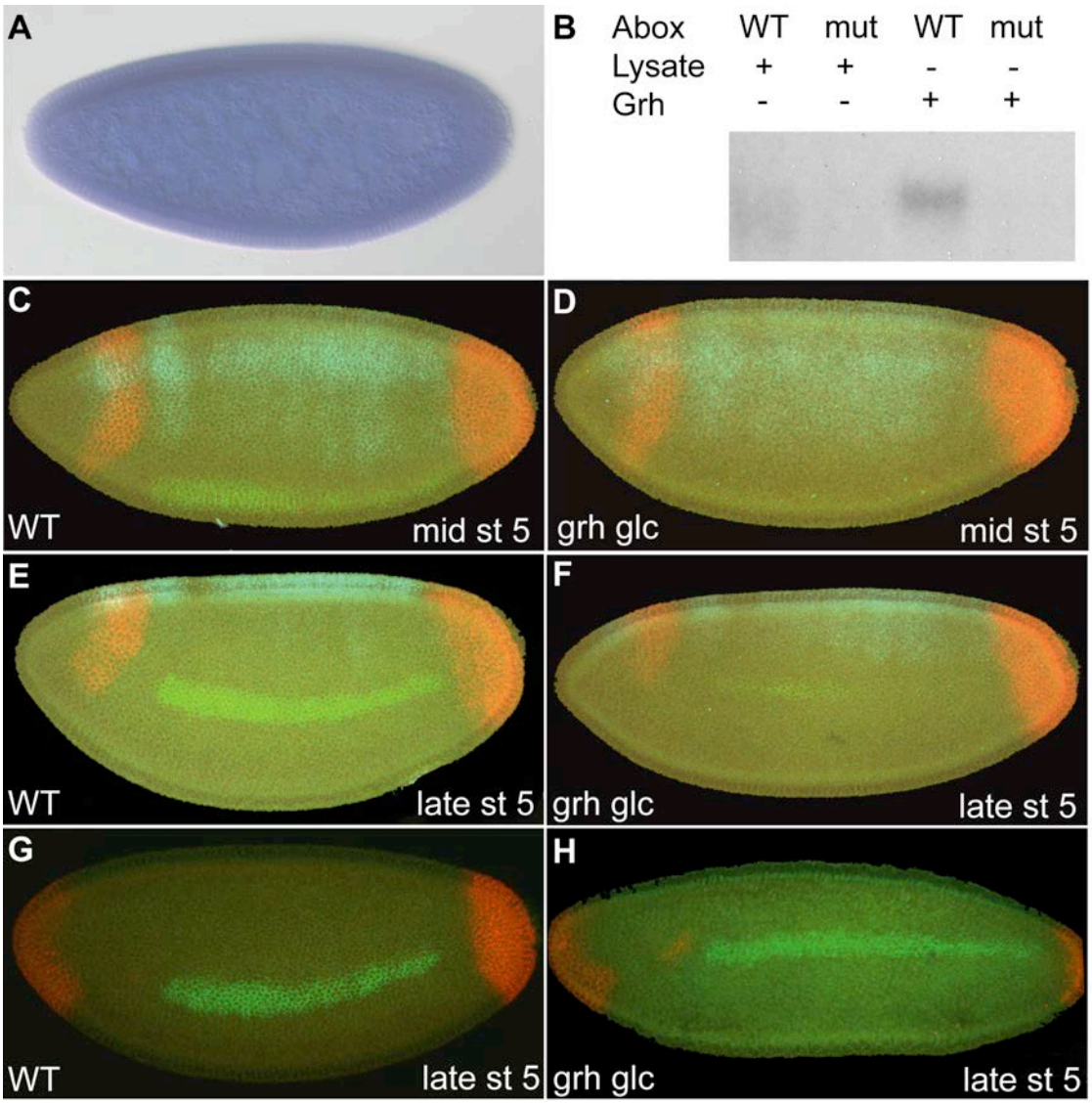


Figure 2.5. Grainyhead binds the A-box element and is involved in activation of *ind*. (A) *Grh* is expressed ubiquitously in embryos as detected by in situ hybridization using a *grh* riboprobe. (B) Grh was expressed in rabbit reticulocytes and EMSA was performed using γ^{32}P-labeled A-box and mut-A-box oligonucleotides. Grh bound the A-box oligo

but did not bind the mutant A-box oligo. Reticulocyte lysate alone was also tested for binding as a control. Expression of *zen* (cyan), *tll* or *hkb* (G and H*)* (red) and *ind* (green) are shown in wild-type (C, E, and G) and *grh* glc-derived embryos (D, F, and H). The embryos in C and D are tilted ventrally to show the broad *zen* expression indicative of mid stage 5. Weak *ind* expression is observed in WT embryos (C) but not in embryos derived from *grh glc* (E). The embryos in E, F, G and H are oriented to show a lateral view and are late stage 5. Strong *ind* expression was detected in wildtype (WT) embryos (E and G) while very faint (F) or thin (H) *ind* expression was detected in embryos derived from *grh* glc females.

To investigate whether Grh is responsible for the repressive function as well as the activation function of the A-box, we assayed whether the A-box could support repression in embryos obtained from *grh* mutant germline clone females. We did not see an effect on the repressor activity supported by an *eve.stripe3/7* CRM flanked by A-box sites in the absence of maternal and zygotic *grh*; the pattern was repressed in dorsal regions even in the absence of *grh* (figure 2.6 E and F).

The loss of *ind* expression in *grh* mutants and retention of dorsally-localized repression was unexpected because we had isolated the Grh protein using the A-box element, which clearly supports repression in dorsal regions of the embryo. Nevertheless, we had observed that mutagenesis of the A-box sites within the *ind* CRM not only caused expansion of the pattern but also caused a reduction in levels of expression of the reporter gene (figure 2.2 B, compare with 2.2 A). Therefore, we reasoned that Grh might function as a transcriptional activator that drives *ind* expression

through the A-box sequence, and hypothesized that yet another factor might bind to the same site, to mediate repression. A recent study shed light on this issue as it presented evidence that the Capicua (Cic) repressor is required to support repression through the A-box and that it is modulated by Egfr signaling (Ajuria et al., 2011).

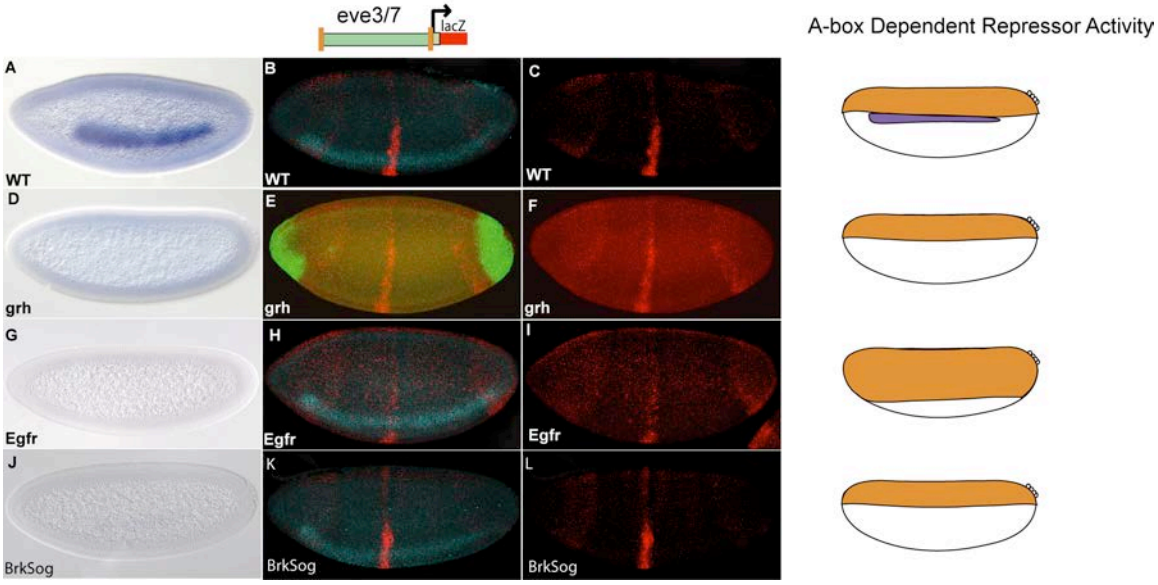


Figure 2.6. Analysis of A-box dependent and A-box independent repression in different mutant backgrounds.

Embryos (stage 5) were analyzed by in situ hybridization for *ind* expression. Multiplex in situ hybridization was used to analyze A-box dependent and A-box independent repression in different mutant backgrounds. The schematic shows the CRMs used to drive expression of *lacZ.* The orange boxes in the schematic correspond to A-box sites. The cartoons to the right of the images show where A-box/Cic dependent (orange) repression is located in WT embryos and in the corresponding mutants; *ind* is only expressed in wild-type embryos (purple). *ind* expression is shown in WT embryos (A), *grh* glc derived embryos (D), *egfr* mutants (G), and *brk sog* double mutants (J). The *A-*

box-eve.stripe3/7-A-box reporter construct was introduced into different mutant backgrounds and analyzed by in situ hybridization; *lacZ* (red), and *vnd* (blue) is shown in a in WT embryo (B), *grh* glc derived embryo shows expression of *hkb* (green) rather than *vnd* and is tilted dorsally relative to the rest of the embryos (E), *egfr* mutant (H) and *brk sog* mutant (K). For clarity *lacZ* expression is shown alone for the corresponding embryos WT (C), *grh* glc (F), *egfr* mutant (I), and *brk sog* mutant (L). The same microscope settings were used to image C, I, and L; different settings were used for F but it was compared to a WT embryo taken under the same settings (not shown).

Loss of Egfr Signaling Expands the A-box Supported Repression Domain Ventrally

To gain insights into the mechanism of repression, we examined *ind* expression as well as A-box mediated repression in *cic* as well as *Egfr* mutants. First we looked at *Egfr* mutants in which it has been shown that *ind* expression is lost (Figure 6E; Von Ohlen and Doe, 2000). Egfr signaling supports *ind* expression either directly by supporting activation through the various ETS sites found in the *ind* CRM (Stathopoulos and Levine, 2005) or indirectly by inhibition of a repressor. If the latter is the case we would expect to see expansion of A-box mediated repression into ventral lateral regions. In *Egfr* mutants, repression of the stripe was expanded ventrally, which we assayed by relating the reporter gene expression to the domain of *vnd* expression (*vnd* is expressed ventral to *ind*, in ventrolateral regions of the embryo) (figure 2.6 H). When the reporter was assayed in a wild-type background (i.e,. *yw*), it extended about 8 cells above the dorsal border of *vnd* (figure 2.6 B). However, in *Egfr* mutants, strong expression of the stripe was only visible up to the ventral border of *vnd* (figure 2.6 H and I) and in some cases

weak expression extends above the dorsal border of *vnd* (data not shown). These results suggested that the repressor binding the A-box element is itself inhibited by Egfr signaling. In the absence of Egfr signaling, repression is unrestrained and expands ventrally toward the ventral border of *vnd*.

Ajuria et al. reported that the *ind* expression domain was slightly expanded in the absence of maternal *cic* transcript (cic^1/cic^1females). We introduced the *A-box.eve.stripe3/7.A-box* reporter into the cic^1/cic^1 mutant background. Reporter expression was expanded into dorsal regions suggesting that repression activity was lost, however anterior-posterior patterning is severely compromised in cic^1/cic^1 mutants (data not shown).

To examine whether Grh-mediated activation and Cic-mediated repression through the A-box might be linked in general, we examined other genes regulated by Cic to determine whether they might also be regulated by Grh. In Ajuria et al., they found that Cic binding sites which are similar to the A-box binding sites are found in several other CRMs and mediate Cic-dependent repression. We looked at one of these genes, *huckebein* (*hkb*), in *grh glc* mutant embryos to test the idea that Grh might act as a general activator for CRMs containing an A-box-like site (figure 2.6 E). We did not see an effect on *hkb* expression, suggesting that Grh activation via the A-box binding site does not act to regulate *hkb* expression (or other activators that support *hkb* expression). Our results suggest that Cic and Grh may work coordinately through the A-box but that they likely have independent binding sites/targets as well (see discussion).

Dpp Signaling Mediates Repression that is Independent of the Repression Mediated by the A-box Elements

We found that Egfr signaling modulates A-box mediated repression, but we also investigated whether Dpp signaling functions through the A-box as previous evidence had shown a relationship between TGF-β signaling and *ind* expression (Mizutani et al., 2006; Von Ohlen and Doe, 2000). If the A-box repressor is a Dpp target gene or is regulated by one of the Dpp target genes we might expect to see a change in our repression activity upon modulation of TGF-β signaling. We introduced the *eve.stripe3/7* CRM flanked by the two A-box sites into *brk sog* double mutants, in order to assay the A-box repressor activity in a background with ectopic Dpp signaling. Brinker (Brk) and Sog both act to restrict Dpp signaling activity to the dorsal most regions of the embryo (Biehs et al., 1996; Jazwinska et al., 1999). The *brk* gene encodes a transcription factor that functions to repress transcription of Dpp target genes; in turn, the *sog* gene encodes an extracellular Dpp binding protein which acts both as a direct Dpp antagonist and is also required for high level Dpp signaling in the dorsal midline. In *brk sog* double mutants, ectopic Dpp is observed in lateral regions of the embryo and at the same time *ind* expression is also diminished (Von Ohlen and Doe, 2000) (figure 2.6 J). If the A-box repressor is a Dpp target gene or is regulated by one of the Dpp target genes, we would expect to see an expansion of the repression domain. However, we did not observe a significant change in the repression activity in this mutant background (figure 2.6 K and L, compare with 2.6 B and C). This suggested that the A-box repressor acts independently of Dpp and its target genes. Dpp and its targets may still play a role in repression of *ind* via other unidentified binding sites.

When we analyzed expression supported by the *eve.stripe3/7-ind-mutant-A-box* reporter construct, we noted repression in the dorsal-most part of the embryo despite the lack of A-box sites (figure 2.1 D and 2.6 M). To investigate whether this particular repression activity was dependent on Dpp signaling, we assayed this reporter in *brk sog* double mutants. If this repression in dorsal-most regions of the embryo is dependent on Dpp signaling, we would expect to see an expansion of the repression into dorsal-lateral regions of the embryo. This was what we observed: the repression supported in *brk sog* mutants was present in a more broad domain, expanded dorsally well beyond its limit in wild-type embryos (figure 2.7 B). These results suggested that this repression in dorsal-most regions is dependent on *Dpp* signaling and is independent of the repression mediated by the A-box elements.

Schnurri is a Dpp target gene that is expressed in dorsal regions of the embryo (Arora et al., 1995; Grieder et al., 1995; Staehling-Hampton et al., 1995). It binds to DNA via the Mad and Medea binding sites forming a Schunurri/Mad/Medea (SMM) protein complex that mediates repression (Dai et al., 2000; Pyrowolakis et al., 2004). A SMM binding site is located in the *ind* CRM; it is possible that Dpp signaling mediates repression of *ind* via this binding site. In order to test this hypothesis we mutated the SMM site (Mad binding component) in an *ind* CRM that contained two mutant A-box sites and found that the expression pattern is further expanded (Figure 2.7 compare D and E).

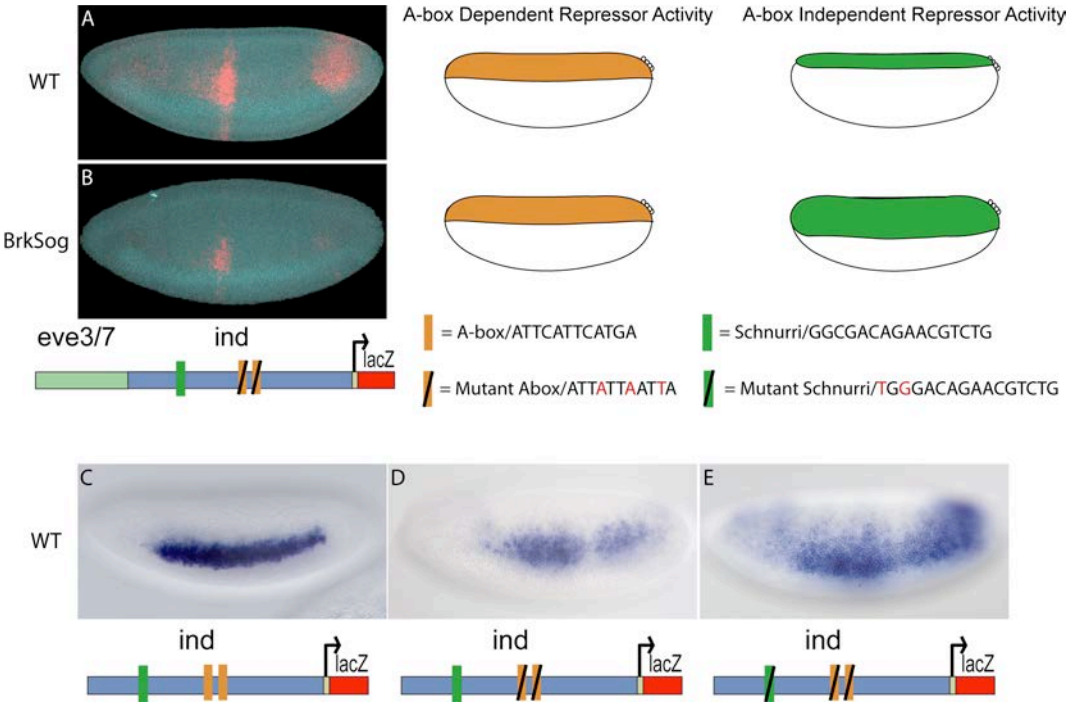


Figure 2.7. Dpp dependent repression is mediated via the Schnurri (SMM) binding site and is independent of A-box repression

A-box independent repression is observed in dorsal-most regions of the embryo in the *eve.stripe3/7-ind-mutant-abox* reporter construct. This construct was introduced into the *brk sog* mutant background and analyzed by in situ hybridization: *lacZ* (red) and *vnd* (blue) are shown in a WT embryo (A) and a *brk sog* mutant (B). The schematic at the bottom of the embryos shows the construct that was used to drive expression. The cartoon embryos on the right show where the designated repressors are possibly expressed. (C) The 1.4 kb *ind* CRM drives expression of a lateral stripe of 5–7 cells in a width comparable to *ind* expression. (D) The 1.4 kb *mut-A-box-ind* CRM drives expression of 7–10 cell width lateral stripe that is diffuse, weak, and expanded compared to the *ind* CRM. (E) The 1.4 kb *mut-A-box-mut-SMM-ind* CRM drives expression of 12–15 cell width lateral stripe that is expanded compared to the *ind* CRM and *mut-A-box-ind* CRM.

Thus, our results suggest that two distinct dorsally-localized repression activities refine *ind,* one dependent on Dpp signaling and the other independent of this signaling. This view is supported by the fact that ectopic Dpp is able to repress *ind* and yet loss of Dpp has no affect on its expression (Von Ohlen and Doe, 2000); we suggest that A-box mediated dorsal repression can compensate in the absence of Dpp. When Dpp signaling is overexpressed in a permissive environment that supports activation of its target genes, its presence is sufficient to repress *ind* in a Dpp-dependent fashion (Mizutani et al., 2006), but when Dpp signaling is lost, repression through a Dpp-independent mechanism (i.e., A-box repressor) is still able to restrict *ind* thus an expanded pattern is not observed.

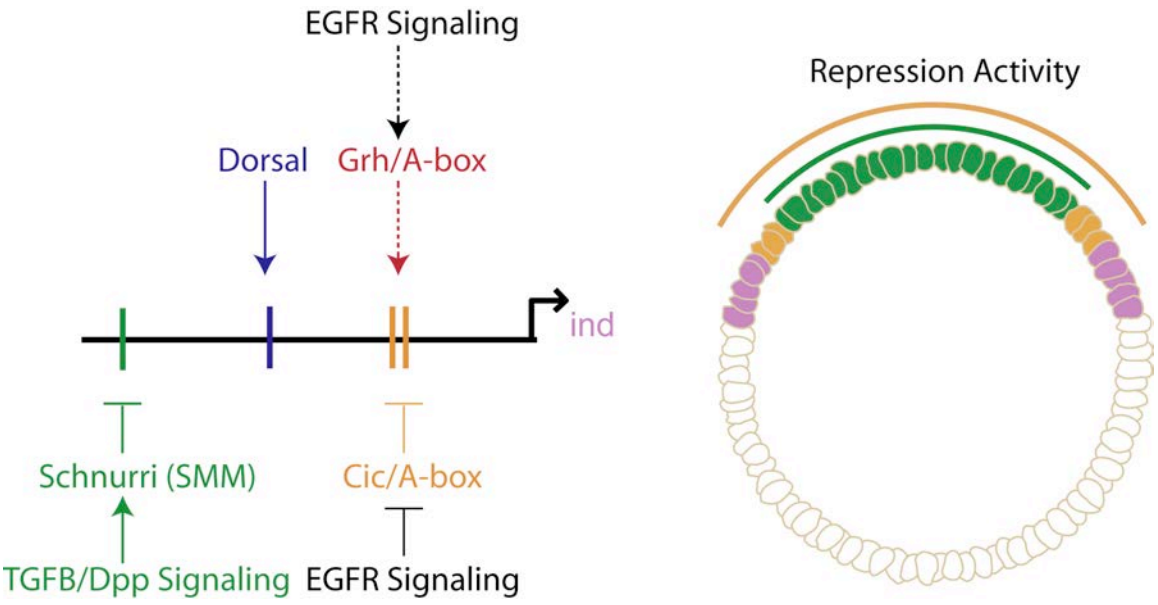


Figure 2.8. Model for transcriptional regulation of *ind* expression.

Our model is based on a compilation of this study and other studies suggesting that several transcription factors and signaling pathways interact to specify the *ind* pattern. This is only a partial model and does not include all the factors that delineate the ventral borders of *ind.* Two tiers of repression establish the dorsal border: one mediated by the

A-box binding site/Cic and the other mediated by a Dpp dependent repressor/Schnurri (SMM). Activation is mediated by Grh via the A-box binding sites and by Dorsal via the Dorsal binding site. The depiction shows the repressor activity relative to *ind* expression. Schnurri repression activity is limited to dorsal-most regions of the embryo. The A-box/Cic activity is found in dorsal and dorsal-lateral regions. The dashed lines indicate interactions that remain unclear.

Discussion

We analyzed the A-box sequence and showed it is both necessary and sufficient for repression of *ind* in dorsal-lateral regions and sufficient for dorsal-most repression. Through DNA affinity chromatography and mass spectrometry, we identified several binding factors many of which are involved in chromatin remodeling. One of the factors we identified, encoded by the *grh* gene, was previously shown to act as an activator as well as a repressor throughout development and during wound response (e.g., Huang et al., 1995; Kim and McGinnis, 2011). We showed Grh protein binds the A-box binding site in vitro. Since mutagenesis of the A-box sites within the *ind* CRM leads to decreased reporter expression and *ind* endogenous expression is also diminished in *grh* mutants, this data suggested that Grh drives activation of *ind* through the A-box; we note however that we cannot dismiss an additional role for Grh through other sequences in the *ind* CRM. We also demonstrated the repressive function of the A-box is restricted by Egfr signaling and is independent of Dpp signaling. In turn, we found, repression mediated by Dpp signaling does impact *ind* in dorsal-most regions of the embryo and possibly acts through the SMM binding site, not the A-box. Collectively, our results show interactions between

several signaling pathways and transcription factors are necessary to establish the *ind* expression pattern (figure 2.8).

Combinatorial Action of Grh and Dorsal Likely Support ind Activation

Other studies have shown combinatorial interactions are necessary to support patterns of gene expression along the DV axis. For instance, one study showed Dorsal and Zelda function together to produce the broad lateral domain of *sog*. Mutation of either the Dorsal sites or the Zelda sites in the *sog* CRM produced a pattern that was narrower than the wild-type expression pattern. It was concluded that both Dorsal and Zelda must be present to produce a proper Sog pattern (Liberman and Stathopoulos, 2009). It is also well appreciated that Dorsal can act cooperatively with the bHLH transcription factor Twist to support expression in ventral and ventrolateral regions of the embryo (e.g., Ip et al., 1992b). We propose that Grh and Dorsal act together to support the *ind* expression pattern. While the *ind* CRM containing a mutant Dorsal site did support some expression, the expression pattern contained a gap and was weaker in posterior regions; in contrast, in Dorsal mutants, *ind* expression is completely absent. This result may be explained if both indirect as well as direct functions for Dorsal are required to support *ind* expression. For instance, Dorsal has other target genes including *rho,* which is required to support Egfr signaling (Ghiglione et al., 2002; Urban et al., 2002). Furthermore, mutation of the A-box/Grh binding site within the *ind* CRM caused expression of the reporter that was expanded dorsally and weak, suggesting this site mediates repression and also activation. Similar to Dorsal mutants, the phenotype we observed when we mutated the A-box sites is different from the phenotype in the Grh

mutants, thus we cannot rule out that Grh may act through other sites as well as the A-box and/or that Grh may act indirectly to influence *ind* expression by regulating the expression of other transcription factors. We propose a model most consistent with the current data which is that *ind* is activated in regions where Dorsal is present as well as optimal levels of Grh (see below); it is then refined by Snail and Vnd in ventral regions and Cic and Schnurri/Mad/Medea (SMM) in dorsal regions (figure 2.8).

Egfr Signaling May Act to Regulate the Activity of the Cic Repressor as well as the Grh Activator to Support ind Expression

grh and *cic* genes are both maternal and ubiquitously expressed, thus, another input is necessary to explain how localized expression of *ind* is supported. This positional information could be provided in part by competition between Grh and Cic proteins for the A-box binding site and in part by ventrolaterally-localized Egfr signaling. A model in which Egfr signaling supports activation of *ind* via inhibition of a ubiquitous repressor (e.g., Cic) is supported by our results that demonstrate A-box mediated repression is expanded in *Egfr* mutants. A recent study also showed expanded expression of an *ind* CRM fragment reporter in *ras cic* double mutants in which neither Egfr signaling or Cic repressor is present, suggesting that Egfr may function by inhibition of an "inhibitor" to promote activation (Ajuria et al., 2011). This data suggests that the putative A-box repressor, Cic, may not be dorsally localized but that its activity is regulated by Egfr signaling which provides the positional information necessary for a sharp boundary. However, the domain of dpERK activation (as detected by anti-dpERK, an antibody to the dual-phosphorylated from of ERK) does not exactly overlap with the

ind expression domain at cellularization (data not shown), as would be expected in the simplest model.

Ajuria et al. suggested that Egfr signaling supports *ind* expression through inhibition of Cic, and we add that it is also plausible Egfr signaling impacts activation of *ind* through Grh. In fact, a recent study showed that Grh activity during wound response is modulated by ERK signaling (Kim and McGinnis, 2011). Specifically, they found both unphosphorylated and phosphorylated Grh can bind DNA and act as an activator. The former is used during normal development of the epidermal barrier and the latter is used to overcome a semi dormant state during wound response. Another study showed the tyrosine kinase Stitcher activates Grh during epidermal wound healing (Wang et al., 2009). In the early embryo Grh may be phosphorylated by Egfr signaling to support activation of *ind* through the A-box binding site. We suggest that phosphorylation of both Grh as well as Cic by Egfr signaling can act as a switch to help fine-tune the expression of *ind*.

Grh and Cic Function Coordinately Through the A-box but Likely also Have Independent Actions at Other Distinct Binding Sites

We investigated whether a relationship between Grh activation and Cic repression was used in regulation of other genes containing A-box or Cic binding sites. We found that one other Cic target gene, *hkb*, was unaffected in Grh mutants. As the A-box site (WTTCATTCATRA) is larger than the Cic consensus binding sequence [T(G/C)AATGAA, complement TTCATT(G/C)A] defined by Ajuria et al., it is possible that Grh needs the full A-box site to bind. The full A-box sequence is not present in the

hkb CRM, but Cic binding may be facilitated by a partial sequence (i.e., TGAATGAA). Alternatively, it is possible that a role for Grh and/or Cic at the A-box is context dependent. For instance, Grh-mediated activation may be a necessary input to support *ind* expression but not for the support of *hkb*, which also receives activation input from Bicoid and Hunchback transcriptional activators and is expressed in the pre-cellularized embryo.

Other studies have suggested that Grh acts to repress transcription of *fushi tarazu (ftz), dpp*, and *tll* in the *Drosophila* embryo (Dynlacht et al., 1989; Huang et al., 1995; Liaw et al., 1995), but our study is the first to identify a role for Grh-mediated gene activation in the early embryo, in support of dorsoventral patterning. Previous studies had shown that Grh can function as an activator at later embryonic stages (Bray and Kafatos, 1991; Dynlacht et al., 1989). One analysis identified Grh (also called NTF-1 or Efl-1) biochemically using an element from the *dpp* early embryonic CRM, however the *dpp* expression domain was unchanged in the *grh* mutants (Huang et al., 1995).

Another recent study also showed Grh binds to sites that are similar to Zelda binding sites (Harrison et al., 2010). Zelda and Grh each showed stronger affinity for different variations of the shared consensus sequence, but in vitro studies showed they also competed for binding. Harrison et al. proposed that increasing levels of Zelda are able to compete against Grh for binding sites and cause activation of the first zygotic genes. Competition at the same binding sites results in a cascading effect in which ubiquitous activators regulate genes in a temporally related manner. They proposed Grh functions first to silence gene expression; while, alternatively, our data is more consistent with a model in which Grh mediated activation follows that of Zelda. *ind* is considered a

"late" response gene as it appears at mid stage 5 (nc 14), at the onset of cellularization, whereas Zelda was shown to support gene expression earlier at nc 10 (Liang et al., 2008).

It is possible that Grh competes for binding to a variety of sites (not only those recognized by Zelda), and that this competition influences gene activation/repression. At the A-box sequence, Cic and Grh may compete to help establish a sharp boundary; unfortunately, the Cic binding to the A-box sequence demonstrated previously *in vitro* was quite weak (Ajuria et al., 2011), so this competition is best examined *in vivo* in future studies.

Tiers of Repression are Likely a Common Mechanism to Ensure Robust Patterning

This study found there is yet another tier of repression activity that is independent of the A-box mediated repression. Analysis of the *eve.stripe3/7-ind-mutant-A-box* reporter construct revealed that, while dorsal-lateral repression was lost, there was still repression in the dorsal-most part of the embryo. This led us to reason that other binding sites in the *ind* CRM, independent of the A-box binding site, mediate repression. Previous research showed ectopic TGF-β/Dpp signaling can repress *ind* expression, and therefore we hypothesized the repression activity we observed in dorsal-most regions of the embryo may be regulated by Dpp signaling. Our results suggested that the Dpp-dependent repression supports repression in the dorsal most part of the embryo and not in dorsal-lateral regions of the embryo. We would not expect to see an expansion of the *ind* domain in the mutants affecting only this dorsal-most repressor, thus we mutated the SMM site in the context of two mutant A-boxes and found that the expression pattern was expanded into dorsal regions of the embryo. However, when we mutated the A-box

sites, we observed expansion of *ind* more dorsally into dorsal-lateral regions but expression was absent in dorsal-most regions. It is possible the embryo can tolerate a slight expansion of *ind* into dorsal lateral regions of the embryo but expansion of *ind* into the non-neurogenic ectoderm is detrimental. Thus, two tiers of repression have developed to insure that expression of *ind* is limited to the neurogenic ectoderm. We suggest that partially redundant repressor mechanisms are more common than appreciated, because in contrast to activation it is difficult to track repression activity.

Chromatin Factors May Play a Role in Regulating ind via the A-box

Epigenetic changes to DNA and chromatin remodeling have been shown to be vital in repression and activation of genes that define structures in late stages of *Drosophila* development. For example, Polycomb group genes silence the homeotic genes of the Bithorax complex, which control differentiation of the abdominal segments (Lanzuolo et al., 2007). To date, little is known regarding how/if chromatin factors play a role in early development of *Drosophila* embryos. Here we presented evidence that several chromatin-related factors bound an A-box affinity column but did not bind a column containing the mutant A-box element (figure 2.S2). Although several of these factors did not bind to the A-box element alone when tested by EMSA, it is possible that they bind indirectly via a larger complex. One of these factors Psq has been implicated in both silencing and activation via the Polycomb/Trithorax response elements (Hodgson et al., 2001; Huang et al., 2002). Independently, Psq was recently found to positively regulate the Torso/RTK signaling pathway in the germline, while being epistatic to *cic* a negative regulator of the Torso signaling (Grillo et al., 2011). It is possible that some of

these factors play a role in regulating *ind* via the A-box element, which would suggest a role for chromatin remodeling early in development—an avenue which is worth pursuing in future studies.

Materials and Methods

Fly Stocks and Mutant Analysis

Drosophila melanogaster flies of the background *yw* were used as wild-type. Transgenic reporters were created by P-element-mediated transformation using standard methods (A-box.*eve.stripe3/7*.A-box) and site-directed transformation into the 86FB strain (all other transgenic lines) *FRT 42D* grh^{IM} and *FRT 42D GFP* fly stocks were used for creating germline clones (Lee, 2004). The grh^{B37} allele was also used (Bray and Kafatos, 1991) and recombined with *FRT 42D* in order to facilitate generation of germline clones. *Df(2R)Pcl7B/Cyoftzlacz* is a deficiency mutant that removes the *grh* locus, and was used to eliminate the possibility that a second-site mutation within the grh^{IM} background was responsible for loss of *ind*. *FRT 42D* grh^{IM}*/Cyoftzlacz; A-box-eve.stripe3/7-A-box* flies were used in the A-box repression assay (figures 2.6 F and H, respectively). The CyO ftzlacZ marked balancer was used to distinguish zygotic genotype in crosses; however we found that the frequency of ftzlacZ+ embryos was very low in the embryos devoid of maternal *grh* therefore assay of zygotic genotype was inconclusive. It is possible that *grh* may be required to support *ftz* expression (M.G. and A.S., unpublished observatins), and other studies have identified a later role for *grh* in supporting *ftz* expression (Dynlacht et al., 1989). The zygotic genotype may relate to the variability observed in the *ind* expression phenotype.

Toll^{rm9}/TM3Ser and *Toll^{rm10}/TM3Sb* fly stocks were used to generate transheterozygous *Toll^{RM9}/Toll^{RM10}* females, and *pipe386/TM3Sb* and *pipe664/TM3Sb* fly stocks were used generate transheterozygous *pipe386/pipe664* females, as previously decribed (Stathopoulos et al., 2002). Homozygous *cic1/cic1* females were obtained from a *cic1/TM3SbSer* stock (Jimenez et al., 2000). Virgin females were obtained from each of these crosssed and mated to males containing the A-box repression reporter (*A-box-eve.stripe3/7-A-box-ep-lacZ)* (this work, see below). *brk^{M68}sog^{Y506}* (Bray and Kafatos, 1991; Jazwinska et al., 1999) mutants were used to create *brk^{M68}sog^{YS06}/FM7ftzlacZ*; *A-box-eve.stripe3/7-A-box* and *Egfr^{f2}* (Nusslein-Volhard et al., 1984) mutants were used to create *Egfr^{f2}/CyoftzlacZ; A-box-eve.stripe3/7-A-box* fly stocks, which were used in the A-box repression assay (figure 2.6).

Plasmid Construction

The *A-box-eve.stripe3/7-A-box* reporter was created by PCR using the following primers:

MG 1 (5'gtgcggccgcAGCGC**ATTCATTCATGA**GGCCAggacacaaggatcctcgaaatcgaga-3') and

MG 2 (5'gtgcggccgcACACT**TCATGAATGAAT**ACATCgaaggaacgagctcgtaaaaacgtgaa-3') and was cloned into pCasper using the Not I site. The chimeric CRM were created by cloning the modified *ind* CRM into a pGemT-easy vector containing the *eve.stripe3/7* CRM using the Spe1 site. The *eve.stripe3/7* CRM (Small et al., 1996) was PCR amplified using MG 48 (ggacacaaggatcctcgaaat) and MG 49 (gaaggaacgagctcgtaaa). A fragment

containing both CRMs in tandem was subsequently cloned into the pLacZattB vector using the Not 1 site.

The mutant CRMs were created by PCR site directed mutagenesis using the following primers: A-box1: MG 87 (caggcagtgcagcgcattattaattaggccaattc) and MG 88 (gaattggcctaattaa-ttaatgcgctgcactgcctg);

A-box2: MG 99 (ctgaagaggttctgcacttcaggatgtattaattaattaagtgtcttccacgcg) MG 100 cgcgtggaagacacttaattaattaatacatcctgaagtgcagaacctcttcag);

Dorsal: MG 106 (caggccca-aagaacctgacccaatttcccagccttgatg) and MG 107 (gtccgggtttcttggactgggttaaagggtcggaactac).

SMM: MG 234 (ggacttatatgcccttgggacagaacgtctggac) and MG 235 (gtccagacgttctgtcccaagggcatataagtcc).

In Situ Hybridization

Embryos were collected, fixed, and subsequently hybridized with dioxygenein-UTP, biotin-UTP or fluorescein-UTP labeled antisense probes as previously described (Jiang et al., 1991; Kosman et al., 2004). Probes were made by PCR from genomic DNA extracted from *yw* male flies. Images were collected using bright field or confocal microscopy.

Preparation of Nuclear Extracts

Nuclear extracts were prepared using 45 grams of 0–6 hour embryos using a modified version of the protocol described in (Zandi et al., 1997). Frozen embryos were ground in liquid nitrogen using a mortar and pestle. The ground embryos were

resuspended in 200 mL of buffer containing 25 mM Hepes pH 7.6, 10 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM EDTA, and 1X Roche proteinase inhibitor. The solution was homogenized using a dounce homoginizer, and subsequently was centrifuged at 10,000 g for 10 minutes. The supernatant was removed and the pellet was resuspended in 150 mL of buffer containing 25 mM Hepes 7.6, 10 mM KCl, 1.5 mM $MgCl_2$, 0.1 mM EDTA, 25% glycerol and 1X Roche Proteinase inhibitor. 15 ml of 5M NaCl was added. The solution was mixed for 20 minutes at 4°C. The solution was centrifuged at 15,000 g for 20 minutes. The resulting supernatant was the nuclear extract.

Affinity Chromatography and Mass Spectrometry

Dialyzed extracts were partially purified by eluting off a heparin column using 0.3 M–1.5 M KCl. Fractions from the heparin column were assayed for A-box binding activity using EMSA. The fractions with activity (i.e., 0.9 M–1.1 M KCl) were combined and dialyzed. Half of the sample was run on an A-box affinity column (gatctgt**attcattcatga**agtgtcttc) and half was run on a mutant A-box affinity column (gatcgcagcgc**attaattaatta**ggc). Columns were prepared and run according to previously described protocols (Kerrigan and Kadonaga, 1998). The fractions were tested for activity using EMSA and binding proteins were identified using GelC/mass spectrometry. Standard in gel trypsin digest with reduction and alkylation was used to process samples for mass spectrometry. A Thermo Finnigan Orbitrap was used for mass spectrometry of samples. The Scaffold program was used to identify targets. Positives were differentiated from false positives by comparing the A-box column list to the mutant A-box column list. The list was also cross-referenced to a list of all previous characterized

transcription factor or factors containing a predicted DNA-binding domain. The list of putative transcription factor was obtained from FlyTF.org (Adryan, 2009).

Elecrophoretic Mobility Shift Assay

The following oligos were used for the EMSA: A-box (gatctgtattcattcatgaagtgtcttc) and mutant A-box (gatctgtattaattaattaagtgtcttc), and standard labeling methods with γ^{32}P-ATP were used. The following buffer and conditions were used for tracking the activity during affinity chromatography: 10 mM Tris pH 7.5, 5% glycerol, 15 M sucrose, 2 mM EDTA, 50 mM DTT, 200 mM KCl, 1% nonidet P-40, 5 μg/ul BSA, 0.3 μg/ul polydIdC 1X Roche complete protease inhibitor, 100 fmol of labeled oligo, and 1 μL of extract in a 25 μL reaction. For the testing of candidate genes 25 mM Hepes pH 7.9, 100 mM KCL, 1 mM DTT, 1% polyvinyl alcohol, 1% nonidet P-40, 0.1% BSA, 10 % glycerol, 0.25 μM calf-thymus DNA, 50 fmol of labeled oligo and 1 μL of reticulocyte in vitro translated protein was added to a reaction of 15 μL total volume. Proteins were prepared using the TNT T7 Quick Coupled Transcription/Translation System from Invitrogen (Grand Island, NY 14072 USA). The reactions were incubated for 30 minutes on ice and then resolved on either 6% or 4% native polyacrylamide gels containing 0.5X TBE.

Chapter 3

Identification of Chromatin Factors that Bind the *ind* Cis-Regulatory Module

Abstract

Analysis of cis-regulatory modules (CRMs) is indispensable when asking questions about factors that act directly to regulate gene transcription. Once CRMs are analyzed and the important regulatory binding sites are identified, it is often desirable to determine the factors that bind to these elements. Here we give a brief overview of two methods used to identify factors that bind to an element of interest: yeast one hybrid and affinity chromatography. We outline in detail our procedure used to identify factors that bind the A-box element of the *intermediate neuroblast defective (ind)* CRM. Lastly, we discuss some of the factors we identified by affinity chromatography.

Introduction

An important question in biology is to understand how genes are regulated. Proteins can affect the expression of a gene either indirectly or directly. Indirect methods for gene regulation include signaling pathway components upstream of a direct effecter and activation or repression of factors that act directly. Direct methods include the binding of repressors and activators to the cis-regulatory modules (CRMs) that control gene expression. These factors can act locally near the promoter or they can act from distances as much as 10 kb away and be located distally to the promoter. Not only do activators and repressors bind to CRMs, factors termed chromatin-remodeling factors can also bind and affect gene transcription by altering the chromatin state.

Detailed analysis of CRMs done by dissecting them into smaller workable fragments and mutating binding sites has proven to be a powerful method for deciphering which components of the CRM are important, and thus, how gene expression is regulated. Sometimes the sequences that are found to be important have known binding factors and mutant analysis of these factors can provide information about how a pattern is specified. Often, the DNA sequences that are identified have no known binding factors and it is necessary to identify the respective DNA-binding proteins.

One of the methods employed to identify binding factors for a given sequence is yeast one-hybrid. In this method, the binding sequence of interest is used to drive the expression of a reporter gene that will allow the yeast to grow on selective media. A library of proteins fused to a yeast activation domain is then transfected into the yeast. One library protein-yeast activation hybrid will enter each cell. If the library protein binds to the sequence of interest, then the yeast activation domain will activate the

reporter gene and the yeast will grow on the minimal media. The library protein can then be identified and further analysis can take place. Recently, a complete *Drosophila* open reading frame transcription factor library was created, this combined with automated plating and tracking, has enabled high-throughput yeast one-hybrid screens (Hens et al., 2011; Ozdemir and Stathopoulos, 2011). One limitation to this method is that it cannot identify proteins that bind as a complex or heterodimer.

Another method used is affinity chromatography. Affinity chromatography can be used to purify a protein out of a nuclear extract using a binding sequence of interest to isolate proteins that bind to it. In this method, the binding element is attached to a resin by a chemical reaction. Nuclear extracts are then added to the immobilized resin. Factors that bind will be immobilized and can be eluted with a high salt buffer. This method is not limited to factors that bind alone, complexes and heterodimers can also be identified. Another strength of this method is that nuclear extracts from select stages of development can be generated to isolate factors from a desired time point. In the past it was necessary to purify protein extracts to the point where a single protein was isolated. This was done by repeatedly running the nuclear extracts through the affinity column. The isolated protein was then sequenced and identified. Today, mass spectrometry can be used to identify proteins found in a mix of proteins, thus purification to a single protein is not necessary. Once the proteins are identified they can be tested individually via electrophoretic mobility shift assay (EMSA) to determine whether they bind to the sequences of interest and mutant analysis can be conducted.

Procedure

Here we used affinity chromatography to purify factors that bind to the A-box sequence (ATTCATTCATGA), which is important in establishing delineated expression of the gene *intermediate neuroblast defective (ind),* which specifies the intermediate neurogenic ectoderm in *Drosophila.* Two A-boxes, which are roughly 100 base pairs apart, are found in the *ind* CRM. When these sites are mutated the expression pattern is expanded dorsally. The *ind* gene is expressed at around 2–3 hours of development. In order to capture proteins within this time window we created nuclear extracts from embryos that were 0–6 hours old. In order to isolate cationic proteins, such as transcription factors, the nuclear extracts were fractionated using a heparin column. The factors that bound were eluted off using a step gradient of KCl from 0.3 M to 1.5 M (figure 3.1). Three 1ml fractions were collected for each step; these fractions were then tested for binding activity by EMSA. The fractions that eluted off of the column with 0.9 M KCl and 1.0 M KCl were found to have activity that was specific to an A-box oligo (gatctgt**attcattcatga**agtgtcttc) and could not be competed with an oligo specific for Dorsal (gatcgtgcg**gggaaatccc**cgtaat) protein or a mutated version of mutant A-box 2 (gatcgcagcgc**attaattaatta**ggc) but was competed with A-box 1 (gatctgt**attcattcatga**agtgtcttc) and mutant A-box 1 (gatctgt**attaattaatta**agtgtcttc) oligo (figure 3.1). The specific activity in the fractions was then challenged by increasing amounts of competitor DNA. We tested several competitor DNAs to find one that would reduce nonspecific binding while maintaining specific binding. We found that while all of the competitor DNAs tested maintained the A-box specific binding calf-thymus DNA was the most effective at reducing the non-specific binding (figure 3.2).

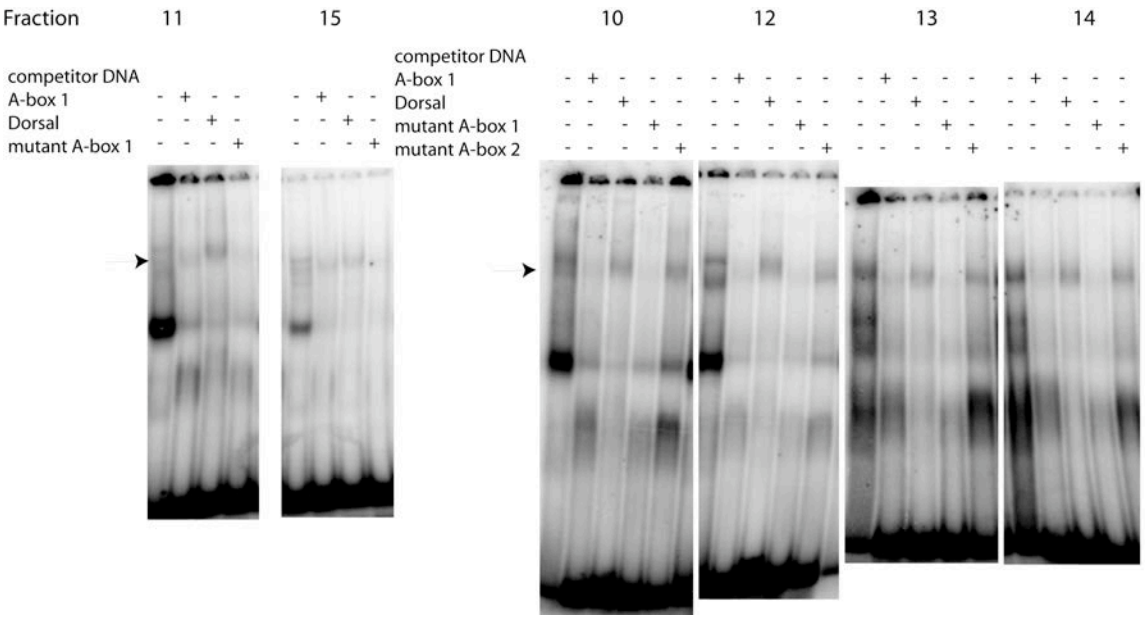


Figure 3.1. Heparin fractionated nuclear extracts show specific binding to an A-box 1 oligonucleotide.

The EMSAs were preformed using γ^{32}P-labeled A-box 1 oligonucleotides on nuclear extract fractions that were fractionated using a heparin column. The presence of competitor DNA is indicated by a +. The arrows indicate A-box specific binding. For fractions 11 and 15 mutant A-box 2 competition was not analyzed.

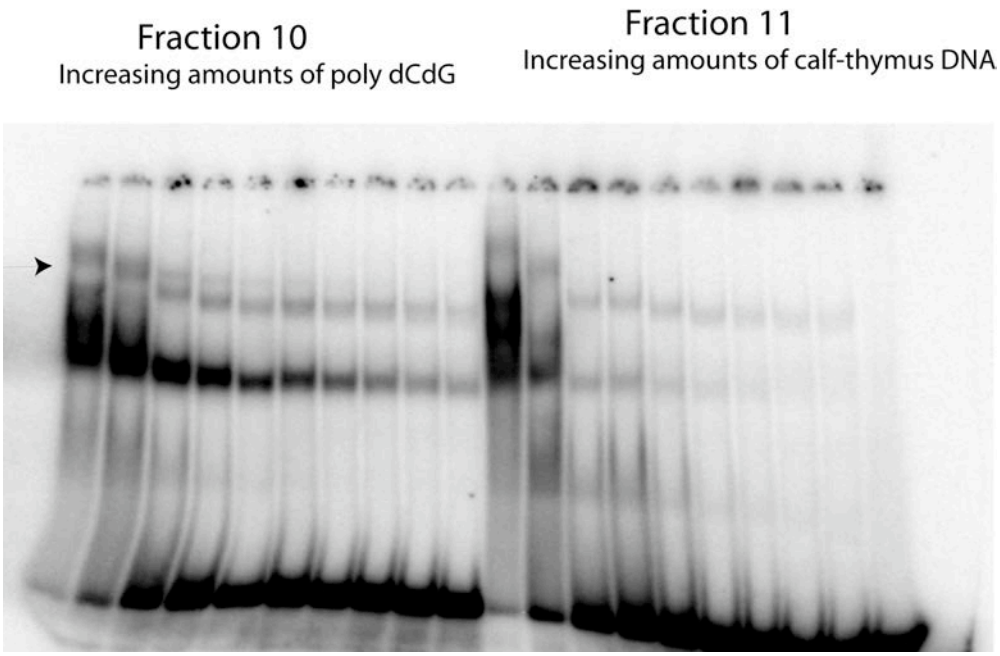


Figure 3.2. Calf-thymus competitor DNA is the most effective at reducing non-specific binding while maintaining A-box specific binding.

The EMSAs were preformed using γ^{32}P-labeled A-box 1 oligonucleotides on nuclear extracts that were fractionated using a heparin column. The arrow indicates the region where the A-box specific activity is located. As an example two competitor DNAs are shown: poly dCdG and calf-thymus DNA. As the competitor DNA is increased the band shifts to a lower position this might be caused by loss of a dimer or a complex component.

Affinity columns were constructed to isolate factors that bound to the A-box element. As a control, a mutant A-box 2 column was constructed so a comparison between the binding factors could be made and A-box specific factors could more easily be identified. Mutant A-box 2 was chosen as a comparison because it was unable to compete the specific binding found to bind to A-box 1 when tested by EMSA. The

fractions that showed specific A-box binding were dialyzed against column buffer and calf-thymus DNA was added. Calculations based on the competitor DNA titrations were used to calculate the amount of calf-thymus DNA that would reduce background while allowing specific binding. The calculation and much of the protocol was based on protocols located in (Kerrigan and Kadonaga, 1998). The fractions were then run on either the A-box column or the mutant A-box column. The factors were eluted using KCl step gradient and fractions were collected in 1ml intervals. EMSA using an A-box 1 oligo was conducted to test the fractions for binding (see chapter 2 figure 2.4 A for details). Binding to A-box oligo was compared and mass spectrometry was used to analyze the fractions that displayed A-box specific binding. The scaffold program was used to identify proteins within the samples.

Results and Discussion

A higher percentage of chromatin factors and transcription factors bound to the A-box column compared to the mutant A-box column (appendix A: chapter 2 supplemental figure 2.S2). It is possible that these factors bind as a complex to affect the chromatin structure of the *ind* CRM and confer a repressive state in dorsal regions of the embryo. We tested several of the factors individually for binding using proteins created using rabbit reticulolysates and found that the only one that bound alone was Grh. Thus, we favor a model where the other factors are only capable of binding in a complex or are recruited to the site by other factors, which were present during the affinity chromatography assay but not EMSA.

Several of the factors that bound to the A-box column specifically are associated with formation and maintenance of the heterochromatin state. Suppressor of variegated [Su(var)3-9] is a known histone methyltransferase which is H3-K9 specific. H3-K9 methylation is found in heterochromatin and is associated with gene silencing. Another gene found on the column, caf1-180 is involved in formation of heterochromatin possibly through interactions with Heterochromatin protein 1 (HP1) (Huang et al., 2010). HP1 is responsible for pericentric heterochromatin formation. Recent studies show a reduction in H3-K9 methylation and a slight reduction of Su(var)3-9 in *caf1-180* mutants although a direct interaction between Caf1-180 and Su(var)3-9 was not found (Huang et al., 2010). However, it was suggested that reduced levels of HP1 in *caf1-180* mutants likely contributes to the reduced levels of Su(var)3-9. Independently to our affinity chromatography study, we conducted a yeast one-hybrid screen in which Su(var)3-7 was found to bind to a 110 bp element containing the A-box sequence (supplemental table 1 for full list of factors). Su(var)3-7 is also involved in heterochromation formation via HP1 and Su(var)3-9 (Delattre et al., 2004; Delattre et al., 2000).

Another group of factors that modify chromatin state are the Polycomb group (PcG) proteins. They confer and maintain silencing memory to keep gene transcription of select genes off during many cell generations. They act through polycomb response elements (PREs), which were discovered during studies of the Bithorax complex (BX-C) along with Trithorax response elements (TRE) that act antagonistically to confer an active chromatin state. PcG and Trithorax group (TrxG) proteins are ubiquitously expressed and it is thought that they act after initiator transcription factors have provided positional information to establish patterns and have degraded (Ringrose and Paro, 2007).

Historically PcG and TrxG proteins have been found to work at later stages of development mainly acting on body segmentation genes, but they are present and have been found to bind DNA much earlier (Orlando et al., 1998). In one study, Trx did not bind to Ubx chromatin collected from 0-2 hour embryos while PC did. Trx did bind to chromatin collected from 2-4 hour embryos and onwards. Likewise, PC binds at these stages but to different regions than Trx (Orlando et al., 1998). This may suggest that PcG silencing is actually functioning at the start of development and the TrxG complex switches chromatin to an active state. There may be several points throughout development where these switches occur.

Many of the studies on PcG complexes have been conducted by analyzing the cis-regulatory elements they bind to. PREs posses interesting properties that make them unique compared to other cis-regulatory elements. When placed proximal to the *mini-white* gene in transgenic *white-* flies they display variegation and pairing-sensitive silencing (Ringrose and Paro, 2004) (figure 3.3). *white-* flies have white eyes because they do not produce eye pigment; addition of the transgne *mini-white* restores the wild-type red eye color. Variegation causes the *mini-white* reporter gene to be silenced in some cell lineages and not others causing a phenotype that displays patches of red and white eye color (figure 3.3 C). Pairing-sensitive silencing is a phenomenon in which homozygous flies for the *mini-white* reporter gene display an eye color that is lighter than that found in heterozygous flies (figure 3.3 B). It is thought that paired PREs can cause repression that is stronger by recruiting a higher-order PcG complex, resulting in repression of the *mini-white* gene and a lighter eye color. PREs have also been found to have high levels of H3K27 methylation and significant but lower amounts of H3K9

methylation. The exact role of this methylation is still unknown although it has been shown that the chromodomain of polycomb preferentially binds to H3K27 methylated tails (Ringrose and Paro, 2004).

PREs contain many defining binding sites including Pleiohomeotic (Pho/Pho1), GAGA factor (Gaf), Pipsqueak (psq), and Zeste (Ringrose and Paro, 2007). PREs overlap with TREs and share defining binding sites, Pho and Psq are PcG proteins and, thus, silence gene expression, while Gaf and Zeste are TrxG proteins and activate gene expression. Several binding sites have also been found to be essential for the function of specific PREs but are absent in most others, these include binding sites for Dorsal switch protein 1 (Dsp1), Grainyhead (Grh) and Sp1/KLF (Blastyak et al., 2006; Brown et al., 2005; Dejardin et al., 2005). Of these proteins Grh and Sp1/KLF are expressed in a temporal and tissue specific manner (Bray and Kafatos, 1991; Brown et al., 2005), which may provide a means for regulation of PcG activity in a temporal- or tissue-dependent manner. In our affinity chromatography analysis of the A-box, we identified Psq, Dsp1, and Grh.

Psq is a transcription factor with a unique DNA binding domain, with no homolog in mammals. It co-purifies with the PcG complex and acts like a PcG protein in genetic studies. Psq and Gaf both bind to the same sequence, $(GA)_n$, it is thought that they act antagonistically to each other competing for binding to recruit either PcG or TrxG complexes. Mutation of the Psq/Gaf binding site in PREs results in loss of silencing (Ringrose and Paro, 2004).

Unlike Psq, Dsp1 is only found in a subset of PREs it was first discovered for its involvement in PcG recruitment during an analysis of the Ab-Fab PRE (Dejardin et al.,

2005). It was determined that PcG mediated repression via this element requires Dsp1 binding sites. When Dsp1 binding was disrupted by mutation of the element or in *dsp1* mutants the element failed to recruit Polycomb. The Dsp1 binding site found in the Ab-Fab PRE is different than the A-box binding site. We conducted an EMSA and determined that Dsp1 does not bind the A-box bindings site alone (data not shown). Previously, it was shown that Dsp1 homologs have weak sequence specificity in vitro (Thomas, 2001). Thus, it is possible that we could not detect Dsp1 binding, because Dsp1 binds the A-box weakly and its binding is enhanced by other factors that were present in the nuclear extract. Of the factors tested by EMSA, including Dsp1 and Psq, Grh was the only one that bound alone. It is possible that binding of Grh in the nuclear extracts facilitated binding of the other factors to the A-box column or that the other factors bound to Grh, rather than to the DNA itself.

We recently showed that Grh is involved in activation of *ind* rather than repression. Initially we thought of the A-box as a repression element because mutation of the A-box binding sites in a chimeric CRM assay lead to expansion of gene expression into dorsal regions of the embryo (chapter 2 figure 2.1 C). Corresponding to a role in activation, when the A-box sites were mutated in the context of the full-length CRM alone the expression level was reduced and patchy, while still expanded (chapter 2 figure 2.2 B). It is possible that the A-box acts as a switch between PcG and TrxG complexes to maintain either silenced or active chromatin states. In one study Grh was shown to interact with a PcG complex via the iab-7 PRE (Blastyak et al., 2006). In vitro, Grh bound to the iab-7 PRE and in *grh* mutants the function of iab-7 was affected. Constructs containing Pho, Gaf/Psq, and Grh binding sites showed homozygous pairing-sensitivity

(PS) that was affected in *grh* mutants only when the Grh binding sites were present (figure 3.3). The affect on PS was increased in *grh pho* doubles. Pho was found to interact with Grh and increased its affinity for DNA (Blastyak et al., 2006).

Further studies are necessary to determine if the A-box and its surrounding DNA are capable of recruiting PcG proteins. We have looked at zygotic mutants of *dsp1* and *psq* but did not see a phenotype for *ind*. It is possibly that they may act at a later stage or it is possible that single mutants would not show a phenotype; individual binding sites have often been found to not be essential for PREs to function (Ringrose and Paro, 2004). Analysis of these factors is complicated by the fact that many are both maternally and zygotically expressed. Thus a phenotype may not be observed in the zygotic mutants, due to the presence of maternal transcripts. We attempted to analyze many of these factors using shRNAi but many resulted in sterility or did not display a phenotype. This data is inconclusive as not all shRNAi lines are expected to successfully reduce RNAi levels; thus, lack of phenotype is not conclusive. Rather than looking at *ind* expression directly it may be possible to determine if the A-box element plays a role in PcG recruitment by analyzing whether it displays variegation and pairing-sensitivity (figure 3.3). If this is the case, we can test mutants of targets found on the column to see if there is an affect on variegation and PS.

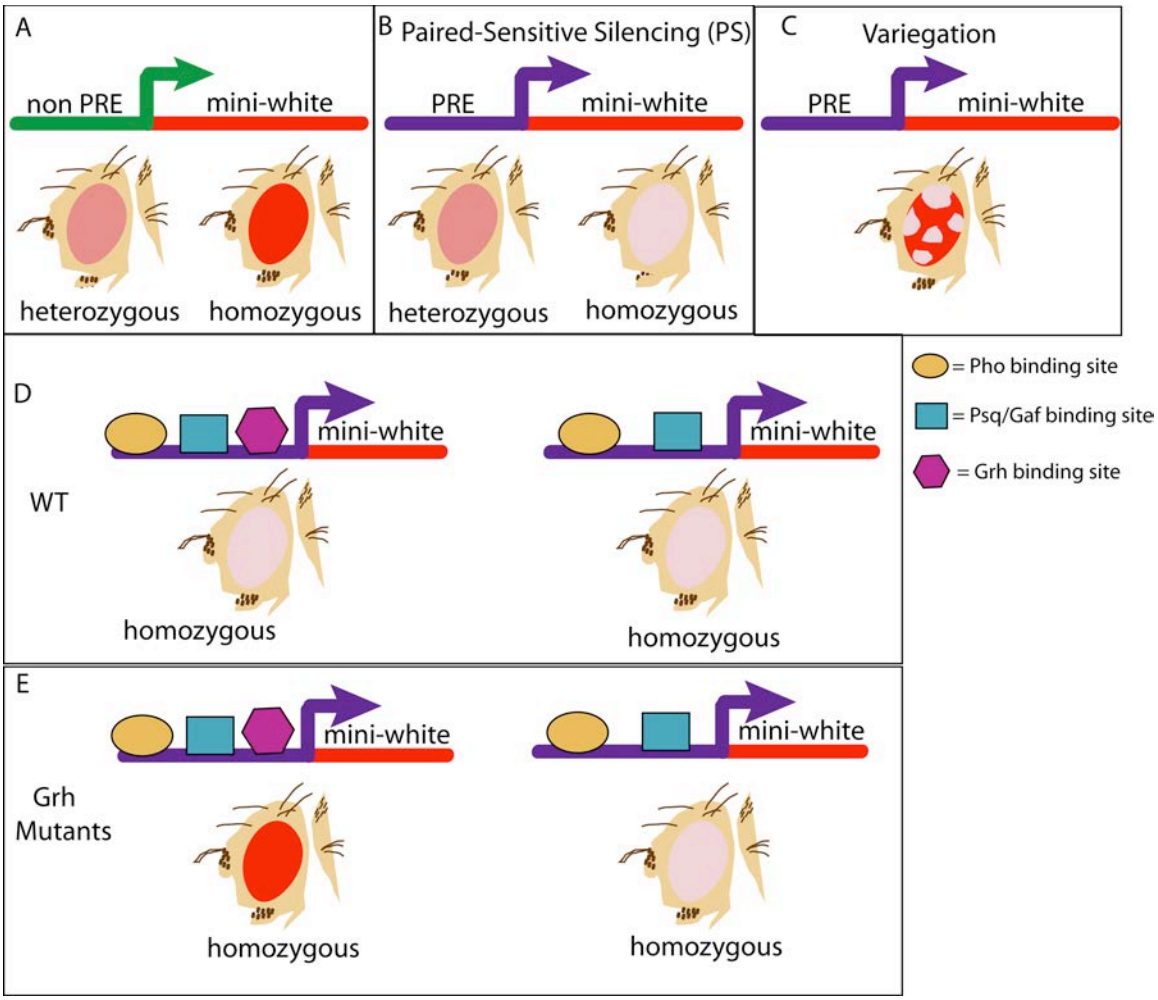


Figure 3.3. Analysis of PREs for paired-sensitive silencing and variegation.

(A) When a non-PRE DNA is placed next to a promoter that drives expression of the *mini-white* gene the resulting heterozygous flies have an eye color that is lighter than the homozygous flies. This is because two copies of the gene result in a darker eye color. **(B)** When a PRE DNA is placed next to a promoter that drives expression of the *mini-white* gene the heterozygous flies display an eye color that is darker than the homozygous flies. This is believed to occur because the paired PREs recruit a higher-order chromatin structure that represses expression of the gene.

(C) Another result that can occur when a PRE is placed next to the promoter is variegation, characterized by light and dark patches in the eye. This is caused by the PRE silencing the mini-white gene is a subset of cells resulting in a patchy eye color.

(D) In WT flies a piece of synthetically constructed DNA containing binding sites for Pho, Psq, and Grh displayed paired-sensitive silencing. Removal of the Grh binding site still resulted in paired-sensitive silencing suggesting the Grh sites where not necessary for the PRE to function.

(E) In Grh mutants the DNA containing the Grh binding site did not display paired-sensitive silencing while the DNA that did not contain the Grh binding site still displayed paired-sensitive silencing. This suggests, the loss of paired-sensitive silencing is directly linked to the presence of the Grh binding site and is not causes by an indirect effect.

Chapter 4

Repressors Play a Role in Establishing the Dorsal and Ventral Borders of Neurogenic Ectoderm Genes

Abstract

Patterning of the *Drosophila melanogaster* neurogenic ectoderm begins early in embryogenesis. The genes that specify the presumptive neurogenic ectoderm are turned on during the blastoderm stage prior to gastrulation. It is important that proper patterning of this tissue occurs since it helps establish the fate of other tissues as well. Previous, studies have shown that combinatorial interactions between transcription activators, such as Dorsal and Twist, provide some positional information to pattern the ventral-lateral neurogenic ectoderm. It is unclear if this is enough to establish the patterns or if other inputs are also necessary. Here we analyze the CRMs of *vnd* and *sog,* using a chimeric CRM assay that aides in the identification of repressive elements. We show that the *vnd* CRM mediates dorsal and ventral repression while the *sog* CRM mediates ventral repression and weak dorsal repression. This suggests that the combinatorial interactions between activators, while important, are not sufficient to insure proper patterning of the neurogenic ectoderm.

Introduction

The neurogenic ectoderm in *Drosophila melanogaster* is specified by the expression of several Dorsal target genes that are expressed in dorsal-ventral stripes across the lateral region of the embryo. These genes are activated by intermediate levels of Dorsal and include *rhomboid (rho), vein (vn), short gastrulation (sog), brinker (brk), ventral neuroblast defective (vnd),* and *intermediate neuroblast defective (ind)*.

CRMs associated with genes expressed in the presumptive ventral neurogenic-ectoderm, *vnd, sog, rho,* and, *vn;* all contain Dorsal, Twist, and Snail binding sites. Dorsal and Twist are activators and Snail is a repressor that keeps expression off in the mesoderm. Even though genes in this domain are essentially regulated by the same factors, they display different responses to these factors and display expression patterns with slight differences. *vnd* expression extends ventrally into regions where *snail* is weakly expressed but both *rho* and *vn* are excluded from this region. While, the dorsal borders of *rho* and *vn* extend two cells more dorsally compared to *vnd*. It is believed that these differences are mediated by combinatorial interactions between Dorsal, Twist, and Snail (Zinzen et al., 2006). The *vnd* CRM contains more Dorsal-Twist-Snail clusters compared to *rho* and *vn* CRMs. Computational modeling shows that the ventral border of *vnd* requires high Snail-Snail cooperatively, and that the dorsal border is likely shifted ventrally due to Twist-Twist cooperatively (Zinzen et al., 2006).

Another possible scenario for definition of the dorsal border of *vnd* is that repressors act on *vnd* to refine the border that is first established by Twist-Twist cooperatively. In fact, evidence that *ind* represses *vnd* in stage 8 embryos has been shown (Von Ohlen and Moses, 2009). Although, such an interaction is not likely to be

important for the initial *vnd* pattern as *ind* is turned on after the *vnd* pattern has been established. BMP signaling, which activates genes in the dorsal most part of the embryo is also a candidate for repression of genes expressed in the neurogenic ectoderm. Many of the CRMs of neurogenic gene contain Schnurri-Mad-Medea Complex (SMM) binding sites, which act downstream of Dpp signaling (Stathopoulos and Levine, 2005). We recently showed that one of these sites acts redundantly with the A-box site to repress the dorsal border of *ind* (chapter 2). Conflicting evidence exists to suggest that Dpp signaling plays a role in the regulation of *vnd*. In one study, ectopic expression of Dpp using a line with 4X Dpp in a *sog* mutant background did not show loss of *vnd* expression at stage 8, the expression pattern was actually slightly expanded dorsally presumably due to a loss of *ind* (Von Ohlen and Doe, 2000). In a different study, *uas-dpp* was ectopically expressed using *kr-gal4*, in this case *vnd* was repressed also at stage 8 (Oh et al., 2002). Thus, it is unclear how the dorsal border of *vnd* is established.

sog is also expressed in lateral regions but it is expressed in a broad lateral region that extends the entire length of the neurogenic ectoderm. Unlike *vnd*, which specifies a distinct set of ventral neuroblasts requiring a restricted expression pattern (Skeath et al., 1994), *sog* acts to inhibit Dpp signaling, such a function requires a broader and graded expression pattern (Biehs et al., 1996; Yu et al., 1996). One of the *sog* CRMs was recently analyzed and the orthologous sequences compared across *Drosophila* species to determine the regulatory logic used to express broad patterns. It was determined that inputs form both Dorsal and the ubiquitous activator Zelda were necessary for the broad lateral pattern; this is in contrast to ventral lateral patterns that require inputs from Dorsal and Twist, although the *sog* CRM does contain Twist binding sites (Liberman and

Stathopoulos, 2009). The dorsal border of *sog* is not as sharp as the border of other genes expressed along the dorsal-ventral axis, nonetheless, it is still possible that repressors act on the dorsal border of *sog* in a graded fashion.

Here we analyze the CRMs of *vnd* and *sog* to determine whether there are repressors acting to establish the dorsal and ventral borders of their expression patterns. We observed strong dorsal and ventral repression in the *vnd* CRM and weak dorsal repression and strong ventral repression in the *sog* CRM. This suggests that dorsal repressors play a larger role in definition of the sharp *vnd* border and a smaller role in specification of the graded dorsal border of *sog.* We also found that the SMM binding site is not necessary for repression in the *vnd* CRM thus another unknown site is responsible for the repression observed and likely acts in parallel to SMM repression.

Results

The vnd CRM is Delineated by Repression in Dorsal and Ventral Regions

In order to gain insights into how *vnd* is regulated we analyzed a 744 bp CRM located within an intron of *vnd,* that recapitulates its expression pattern *(Stathopoulos et al., 2002)*. Using an approach similar to that used to analyze the *ind* CRM we created chimeric CRMs where the *eve.stripe3/7* CRM was fused to the *vnd* CRM (Stathopoulos and Levine, 2005). We found that repression elements located in the *vnd* CRM were capable of repressing *eve.stripe3/7* in both dorsal and ventral regions (figure 4.1 A). Using Gene Palette we identified binding sites for known transcription factors including Snail, Dorsal, Schnurri-Mad-Medea (SMM), ETS, and Su(H). The repression we observe

in ventral regions is most likely mediated by Snail, as Snail is known to repress *vnd* (Cowden and Levine, 2003).

We segmented the *vnd* CRM into four approximately 200 bp modules and created chimeric *eve.stripe3/7-vnd-moudle* constructs used to drive expression of the *lacZ* reporter gene. The bulk of the known binding sites were found in the first 200 bp of the CRM, although this module alone cannot recapitulate the *vnd* expression pattern (figure 4.1 A and B). Dorsal and ventral repression was only observed in module A (figure 4.1 B). Modules B and C both contain Snail binding sites but there was no repression of *eve.stripe3/7* in these modules (figure 4.1 C and D). While module A contains two Snail binding sites the ventral repression observed in this module was significantly weaker than that observed when the full-length CRM was used (figure 4.1 compare A and B). This suggest that all of the Snail binding sites are needed for proper specification of the ventral border, alternatively, unknown binding sites may be present in the other fragments besides module A that specify the ventral border.

Unlike the ventral border, there are no known factors that specify the dorsal border of *vnd*. SMM sites are present in the *vnd* CRM, but when we mutated them in the *eve.stripe3/7-vnd* chimeric CRM we did not observe a change in the expression pattern (figure 4.2 D). This is similar to what we observed with *ind* (figure 4.2 B). Recently we found that the SMM site is functional in the *ind* CRM but its effects are only observed when the A-box site is also mutated (chapter 2; figure 2.2 B). It is possible that the SMM site is functional in the *vnd* CRM but another repressor acting on the dorsal border also works in parallel.

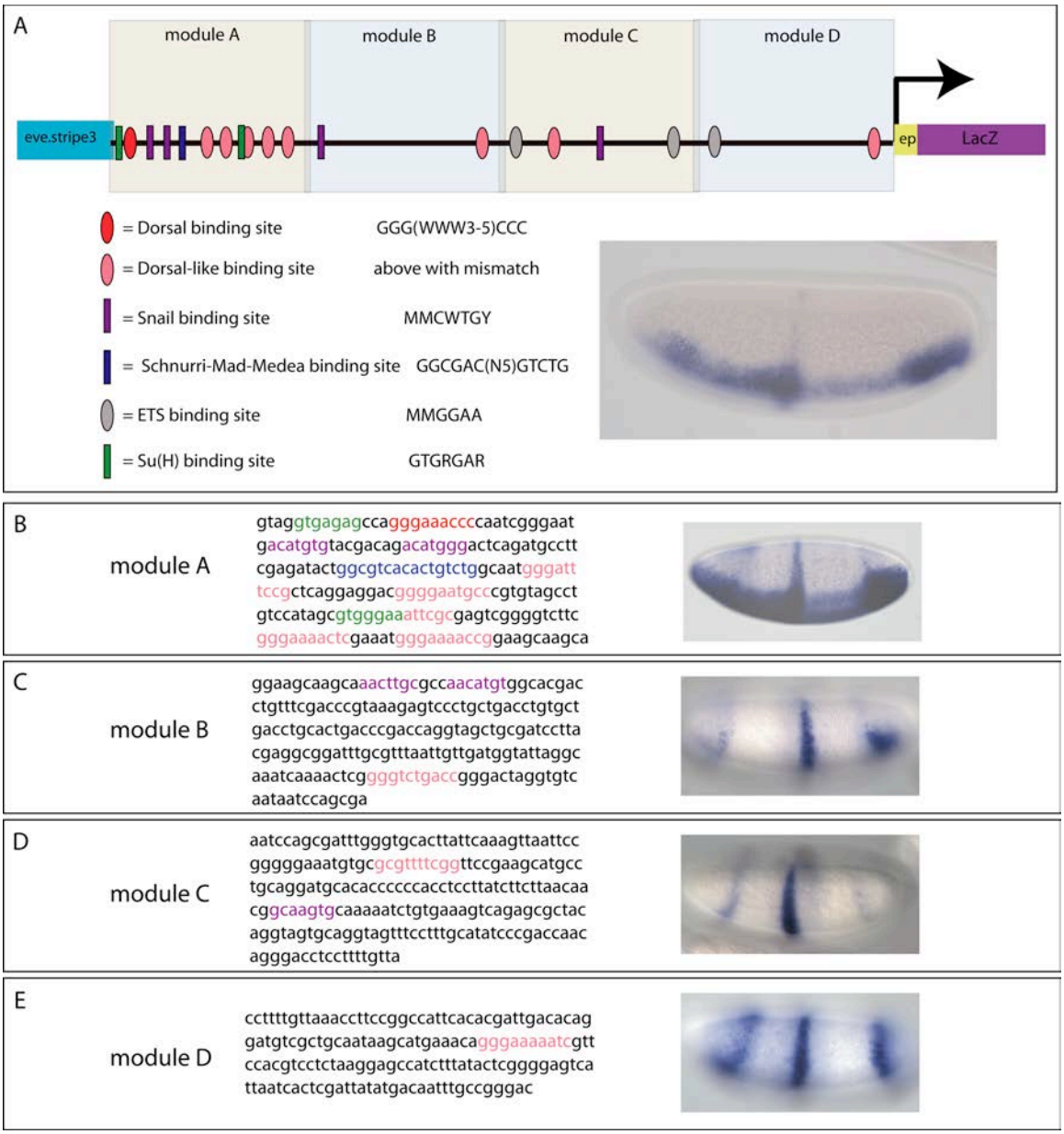


Figure 4.1. The *vnd* CRM contains binding sites that mediate ventral and dorsal repression.

We conducted a chimeric CRM assay in which the *eve.stripe3/7* CRM was placed proximal to the *vnd* CRM, this chimeric CRM along with the *eve* promoter was used to drive expression of a *lacZ* reporter gene. The schematic depicts the chimeric CRM construct along with the binding sites that are present in the *vnd* CRM. The embryo in

part A was stained by in situ hybridization using an RNA riboprobe for the reporter gene *lacZ.* The full length CRM mediates dorsal and ventral repression (A). The *vnd* CRM was divided into four modules. Module A shows weak dorsal and ventral repression (B). Modules B, C, and D show no repression (C–E).

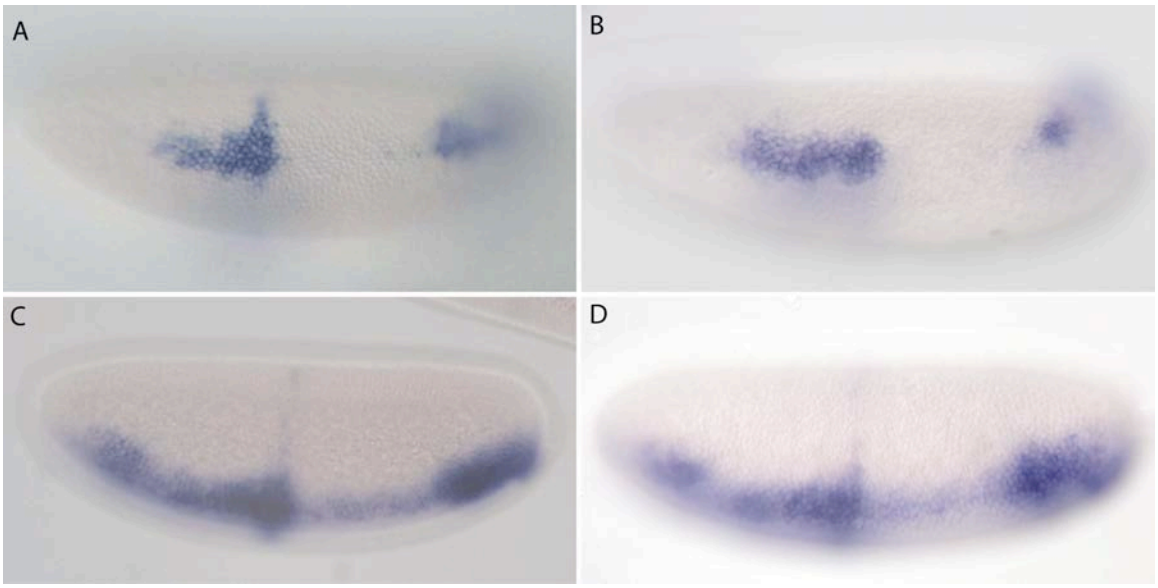


Figure 4.2. Mutation of SMM binding sites in *ind* and *vnd* CRMs does not have an affect on dorsal repression.

We analyzed the function of the SMM binding sites in the *ind* and *vnd* CRMs by mutating these sites in the context of the chimeric CRMs where the *eve.stripe3/7* CRM was fused to either the *ind* or *vnd* CRMs. The *ind* CRM fused to the *eve.stripe3/7* CRM shows repression in dorsal regions (A). When the SMM binding site is mutated in the *ind CRM* repression is still observed (B). The *vnd* CRM fused to the *eve.stripe3/7* CRM shows dorsal repression (C). Mutation of the SMM binding site does not affect repression (D).

Similar to ventral repression, dorsal repression was only mediated by module A of the *vnd* CRM (figure 4.1 B). Likewise, the repression we observed with this module was significantly weaker than that observed with the full-length CRM suggesting important binding sites are located elsewhere in the CRM. The repression mediated by this module appears as a tapering of the *eve.stripe3/7* stripe and not the complete loss of expression we observed with the full-length CRM (compare figure 4.1 A and B). To determine which part of module A was important for dorsal repression we divided module A into two approximately 100 bp fragments called module A1 and module A2 and analyzed them using the chimeric CRM assay. Module A1 contains a Su(H) site , a Dorsal site , two Snail sites, and a SMM site. Module A2 contains four Dorsal-like sites and a Su(H) site that overlaps with one of the Dorsal-like sites. Repression was mediated by module A2 and not module A1 (figure 4.3). This was surprising because we expected the SMM site to contribute to the repression. It is possible that a repressor binds to the Dorsal-like sites possibly competitively against Dorsal. Another possibility is that the Su(H) site mediates repression. The Su(H) site is found in most of the CRMs of genes expressed along the dorsal-ventral axis. It has been implicated in repression of the *snail* CRM (A. Ozdemir and A. Stathopoulos, unpublished observation). It is possible that Su(H) acts in all dorsal-ventral genes to mediate or assist in repression.

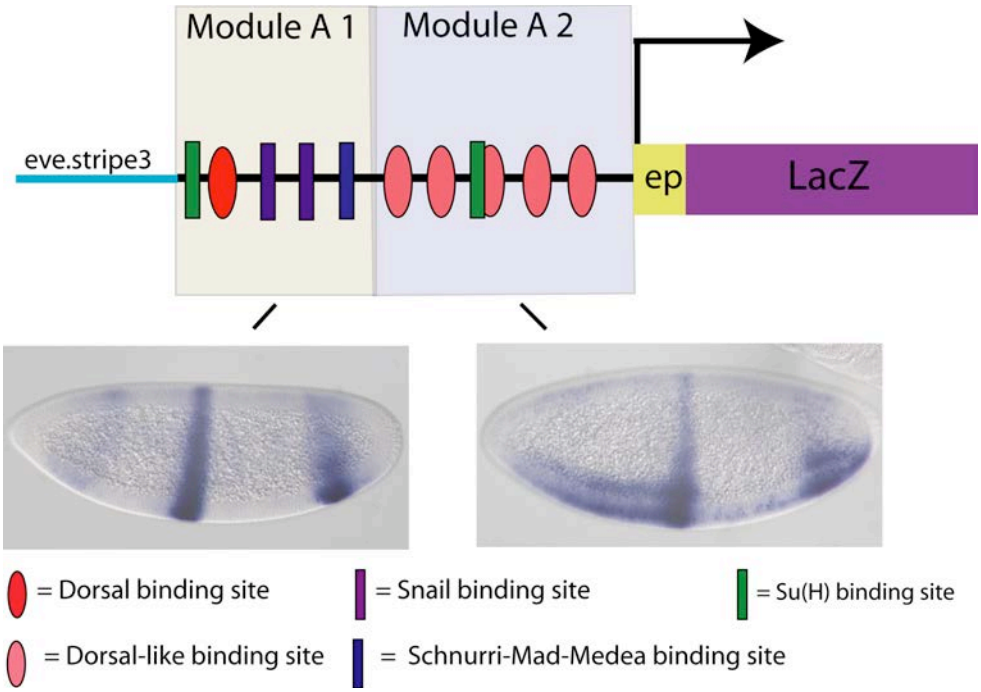


Figure 4.3. Dorsal repression is mediated by module A2.

The schematic shows the division of module A into module A1 and A2. The embryos correspond to each module in the absence of the other module. Repression is seen only when module A2 is placed in between the *eve.stripe3/7* CRM and the *lacZ* reporter gene.

The sog CRM Mediates Ventral Repression and Weak Dorsal Repression

We used the *eve.stripe3/7* chimeric CRM assay to analyze a 392 bp *sog* CRM located within the *sog* intronic region (Markstein et al., 2004). We were interested in whether the dorsal and ventral borders of *sog* are delineated by repressors. We analyzed both the full-length *sog* CRM as well as three partial sequences called modules A, B, and C. We found that ventral repression was mediated with the full-length CRM and weak dorsal repression was observed (figure 4.4 A). The *sog* CRM contains two Snail binding sites that likely mediated repression as it has been previously shown that loss of *sna* affects *sog* expression (Cowden and Levine, 2003). The module that mediated the

strongest ventral repression was module A, which contains two Snail binding sites (figure 4.4 B). Module C showed weak ventral repression, with one Snail binding site (figure 4.4 D). Module B mediated weak dorsal repression (figure 4.4 C). There are no known binding sites that would be capable of mediating dorsal repression. It is still unclear how the dorsal border of *sog* is determined, but the fact that the repression we observed is very weak suggests that repressors play a minor role in *sog* regulation as a sharp boundary is not necessary.

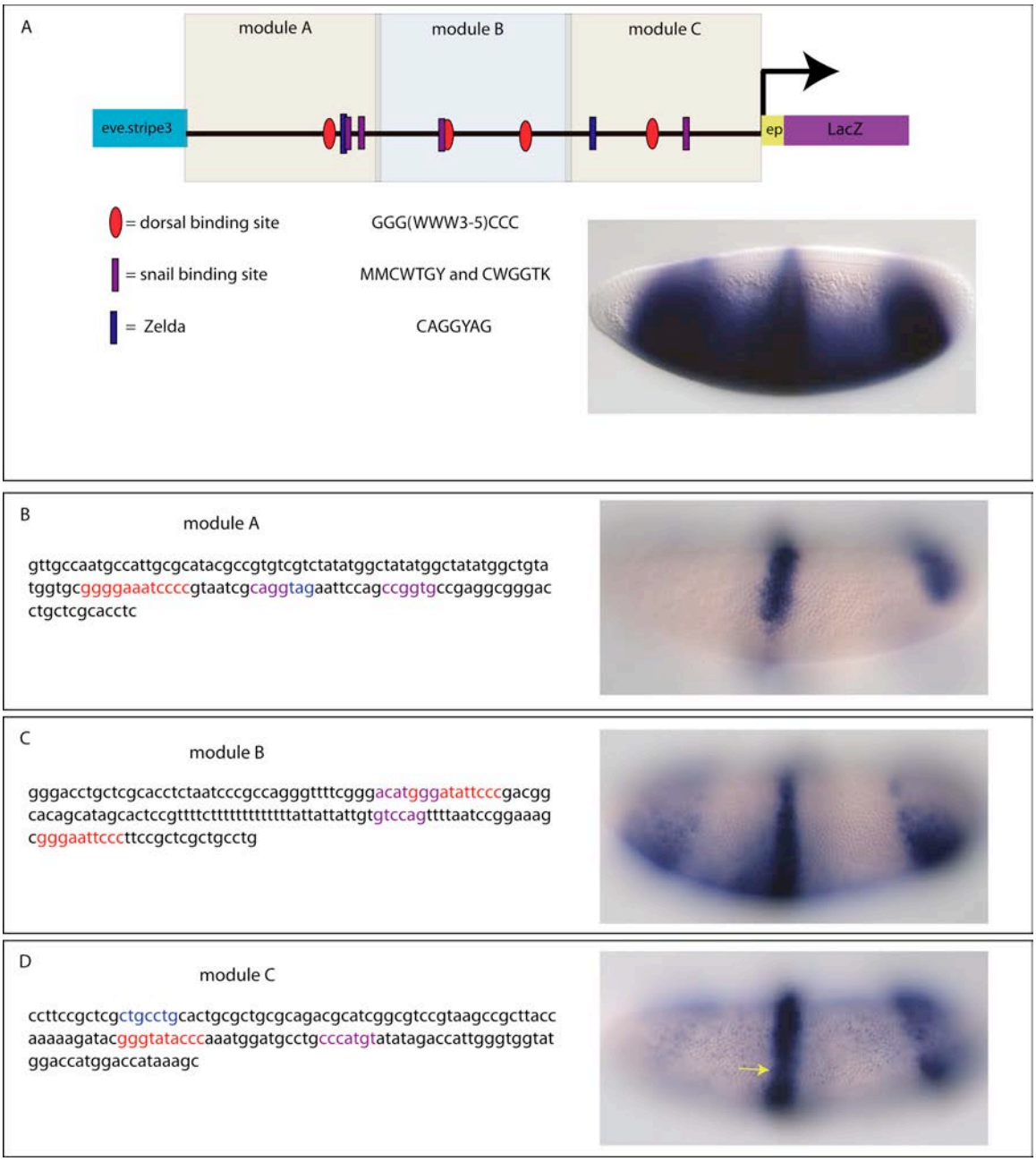


Figure 4.4. The *sog* CRM contains binding sites that mediate ventral repression and weak dorsal repression.

We conducted a chimeric CRM assay in which the *eve.stripe3/7* CRM was placed proximal to the *sog* CRM, this chimeric CRM along with the *eve* promoter was used to drive expression of a *lacZ* reporter gene. The schematic shows the construct used to

drive expression of *lacZ*. The binding sites present in the *sog* CRM are depicted in the schematic. The embryos were stained by in situ hybridization and show the expression of *lacZ*. The full-length CRM shows weak dorsal repression (A). The CRM was divided in to three modules. Module A mediates ventral repression (B). Module B mediates weak dorsal repression (C). Module C mediates weak ventral repression indicated by the yellow arrow (D).

Discussion

Combinatorial interactions between activators during early embryogenesis are important since they allow the same set of activators to be used by genes that are expressed in different patterns. For instance, *vnd* and *rho* are both activated by Twist and Dorsal but *vnd* is shifted ventrally and is about two cells narrower. It is believed that this is due to *vnd's* requirement for Twist-Twist interactions (Zinzen et al., 2006). It is possible that combinatorial interactions act early to establish activation domains that are later refined and maintained by repressors. This might explain why expression patterns first appear less sharp than their final expression pattern.

We analyzed the CRMs of *vnd* and *sog* for the presence of repressors. As expected we found that both CRMs repressed *eve.stripe3/7* in ventral regions in a chimeric CRM assay. Snail binding sites that are present in both CRMs likely mediate this repression. The dorsal borders of *vnd* and *sog* are very different in both location and sharpness, thus it was not surprising that the *vnd* CRM mediated strong repression while the *sog* CRM mediated weak repression of *eve.stripe3/7* in dorsal regions.

We have not determined which sites are necessary for dorsal repression of *vnd* but we did isolate a 100 bp module that mediates repression. This module contains five dorsal-like sites, with one of the sites overlapping a Su(H) binding site. One possibility is that a repressor binds the dorsal-like binding sites and possibly competes with Dorsal for binding. Another possibility is that Su(H) binding sites can mediate repression.

Su(H) is ubiquitously expressed in the early embryo and is the downstream effector of Notch signaling. Notch signaling is activated in the mesectoderm and is essential for the one cell expression pattern of *single-minded (sim) (Morel and Schweisguth, 2000)*. Notch signaling has not been shown to have an effect on any of the other dorsal-ventral patterning genes. A recent study that sought to find a regulatory code that would specify the neurogenic ectoderm and could be used to identify other CRMs, defined clusters of Dorsal, Twist, and Su(H) as being important in identifying CRMs (Markstein et al., 2004). In this study mutation of both Su(H) sites in the *brk* CRM resulted in severe reduction of the expression pattern. In contrast, when the Su(H) binding sites were mutated in the *vnd* CRM only a slight reduction occurred. The authors suggested that Su(H) is involved in activation of genes expressed in the neurogenic ectoderm.

Contrary to a role for Su(H) binding sites in activation there is evidence to suggest that they play a role in repression as well. In a *eve.stripe3.7-snailCRM* chimeric assay modules containing Su(H) binding sites were found to mediate repression. When the Su(H) site was mutated the repression was lost, but mutation of the Su(H) sites did not have an affect on the *sna* expression pattern alone (A. Ozdemir and A. Stathopoulos, unpublished observations). It is possible that other unknown repressor binding sites

present in the *sna* CRM masked the effects of *Su(H),* thus when the sites were mutated in the full-length CRM alone expansion of the *sna* expression domain was not observed. Another possibility is that *Su(H)* does not mediate repression itself but acts to facilitate long-range repression, accordingly, when the sites were mutated in the chimeric assay repressors in the *sna* CRM could no longer act on the *eve.stripe3/7* CRM.

Similarly, Su(H) was shown to play a role in repression of *sim*. The researchers found that Su(H) represses *sim* independent of Notch and also activates *sim* downstream of Notch signaling (Morel and Schweisguth, 2000). It is believed that Notch signaling converts Su(H) from a repressor to an activator. The precise role of Su(H) in regulation of other dorsal-ventral patterning genes is still unclear. Future experiments such as mutating the Su(H) binding site in the *vnd* CRM and looking at *vnd* expression in Su(H) mutants will be necessary to determine its role. A pervasive role for Su(H) in supporting repression of the dorsal-ventral patterning genes would change the patterning paradigm. Su(H) was previously believed to act as an activator for dorsal-ventral patterning genes and was only shown to have a repressive effect on *sim.*

Genomic analysis such as Chip-Chip and Chip-Seq experiments have aided in the identification of CRMs, recently the number of known CRMs has increased dramatically (Ozdemir et al., 2011). Many alternate CRMs have been identified for genes with known CRMs. Superficially, these, CRMs appear to regulate gene expression in patterns similar to the previously identified CRMs. Detailed analysis of some of these CRMs has shown that they differ and are both are necessary for proper patterning (Dunipace et al., 2011). For example there are two *sna* CRMs that appear to drive expression in very similar patterns, although one has a sharper dorsal border (Ozdemir A. and Stathopoulos A.

unpublished data). Analysis of the two *sna* CRMs in a chimeric CRM analysis has shown that the one with the sharper border contains dorsal repression elements. It is possible that the CRMs that are used to initiate patterning are more permissive and contain fewer repression elements, while the CRMs used to refine the patterns contain repression elements. An alternate *sog* CRM has also been identified, it is possible that this CRM will contain stronger dorsal repression than the intronic *sog* CRM that we tested in this study. In future studies we would like to focus on analyzing both sets of CRMs where they exist to see if there is a trend, where initiator intronic CRMs are more permissive and the bulk of repressors are located in more distant CRMs.

Chapter 5

Discussion

The main goal of this thesis was to understand the inputs that are necessary to determine specific expression patterns, in doing so; we strove to gain knowledge about how the embryo is patterned in general and how axis specification is determined. We analyzed the CRMs of *ind*, *vnd,* and *sog*. In doing so we hoped to gain knowledge about the regulatory inputs that pattern the dorsal-ventral axis of the early embryo. Our analysis was motivated by the lack of known factors capable of defining the dorsal borders of many of the genes expressed along the dorsal-ventral axis.

Our analysis of the *ind* CRM showed that regulation of this gene is more complex than once believed. Previously, it was thought that the dorsal border of the *ind* pattern was established by limiting amounts of the activators Dorsal and Egfr signaling and possibly repression by Dpp signaling. In chapter 2 we showed that another activator, Grh, is also necessary for *ind* expression and that Egfr signaling activates *ind* indirectly, by releasing it from repression of the A-box binding factor, likely Cic. We also provide the first direct evidence to implicate SMM binding in repression of a neurogenic ectoderm gene (*ind*) at early stages of development, thus establishing a role for Dpp signaling in patterning the neurogenic ectoderm.

We also provide evidence for the possibility of chromatin-remodeling factors acting early in embryonic development to pattern the neurogenic ectoderm. In chapter 3 we described the functions of the factors that were found to bind the A-box affinity column and discussed how they might function in regulating *ind*. If they do function in the early embryo, this would be the earliest documented use of chromatin-remodeling in patterning, as repression by chromatin factors has only been shown to act at later stages

in embryonic development. Further analysis is needed to solidify a role for chromatin remodeling factors at this stage in development.

Our analysis of the *vnd* and *sog* CRMs in chapter 4 revealed that similar to *ind,* expression of *vnd* is likely regulated by a as of yet unknown repressor. This repressor may bind to dorsal-like binding sites. Similar to *ind* there is the possibility that Dpp signaling acts as a second tier of repression to define the *vnd* dorsal border. Analysis of the *sog* CRM did not show significant dorsal repression which is compatible with its broad diffuse expression pattern. Together our results indicated that while the Dorsal morphogen gradient is instrumental in establishing the initial domains of dorsal-ventral gene expression patterns, other factors contribute in order to refine the patterns and assure that proper patterning is maintained.

Analysis of CRMs is Important in Understanding Gene Regulation and Patterning

CRM analysis has been used as a method to understand gene regulation for a number of years. Our analyses show that this method still shows merit and there is still a lot of information to be learned from CRMs. It is unlikely that Grh would have been identified as an activator of *ind* by any other method. Also, for years, it was believed that BMP/Dpp signaling patterns the neurogenic ectoderm of the early embryo and yet the strongest evidence to support this was ectopic expression of Dpp in a highly manipulated background, where embryos were lateralized, through the use of an activated Toll receptor, in a *brk sog* background to eliminated repression of Dpp signaling (Mizutani et al., 2006)(chapter 1 figure 1.1 D). We provided the first direct evidence that SMM sites are functional at this early stage of development, directly linking Dpp signaling to the

repression of a Dorsal target gene. The analysis of other CRMs that control gene expression along the dorsal-ventral axis is likely to provide more clues into how patterns are specified. The alternate CRMs, which have been identified for several dorsal-ventral patterning genes are excellent candidates for future chimeric CRM analysis.

How Can we Study Chromatin-Remodeling Factors in the Early Embryo?

The study of chromatin-remodeling factors in the early embryo is complicated by the fact that many of them are maternally deposited and used during development of the oocyte. Recently, the creation of shRNA lines has made if feasible to knockdown the expression of RNAs in the oocyte and also knockdown RNAs that are maternally deposited, without having to generate germ-line clones (Ni et al., 2011). Thus, larger sets of genes can be screened. The problem with this technique is that several of the factors we are interested in have resulted in sterility and thus the embryo cannot be assayed. Although, we have not tested Dsp1 and Caf1-180 because shRNA lines were not available, thus it is possible we may see an effect in these lines. We are currently in the processes of generating these lines.

One possibility is to test the A-boxes and their flanking regions, perhaps by taking the entire 110 bp fragment that contains both A-box sites and testing to see if it displays properties of a PRE such as pair-sensitive silencing and variegation. The limitation of this assay is that it is conducted in the adult fly by looking for an eye color phenotype. It would not tell us at what stage of development this PRE is functional, thus we would not be able to link chromatin-remodeling to the early embryo, but it would give us information about whether it is even possible for this region of DNA to recruit PcG

proteins. We can use this information to determine how to pursue further experiments in the embryo.

Are Tiers of Dorsal Repression a General Mechanism for Dorsal-Ventral Patterning?

When looking at specific genes an important question to ask is if what is seen for a specific case can be generally applied. The presence of, seemingly non-functional, SMM binding sites in several dorsal-ventral CRMs suggests that tiers of dorsal repressors may not be unique to *ind* and it is likely a general mechanism in patterning dorsal-ventral genes with sharp boundaries. It is possible that the SMM binding sites do function at this early stage of development but independent repressors acting in parallel do not allow us to see a phenotype when the sites are mutated or zygotic *schnurri* mutants are analyzed. The question that remains is if the Dpp independent repression is the same for all dorsal-ventral pattering genes or if each gene has a specific repressor. Since many of the genes expressed along the dorsal-ventral axis do not share borders, it is possible that many have dedicated repressors. It is also possible that combinatorial interactions between a shared set of repressors and activators define the different borders. A search for A-box-like sites in the CRMs of dorsal-ventral patterning genes did not result in any hits, but a Cic-like binding site was found in the *brk* CRM. It is possible that Cic and SMM act in parallel to refine the dorsal border of *brk;* in order to test this idea it will be necessary to create *cic*, *shn* double mutants. Alternatively, the Cic and SMM binding sites can be mutated in the *brk* CRM to see if the expression pattern is affected.

How are Patterning and Scaling Related?

Another important question to consider is if the patterns that are expressed along the dorsal-ventral axis scale with the size of the embryo. Whether they do or not can give us information about whether patterns need to be fixed to a certain number of cells or if a larger embryo can tolerate having more of a certain gene expressed along its axis. It is possible that different mechanisms are used to pattern genes that are fixed versus genes that scale.

Recent work has shown that the Dorsal gradient strictly scales along the dorsal-ventral axis with the size of the embryo (M. Nahmad, G. T. Reeves, M. Garcia, and A. Stathopoulos; in preparation). The mesoderm specific gene *sna* also scales. In the neurogenic ectoderm, both borders of *vnd* and the ventral border of *sog* scale but not as strictly as Dorsal and *sna.* While, the dorsal border of *sog* and both of the *ind* borders under-compensate (Nahmad M. and Stathopoulos A. unpublished data). It is interesting that *ind* has a fixed width in embryos regardless of size. It will be interesting to see if one of the tiers of dorsal repression has an affect on this behavior; it is possible that in the absence of one of the tiers of repression *ind,* will scale. Two tiers of repression may have been developed because it was important to tightly restrict *ind* regardless of the size of the embryo. We are currently conducting mutant analysis to see if the scaling properties of the genes can be affected.

Future Directions

We have learned a great deal from studying individual CRMs, in future experiments it would be interesting to study CRMs with a focus on finding trends that

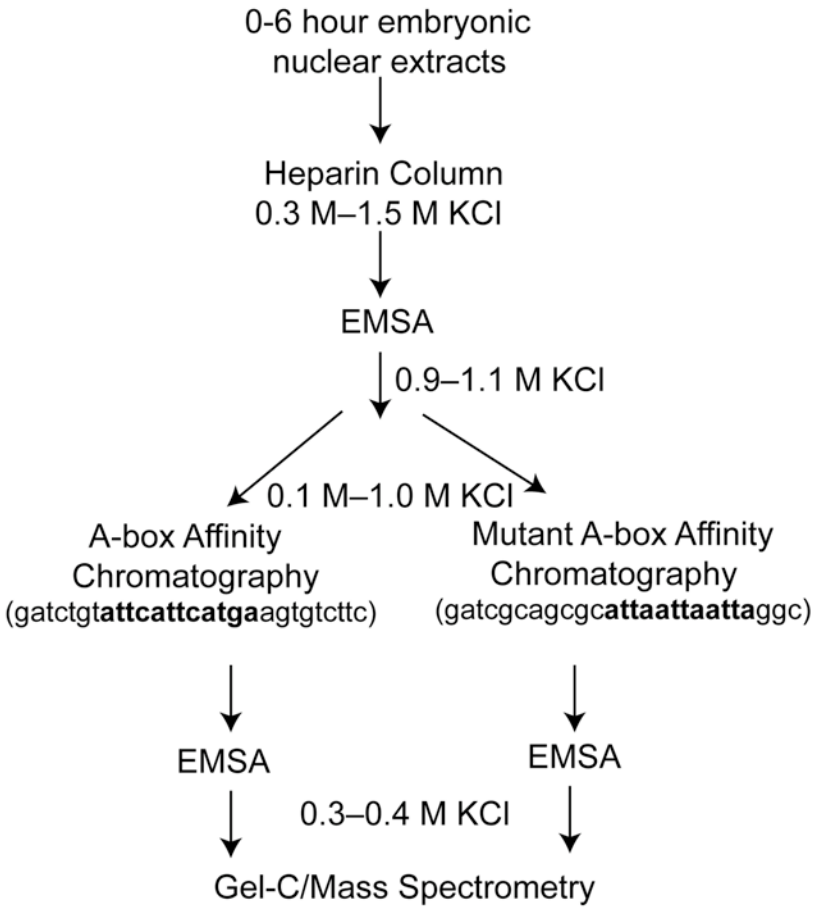
can be applied generally to dorsal-ventral pattern. We have already uncovered some trends and we would first need to follow up with these to see how widespread they are. For example it is possible that multiple genes use two tiers of repression to define their dorsal borders. It will be noteworthy to investigate if these genes use dedicated repressors or if shared repressors are used. We already have an indication that Su(H) may act as a shared repressor, studies are currently being conducted to determine how pervasive its repression is (Ozdemir A. and Stathopoulos A.).

Another trend we have observed is the presence of multiple CRMs with seemingly similar patterns but with different abilities for mediating repression, such as the *sna* CRMs. In order to determine if this is a general trend we will need to analyze more CRMs. Together these analyses will tell us more about the mechanisms used to pattern the dorsal-ventral axis.

Appendices

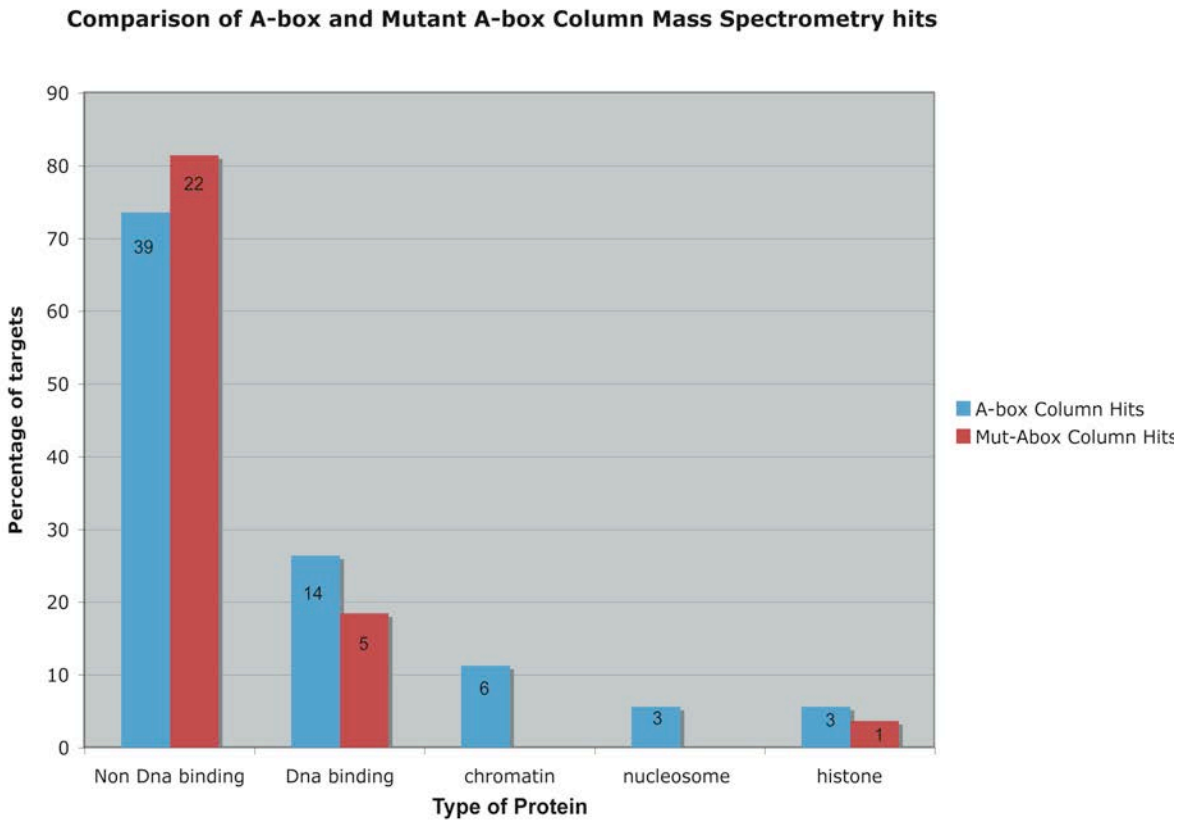
Appendix A

Supplemental Information: Chapter 2

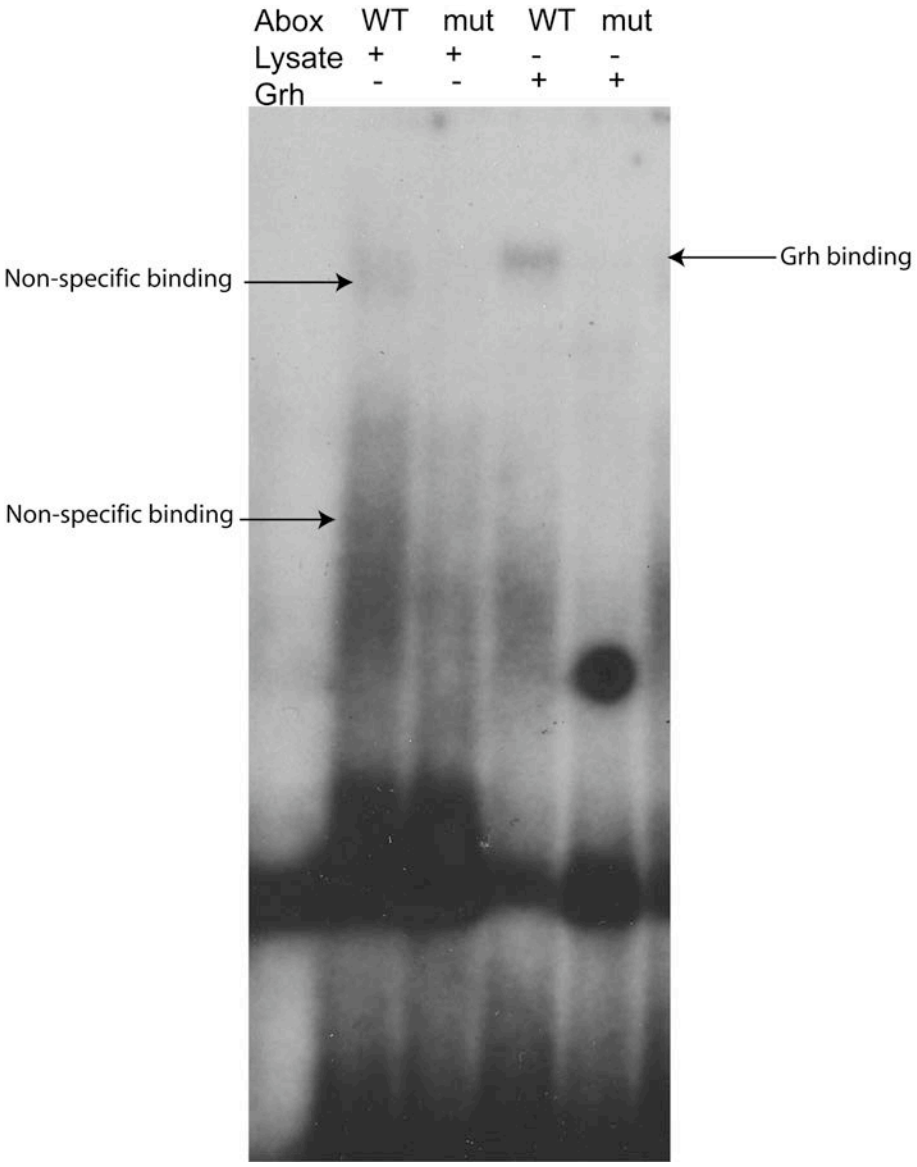


Supplemental Figure 2.S1: Flow-chart outlining the protocol used to purify factors that bind the A-box element.

First we created nuclear extracts from 0–6 hour embryos. Then we fractionated the sample using a heparin column and tested the fractions for A-box specefic binding. We affinity purified the fractions that contained A-box specific activity using an A-box column and a mutant A-box column. We again tested for A-box specific binding and identified factors bound to both columns using mass spectrometry.



Supplemental Figure 2.S2. Overabundance of Chromatin remodeling and histone modifying factors found binding to A-box column versus the mutant A-box column. The percentage was calculated by dividing the number of factors in each specific category by the total number of factors found to bind only the A-box column or mutant A-box column. The number on the bar corresponds to the number of factors in each specified category.



Supplemental Figure 2.S3. EMSA shows binding of Grh to the A-box binding site. Rabbit reticulolysates were used to *in-vitro* translate the Grh protein and EMSA was preformed using γ^{32}P-labeled A-box oligonucleotides. Non-specific binding (indicated by the black arrows to the left) was detected in the lysate alone. This binding was diffuse throughout the column. Grh binding was strong and sharp (indicated by the black arrow to the right) and was only seen when the A-box oligonuleotide was used and not the mutant A-box oligonucleotide.

Appendix B

Supplementary Information: Chapter 3

Yeast One-Hybrid Analysis	
A-box	**ind 110**
CG14655	drm
Neu2	bap
Wor	Hr96
Aef1	CG1832
hkb	gsb-n
CG7928	CG15182
side	side
Rbf2	lola
CG31385	CG15336
Hey	nf1
mor	cg11294
gcm2	rx
CG13204	dpn
CG11695	nc2
dsf	cg8301
dys	sc
NA	cg17801
NA	cg7987
NA	cg13441
NA	cg31670
NA	cg15782
NA	cg6470
NA	hlh106
NA	blimp-1
NA	ches-1-like
NA	su(var)3-7
NA	Opa
NA	salr
NA	woc
NA	twi

Table 1. Targets that bound to an A-box element or a 110 bp *ind* CRM fraction containing both A-boxes

Yeast one hybrid analysis was conducted using either an A-box element or a 110 bp *ind* CRM fraction. In no particular order the transcription factors that were found to bind are listed. (Conducted by Ozdemir A.)

Bibliography

Adryan, B. (2009) FlyTF.org: The *Drosophila* transcription factor database. In, equivalent to release v1 - trusted TFsVol. 2009.

Ajuria, L., Nieva, C., Winkler, C., Kuo, D., Samper, N., Andreu, M.J., Helman, A., Gonzalez-Crespo, S., Paroush, Z., Courey, A.J. and Jimenez, G. (2011) Capicua DNA-binding sites are general response elements for RTK signaling in Drosophila. Development 138, 915-24.

Arora, K., Dai, H., Kazuko, S.G., Jamal, J., O'Connor, M.B., Letsou, A. and Warrior, R. (1995) The Drosophila schnurri gene acts in the Dpp/TGF beta signaling pathway and encodes a transcription factor homologous to the human MBP family. Cell 81, 781-90.

Ashe, H.L. (2005) BMP signalling: synergy and feedback create a step gradient. Curr Biol 15, 375-77.

Ashe, H.L. and Briscoe, J. (2006) The interpretation of morphogen gradients. Development 133, 385-94.

Balaskas, N., Ribeiro, A., Panovska, J., Dessaud, E., Sasai, N., Page, K.M., Briscoe, J. and Ribes, V. (2012) Gene regulatory logic for reading the Sonic Hedgehog signaling gradient in the vertebrate neural tube. Cell 148, 273-84.

Belvin, M.P., Jin, Y. and Anderson, K.V. (1995) Cactus protein degradation mediates Drosophila dorsal-ventral signaling. Genes Dev 9, 783-93.

Biehs, B., Francois, V. and Bier, E. (1996) The Drosophila short gastrulation gene prevents Dpp from autoactivating and suppressing neurogenesis in the neuroectoderm. Genes Dev 10, 2922-34.

Bier, E., Jan, L.Y. and Jan, Y.N. (1990) rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in Drosophila melanogaster. Genes Dev 4, 190-203.

Blastyak, A., Mishra, R.K., Karch, F. and Gyurkovics, H. (2006) Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic. Mol Cell Biol 26, 1434-44.

Bray, S.J. and Kafatos, F.C. (1991) Developmental function of Elf-1: an essential transcription factor during embryogenesis in Drosophila. Genes Dev 5, 1672-83.

Brown, J.L., Grau, D.J., DeVido, S.K. and Kassis, J.A. (2005) An Sp1/KLF binding site is important for the activity of a Polycomb group response element from the Drosophila engrailed gene. Nucleic Acids Res 33, 5181-9.

Canty, E.G., Garrigue-Antar, L. and Kadler, K.E. (2006) A complete domain structure of Drosophila tolloid is required for cleavage of short gastrulation. J Biol Chem 281, 13258-67.

Chen, Y. and Schier, A.F. (2001) The zebrafish Nodal signal Squint functions as a morphogen. Nature 411, 607-10.

Chu, H., Parras, C., White, K. and Jimenez, F. (1998) Formation and specification of ventral neuroblasts is controlled by vnd in Drosophila neurogenesis. Genes Dev 12, 3613-24.

Cornell, R. and Von Ohlen, T. (2000) *Vnd/nkx, ind/gsh, and msh/msx*: conserved regulator of dorsalventral neural patterning? Current Opin Neurbiol 10, 63-71.

Cowden, J. and Levine, M. (2003) Ventral dominance governs sequential patterns of gene expression across the dorsal-ventral axis of the neuroectoderm in the Drosophila embryo. Dev Biol 262, 335-49.

Dai, H., Hogan, C., Gopalakrishnan, B., Torres-Vazquez, J., Nguyen, M., Park, S., Raftery, L.A., Warrior, R. and Arora, K. (2000) The zinc finger protein schnurri acts as a Smad partner in mediating the transcriptional response to decapentaplegic. Dev Biol 227, 373-87.

Dejardin, J., Rappailles, A., Cuvier, O., Grimaud, C., Decoville, M., Locker, D. and Cavalli, G. (2005) Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. Nature 434, 533-8.

Delattre, M., Spierer, A., Jaquet, Y. and Spierer, P. (2004) Increased expression of Drosophila Su(var)3-7 triggers Su(var)3-9-dependent heterochromatin formation. J Cell Sci 117, 6239-47.

Delattre, M., Spierer, A., Tonka, C.H. and Spierer, P. (2000) The genomic silencing of position-effect variegation in Drosophila melanogaster: interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. J Cell Sci 113 4253-61.

Doyle, H.J., Kraut, R. and Levine, M. (1989) Spatial regulation of zerknullt: a dorsal-ventral patterning gene in Drosophila. Genes Dev 3, 1518-33.

Driever, W., Thoma, G. and Nusslein-Volhard, C. (1989) Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. Nature 340, 363-7.

Dunipace, L., Ozdemir, A. and Stathopoulos, A. (2011) Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. Development 138, 4075-84.

Dynlacht, B.D., Attardi, L.D., Admon, A., Freeman, M. and Tjian, R. (1989) Functional analysis of NTF-1, a developmentally regulated Drosophila transcription factor that binds neuronal cis elements. Genes Dev 3, 1677-88.

Ferguson, E. and Anderson, K. (1992) Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo. Cell 71, 451-61.

Ghiglione, C., Bach, E.A., Paraiso, Y., Carraway, K.L., 3rd, Noselli, S. and Perrimon, N. (2002) Mechanism of activation of the Drosophila EGF Receptor by the TGFalpha ligand Gurken during oogenesis. Development 129, 175-86.

Green, J.B., New, H.V. and Smith, J.C. (1992) Responses of embryonic Xenopus cells to activin and FGF are separated by multiple dose thresholds and correspond to distinct axes of the mesoderm. Cell 71, 731-9.

Green, J.B. and Smith, J.C. (1990) Graded changes in dose of a Xenopus activin A homologue elicit stepwise transitions in embryonic cell fate. Nature 347, 391-4.

Grieder, N.C., Nellen, D., Burke, R., Basler, K. and Affolter, M. (1995) Schnurri is required for Drosophila Dpp signaling and encodes a zinc finger protein similar to the mammalian transcription factor PRDII-BF1. Cell 81, 791-800.

Grillo, M., Furriols, M., Casanova, J. and Luschnig, S. (2011) Control of germline torso expression by the BTB/POZ domain protein Pipsqueak is required for embryonic terminal patterning in Drosophila. Genetics 187, 513-20.

Harrison, M.M., Botchan, M.R. and Cline, T.W. (2010) Grainyhead and Zelda compete for binding to the promoters of the earliest-expressed Drosophila genes. Developmental Biology 345, 248-55.

Hens, K., Feuz, J.D., Isakova, A., Iagovitina, A., Massouras, A., Bryois, J., Callaerts, P., Celniker, S.E. and Deplancke, B. (2011) Automated protein-DNA interaction screening of Drosophila regulatory elements. Nat Methods 8, 1065-70.

Hodgson, J.W., Argiropoulos, B. and Brock, H.W. (2001) Site-Specific Recognition of a 70-Base-Pair Element Containing d(GA)n Repeats Mediates bithoraxoid Polycomb Group Response Element-Dependent Silencing. Molecular and Cellular Biology 21, 4528-43.

Holley, S.A., Jackson, P.D., Sasai, Y., Lu, B., De Robertis, E.M., Hoffmann, F.M. and Ferguson, E.L. (1995) A conserved system for dorsal-ventral patterning in insects and vertebrates involving sog and chordin. Nature 376, 249-53.

Huang, A.M., Rusch, J. and Levine, M. (1997) An anteroposterior Dorsal gradient in the Drosophila embryo. Genes Dev 11, 1963-73.

Huang, D.H., Chang, Y.L., Yang, C.C., Pan, I.C. and King, B. (2002) pipsqueak Encodes a Factor Essential for Sequence-Specific Targeting of a Polycomb Group Protein Complex. Molecular and Cellular Biology 22, 6261-71.

Huang, H., Yu, Z., Zhang, S., Liang, X., Chen, J., Li, C., Ma, J. and Jiao, R. (2010) Drosophila CAF-1 regulates HP1-mediated epigenetic silencing and pericentric heterochromatin stability. J Cell Sci 123, 2853-61.

Huang, J.D., Dubnicoff, T., Liaw, G.J., Bail, Y., Valentine, S.A., Shirokawa, J., Lengyel, J.A. and Courey, A.J. (1995) Binding sites for transcription factor NTF-1/Elf-1 contribute to the ventral repression of *decapentaplegic* Genes Dev 9, 3177-89.

Ingham, P.W. and Fietz, M.J. (1995) Quantitative effects of hedgehog and decapentaplegic activity on the patterning of the Drosophila wing. Curr Biol 5, 432-40.

Ip, Y.T., Park, R.E., Kosman, D., Bier, E. and Levine, M. (1992a) The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the Drosophila embryo. Genes Dev 6, 1728-39.

Ip, Y.T., Park, R.E., Kosman, D., Yazdanbakhsh, K. and Levine, M. (1992b) dorsal-twist interactions establish snail expression in the presumptive mesoderm of the Drosophila embryo. Genes Dev 6, 1518-30.

Irish, V.F. and Gelbart, W.M. (1987) The decapentaplegic gene is required for dorsal-ventral patterning of the Drosophila embryo. Genes Dev 1, 868-79.

Isshiki, T., Takeichi, M. and Nose, A. (1997) The role of the msh homeobox gene during Drosophila neurogenesis: implication for the dorsoventral specification of the neuroectoderm. Development 124, 3099-109.

Jazwinska, A., Rushlow, C. and Roth, S. (1999) The role of brinker in mediating the graded response to Dpp in early Drosophila embryos. Development 126, 3323-34.

Jiang, J., Cai, H., Zhou, Q. and Levine, M. (1993) Conversion of a *dorsal*-dependent silencer into and enhancr: evidence for *dorsal* corepressors. EMBO J 12, 3201-9.

Jiang, J., Kosman, D., Ip, Y.T. and Levine, M. (1991) The dorsal morphogen gradient regulates the mesoderm determinant twist in early Drosophila embryos. Genes Dev 5, 1881-91.

Jiang, J., Rushlow, C.A., Zhou, Q., Small, S. and Levine, M. (1992) Individual Dorsal morphogen binding sites mediate activation and repression in the *Drosophila* embryo. EMBO J 11, 3147-54.

Jimenez, F., Martin-Morris, L.E., Velasco, L., Chu, H., Sierra, J., Rosen, D.R. and White, K. (1995) vnd, a gene required for early neurogenesis of Drosophila, encodes a homeodomain protein. EMBO J 14, 3487-95.

Jimenez, G., Guichet, A., Ephrussi, A. and Casanova, J. (2000) Relief of gene repression by torso RTK signaling: role of capicua in Drosophila terminal and dorsoventral patterning. Genes Dev 14, 224-31.

Kerrigan, L.A. and Kadonaga, J.T. (1998) Purification of Sequence-Specific DNA-Binding Proteins by Affinity Chromatography. In, Curr Protoc Prot Sci, John Wiley and Sons, Inc.

Kim, M. and McGinnis, W. (2011) Phosphorylation of Grainy head by ERK is essential for wound-dependent regeneration but not for development of an epidermal barrier. Proc Natl Acad Sci U S A 108, 650-5.

Kosman, D., Mizutani, C., Lemons, D., Cox, G., McGinnis, W. and Bier, E. (2004) Multiplex Detection of RNA Expression in *Drosophila* Embryos. Science 305, 846.

Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F. and Orlando, V. (2007) Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. . Nat Cell Biol 9, 1167-84.

Lee, H. (2004) The grainy head transcription factor is essential for the function of the frizzled pathway in the Drosophila wing. Mech Dev 121, 37-49.

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, 400-3.

Liaw, G.J., Rudolph, K.M., Huang, J.D., Dubnicoff, T., Courey, A.J. and Lengyel, J.A. (1995) The torso response element binds GAGA and NTF-1/Elf-1, and regulates tailless by relief of repression. Genes Dev 9, 3163-76.

Liberman, L.M., Reeves, G.T. and Stathopoulos, A. (2009) Quantitative imaging of the Dorsal nuclear gradient reveals limitations to threshold-dependent patterning in Drosophila. Proc Natl Acad Sci U S A 106, 22317-22.

Liberman, L.M. and Stathopoulos, A. (2009) Design flexibility in cis-regulatory control of gene expression: Synthetic and comparative evidence. Dev Biol 327, 578-89.

Luschnig, S., Moussian, B., Krauss, J., Desjeux, I., Perkovic, J. and Nusslein-Volhard, C. (2004) An F1 genetic screen for maternal-effect mutations affecting embryonic pattern formation in Drosophila melanogaster. Genetics 167, 325-42.

Markstein, M., Zinzen, R., Markstein, P., Yee, K.P., Erives, A., Stathopoulos, A. and Levine, M. (2004) A regulatory code for neurogenic gene expression in the Drosophila embryo. Development 131, 2387-94.

Mizutani, C.M., Meyer, N., Roelink, H. and Bier, E. (2006) Threshold-Dependent BMP-Mediated Repression: A Model for a Conserved Mechanism That Patterns the Neuroectoderm. Plos Biol 4, 1777-89.

Morel, V. and Schweisguth, F. (2000) Repression by Suppressor of Hairless and activation by Notch are required to define a sing row of single-minded expressing cells in the Drosophila embryo. Genes Dev 14, 388-99.

Morisato, D. (2001) Spatzle regulates the shape of the Dorsal gradient in the Drosophila embryo. Development 128, 2309-19.

Neuman-Silberberg, F.S. and Schupbach, T. (1993) The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein. Cell 75, 165-74.

Ni, J.Q., Zhou, R., Czech, B., Liu, L.P., Holderbaum, L., Yang-Zhou, D., Shim, H.S., Tao, R., Handler, D., Karpowicz, P., Binari, R., Booker, M., Brennecke, J., Perkins, L.A., Hannon, G.J. and Perrimon, N. (2011) A genome-scale shRNA resource for transgenic RNAi in Drosophila. Nat Methods 8, 405-7.

Nusslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984) Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster. I. Zygotic loci on the second chromosome. Rouxs Arch. Dev. Biol. 193, 267-82.

Oh, C.T., Kwon, S.H., Jeon, K.J., Han, P.L., Kim, S.H. and Jeon, S.H. (2002) Local inhibition of Drosophila homeobox-containing neural dorsoventral patterning genes by Dpp. FEBS Lett 531, 427-31.

Orlando, V., Jane, E.P., Chinwalla, V., Harte, P.J. and Paro, R. (1998) Binding of trithorax and Polycomb proteins to the bithorax complex: dynamic changes during early Drosophila embryogenesis. EMBO J 17, 5141-50.

Ozdemir, A., Fisher-Aylor, K.I., Pepke, S., Samanta, M., Dunipace, L., McCue, K., Zeng, L., Ogawa, N., Wold, B.J. and Stathopoulos, A. (2011) High resolution mapping of Twist to DNA in Drosophila embryos: Efficient functional analysis and evolutionary conservation. Genome Res 21, 566-77.

Ozdemir, A. and Stathopoulos, A. (2011) Exciting times: bountiful data to facilitate studies of cis-regulatory control. Nat Methods 8, 1016-7.

Perrimon, N. (1998) Creating mosaics in Drosophila. Int J Dev Biol 42, 243-7.

Pyrowolakis, G., Hartmann, B., Muller, B., Basler, K. and Affolter, M. (2004) A simple molecular complex mediates widespread BMP-induced repression during Drosophila development. Dev Cell 7, 229-40.

Ray, R.P., Arora, K., Nusslein-Volhard, C. and Gelbart, W.M. (1991) The control of cell fate along the dorsal-ventral axis of the Drosophila embryo. Development 113, 35-54.

Reeves, G.T. and Stathopoulos, A. (2009) Graded Dorsal and Differential Gene Regulation in the Drosophila Embryo. In, Perspectives on Generation and Interpretation of Morphogen Gradients, Cold Spring Harbor Laboratory Press.

Reeves, G.T., Trisnadi, N., Truong, T.V., Nahmad, M., Katz, S. and Stathopoulos, A. (2012) Dorsal-ventral gene expression in the Drosophila embryo reflects the dynamics and precision of the dorsal nuclear gradient. Dev Cell 22, 544-57.

Ringrose, L. and Paro, R. (2004) Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet 38, 413-43.

Ringrose, L. and Paro, R. (2007) Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development 134, 223-32.

Rogers, K.W. and Schier, A.F. (2011) Morphogen gradients: from generation to interpretation. Annu Rev Cell Dev Biol 27, 377-407.

Rushlow, C., Colosimo, P.F., Lin, M.C., Xu, M. and Kirov, N. (2001) Transcriptional regulation of the Drosophila gene zen by competing Smad and Brinker inputs. Genes Dev 15, 340-51.

Rutledge, B.J., Zhang, K., Bier, E., Jan, Y.N. and Perrimon, N. (1992) The Drosophila spitz gene encodes a putative EGF-like growth factor involved in dorsal-ventral axis formation and neurogenesis. Genes Dev 6, 1503-17.

Schnepp, B., Grumbling, G., Donaldson, T. and Simcox, A. (1996) Vein is a novel component in the Drosophila epidermal growth factor receptor pathway with similarity to the neuregulins. Genes Dev 10, 2302-13.

Sen, J., Goltz, J.S., Stevens, L. and Stein, D. (1998) Spatially restricted expression of pipe in the Drosophila egg chamber defines embryonic dorsal-ventral polarity. Cell 95, 471-81.

Serpe, M., Ralston, A., Blair, S.S. and O'Connor, M.B. (2005) Matching catalytic activity to developmental function: tolloid-related processes Sog in order to help specify the posterior crossvein in the Drosophila wing. Development 132, 2645-56.

Shimmi, O., Umulis, D., Othmer, H. and O'Connor, M.B. (2005) Facilitated transport of a Dpp/Scw heterodimer by Sog/Tsg leads to robust patterning of the Drosophila blastoderm embryo. Cell 120, 873-86.

Skeath, J.B. (1998) The Drosophila EGF receptor controls the formation and specification of neuroblasts along the dorsal-ventral axis of the Drosophila embryo. Development 125, 3301-12.

Skeath, J.B., Panganiban, G.F. and Carroll, S.B. (1994) The ventral nervous system defective gene controls proneural gene expression at two distinct steps during neuroblast formation in Drosophila. Development 120, 1517-24.

Small, S., Blair, A. and Levine, M. (1996) Regulation of two pair-rule stripes by a single enhancer in the Drosophila embryo. Dev Biol 175, 314-24.

Staehling-Hampton, K., Laughon, A.S. and Hoffmann, F.M. (1995) A Drosophila protein related to the human zinc finger transcription factor PRDII/MBPI/HIV-EP1 is required for dpp signaling. Development 121, 3393-403.

Stathopoulos, A. and Levine, M. (2004) Whole-genome analysis of Drosophila gastrulation. Curr Opin Genet Dev 14, 477-84.

Stathopoulos, A. and Levine, M. (2005) Localized repressors delineate the neurogenic ectoderm in the early Drosophila embryo. Dev Biol 280, 482-93.

Stathopoulos, A., Van Drenth, M., Erives, A., Markstein, M. and Levine, M. (2002) Whole-genome analysis of dorsal-ventral patterning in the Drosophila embryo. Cell 111, 687-701.

Stein, D., Roth, S., Vogelsang, E. and Nusslein-Volhard, C. (1991) The polarity of the dorsoventral axis in the Drosophila embryo is defined by an extracellular signal. Cell 65, 725-35.

Teleman, A.A., Strigini, M. and Cohen, S.M. (2001) Shaping morphogen gradients. Cell 105, 559-62.

Thomas, J.O. (2001) HMG1 and 2: architectural DNA-binding proteins. Biochem Soc Trans 29, 395-401.

Urban, S., Lee, J.R. and Freeman, M. (2002) A family of Rhomboid intramembrane proteases activates all Drosophila membrane-tethered EGF ligands. EMBO J 21, 4277-86.

Von Ohlen, T.L. and Doe, C.Q. (2000) Convergence of dorsal, dpp, and egfr signaling pathways subdivides the drosophila neuroectoderm into three dorsal-ventral columns. Dev Biol 224, 362-72.

Von Ohlen, T.L. and Moses, C. (2009) Identification of Ind transcription activation and repression domains required for dorsoventral patterning of the CNS. Mech Dev 126, 552-62.

Wang, S., Tsarouhas, V., Xylourgidis, N., Sabri, N., Tiklova, K., Nautiyal, N., Gallio, M. and Samakovlis, C. (2009) The tyrosine kinase Stitcher activates Grainy head and epidermal wound healing in Drosophila. Nat Cell Biol 11, 890-5.

Weber, U., Siegel, V. and Mlodzik, M. (1995) pipsqueak encoded a novel nuclear protein required downstream of seven-up for the development of photoreceptors R3 and R4. EMBO J 14, 6247-57.

Weiss, J.B., Von Ohlen, T., Mellerick, D.M., Dressler, G., Doe, C.Q. and Scott, M.P. (1998) Dorsoventral patterning in the Drosophila central nervous system: the intermediate neuroblasts defective homeobox gene specifies intermediate column identity. Genes Dev 12, 3591-602.

Wolpert, L. (1996) One hundred years of positional information. Trends Genet 12, 359-64.

Yu, K., Sturtevant, M.A., Biehs, B., Francois, V., Padgett, R.W., Blackman, R.K. and Bier, E. (1996) The Drosophila decapentaplegic and short gastrulation genes function antagonistically during adult wing vein development. Development 122, 4033-44.

Zandi, E., Tran, T.N., Chamberlain, W. and Parker, C.S. (1997) Nuclear entry, oligomerization, and DNA binding of the Drosophila heat shock transcription factor are regulated by a unique nuclear localization sequence. Genes Dev 11, 1299-314.

Zinzen, R.P., Senger, K., Levine, M. and Papatsenko, D. (2006) Computational models for neurogenic gene expression in the Drosophila embryo. Curr Biol 16, 1358-65.