



Published in final edited form as:

*Mech Dev.* 2009 July ; 126(7): 552–562. doi:10.1016/j.mod.2009.03.008.

## Identification of Ind transcription activation and repression domains required for dorsoventral patterning of the CNS

Tonia L. Von Ohlen<sup>1</sup> and Cade Moses

### Abstract

Specification of cell fates across the dorsoventral axis of the central nervous system in *Drosophila* involves the subdivision of the neuroectoderm into three domains that give rise to three columns of neural precursor cells called neuroblasts. Ventral nervous system defective (*Vnd*), Intermediate neuroblasts defective (*Ind*) and Muscle segment homeobox (*Msh*) are expressed in the three columns from ventral to dorsal, respectively. The products of these genes play multiple important roles in formation and specification of the embryonic nervous system. *Ind* for example is known to play roles in two important processes. First, *Ind* is essential for formation of neuroblasts conjunction with SoxB class transcription factors. Sox class transcription factors are known to specify neural stem cells in vertebrates. Second, *Ind* plays an important role in patterning the CNS in conjunction with, *vnd* and *msh*, which is also similar to how vertebrates pattern their neural tube. This work focuses two important aspects of *Ind* function. First, we used multiple approaches to identify and characterize specific domains within the protein that confer repressor or activator ability. Currently, little is known about the presence of activation or repression domains within *Ind*. Here we show that transcriptional repression by *Ind* requires multiple conserved domains within the protein, and that *Ind* has a transcriptional activation domain. Specifically, we have identified a novel domain, the Pst domain, that has transcriptional repression ability and appears to act independent of interaction with the co-repressor Groucho. This domain is highly conserved among insect species, but is not found in vertebrate Gsh class homeodomain proteins. Second, we show that *Ind* can and does repress *vnd* expression, but does so in a stage specific manner. We conclude from this that the function of *Ind* in regulating *vnd* expression is one of refinement and maintenance of the dorsal border.

### 1. Introduction

Formation of the nervous system in *Drosophila melanogaster* involves the initial subdivision of the neuroectoderm into three domains, delineated by the expression of three homeodomain transcription factors. These are from ventral to dorsal; *ventral nervous system defective (vnd)*, *intermediate neuroblasts defective (ind)* and *muscle segment homeobox (msh)* (Chu et al., 1998; Isshiki et al., 1997; Mc Donald et al., 1998; Weiss et al., 1998). Expression of these three homeodomain proteins in stripes running the length of the ventral neuroectoderm is a critical step in formation of the embryonic nervous system. Specifically, loss of either *vnd* in the ventral column, or *ind* in the intermediate column, results in failed formation of ventral or intermediate neuroblasts, respectively (Chu et al., 1998; Mc Donald et al., 1998; Weiss et al., 1998). Also, loss of *msh* results in mis-specification of the dorsal neuroblasts to more ventral fates (Isshiki

<sup>1</sup>Author for Correspondence: Division of Biology, Kansas State University, 104 Ackert Hall, Manhattan, KS 66506 (785) 532-6149, email E-mail: tonia@ksu.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

et al., 1997). Thus, proper formation and maintenance of these domains of gene expression is an essential step in formation of the columns of neural precursor cells.

These transcription factors work in conjunction with other proteins to regulate expression of themselves and each other. Both Vnd and Ind have been specifically shown to interact with the co-repressor Groucho (Uhler et al., 2007; Von Ohlen et al., 2007a). Msh also has a putative Groucho interaction domain, suggesting a similar interaction (Smith and Jaynes, 1996). Vnd and Ind both interact genetically and physically with the Sox domain protein Dichaete (Buescher et al., 2002; Overton et al., 2002; Zhao et al., 2007a; Zhao and Skeath, 2002). Thus, the ability of Vnd and Ind to regulate gene expression appears to be intimately associated with their ability to interact with other co-regulators. Furthermore, both Vnd and Ind are required to maintain their own expression (Saunders et al., 1998; Von Ohlen et al., 2007b). The auto-regulatory role for both Vnd and Ind appears to be a maintenance role and not an initiation function. Both Vnd and Ind act on enhancers that control later aspects of expression as opposed to the enhancer elements that are essential for initiation. Therefore, both of these proteins can act as either transcriptional activators or repressors depending on the gene they are regulating and association with interacting factors, such as Groucho and Dichaete. It is also important to note that Ind cannot exclusively be acting as a Groucho dependent repressor. In the ventral and intermediate neurectoderm the co-repressor activity of Groucho is inhibited by MapKinase (Cinnamon et al., 2008).

How the protein products of the DV homeobox genes interact is also in question. One hypothesis is that there is a ventral dominance mechanism controlling formation of the stripes of homeodomain protein expression (Cowden and Levine, 2003). This hypothesis is based on the observation that Vnd represses *ind*, and in turn, Ind represses *msh* (Mc Donald et al., 1998; Von Ohlen et al., 2007a; Weiss et al., 1998). The Cowden and Levine study also claims that ectopic expression of Ind is not able to repress Vnd. It is interesting to note however, that this study only examined embryos at stages 5–7 and not later stages of development. An alternate hypothesis suggests there could be cross-repressive interactions occurring between the DV restricted homeodomains. Specifically, Zhao et al. (2007) observed expansion of Vnd in *ind* mutant embryos after stage 9 (Zhao et al., 2007b). Thus, there appears to be stage specific difference in the regulation of *vnd* expression. These data suggest that there is a mechanism in place to refine or maintain the boundaries of homeodomain gene expression in the developing CNS.

Understanding the mechanisms by which Ind regulates gene expression will provide important information about the function of this gene in formation of the embryonic CNS. Here we investigate the function of conserved and non-conserved domains within Ind protein. This work presents novel data that there is a previously uncharacterized conserved domain, in addition to the previously described Groucho interaction domain (Von Ohlen et al., 2007a), that plays an essential role in the ability of Ind to confer maximal repression on target genes. An assay developed to identify functional domains of Ind also revealed the presence of a transcriptional activation domain. Lastly, data is presented demonstrating that Ind can repress *vnd* transcription, but does so in a stage specific manner.

## 2. Results

### 2.1 Identification of evolutionarily conserved domains within the Ind protein sequence

In order to identify regions of the Ind protein that are potentially important for the proper function of Ind, we chose to look for evolutionarily conserved sequences at the protein level. An alignment of the Ind protein sequences from three insect species (*Drosophila*, *Tribolium castaneum* and *Anopheles gambiae*) revealed that outside the homeodomain there are two additional regions of highly conserved sequence similarity; both are located in regions close

to the amino-termini of the proteins. The sequence of *Anopheles* Ind was found by BLAST searches of AnoBase for sequence similarity to the *Drosophila* Ind protein, while sequence information for *Tribolium* Ind was provided by Dr. James Skeath (Washington University, St Louis MO; (Wheeler et al., 2005). We used Vector NTI software to align the protein sequences for all three insects (Figure 1). Based on sequence information we predicted that the function of the first highly conserved region, highlighted in blue, probably represents the Gro interaction domain. Because this domain was first identified in the transcription factor engrailed we refer to this domain as the engrailed homology domain 1 (Eh1) (Cowden and Levine, 2003; Smith and Jaynes, 1996). This was shown to be the case in our previous work (Von Ohlen et al., 2007a). The second domain, highlighted in red, is also highly conserved among the insect species. However, this domain is not found among vertebrate species or at least is not as highly conserved at the sequence level. The second conserved domain is described as the Pst domain because PstI restriction sites flank the coding region for this domain. This allowed us to make an in-frame deletion from the cDNA. Thus, we have identified potentially important regions of the Ind protein and have tested their function using an in vivo ectopic expression assay in which we are expressing modified versions of Ind lacking specific domains.

Our previous data have shown that ectopic expression of full length Ind is capable of repressing *msh* and *ac* (Figure 2C; (Von Ohlen et al., 2007a). Thus, to assay for function of the modified Ind proteins we examined their effect on expression of the known Ind target, *msh*. Expression of the deleted proteins across the DV axis was accomplished using the Krüppel (Kr) Gal4 driver in an otherwise wild type background. The presence of the endogenous Ind protein in the intermediate column should have no effect on the interpretation of the experiments because we are assaying function of the modified protein in the lateral column where endogenous Ind is not expressed. Previously, we demonstrated that the Eh1 domain is essential for transcriptional repression activity of Ind (Figure 2D; (Von Ohlen et al., 2007a). This data is included here for comparison purposes. The second conserved domain, Pst, bears no significant homology to a known functional domain. This was the first attempt to define a function for this protein domain. Ectopic expression of the Ind $\Delta$ Pst protein product had a reduced ability to repress *msh* transcription (Figure 2E). Thus, this domain appears to contribute to the transcriptional repression activity of Ind. Finally, because the region of Ind C-terminal to the homeodomain was shown to be important for binding to Groucho in our Co-IP assays (Von Ohlen et al., 2007a) we have built an additional deletion construct that produces a truncated form of the protein, Ind $\Delta$ C. To confirm that the protein products are made and persist we tested for production of the Ind protein product. Expression of Ind $\Delta$ C across the DV axis resulted in repression of Msh expression in the Kr domain (Figure 2F). However, this repression ability Ind $\Delta$ C does not appear to be as strong as the full-length protein. This is particularly obvious when examining the ability to repress *vnd* relative to Msh (Compare figure 2F to 2K). This demonstrated that removal of the C-terminal portion may partially influence the ability of Ind to repress expression of the target gene Msh. This suggests that, the C-terminal domain, which facilitates interaction with Groucho in vitro, may also be required for this interaction in vivo. Taken together, these results suggest that Ind contains multiple domains that are each required for conferring maximal transcriptional repression on target genes, including *msh*.

## 2.2 Ind also has a transcriptional activation domain

The above data suggest that the Pst domain has repressor function independent of an interaction with Groucho. Therefore, we expected that this domain would be capable of conferring transcriptional repression activity on an alternative DNA binding domain. To test this hypothesis, required an assay to test for the function of the Pst domain independent of other functional domains within the Ind protein. To do this, we built Gal4 DNA-binding domain (DBD)-Ind fusion proteins. These were expressed during embryogenesis under the control of the *even-skipped* stripe two enhancer (EveSt2) (Figure 3A–D; (Arnosti et al., 1996). We built

four constructs containing different regions of the protein. IndNT includes the entire N-terminal including the Eh1 domain and the Pst domain. IndPst includes the Pst domain. IndP-HD includes the region of Ind between the Pst domain and the homeodomain and IndHD includes the homeodomain and C-terminal. These four constructs encompassed the entire protein-coding region of Ind. Expression of each of the fusion proteins was examined using an anti-Gal4 antibody (see materials and methods). We found that there were varying levels of expression for each of the fusion protein constructs (Figure 3B–D). However, each fusion protein was expressed in embryos at the pattern and at the appropriate stage to determine if they could influence expression of the reporter construct. Thus, because this is a qualitative assay, if they had transcriptional repression or activation ability we could assess that activity in the reporter system used. The reporter construct, which expressed *lacZ* under control of a compound enhancer is described in: (Sutrias-Grau and Arnosti, 2004). Briefly, the CRT4U(–55)*lacZ* reporter transgene includes a combination of Twist binding sites in conjunction with two copies of the *rhomboïd* minimal enhancer and four consensus Gal4 binding sites, was used to test for transcriptional activity of the Gal4-Ind fusions proteins. This construct, CRT4U(–55)*lacZ*, expresses *lacZ* message in the ventral regions of the embryo, consistent with activation by Twist and other transcription factors essential for *rhomboïd* expression (Figure 3F; (Sutrias-Grau and Arnosti, 2004). The presence of the Gal4 binding sites allows for binding of the fusion proteins via the Gal4 DNA binding domain. Once we had established that our Gal4-Ind fusion proteins were expressed in the early embryo in the correct pattern then the lines with the highest levels expression for each fusion construct were crossed to the CRT4U(–55)*lacZ* reporter construct carrying lines. In each case the embryo progeny of the crosses were analyzed for transcription of *lacZ*. For the crosses with the IndNT and IndPst fusion protein we observed that both of these fusion proteins exhibited a weak ability to repress transcription of the CRT4U(–55)*lacZ* reporter construct (Figure 3G & H). In spite of the fact that the observed repression did not completely eliminate *lacZ* expression, the data support the hypothesis that the Pst domain was sufficient to confer repression activity independent of the rest of the protein. In addition, we also found that IndP-HD had transcriptional activation activity. That is, there was ectopic expression of the *lacZ* message in a domain that corresponded to the EveSt2 domain (Figure 3I). The identification of an activation domain is not entirely surprising because Ind is necessary to maintain its own expression (Von Ohlen et al., 2007b). Finally, the fragment of Ind including the homeodomain and C-terminus had no obvious effect on reporter gene expression. The results of these experiments support the hypothesis that Ind can function as both a transcriptional activator and a repressor.

### 2.3 The Pst domain represses target gene expression independent of Groucho binding

The data presented thus far suggest that the Pst domain mediates repression, but does not directly address whether this repression involves the co-repressor, Groucho. We previously reported that full length Ind acts as a potent repressor in transient transfections. Thus, we asked whether the repressor activity of Ind in transient transfections is affected by the absence of the Pst domain. The *ind* constructs were tagged at the amino terminus with the Gal4 DNA binding domain (DBD) using the pBind expression vector. Constructs encoding the Gal4 DBD alone the Gal4-Ind wild type and mutant chimera were transfected into Hek293 cells. Vertebrate rather than *Drosophila* cells were used because of the extremely low transfection efficiency of the latter.

We compared the capacity of the Ind chimeras to regulate expression of a firefly luciferase reporter downstream of a Gal4 UAS in transient transfection assays, while levels of renilla luciferase that was co-expressed by the expression vector was used to monitor transfection efficiency (Fig. 4A). As previously reported, the full length Ind-Gal4 DBD fusion protein acted as a strong repressor relative to the Gal4 DBD alone. Deletion of the Eh1 domain at the N terminus of Ind reduced the repression capacity of the protein 24% similar to the effects of

deleting the Vnd Eh1 domain in this assay. Whereas deletion of the Pst domain caused a significant effect on the repression activity of Ind, reducing the repression 5 fold (Fig. 4B). Thus, both the embryonic and transient transfection data indicate that the Pst domain has repressor activity. Next, we directly addressed whether the Pst domain is necessary for binding of Ind to the co-repressor, Groucho. We generated cell extracts containing either Myc-tagged Groucho or the Gal4-Ind fusion proteins following transient transfection. Then we immunoprecipitated Myc-tagged Groucho and incubated equal amounts of the immunoprecipitate with extracts of the Gal4 DBD Ind fusion proteins. Consistent with the ability of Ind to function as a potent repressor, full length Ind pulled down Groucho strongly relative to the Gal4 DBD alone (Fig. 4C). As previously reported, deletion of the Eh1 domain significantly reduced Ind's capacity to bind Groucho relative to the full length Ind chimera. While deletion of the Pst domain had little effect on the binding of Ind to Groucho, given the difference in the expression levels of the individuals Ind proteins. Thus, these co-immunoprecipitation analyses demonstrate that the Pst domain has minimal effects on the capacity of Ind to bind Groucho, and further suggest that the Pst domain mediates repression independent of Groucho.

#### 2.4 Zebrafish Gsh1 lacks a “Pst domain” and is functionally similar to Ind $\Delta$ Pst protein

Identification of the Pst domain in *Drosophila* and other insect species but not in vertebrates, suggested that expression of a Gsh/Ind protein lacking the Pst domain might not effectively repress transcription in *Drosophila*. To test this hypothesis, we obtained a full length *Zebrafish Gsh1* cDNA (Cheesman and Eisen, 2004). Our data that deletion of the Pst domain from the *Drosophila* gene markedly, but not completely reduces the ability of Ind to repress *msh* expression, suggests that Gsh1 should repress *msh* expression, but not as effectively as full length Ind. Gsh1 should still bind Groucho and mediate Groucho-dependent repression. Initially, we compared expression levels of the Gsh1 transgenic lines to full length Ind line using the Gsh1 antibody. Once we had identified lines with similar expression levels we then tested those lines for repression activity. The data in Figure 5 suggest that this transgene represses transcription at a similar level as the  $\Delta$ Pst construct but not the full length Ind (Figure 5C). Next, we wanted to rule out the possibility that the phenotype observed when Gsh1 is over expressed was simply due to different DNA binding affinity on the Msh regulatory elements. Therefore, we constructed a fusion protein in which we fused the N-terminal region of zGsh1 to the homeodomain and C-terminus of Ind was constructed (Gind; Figure 5A). The fusion protein also partially repressed Msh expression (Figure 5D).

Our hypothesis that the Pst domain confers additional repression activity on Ind independent of the Eh1 domain predicts that addition of the Pst domain from Ind to zGsh1 would render Gsh1 a stronger repressor. To test this we built two additional transgenes, IndGsh1 fusion 1 includes the entire N-terminal region up to the homeodomain of Ind fused to the homeodomain and C-terminus of zGsh1. IndGsh1 fusion 2, includes the N-terminal of Ind including the Eh1 domain and Pst domain fused to Gsh1 minus its first 42 amino acids. We again used KrGal4 to ectopically express both fusion 1 and fusion 2 across the DV axis. In either case we found that addition of the portion of Ind, including the Pst domain, to zGsh1 rendered the protein a stronger repressor of Msh expression than the unaltered Gsh1 (Figure 5E & F). Therefore, the Pst domain of Ind assists in conferring maximal repression activity of Ind/Gsh1 proteins in *Drosophila*.

#### 2.4 Ind confers stage specific repression on vnd

There is some discrepancy as to whether Ind can repress expression of *vnd*. (Cowden and Levine, 2003) reported Ind is not able to repress *vnd*. However, (Zhao et al., 2007b) reported that *vnd* is expanded dorsally in *ind* mutant embryos. To resolve this discrepancy we chose to investigate the ability of ectopically expressed Ind to repress *vnd* expression. At early stages, prior to stage 9, Ind is not capable of repressing *vnd*. However, after stage 9 there is moderate



to strong repression of *vnd* message in the Kr Gal4 expression domain (Figure 6). These data suggested that Ind was capable of repressing *vnd* but does so in a stage-specific way. Interestingly, there was not strong repression of *vnd* at the protein level. The repression is less obvious when we use the anti-Vnd antibody (data not shown); this result suggests that the Vnd protein lasts longer than the message and supports the hypothesis that regulation of *vnd* by Ind is essential for maintenance of the dorsal boundary, but Ind is not essential for initial establishment of that boundary.

To further investigate the ability of Ind to repress Vnd transcription we examined the effect of Ind expression on *vnd-lacZ* reporter constructs. Regulation of *vnd* expression is controlled by at least two separable enhancer elements. An early neuroectodermal enhancer (NEE) located within the first intron is known to control initiation of *vnd* expression (Stathopoulos et al., 2002). An additional element located upstream is responsible for later aspects of *vnd* expression including expression in neuroblasts and repression in the midline (Estes et al., 2001; Saunders et al., 1998; Shao et al., 2002). Based on the result that Ind represses *vnd* expression at later stages of development, we predicted that ectopic expression of Ind would result in repression of the upstream reporter constructs but not the VndNEE construct. For these experiments multiple versions of the upstream enhancer were used with similar results obtained for each reporter construct tested. Indeed, while Ind can repress transcription of the reporter gene from the upstream element ( $-4.7/-2.8$  *vndlacZ*), Ind fails to be able to repress the VndNEE element (Figure 7 compare B & D). Thus, as predicted the function of Ind in regulating *vnd* expression seems to be acting on the late element and not the early element.

There are temporal differences in regulation by Ind for the two target genes *msh* and *vnd*. Ind represses *msh* from very early stages of development, however Ind can only repress *vnd* at stage 10 and later. Therefore, we wanted to determine if there were differences in the ability of the modified Ind proteins to regulate *vnd* as compared to their ability to regulate *msh*. Thus, *vnd* mRNA was assayed in each of the full length and modified Ind mis-expression contexts. We found that full length Ind is able to repress *vnd* after stage 9, but the modified proteins had varying ability to repress *vnd*. Ind $\Delta$ Eh1 failed to inhibit *vnd* expression and the Ind $\Delta$ Pst had a reduced ability to repress *vnd* (Figure 2G–K). These data suggest that the ability of Ind to repress *vnd* is exclusively a function Ind binding to the different enhancer elements acting at different stages of development and not because of differential repressor activity of Ind itself. Thus, we propose that while a ventral dominance mechanism governs initiation of the stripes of homeodomain gene expression, at later stages cross repressive interactions are required to maintain the precise boundaries between these genes.

### 3. Discussion

The function of Ind in development of the embryonic nervous system is multifold. Initially, Ind serves to define the intermediate column of the neuroectoderm, this subsequently leads to formation of the corresponding neuroblasts. Here we present data that transcriptional repression activity by Ind involves at least two transcriptional repression domains, suggesting that Ind represses transcription via Groucho dependent and Groucho independent mechanisms. There are two highly conserved domains in the N-terminal region of the Ind protein. Both appear to be essential for maximal repression activity of Ind. In addition, we have identified a third domain that is capable of conferring transcriptional activation ability on a heterologous DNA binding domain. Also, we present data demonstrating that Ind functions to define and maintain this domain via transcriptional repression of other columnar genes *vnd* and *msh*. Suggesting that, depending on which enhancer it is bound to and possibly association with co-factors, Ind can act as either a transcriptional repressor or as activator. Finally, data from our previous manuscript and this work demonstrate that an Ind protein lacking the Eh1 domain but retaining the Pst domain fails to physically interact with purified Groucho protein (Von Ohlen

et al., 2007a). Furthermore, the Gal4-Ind $\Delta$ Eh1 protein was still a strong repressor of transcription in cultured cells (Von Ohlen et al., 2007a). These results in conjunction with data presented here strongly support the hypothesis that the Pst domain confers repressor activity independent of Groucho interaction. However, we cannot rule out the possibility that the Pst domain also plays a role in stabilizing the interaction with Groucho or association with other co-factors in vivo.

It is not surprising that Ind has incorporated additional repressor activities that are independent of Groucho activity. Formation of the intermediate column of neuroblasts is also dependent on the activity of the Egfr signaling pathway (Skeath, 1998). Specifically, in *egfr* mutant embryos the intermediate column of neuroblasts fails to form because Ind is not expressed (Von Ohlen and Doe, 2000). The readout for activation of the Egfr pathway is the presence of the activated form of MapKinase (dpErk). DpErk is detected in the ventral and intermediate columns of the neurectoderm at the early stages of development (Gabay et al., 1997; Skeath, 1998). Interestingly, the activation of DpErk appears to correlate with down regulation of Groucho activity. Specifically, MapKinase directly phosphorylates Groucho and this phosphorylation of Groucho results in reduced co-repressor activity (Cinnamon et al., 2008). Since Groucho activity is down-regulated in the region where Ind is expressed and Ind is a Groucho dependent transcriptional repressor, additional repression activity may be necessary to overcome the effects of Egfr signaling on Groucho activity.

Formation of the proper complement of neuroblasts in the embryonic nervous system of *Drosophila* and other insects is essential for the proper development of the organism. Initially the neuroblasts form in three columns that correspond to the domains of *vnd*, *ind* and *msh* expression. Therefore, formation of the stripes of homeobox gene expression is essential for the ultimate formation of the CNS. While there is an apparent ventral dominance mechanism in place to initiate the expression of these genes, there is also a cross repressive relationship that is essential for maintaining the boundaries between the domains of gene expression. Ind only represses *vnd* at stages 9 and 10 of embryonic development and not earlier. This coincides with differences in the ability of Ind to repress *vnd* reporter gene expression. Thus, Ind can repress transcription from enhancer elements located upstream that regulate expression in neuroblasts. However, Ind is unable to repress transcription of *lacZ* message from reporter constructs that include the *vnd* NEE, which is essential for initiation of *vnd* expression. In conclusion, the temporal differences in the ability of Ind to repress transcription of *vnd* reflect a role for Ind maintaining the boundary between Vnd and Ind domains. However, Ind was not required for establishing the dorsal border of *vnd* expression at the earliest stages of embryogenesis.

## 4. Experimental Procedures

### 4.1 Fly lines used

UAS *ind* (Von Ohlen et al., 2007b).  $-8.4/+0.35$  NK2 *lacZ* and  $-4.7/-2.8$  NK2 *lacZ* from M. Nirenberg (NIH) (Shao et al., 2002). UAS *msh* (Isshiki et al., 1997). UAS *ind* $\Delta$ Eh1 and *KrGal4* (Von Ohlen et al., 2007a).

### 4.2 Construction of Deleted Ind and zGsh1 transgenes

**Ind $\Delta$ Pst**—An internal inframe deletion was generated by cloning Ind from pUAST (Von Ohlen et al., 2007b) into pBluescriptKS at the Kpn we and Xho we sites. This construct was then cut with PstI and the large linear fragment was recircularized. This was then inserted back into the pUAST vector (Brand and Perrimon, 1993) with Kpn we and Xho we restriction enzymes.

**IndΔC**—The C-terminal deletion was made using PCR primers as follows: indC F ccccgaaaccaagatg, and indC R ctagcccttcttctgcttgac.

This PCR product was cloned into pUAST at KpnI and XhoI restriction sites.

**Gsh1**—Full-length zebrafish *Gsh1* cDNA (Genbank accession number [AY486348](#)) was kindly provided by Sarah Cheesman (Cheesman and Eisen, 2004). *Gsh1* cDNA was cloned into pUAST at the EcoRI site.

*Gind* fusion: the 5' end of the *Gsh1* cDNA was amplified using the following primers: Gshind F; CCGCTCGAGCTAACAGACATCAGG and Gshind R; AGTGC GGATCCTTTTGCTGCTCTGAAG. PCR product was then cloned into *ind* cDNA in pBluescript between the XhoI and BamHI sites. The fusion protein was then ligated into pUAST as an XhoI KpnI fragment. Expression of protein from each of these constructs was confirmed using an anti-Ind antibody. We have recently developed a Rabbit anti-Ind antibody that is directed against the homeodomain of Ind there is capable of recognizing all the versions of Ind protein used in this manuscript (data not shown).

**IndGsh1 fusion 1**—The Homeodomain and C-terminal of ZGsh1 was PCR amplified from the cDNA with the following primers: IndGshF1 AGAAGGATCCGCACTGCGTTCACCAGC and Ind GshR CGGGGTACCATGTTACGGGCTGACG. The PCR product was inserted into the Ind cDNA between the BamHI site and the KpnI site, this replaced the Ind homeodomain and C-terminal sequences with the Gsh1 sequences.

**IndGsh1 fusion 2**—A larger fragment of the Gsh1 cDNA was amplified using the IndGshR primer and the following forward primer: ATAAGAATGCGGCCCGCCAGGCTCCTGCCACTCTCG. This PCR product was inserted into the Ind cDNA between the NotI and KpnI sites, this replaced most of the coding sequences of the Ind protein with Gsh1 sequences. Each fusion construct was then sequenced to ensure no frame shifts were introduced. Then the full fusion coding sequences were cloned in to pUAST between the XbaI and KpnI sites.

**VndNEE lacZ**—PCR primers to amplify VndNEE were designed as described in (Markstein et al., 2004). The product was amplified from gDNA isolated from wild type adult flies, and cloned in the pCaSpeR4lacZ vector.

### 4.3 In vivo assay

Gal4 DNA binding domain fusion proteins were built as described in (Von Ohlen et al., 2007a), using the pBIND vector (Invitrogen). The constructs used in (Von Ohlen et al., 2007a) were PCR amplified from the pBIND vector. The Gal4-ind fusions were then cloned into the pCaSpeR4 at the Stu I site. A construct containing two copies of the EveSt2 enhancer was obtained from Hilary Ashe (Manchester UK) (Kosman and Small, 1997). The two EveSt2 elements were cloned in the CaSpeR4 at the Not I site. The Hsp70 minimal promoter was excised from pCaSHsp-43 *lacZ* with XhoI. For the reporter line, transgenic flies carrying the CRT4U (−55) *lacZ* transgene were obtained from Dr. David Arnosti (Michigan State). This construct is described in (Sutrias-Grau and Arnosti, 2004). Briefly, the transgene contains two copies of the *rhomboid* neurectodermal enhancer (R), multiple binding sites for the mesoderm specific transcriptional activator Twist (T) and four consensus Gal4 binding sites (4U) located approximately 55bp upstream of the *lacZ* transcription start site. Expression was tested using a rabbit anti-Gal4 antibody (Santa Cruz Biotechnology)



All transformation constructs were injected at the Non-Mammalian model systems unit at Duke University. Multiple independent transformants were tested for each construct. For all the UAS misexpression constructs the lines that had expression levels comparable to the full length UAS *ind* construct were used in the experiments shown. For the Gal4 DBD fusion proteins all lines were tested for expression with the anti-Gal4 antibody. The lines that gave the highest levels of expression for each construct are shown.

#### 4.4 Transfections, luciferase assays, and Western analyses

Hek 293 cells were maintained in DMEM containing 10% fetal calf serum and 1% penicillin/streptomycin. For transient transfections, cells were plated at 60% confluency into 24-well plates and transfected the next day with Fugene-6 (Roche) according to the manufacturer's instructions. The total amount of DNA added was 0.4  $\mu\text{g}$  per well. Cells were harvested approximately 48 h post-transfection. Luciferase activity was measured using the Dual Luciferase Kit (Promega) according to the manufacturer's protocol. The pG5luc reporter vector (Promega) and pBind-*ind* expression vectors were added at a molar ratio of 1:1. The pGEM3Z vector (Promega) was added to bring the final amount of DNA to 0.4  $\mu\text{g}$  per well. Each transfection was performed in triplicate. For Western analyses, samples were pooled and subjected to SDS-PAGE and Western blotted. The Gal4 DNA binding domain alone or fused to Ind sub-domains was detected using a Gal4 specific antibody (Santa Cruz). Binding of peroxidase-conjugated secondary antibodies was detected by chemiluminescence, using the Lightning kit (Perkin-Elmer). Results are expressed as mean values  $\pm$  the standard error of the mean (SEM).

#### 4.5 Co-immunoprecipitation of Gal4-Ind chimeras with Myc-tagged Groucho and Flag-tagged Dichaete

Immunoprecipitations were performed as described in Von Ohlen et al. (2007) Briefly, dishes (100 mm) containing Hek 293 cells were independently transfected with 6  $\mu\text{g}$  of pBind (Promega) that encodes the Gal4 DNA binding domain, the pBind-*ind* constructs or GS2-*groucho* using 18  $\mu\text{l}$  of Fugene. Cell lysates were prepared 48 h after transfection using immunoprecipitation (IP) buffer (20 mM Tris-HCL, 100 mM NaCl, 10 mM NaF, 1 mM Na<sub>3</sub>VPO<sub>4</sub>, 1 mM PMSF), and proteinase inhibitor cocktail (Roche) containing 0.5% Triton X-100. Ind Gal4 DNA binding domain chimeras were precipitated from cell lysates using an antibody that recognizes the Gal4 DNA binding domain (Santa Cruz Biotechnology) rotating overnight at 4°C. The Gal4 DNA binding domain alone was also precipitated to monitor non-specific binding of protein to beads. Following incubation with protein A/G PLUS agarose (Santa Cruz Biotechnology) at 4°C for 2 h, beads were precipitated by centrifugation, and washed three times with IP buffer containing 0.1% Triton X-100. Then, beads were divided into two aliquots, one of which was subjected to Western analysis directly, the second incubated with equal amounts of cell lysate containing Myc-tagged Groucho. The co-immunoprecipitates were rotated for 2 h at 4°C, precipitated by centrifugation, and washed three times with IP buffer containing 0.1% Triton X-100. Immunoprecipitates were separated by SDS-PAGE electrophoresis, transferred to Immobilon-P (Millipore) and Western blotted. Duplicate blots were incubated with anti-Myc antibody (Sigma) to detect the Myc-tagged Groucho or a Gal4-specific antibody (Santa Cruz Biotechnology) to detect Gal4-Ind chimeras. Binding of peroxidase-conjugated secondary antibodies was detected by chemiluminescence using the Lightning kit (Perkin-Elmer).

### Acknowledgments

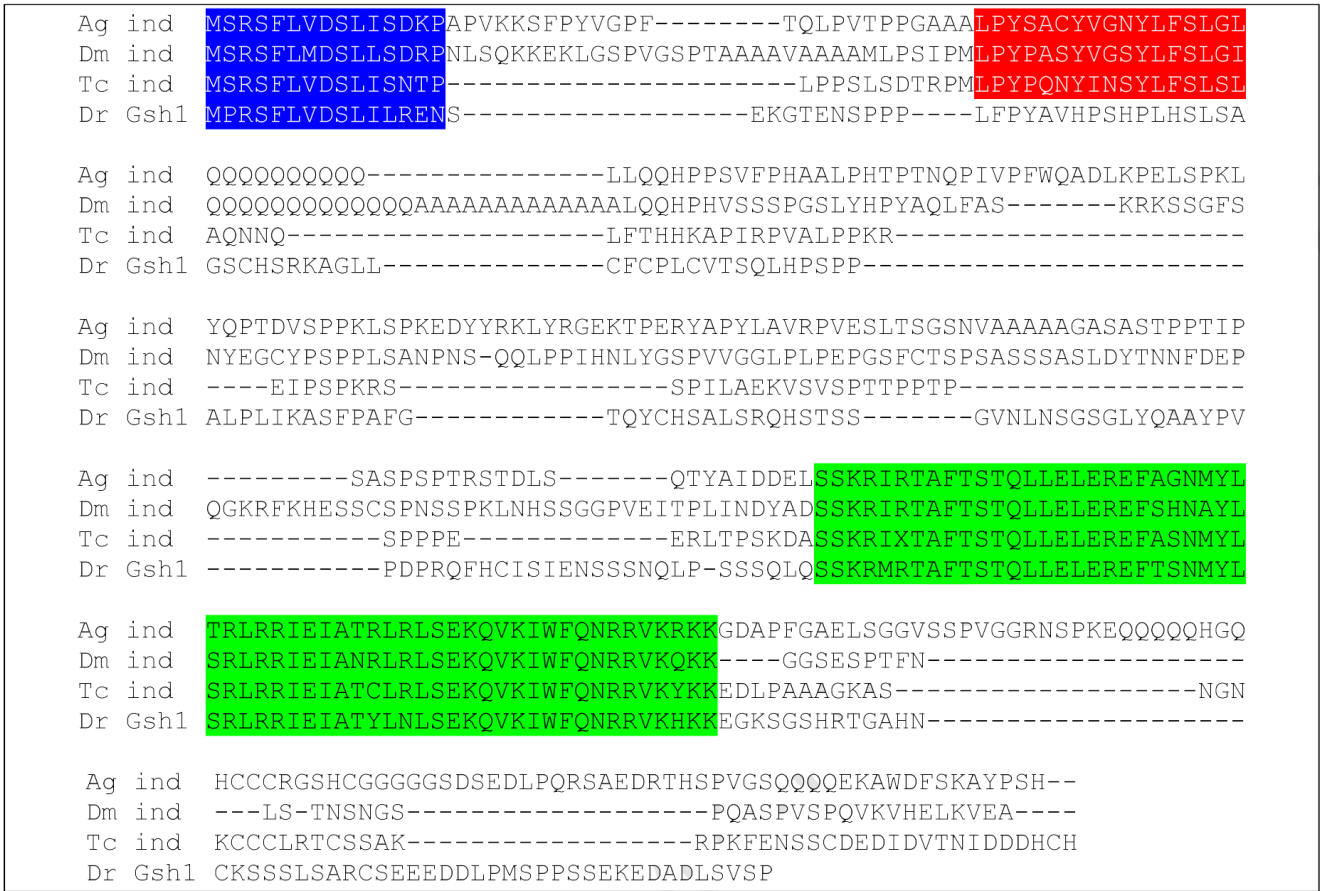
We sincerely thank Dervla Mellerick for technical assistance and critical reading of the manuscript. Additional thanks go to S. Keith Chapes and Chris Doe for critical reading of the manuscript and to Sarah Cheesman and Judith Eisen for the Gsh1 cDNA. This publication was in part, made possible by Grant Number P20 RR016475, to TVO, from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). The contents

of this publication are solely the responsibility of the author and do not necessarily represent the official views of NCRN or NIH. An Innovative Research Award, to TVO, from the Terry C. Johnson Center for Basic Cancer Research at Kansas State University, provided additional funding.

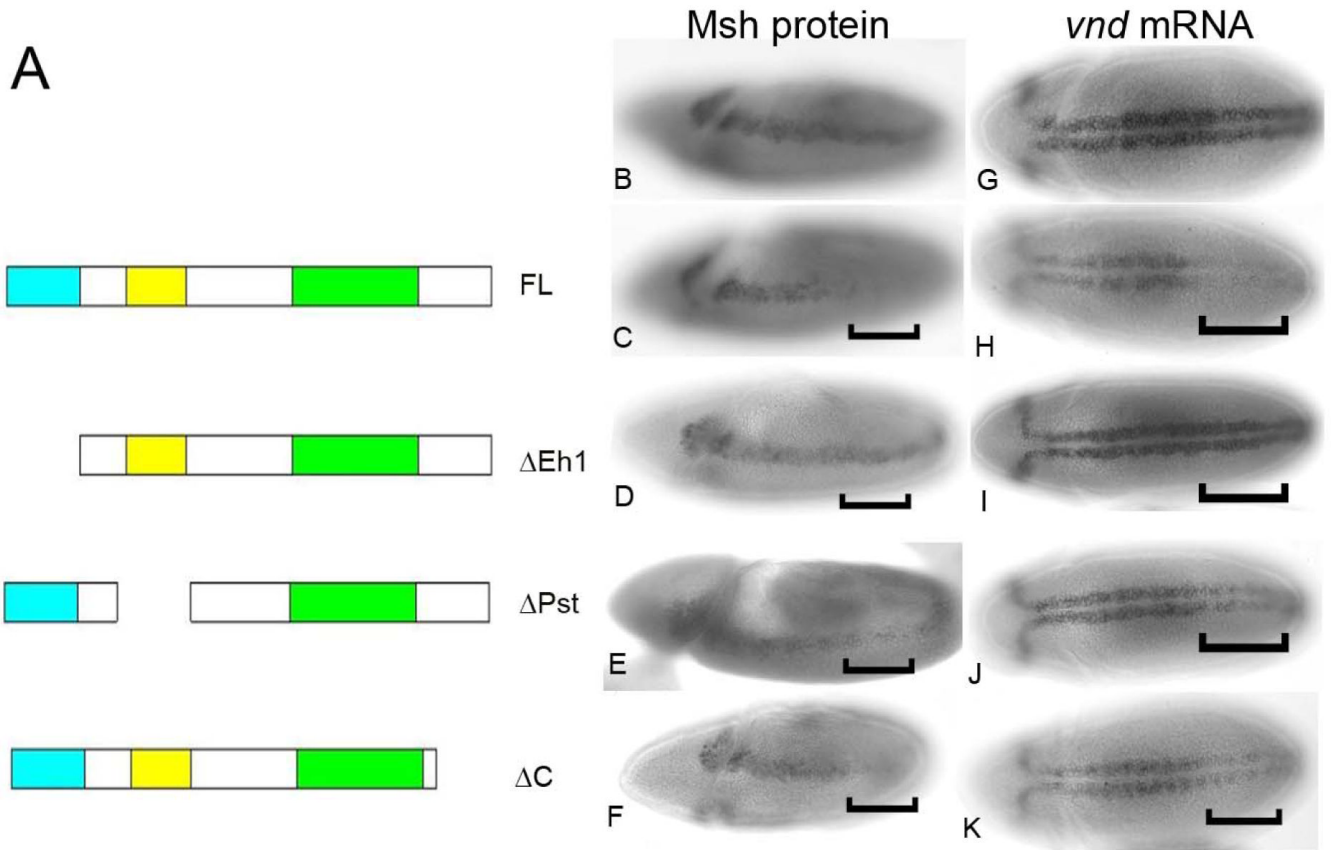
## Literature Cited

- Arnosti DN, Barolo S, Levine M, Small SJ. The eve strip 2 enhancer employs multiple modes of transcriptional synergy. *Development* 1996;122:205–214. [PubMed: 8565831]
- Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 1993;118:401–415. [PubMed: 8223268]
- Buescher M, Hing FS, Chia W. Formation of Neuroblasts in the embryonic central nervous system of *Drosophila melanogaster* is controlled by *SoxNeuro*. *Development* 2002;129:4193–4203. [PubMed: 12183372]
- Cheesman SE, Eisen JS. *gsh1* demarcates hypothalamus and intermediate spinal cord in zebrafish. *Gene Expr Patterns* 2004;5:107–12. [PubMed: 15533825]
- Chu H, Parras C, White K, Jimenez F. Formation and specification of ventral neuroblasts is controlled by *vnd* in *Drosophila* neurogenesis. *Genes Dev* 1998;12:3613–3624. [PubMed: 9832512]
- Cinnamon E, Helman A, Ben-Haroush Schyr R, Orian A, Jimenez G, Paroush Z. Multiple RTK pathways downregulate Groucho-mediated repression in *Drosophila* embryogenesis. *Development* 2008;135:829–37. [PubMed: 18216172]
- Cowden J, Levine M. Ventral dominance governs sequential patterns of gene expression across the dorsal-ventral axis of the neuroectoderm in the *Drosophila* embryo. *Dev Bio* 2003;262:335–349. [PubMed: 14550796]
- Estes P, Mosher J, Crews ST. *Drosophila* single-minded represses gene transcription by activating the expression of repressive factors. *Dev Biol* 2001;232:157–75. [PubMed: 11254355]
- Gabay, I; Seger, R.; Shilo, BZ. MAP kinase in situ activation atlas during *Drosophila* embryogenesis. *Development* 1997;124:3535–3541. [PubMed: 9342046]
- Isshiki T, Takeichi M, Nose A. The role of the *msh* homeobox gene during *Drosophila* neurogenesis: implication for the dorsoventral specification of the neuroectoderm. *Development* 1997;124:3099–3109. [PubMed: 9272951]
- Kosman D, Small S. Concentration-dependent patterning by an ectopic expression domain of the *Drosophila* gap gene *knirps*. *Development* 1997;124:1343–54. [PubMed: 9118805]
- Markstein M, Zinzen R, Markstein P, Yee KP, Erives A, Stathopolous A, Levine M. A regulatory code for neurogenic gene expression in the *Drosophila* embryo. *Development* 2004;131:2387–2394. [PubMed: 15128669]
- Mc Donald JA, Holbrook S, Isshiki T, Weiss J, Doe CQ, Mellerick DM. Dorso-ventral patterning in the *Drosophila* CNS: the *vnd* homeobox gene specifies ventral column identity. *Genes Dev* 1998;12:3606–3612.
- Overton PM, Meadows LA, Urban J, Russell S. Evidence for differential and redundant function of the *Sox* genes *Dichaete* and *SoxN* during CNS development in *Drosophila*. *Development* 2002;129:4219–4228. [PubMed: 12183374]
- Saunders HMH, Koizumi K, Odenwald W, Nirenberg M. Neuroblast pattern formation: regulatory DNA that confers the *vnd/NK2* homobox gene pattern on a reporter gene in transgenic lines of *Drosophila*. *PNAS* 1998;95:8316–8321. [PubMed: 9653184]
- Shao X, Koizumi K, Nosworthy N, Tan DP, Odenwald W, Nirenberg M. Regulatory DNA required for *vnd/NK-2* homeobox gene expression pattern in neuroblasts. *Proc Natl Acad Sci U S A* 2002;99:113–7. [PubMed: 11752402]
- Sketh JB. The *Drosophila* EGF receptor controls the formation and specification of neuroblasts along the dorsal-ventral axis of the *Drosophila* embryo. *Development* 1998;125:3301–3312. [PubMed: 9693134]
- Smith ST, Jaynes JB. A conserved region of engrailed, shared among all en- *gsc*- *Nk1*-, *Nk2* and *msh*-class homeoproteins, mediates active transcriptional repression in vivo. *Development* 1996;122:3141–3150. [PubMed: 8898227]

- Stathopoulos A, Van Drenth M, Erives A, Markstein M, Levine M. Whole-genome analysis of Dorsal-ventral patterning in the *Drosophila* embryo. *Cell* 2002;111:687–701. [PubMed: 12464180]
- Sutrias-Grau M, Arnosti DN. CtBP contributes quantitatively to Knirps repression activity in an NAD binding-dependent manner. *Mol Cell Biol* 2004;24:5953–66. [PubMed: 15199149]
- Uhler J, Zhang H, Syu LJ, Mellerick DM. The Nk-2 box of the *Drosophila* homeodomain protein, Vnd, contributes to its repression activity in a Groucho-dependent manner. *Mech Dev* 2007;124:1–10. [PubMed: 17070676]
- Von Ohlen T, Doe CQ. Convergence of Dorsal, Dpp and Egfr signaling pathways subdivides the *Drosophila* neuroectoderm into three dorsal-ventral columns. *Dev Bio* 2000;224:362–372. [PubMed: 10926773]
- Von Ohlen T, Syu LJ, Mellerick DM. Conserved properties of the *Drosophila* homeodomain protein, Ind. *Mech Dev* 2007a;124:925–934. [PubMed: 17900877]
- Von Ohlen TL, Harvey C, Panda M. Identification of an upstream regulatory element reveals a novel requirement for Ind activity in maintaining ind expression. *Mech Dev* 2007b;124:230–236. [PubMed: 17224261]
- Weiss JB, Von Ohlen T, Mellerick D, Dressler G, Doe CQ, Scott MP. Dorsoventral patterning in the *Drosophila* Central Nervous System: The intermediate neuroblasts defective Homeobox gene specifies intermediate column identity. *Genes Dev* 1998;12:3591–3602. [PubMed: 9832510]
- Wheeler SR, Carrico ML, Wilson BA, Skeath JB. The *Tribolium* columnar genes reveal conservation and plasticity in neural precursor patterning along the embryonic dorsal-ventral axis. *Dev Bio* 2005;279:491–500. [PubMed: 15733674]
- Zhao G, Boekhoff-Falk G, Wilson BA, Skeath JB. Linking pattern formation to cell-type specification: Dichaete and Ind directly repress achaete gene expression in the *Drosophila* CNS. *Proc Natl Acad Sci U S A* 2007a;104:3847–52. [PubMed: 17360441]
- Zhao G, Skeath JB. The Sox-domain containing gene *Diachaete/fish-hook* acts in concert with vnd and ind to regulate cell fate in the *Drosophila* neuroectoderm. *Development* 2002;129:1165–1174. [PubMed: 11874912]
- Zhao G, Wheeler SR, Skeath JB. Genetic control of dorsoventral patterning and neuroblast specification in the *Drosophila* Central Nervous System. *Int J Dev Biol* 2007b;51:107–15. [PubMed: 17294361]

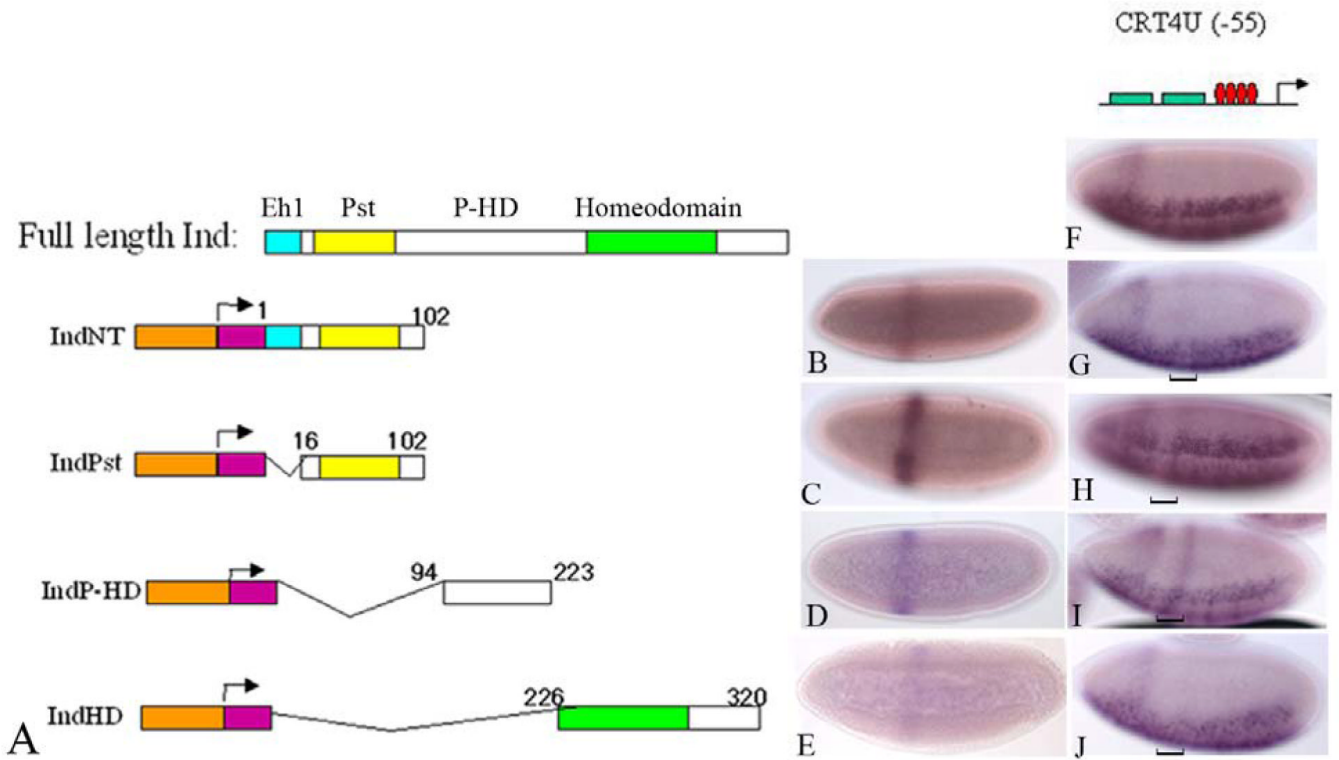


**Figure 1.** Alignment of Ind protein sequences for *Anopheles*, *Drosophila* and *Tribolium* reveals regions of high sequence identity. This predicts an important function for these regions. Sequences for Zebrafish Gsh1 have been added for comparison to vertebrate Gsh family members. Homeodomains are highlighted in green. The domain highlighted in blue resembles a Gro interaction domain. The domain highlighted in red is a novel domain found in insect Ind proteins.



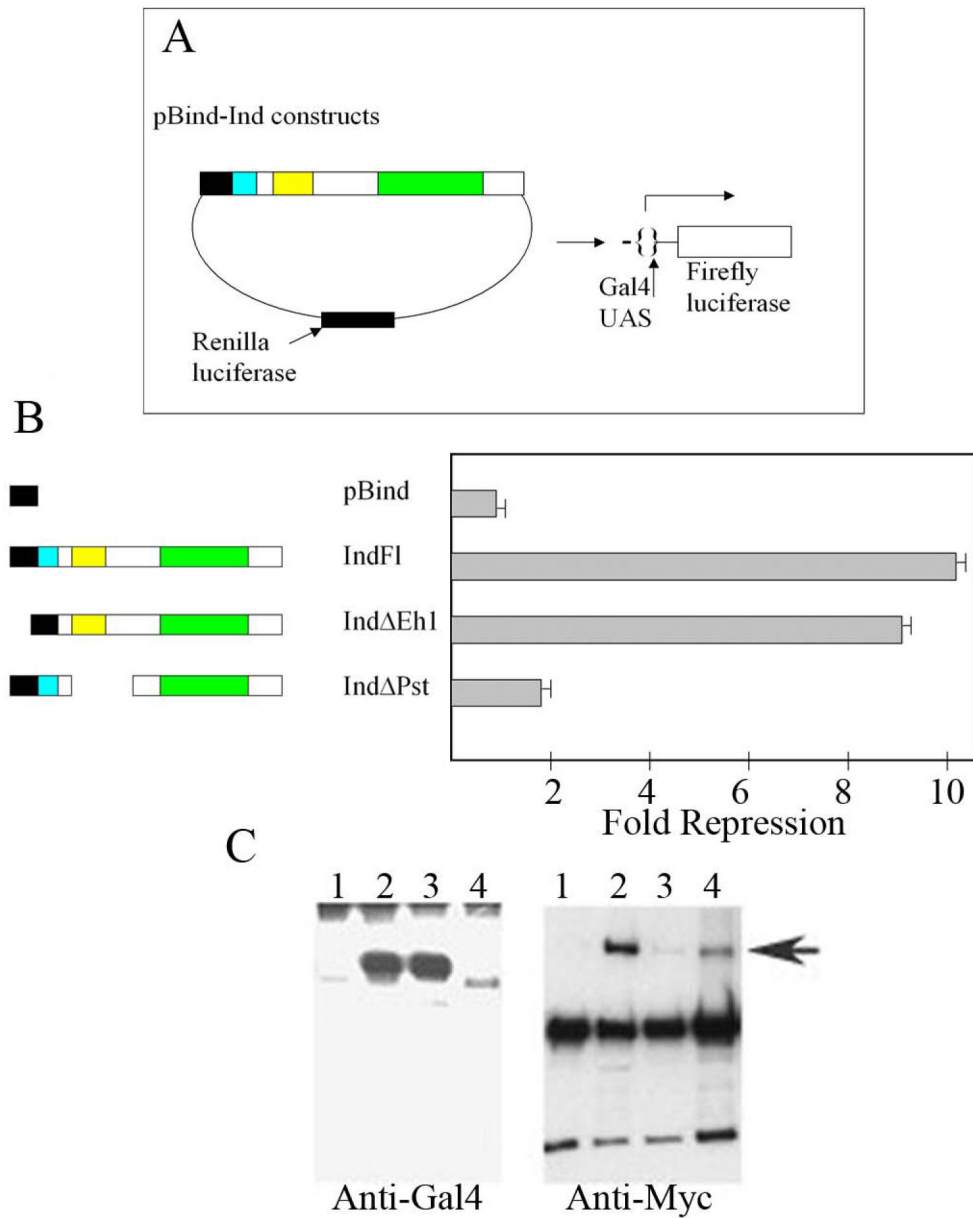
**Figure 2.** Modified Ind proteins have altered ability to repress expression of Ind target *msh* *in vivo*. A) Ind deletion constructs. Deletion constructs of each of the conserved domains with significant sequence similarity among the Ind genes from different species were generated. Ind $\Delta$ Eh results in removal of the N-terminal, domain. Ind $\Delta$ Pst removes the second domain showing significant sequence identity. B–F) Antibody stains showing expression of Msh protein on stage nine embryos. Brackets indicate approximate location of Kr Gal4 expression domain B) Wild type. C) KrGal4 UAS ind, Full length Ind strongly represses Msh expression, bracket indicates Kr expression domain. D) KrGal4 UAS ind $\Delta$ Eh no repression of Msh is observed. E) KrGal4 UAS ind $\Delta$ Pst, has reduced ability to repress Msh. F) KrGal4 UAS ind $\Delta$ C, retains ability to repress Msh expression. G–K) *vnd* mRNA expression, stage 10 embryos brackets indicate approximate location of Kr-gal4 expression domain. G) Wild type *vnd* mRNA. H) KrGal4 UAS ind, Full length Ind strongly represses *vnd* expression. I) KrGal4 UAS ind $\Delta$ Eh no repression of *vnd* is observed. J) KrGal4 UAS ind $\Delta$ Pst, has reduced ability to repress *vnd*. K) KrGal4 UAS ind $\Delta$ C, retains ability to repress *vnd* expression.





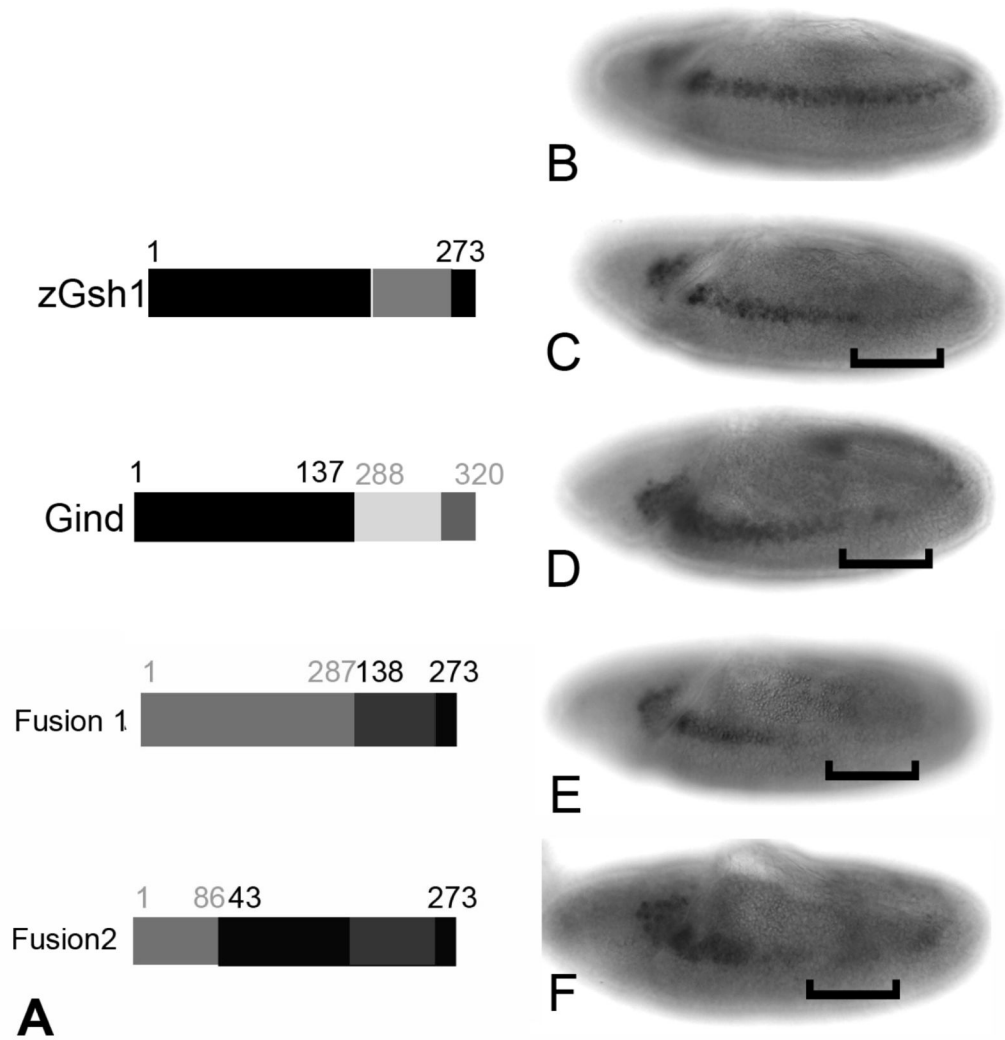
**Figure 3.**

Invivo assay for Ind functional domains reveals two repression domains and an activation domain. A) Schematic of expression constructs. Orange is Even-skipped stripe two enhancer. Purple indicates the Gal4 DNA binding domain. Turquoise indicates the Eh1 domain. Yellow is Pst domain and green is the homeodomain. Numbers correspond to the amino acid positions within each construct, based on the full length sequence. B–E) Anti Gal4 antibody stains on expression construct lines. B) IndNT, C) IndPst, D) IndP-HD, E) IndHD. F–J) lacZ message on CRT4U(-55) lacZ lines crossed to Ind Gal4 DBD fusion lines. F) WT CRT4UlacZ. G) crossed to IndNT. H) IndPst, I) IndP-HD, J) IndHD. Brackets indicate the location of the Eve St2 domain where the proteins are expressed.

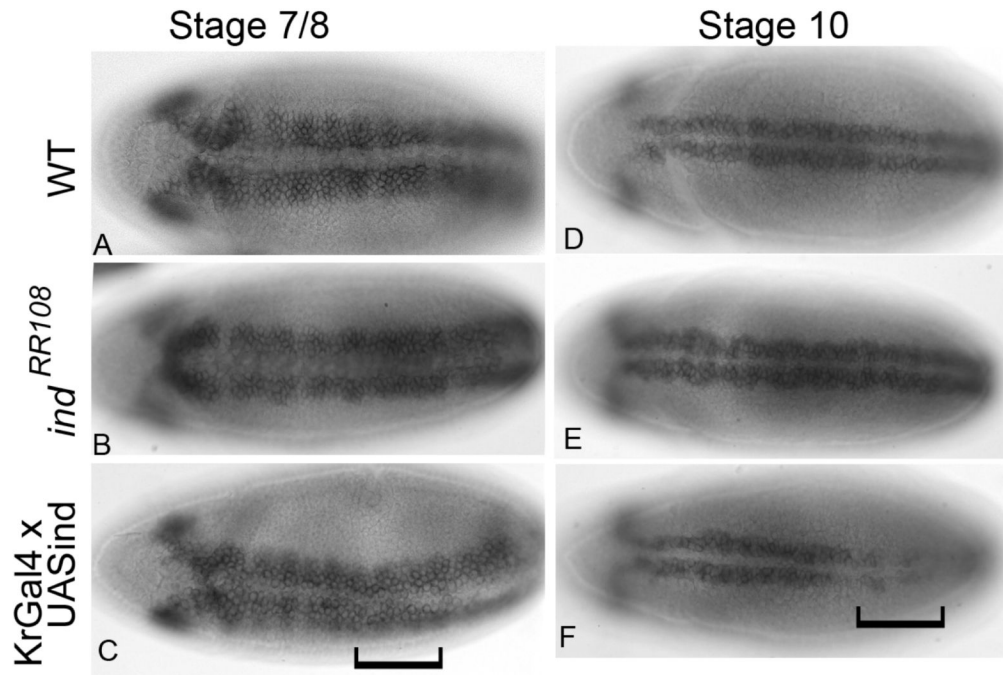


**Fig. 4.** The Pst domain is required for efficient repression in tissue culture cells, but not for Groucho binding. (A) Schematic showing the reporter assays. Note that the expression vector encodes renilla luciferase; wholes levels are used to measure transfection efficiency, while the GAL4 UAS drives expression of firefly luciferase. (B). Left. Schematics showing the *ind* Gal4 DBD constructs, 1–4, tested for regulatory activity. 1. is pBind that encodes the GAL4 DNA binding domain, 2 is full length (FL) ind, 3. and 4. are the  $\Delta$ EH1 and the  $\Delta$ Pst1 constructs, respectively. Right. The repression activity is expressed as fold repression relative to the empty pBind vector (1). Note the  $\Delta$ Pst1(4) repression activity is significantly compromised relative to the full length ind construct (2) (C). Left. Western blots of cell extracts containing constructs 1–4 (according to B above) following immunoprecipitation. Note that the level of the  $\Delta$ Pst1 protein is significantly less than the full-length protein or the  $\Delta$ EH1 protein. Right. Blot of immunoprecipitates incubated

with an anti-Myc antibody, which recognizes Myc-tagged Groucho, highlighted with arrow. The Gal4 DBD alone fails to bind Groucho (1). Thus, non-specific bead binding is not an issue. Non-specific bands in lane 1–5 correspond to antibody cross-reaction with immunoprecipitate. Full length Ind pulls down Groucho (2), in contrast to the construct lacking only the Eh1 domain, which pulls down minimal amounts of the co-repressor (3). Deletion of the Pst domain has minimal effects on Groucho binding.

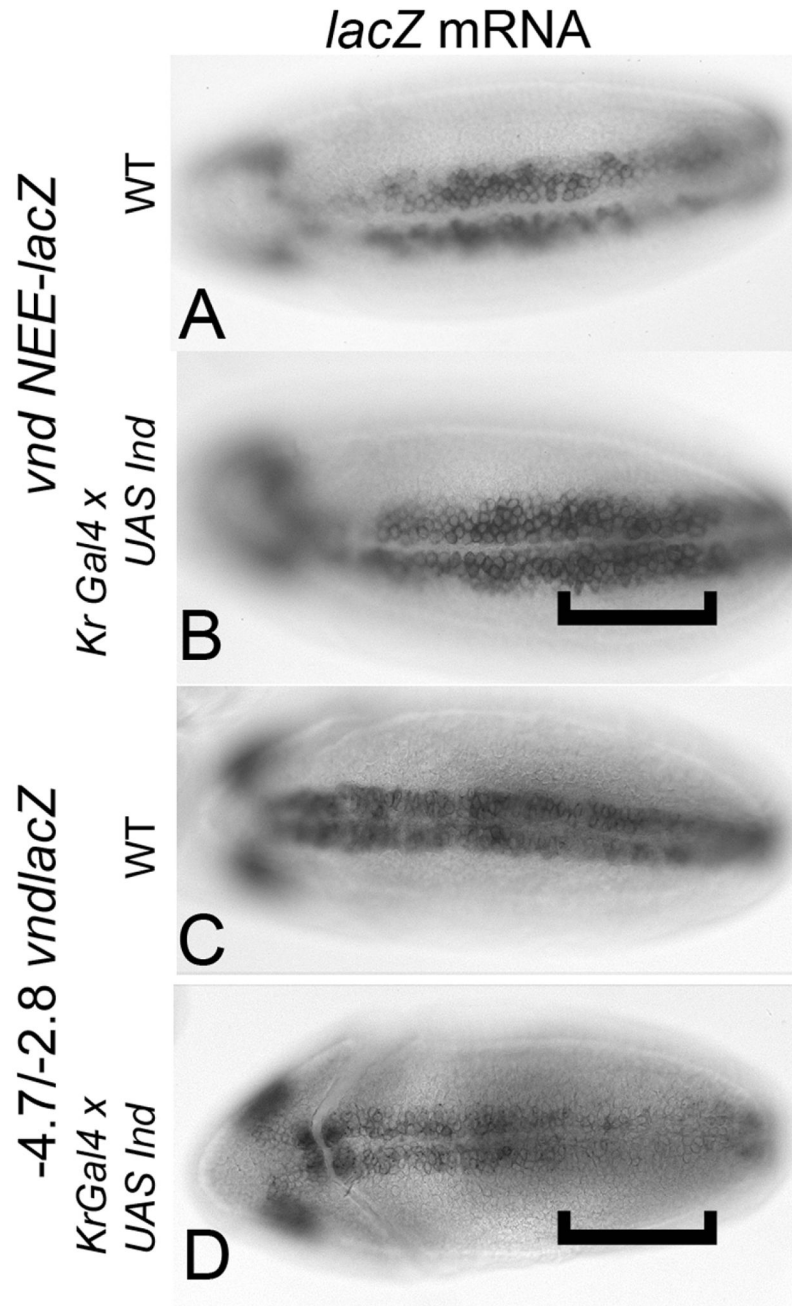


**Figure 5.** Expression of Gsh1 in *Drosophila* embryos partially reproduces effect of Ind expression. A) Graphic depiction of proteins made by expression constructs. Top is full-length zebrafish Gsh1, amino acid positions are shown. Bottom is Gsh1-Ind fusion protein. The N-terminal 137 amino acids are fused to the homeodomain and C-terminal amino acids of Ind. The numbers of the corresponding amino acids in Ind are shown in grey. B) Wild type Msh expression. C) Msh protein in a KrGal4UASGsh1 embryo. D) Msh protein in KrGal4UASGind embryo. E) Msh protein in KrGal4UASIndGsh1 fusion 1. F) Msh protein in KrGal4UASIndGsh1 fusion 2. All embryos are stage 9 and anterior is to the left.



**Figure 6.** Repression of Vnd by Ind occurs in a stage specific manner. A–F) *vnd* mRNA A–C) stage 7–8 embryos D–F) stage 10 embryos. A and D) wild type. B and E) *ind<sup>RR108</sup>* embryos. C and F) KrGal4 x UASind embryos. Anterior is to the left and all are ventral views. Brackets indicate domain of KrGal4 expression. Ind fails to repress Vnd expression in early embryos (prior to stage 9) at later stages Ind is capable of efficiently repressing Vnd expression.





**Figure 7.** Ind differentially represses Vnd reporter gene expression. Bracket indicates position of Kr domain. A) Expression of VndNEE lacZ in wildtype embryo. B) Expression of VndNEE lacZ in Kr-Ind embryo. C) Expression of Vnd-Upstream element (-4.7/-2.8 vndlacZ (Saunders et al., 1998) in a wild type embryo. D) Expression of Vnd-Upstream lacZ in Kr-Ind embryo. All embryos are stage 9, ventral views. Anterior is to the left.