

***Drosophila* Paired regulates late *even-skipped* expression through a composite binding site for the paired domain and the homeodomain**

Miki Fujioka¹, Pawel Miskiewicz², Lakshmi Raj², Alyssa A. Gulledge^{1,†}, Michael Weir² and Tadaatsu Goto^{1,*}

¹Kimmel Cancer Institute, Thomas Jefferson University, 1020 Locust Street, Philadelphia, PA 19107, USA

²Department of Biology, Wesleyan University, Middletown CT 06459, USA

*Author for correspondence (e-mail: T_Goto@lac.jci.tju.edu)

†Current address: Curriculum in Genetics, University of North Carolina, Chapel Hill, NC 27599, USA

SUMMARY

The *even-skipped* (*eve*) pair-rule gene plays a key role in the establishment of the anterior-posterior segmental pattern of the *Drosophila* embryo. The continuously changing pattern of *eve* expression can be resolved into two phases. Early expression consists of seven broad stripes in the blastoderm embryo, while late expression, which occurs after cellularization, consists of narrow stripes with sharp anterior borders that coincide with the odd-numbered parasegment boundaries. Previous studies have shown that these two phases are controlled by separate classes of *cis* elements in the *eve* promoter. Early stripes are expressed by multiple stripe-specific elements under the control of maternal-effect genes and gap genes, while late stripes are expressed by a single regulatory element, the 'late element', under the control of pair-rule genes including *eve* itself. We report here that *paired* (*prd*), a pair-rule gene which had

been considered to be below *eve* in the regulatory hierarchy of pair-rule genes, in fact plays a critical role in the regulation of late *eve* expression. Transgenic analysis shows that this regulation is largely mediated by an evolutionarily conserved sequence within the late element termed PTE (Paired Target Element). In vitro analysis shows that the Prd protein binds strongly to this sequence. Interestingly, PTE contains juxtaposed binding sites for the two DNA-binding domains of the Prd protein, the paired domain and the homeodomain. Mutagenesis of either binding site leads to significant reduction in the activity of the late element, indicating that both DNA-binding domains in the Paired protein are required for regulation.

Key words: *Drosophila* segmentation, homeodomain, paired domain, Paired, Pax protein, *even-skipped*

INTRODUCTION

During early *Drosophila* development, a cascade of regulatory interactions between the segmentation genes results in a progressive refinement of spatial information (Nusslein-Volhard and Wieschaus, 1980; Akam, 1987; Ingham, 1988). This refinement eventually gives rise to the reiterated segmental domains defined by the striped expression of the pair-rule and segment-polarity genes. In this cascade, reiterated spatial information first arises with the striped expression of the primary pair-rule genes, *hairy*, *runt* and *even-skipped* (*eve*). According to the hierarchical model of pair-rule genes, the coarse striped information of the primary pair-rule genes is then transformed into refined stripes of the secondary pair-rule and segment-polarity genes. Understanding the molecular regulatory mechanisms of the establishment of the stripe patterns by primary pair-rule genes, and the refinement by the secondary pair-rule genes and segment-polarity genes, is key to understanding how the segmentation gene cascade defines positional information in the early embryo.

The promoters of the primary pair-rule genes, *eve* and *hairy*, have been shown to have separate *cis*-regulatory elements for

each of the individual stripes (Howard et al., 1988; Goto et al., 1989; Harding et al., 1989; Howard and Struhl, 1990; Pankratz et al., 1990; Riddihough and Ish-Horowitz, 1991), reflecting the fact that the expression stripes of the primary pair-rule genes are each regulated by different combinations of maternal coordinate and gap gene expression. Furthermore, molecular studies of some of these stripe-specific *cis* elements have identified binding sites for the products of the maternal coordinate and gap genes (Stanojevic et al., 1989; Small et al., 1991). In contrast, the promoters of the secondary pair-rule genes and segment-polarity genes need not, in principle, have stripe-specific elements, since the striped regulatory inputs for these genes are already repetitive in nature. Instead, expression of these genes is thought to occur through *cis*-regulatory elements responsible for multiple stripes. Examples of these multistripe elements include the 'zebra' element of the secondary pair-rule gene, *fushi tarazu* (*ftz*), responsible for all seven *ftz* stripes (Hiromi et al., 1985), and the even-parasegment *cis* element of the segment-polarity gene *engrailed* (*en*) (DiNardo et al., 1988). Regulation of multistripe *cis* elements is likely to involve combinatorial mechanisms. However, little is known as to how this is achieved at the molecular level.

In the past few years, it has become clear that the picture presented thus far, equating the primary pair-rule genes with establishment of stripes and secondary pair-rule genes with refinement, is over-simplified in that the refinement of stripe information also occurs within the promoters of individual genes through separate *cis* elements responsible for different temporal phases of expression. For example, transgenic studies of *eve* have shown that this gene has two regulatory programs, early and late, regulated by two distinct classes of *cis*-acting elements in the *eve* promoter (Goto et al., 1989). The 'early elements' (E) are stripe-specific and define the initial (early) expression of each of its stripes, which are located in the odd-numbered parasegments. In contrast, the 'late element' (L), responsible for later expression of *eve* at gastrulation, is a single multistripe element responsible for the regulation of all seven late *eve* stripes in the odd-numbered parasegments (*eve* also has secondary minor stripes in the even-numbered parasegments whose regulation is not understood and not considered here). Since the early elements (E) give rise to broad stripes whereas the late element (L) gives narrow stripes, this results in the spatial refinement of *eve* stripes over time: their anterior borders become sharply demarcated with concomitant loss of expression from the posterior. Other segmentation genes also exhibit refinement of their expression stripes over time suggesting that they may also have separate *cis* elements for different phases of their expression (Carroll and Scott, 1986; Coulter et al., 1990; Grossniklaus et al., 1992; Klingler and Gergen, 1993; Gutjahr et al., 1994; Yu and Pick, 1995).

The late element (L) of *eve* provides an accessible multistripe control element for molecular dissection. This late element, located more than 5 kb upstream of the transcription start site, is a *cis*-acting site for regulation by other pair-rule genes (Goto et al., 1989; Warrior and Levine, 1990) as well as for *eve* autoregulation (Goto et al., 1989; Harding et al., 1989; Jiang et al., 1991). Recent work has shown (Fujioka et al., 1995) that the pair-rule gene *paired* (*prd*) is negatively regulated by early *eve* stripes in a concentration-dependent manner. This study also suggested the possibility that the *prd* gene product (Prd) may in turn activate late *eve* expression, although this was not predicted by previous genetic studies that had placed *prd* at the bottom of the regulatory hierarchy of the pair-rule genes (Baumgartner and Noll, 1991). Here, we report that late *eve* expression is, indeed, positively regulated by Prd through a direct interaction with a conserved sequence in the late element. This is one of the first identified *in vivo* target sequences of Prd, a Pax protein with two DNA-binding domains, a paired domain and a homeodomain. Strikingly, the site is apparently a composite of a paired domain and a homeodomain binding site.

MATERIALS AND METHODS

Drosophila strains

Flies of the genotype *y w; +/+; +/+* were used as host for injection of C3D-based constructs (see below). The injection procedures are as described previously (Rubin and Spradling, 1982). The allele of *prd* used for mutant analysis was *prd^Δ* (= *prd^{2.45.17}*), which has a 1.1 kb insertion into the paired box (Frigerio et al., 1986) and shows no protein staining with a polyclonal antibody directed against the C-terminal half of the Prd protein (Bertuccioli et al., 1996). The *lacZ*-

marked balancer chromosome *CyO^{ptb}* used to distinguish mutant embryos was generated (Morrissey et al., 1991) by mobilizing a *Hunchback*-promoter/*lacZ* P-element (Driever et al., 1989) onto the *CyO* chromosome.

Construction of transgenes and preparation of DNA for injection

A transgene vector, called C3D, was constructed by removing restriction sites for *Pst*I, *Sal*I, *Xba*I and *Bam*HI within the polylinker sequence of pCaSpeR3 (Thummel et al., 1988). A fragment containing the *eve* promoter fragment (from -275 to +166), *lacZ* and the poly(A) attachment site of the *tubulin* α subunit gene was cloned into the *Bg*III and *Eco*RI sites of C3D in such a way that the directions of the transcription of the *w* (*white*) marker gene and the β -galactosidase gene are opposite. This 'injection plasmid', called C3DZ, contains *Xba*I and *Not*I sites just upstream of the *eve-lacZ* fusion gene, which provides a cloning site for the late element (L) or its deletion derivatives.

The LmLd fragment, from the *Eco*RV site (-5653; converted to a *Bg*III site) to *Kpn*I site (-4799), was cloned into a modified pSP72 vector (the *Sma*I site in the polylinker was converted to a *Bg*III site and the *Bg*III site to a *Not*I site). This plasmid (RVKp) therefore contains the LmLd sequence flanked by *Xba*I and *Not*I sites and was used for creating deletions shown in Fig. 2. Most endpoints were created by cleavage with restriction enzymes identified in Fig. 2. Three endpoints (-5532, -5446, -5349) were created by *Bal*31 exonuclease digestion. One endpoint (-5541) was created by polymerase chain reaction (PCR) using the synthesized primer EL12 (5'-TCGAAGGATCCACTCACCGTGGCTAATTGC-3'), which creates a *Bam*HI site just upstream of the PTE.

The *in vitro* mutagenesis of the PTE was also carried out with the RVKp plasmid and synthetic primers using standard PCR procedures (Sambrook et al., 1989). For the PTE-pd and PTE-phd mutations, 58-mer oligonucleotides containing mutated PTE and the surrounding sequences were synthesized: PTE-pd primer, (5'-CTCCCCGGCCCTTTATTGATGGGCAATTAGCCGCTCGAGCGGCGTTTCGAGTGAG-3'); PTE-phd primer, (5'-AGGAATTCCTCACTGCCCGGGCCCTTTATTGATGGGCGAGTCCGACACGGTGAGTGGCG-3'). The PCR reaction was performed with one of these primers plus the SP6 promoter primer with the RVKp template. Each product was cleaved with either *Xba*I and *Sma*I (PTE-pd) or *Xba*I and *Eco*RI (PTE-phd), and was cloned into the RVKp plasmid cut with the same set of enzymes. The LmLd sequence was then transferred to the injection plasmid C3DZ. For creating the PTE-hd mutation, two complementary primers containing the mutated PTE and the surrounding sequences were used: (5'-CCGTGGCTCGAGGCCCATCAATAAAGGGC-3') and (5'-GATGGGCCCTCGAGCCCGGTGAGTGGCG-3'). The first primer was combined with the T7 promoter primer, the second primer with the SP6 promoter primer, and PCR reactions were carried out with the RVKp template. Products from this first round were mixed and the second round PCR reaction was performed with SP6 and T7 promoter primers. The product of the second round was cleaved with *Xba*I and *Eco*RI, and cloned into the RVKp plasmid as described above. Fidelity of all PCR reactions and subsequent cloning steps were confirmed by DNA sequencing. To create internal deletion 12 (-5541 to -5503; Fig. 2A), the synthetic oligonucleotide EL22 (5'-TGAGTGGATCCTTCGAGTGCAGG-3') was used in conjunction with EL12 (above) to create a *Bam*HI site at -5541. The *Sma*I site (-5503) of this plasmid was converted to a *Bam*HI site by ligation of *Bam*HI linker to *Sma*I-cleaved plasmid. The sequence between these two *Bam*HI sites were removed by *Bam*HI digestion before reclosure. Plasmid DNAs for injection were purified with polyethylene glycol as described by Sambrook et al. (1989) with minor modifications. DNA was filtered through a Probind filter (Millipore) before injection.

Cloning of the *D. hydei* *eve* gene

Radioactive probes homologous to the *D. melanogaster* *eve* gene were

made using random primers and restriction fragments from the late element and the coding sequence. A phage genomic library of *D. hydei* (gift of D. Maier) (Maier et al., 1990) was screened with the probes using low stringency hybridization procedures. Hybridization was carried out with 2× SSC, 0.1% SDS, 1× blocking reagent (Boehringer Mannheim) at 50°C overnight, and the filters were washed with 0.5× SSC, 0.5% SDS at 60°C for three times, 15 minutes each. Three independent positive clones were isolated for further analysis. DNA was made from these clones using plate lysates (Sambrook et al., 1989). Southern blot hybridization analysis shows that they together contain 24 kb of the putative *D. hydei eve* locus. *EcoRI* fragments of the inserts were cloned into the pSP72 vector. Southern blot hybridization was used to identify the locations of the late element and the coding sequence in the subclones. DNA sequencing was carried out using either the Sequenase sequencing system (US Biochemicals) or an automated sequencer (Applied Biosystems). Comparative analysis of DNA sequences (Fig. 1) was performed using the GCG software (Genetics Computer Group). The nucleotides are numbered from an upstream *Clal* site. To test the transcriptional activity in vivo of the *D. hydei* DNA, two DNA fragments (9 to 1594, and 788 to 1594, corresponding to L and LmLd in *D. melanogaster*, respectively) were cloned into C3DZ as described above. Their activities in vivo were found to be equivalent; an example of the stripe-forming activity of the shorter version (788 to 1594) is shown in Fig. 2F.

mRNA and protein localization

In situ hybridization to whole-mount embryos using digoxigenin-(DIG-) labeled probes was performed as described (Tautz and Pfeifle, 1989). After hybridization, embryos were washed with ethanol to reduce background (Manoukian and Krause, 1992). DIG-labeled antisense mRNA and DNA probes were used for hybridization and were visualized using alkaline phosphatase-conjugated anti-DIG antibody (Boehringer Mannheim). Embryos were mounted in Fluoromount (Southern Biotechnology) or Aqua-Poly/Mount (Polysciences).

Experiments with *hs-prd* constructs were performed as described (Morrissey et al., 1991) using two 10 minute heat treatments. To obtain homozygous *prd⁻* embryos with one copy each of *hs-prd* and *L-lacZ*, we examined embryos from the cross: *prd⁴/CyO^{phb}; hs-prd X prd⁴/CyO^{phb}; L-lacZ*.

Electrophoretic mobility shift analysis

³²P-labelled probes used in EMSA were 26-mer fragments containing the wild-type or mutant PTE sequences. A histidine-tagged Prd protein fragment (amino acids 27-273) was prepared from bacteria as described (Novagen, Madison WI). 25 minute binding reactions with 0.1-1 ng of labelled DNA were performed at 20°C in 15 mM Tris-HCl pH 7.5, 60 mM KCl, 0.5 mM DTT, 0.25 mg/ml BSA, 0.05% NP-40 and 7.5% glycerol in a volume of 20 µl. We included 100 ng/reaction of poly(dIdC) (Sigma) to minimize non-specific binding by the PD and HD, based on preliminary tests with single PD (Jun and Desplan, 1996) and HD (P2, Wilson et al., 1993) sites using proteins with a mutated HD or PD (Miskiewicz et al., 1996) (data not shown). Binding reactions were electrophoresed on an 8% native polyacrylamide gel (29:1 acrylamide:bis) using 0.25× TBE at 4°C and subjected to autoradiography.

RESULTS

L contains evolutionarily conserved sequences

A regulatory element from the *NdeI* site of the *eve* gene at -6415 (measured from the transcriptional start site) to the *KpnI* site at -4799 was shown previously to be fully functional in driving the expression of a linked reporter gene in seven

narrow stripes, mimicking late *eve* expression (Goto et al., 1989). We refer to this region as the late element (L). In order to identify potentially important sites within L, we undertook a cross-species comparison of sequences based on the assumption that regulatory regions crucial for gene function would diverge more slowly than other non-coding regions in the course of evolution (Wilson et al., 1987). The *eve* gene from *D. hydei* was cloned using the coding and L sequences of the *D. melanogaster* gene as probe. The L sequences of *D. hydei* and *D. melanogaster* share a number of islands of conservation. These conserved islands form a cluster from -6080 to -4875, as shown in Fig. 1. Based on the pattern of the conservation, we divided the *D. melanogaster* L into three regions, Lu (Lu), Lmiddle (Lm) and Ldown (Ld), with the *EcoRV* site (-5653) marking the Lu/Lm boundary, and the *PstI* site (-5306) the Lm/Ld boundary. Lm contains the Minimal Autoregulatory Sequence (MAS), consisting of two in vitro Eve-binding regions (EVE-D and EVE-P) and two nuclear factor-binding sites (DENF-1 and DENF-2), which have previously been shown to be important for *eve* autoregulation (Jiang et al., 1991). The EVE-D and EVE-P sequences are well conserved. The DENF-1 and DENF-2 sites were previously found to be essential for *eve* autoregulatory activity, since deletion of either site eliminated the stripe-forming activity of an MAS multimer (Jiang et al., 1991). It has been proposed that the proteins that bind to the DENF-1 and DENF-2 sites are GAGA factor and Tramtrack, respectively (Jiang et al., 1991). However, the consensus binding sequence for GAGA factor within the DENF-1 site is not conserved, and the DENF-2 site is absent, apparently deleted, in *D. hydei*. Similar findings have been made in a parallel study of sequence conservation between *D. melanogaster* and the Hawaiian *Drosophila*, *D. picticornis* (Sackerson, 1995). No function is known for the other observed islands of conservation in Fig. 1. The importance of the conserved nucleotides in vivo is underscored by the finding that the *D. hydei* L apparently has full late activity in the *D. melanogaster* embryo (see below and Fig. 2F).

Deletion analysis of the late element reveals a critical site within L

In addition to this evolutionary comparison, we initiated a dissection of the late element using transgenes in which altered late elements were placed upstream of the *eve* promoter-proximal region (-275 to +166) contiguous to the *lacZ* reporter. The in vivo activities of altered L elements were tested by staining whole-mount embryos of transgenic lines for β-galactosidase (β-gal) mRNA. We found that the *EcoRV-KpnI* fragment (LmLd) is fully active as a late element, as judged by the staining intensity of the embryos (see Fig. 2B), as previously reported (Goto et al., 1989). We also found that Lu+Lm is active, whereas Lu+Ld is not (data not shown). Thus, Lu and Ld are redundant in that either one can augment Lm for activity. Despite this redundancy, a comparison of the conserved sequences in Lu and Ld (Fig. 1) does not reveal obvious homology, although short stretches of sequence similarity can be noted.

To further dissect late function, we created deletions in the LmLd fragment and tested their activities in vivo (Fig. 2A). Examples of the activities of several constructs are shown in Fig. 2B-E. The presence or absence of activity was scored by

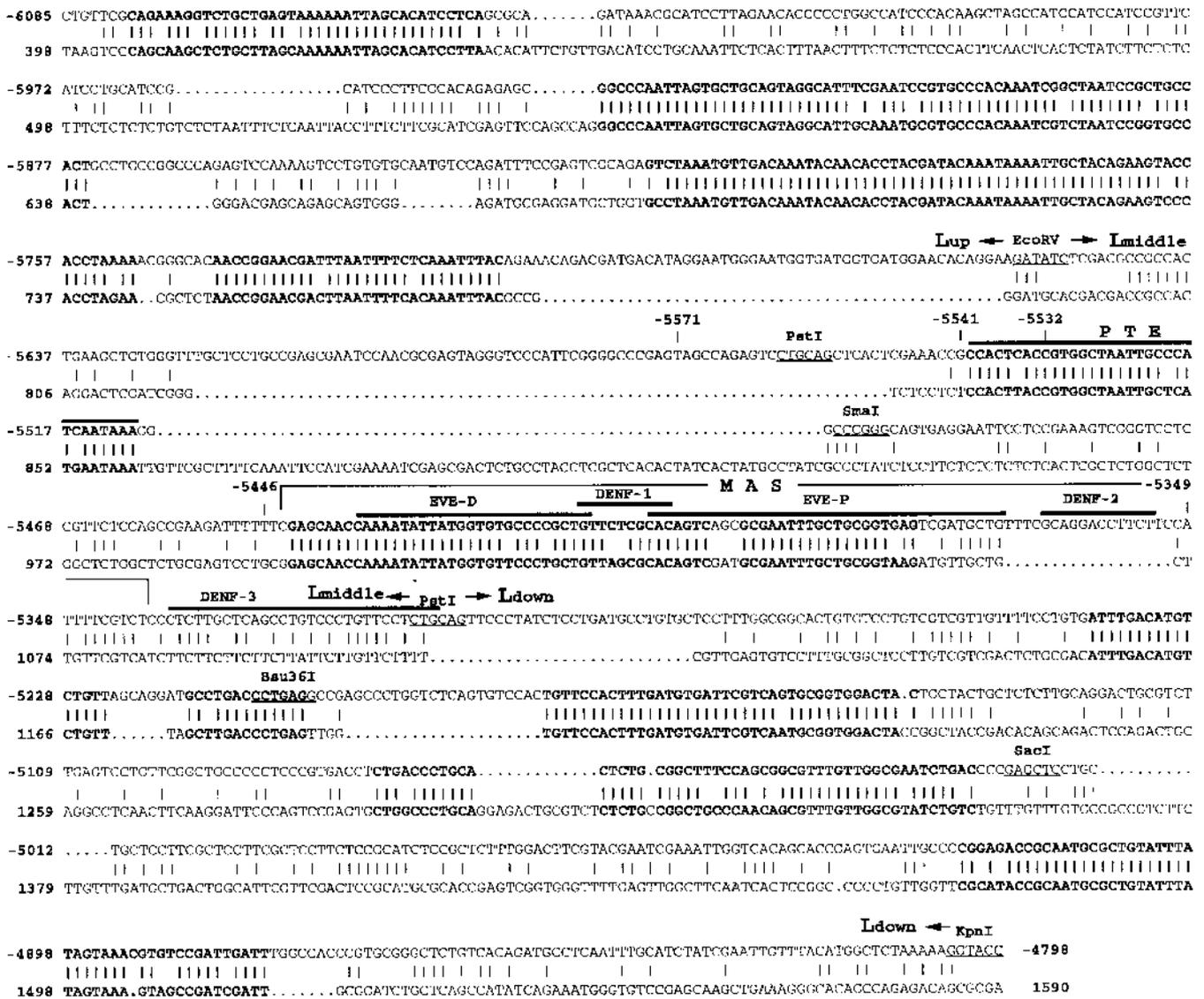


Fig. 1. Several clusters of sequence in the late element are conserved between *D. melanogaster* and *D. hydei*. The *D. melanogaster* sequence (top) and the *D. hydei* sequence (bottom) are aligned. Vertical lines indicate matches of nucleotides. Dots denote gaps in alignment. Stretches of 10 or more bases with at least 80% identity are highlighted in bold. The sequence shown includes all the islands of conservation between the *NdeI* site (–6415 of the *D. melanogaster* coordinate; not shown) and the *KpnI* site (–4799) that meet this criterion. The *D. hydei* nucleotides are numbered from an upstream *Clal* site (not shown). The MAS and the PTE (see text) are identified. Only the MAS has been characterized previously: a hexamer of the MAS can drive the expression of a linked reporter gene in stripes (Jiang et al., 1991). The MAS contains two Eve-binding regions, EVE-D and EVE-P, and binding sites for two embryonic nuclear factors, DENF-1 and DENF-2. A mutation in either the DENF-1 or DENF-2 site eliminates the stripe forming activity of the MAS. A combination of mutations in the EVE-D and EVE-P sites also eliminates the activity, but mutation in either one alone has no effect, suggesting that they may have redundant functions (Jiang et al., 1991). The binding site for another nuclear factor, DENF-3, is also shown, but this site has been shown to be dispensable.

the intensity of staining. We note that some of the constructs scored as ‘negative’ (marked with an asterisk in Fig. 2A) do express extremely weak stripes. These weak stripes are sometimes incomplete and are typically detectable at only later stages due in part to accumulation of the stable β -gal product. However, as illustrated in Fig. 2, the levels of expression by these ‘negative’ constructs (e.g. embryos Fig. 2D,E) were extremely weak compared to those of transgenes scored as active (embryos Fig. 2B,C). Moreover, these scoring criteria were consistent with the results of a second, independent ‘localized rescue’ assay. In this assay, *eve* regulatory regions

were used to drive the *eve* coding sequence and the ability of late elements to augment activity of early expression elements in rescuing the *eve* deficiency phenotype were assayed (see Fig. 2 legend and Fujioka et al., 1995). Analysis of a subset of the constructs in Fig. 2A revealed that elements scored as positive in the first assay (e.g. LmLd) had full rescue activity in the second assay, whereas elements with very weak β -gal stripes (e.g. Lm) showed only partial rescue. We note that the very weak β -gal stripes of Lm were previously scored as positive by less stringent criteria (Harding et al., 1989).

In our deletion analysis (Fig. 2), the absence of activity by

deletion 1 confirms that the MAS, and possibly surrounding sequences, is critical for late activity. The activities of the next three deletions (deletion 2-4) indicate that at least two subfragments of Ld (-5349 to -5206, or -5018 to -4799) can independently augment Lm for activity. Deletion 5 shows that the upstream-most 97 bp of Lm can be deleted with minimal loss of activity. A further deletion of 24 bp (deletion 6) drastically reduces activity, identifying the location of a critical sequence between -5556 and -5532. The next two deletions (deletions 7 and 8) show that this conclusion is not dependent on the choice of Ld subfragments (Fig. 2C,D). The activity of deletions 9 and 10 further narrows the upstream limit of the newly identified critical sequence to -5541. The importance and uniqueness of the sequence in question are also underscored by the elimination of late activity by the small deletions 11 and 12 (deletion 12 is shown in Fig. 2E). Deletion 13 sets the downstream limit of the critical sequence at -5503, and shows further that the interval between the newly identified critical sequence and the MAS can be deleted without effect. Thus, this deletion analysis identifies three separate regions

within LmLd to be critical for the *cis*-activity: the 38 bp upstream element between -5541 and -5503, the MAS, and either of the redundant Ld subfragments.

Prd binds to the PTE sequence

As shown in Fig. 1, the functionally critical 38 bp element between -5541 and 5503 contains a stretch of 32 nucleotides of which 29 are conserved between *D. melanogaster* and *D. hydei*. (These nucleotides are also conserved between *D. melanogaster* and *D. piticornis*; Sackerson, 1995) As shown in Fig. 3, the conserved region (labeled as PTE in the figure) is also found to be homologous to e5, a site within the *eve* promoter that has been shown to be bound by Prd and Eve in vitro (Hoey and Levine, 1988; Treisman et al., 1991). The e5 sequence is also conserved between *D. melanogaster* and *D. hydei*. Significantly, many of the bases shared by the 32 bp region and e5 are among those conserved. Prd contains two DNA-binding domains, a Paired-class homeodomain (HD) and a paired domain (PD) (Bopp et al., 1986; Frigerio et al., 1986), and footprinting analysis has shown that the e5 site has two

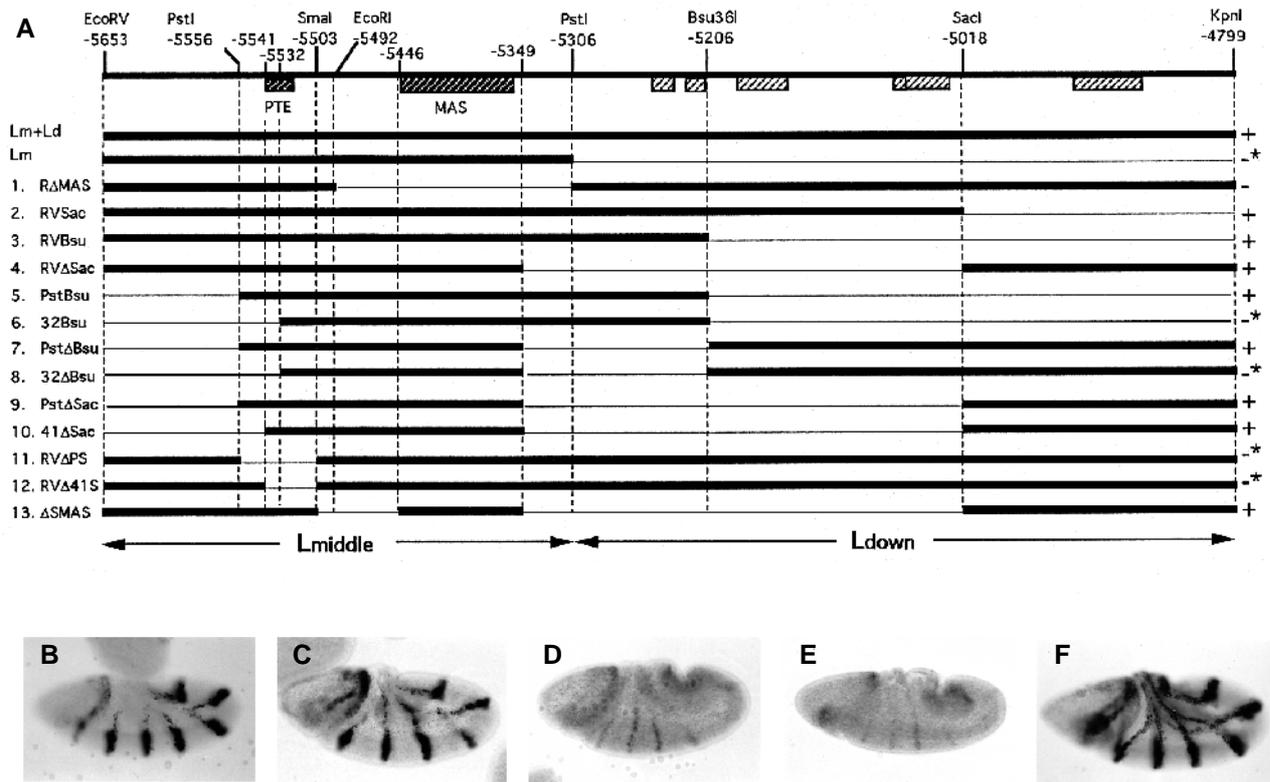


Fig. 2. Deletion analysis of the late element shows the importance of the PTE. (A) The map of the Lmiddle and Ldown (LmLd) portion of the late element is shown at the top. Blocks of conserved sequences (Fig. 1) are indicated by hatched boxes. Below the map are structures of transgenes used for the present study. Portions of the sequence present in each construct are indicated by thick lines while deleted portions are in thin lines. Endpoints of these constructs are identified at the top by the nucleotide number relative to the transcription start site of the *eve* locus. These deletion fragments were placed upstream of the *SfiI* site (-275) of the *eve* promoter-proximal region contiguous to the *lacZ* reporter gene. Their activities are noted on the right. (B-F) Examples of embryos stained at the gastrulation stage for β -galactosidase. (B) LmLd. This activity is comparable to that of the entire L. (C) Deletion 7. (D) Deletion 8. (E) Deletion 12. While the differences in activities of the constructs marked with + and -, as judged by the staining intensities, are clear-cut (C versus D), those marked with asterisk (*) in A do initiate and maintain very weak stripes (D,E). However, when tested with a localized rescue assay (Fujioka et al., 1995), the levels of Eve expressed by these constructs are not high enough to fully rescue *eve* deficiency cuticle and *en* expression phenotypes. A subset of the constructs were tested in the rescue assay: deletions 11, 12 and Lm showed only partial rescue; deletions 2, 3, 4, 10 and LmLd showed full rescue. (F) LmLd of *D. hydei*. The fragment used in this construct (nucleotide 788 to 1594 in the *D. hydei* coordinate shown in A) corresponds to LmLd of *D. melanogaster*. Multiple independent transformant lines were tested for each construct.

PTE <i>mel.</i>	CCACTCACCGTGGCTAATTGCCCATCAATA
PTE <i>hydei</i>	CCACTTACCGTGGCTAATTGGCTCATGAATA
e5 <i>mel.</i>	AGCGGAACCGTGG-CTAATCGCGGTGCTG
e5 <i>hydei</i>	TGAACGATGGTGG-CTAATGTTGTGCTGCTG
PD opt	CTCACCGTGACGA
HD opt	GATTAATCGATTATC
PD/HD opt (PH0)	AAGCGTGACTAATTG
PTE <i>mel.</i>	CCACTCACCGTGGCTAATTGCCCATCAATA
PTE-pd mut.	..G...GAG.C.....
PTE-hd mut.CGAG.....
PTE-prd mut.TCG.C.....

Fig. 3. Sequence comparison and conservation of Prd-binding sites, and mutagenesis of the PTE. The PTE and e5 sequences from *D. melanogaster* and *D. hydei* are shown at the top. Since the putative PD- and HD-binding sites in e5 are one base closer to each other than those in the PTE, a gap of one nucleotide, denoted by a hyphen (-), was introduced to align the PTE and e5. The PD opt sequence is a consensus of the sequences bound efficiently *in vitro* by the PD moiety of Prd alone (i.e., in the absence of the HD domain) (Jun and Desplan, 1996). A similar study found that the HD of Prd binds efficiently to the palindromic sequence, labeled HD opt (Wilson et al., 1993). The PD/HD opt is the sequence (PH0) that was selected *in vitro* for binding to a fragment of Paired containing both the PD and HD (Jun and Desplan, 1996). Bases common to the *D. melanogaster* PTE and PH0 are highlighted in bold. Base positions in PD opt and HD opt predicted to be contacted by PD and HD protein respectively (Wilson et al., 1995; Xu et al., 1995) are underlined. The last three lines identify mutations created in the PTE. Bases denoted by dots (.) are unmutagenized bases, identical to those of the intact PTE shown above the mutations.

adjacent half-sites that can be bound either simultaneously or individually by the two domains (Treisman et al., 1991). These half-sites can be aligned with optimized binding sites, obtained by PCR selection, for the isolated Prd HD (HD opt in Fig. 3) (Wilson et al., 1993) and the isolated PD (PD opt) (Xu et al. 1995; Jun and Desplan, 1996). Comparison of the 32 bp conserved region with these PCR-optimized sites suggests that it also contains adjacent half-sites for both domains.

To test whether Prd can bind to the conserved sequence, we performed an electrophoretic mobility shift analysis (EMSA) using DNA fragments containing the sequence and a histidine-tagged fragment of the Prd protein containing both the PD and the HD (amino acids 27-276). As illustrated in Fig. 4, the Prd protein shifted the conserved sequence (lanes 1-5). This binding was effectively competed by adding cold probe, but not random salmon sperm DNA sequences (not shown). This indicates that the conserved sequence is indeed a Paired binding site. Interestingly, with higher concentrations of Prd protein, lower mobility band shifts of PTE and e5 are observed (not shown), consistent with the fact that the protein has two DNA-binding activities and that the DNA probes have two binding sites that can be bound by the HD and PD of separate Prd molecules (Treisman et al., 1991).

In a parallel study, Jun and Desplan (1996) recently determined a PCR-optimized Prd-binding sequence using a Prd peptide containing both the PD and HD. As shown in Fig. 3, the conserved sequence, PH0, matches this optimal Prd sequence remarkably well (12 out of 15 bases). The sequences match at all 6 base pairs predicted to contact the HD (Wilson et al., 1995), and at 6 of the 11 base pairs thought to contact the PD (Xu et al., 1995). In the following sections, this 32 bp

conserved sequence will be referred to as the PTE (the Paired Target Element).

prd mutation affects late *eve* expression

Prd is an activator of transcription in cell culture assays (Han et al. 1989; Ananthan et al., 1993; Cai et al., 1994), a property expected for a factor interacting with the PTE, a positive regulatory site. In addition, *prd* stripes overlap with the anterior portion of early *eve* stripes, where late *eve* stripes are activated. Therefore, Prd may directly interact with the PTE to activate late *eve* expression. However, previous studies have found that *prd* is at the bottom of the pair-rule gene hierarchy (Baumgartner and Noll, 1991), and no alterations in the *eve* expression pattern were previously observed in *prd* mutants (Frasch and Levine, 1987). However, this and other early studies of segmentation gene interactions did not make use of *lacZ*-tagged balancer chromosomes, which allow unequivocal identification of mutant embryos. We therefore re-examined *eve* expression in *prd* mutants making use of a *CyO* balancer chromosome with a *hunchback/lacZ* P-element insertion (see Fig. 5 legend). The patterns of *eve* expression in *prd* mutant and in wild-type embryos are shown in Fig. 5. At the blastoderm stage, there is no appreciable difference between the two (Fig. 5A,C), but during gastrulation, *eve* expression starts to fade in the *prd* mutant (Fig. 5D), particularly in the ventral ectoderm, and becomes undetectable as germband elongation proceeds (not shown). In contrast, the wild-type sibling embryos at the same morphological stages (and stained simultaneously) express strong, refined *eve* RNA stripes (Fig. 5B). Thus, early *eve* stripes are normal in *prd* mutants, but late stripes are absent. To test whether the *prd* mutation exerts its effect through the late element, we examined the expression of late element reporter constructs in *prd* mutants. The *L-lacZ* and

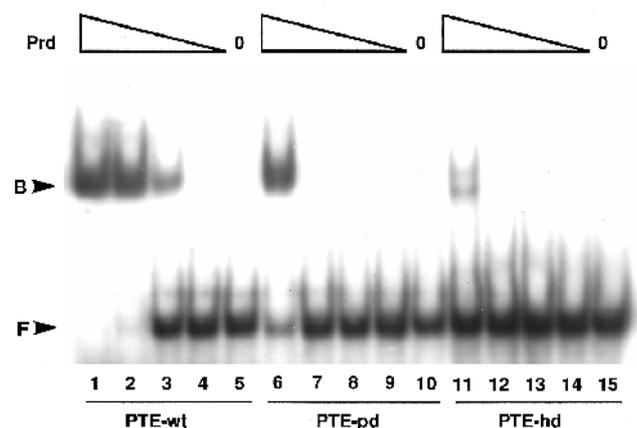


Fig. 4. The Prd protein binds to the PTE. EMSA (gel shift analysis) was used to measure binding by wild-type Prd (amino acids 27-273) to wild-type PTE (PTE-wt) and two mutations of PTE (PTE-pd and PTE-hd; see Fig. 3 and Results for descriptions of mutations). For each target, four-fold decreasing concentrations of protein were tested in adjacent lanes (starting concentration = 1.6×10^{-6} M). Prd binds wild-type PTE more strongly than PTE-pd (by greater than 4-fold), or PTE-hd (>16-fold). As illustrated in lane 11, we observe a second shifted band of slightly higher mobility which is a property of binding to PTE-hd. F, free probe; B, bound probe.

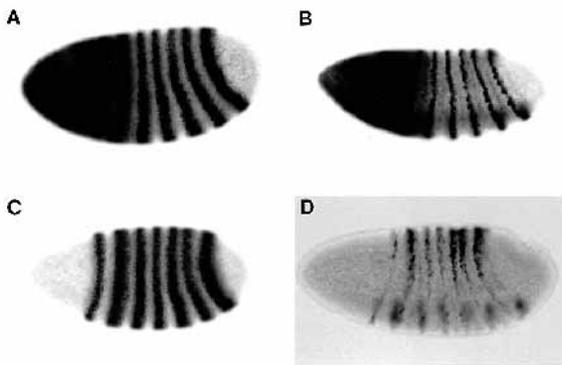


Fig. 5. *eve* stripes are not maintained in *prd* mutants. The *prd* mutation in the strain used in this experiment is balanced by a *CyO* balancer chromosome with a *hunchback/lacZ* P-element (Morrissey et al., 1991). Embryos at blastoderm (A,C) and gastrulation (B,D) are stained for *eve* and β -galactosidase messages. *prd*⁻ (C,D) embryos can be distinguished from *prd*⁺ embryos (A,B) by the absence of β -galactosidase head staining due to the *hunchback/lacZ* in the balancer chromosome. In *prd*⁺ embryos, early broad stripes (A) and late narrow stripes (B) are expressed strongly. In *prd*⁻ embryos, early stripes are present (C), but the late stripes fade prematurely (D), indicating a role for *prd* in the late *eve* expression. The expression in the mesoderm seems to be less affected by the *prd* mutation (D).

the *LmLd-lacZ* transgenes were crossed into a *prd* mutant line and their expression patterns examined. Homozygous *prd*⁻ embryos showed no *L-lacZ* (Fig. 6B) or *LmLd-lacZ* (not shown) expression stripes, unlike *prd*⁺ siblings which showed the characteristic stripe expression (Fig. 6A). Thus, *prd* is required for L to activate late stripes.

Ectopic expression of Prd regulates late *eve* expression

If *prd* is a regulator of late *eve* expression, then ectopic expression of an introduced *prd* transgene would also be expected to affect *eve* expression. If this were the case, it would provide an independent way to examine interaction of Prd with the late element. It was previously shown that induction of ectopic Prd driven by a heat-shock promoter leads to a posterior expansion of odd-numbered *en* stripes (Morrissey et al., 1991). Moreover, the loss of odd-numbered *en* stripes in *prd*⁻ embryos can be rescued by heat treatment of *prd*⁻ *hs-prd* embryos (Miskiewicz et al., 1996). We therefore examined the expression patterns of endogenous *eve* and *L-lacZ* following ectopic Prd induction. In heat-treated *prd*⁻ embryos containing a *hs-prd* transgene, we observed rescue of *L-lacZ* striped expression (Fig. 6C) and of late *eve* RNA expression (not shown), confirming that Prd is an activator of late *eve* expression. Moreover, in both *prd*⁻ and *prd*⁺ embryos, ectopic Prd caused the posterior borders of both *L-lacZ* (Fig. 6C) and late *eve* RNA (not shown) stripes to be shifted posteriorly. This suggests that Prd is at least partly responsible for determining the posterior borders of late *eve* stripes.

A weakened interaction between Prd and the PTE leads to decreased activity of L

The results described above show that Prd interacts with L in vivo to activate late *eve* stripes. We next asked whether this interaction is a result of Prd binding to the PTE. As noted

above, the PTE contains juxtaposed binding sites for a *paired*-class HD and a PD. If binding of Prd to the PTE regulates the activity of the late element, then we would expect that a targeted mutation in either of these half-sites, which reduces Prd-binding in vitro, would weaken the in vivo activity of the late element. We created specific mutations within the PTE to weaken or eliminate Prd binding and tested their effects on maintenance of late stripes *in vivo*. These experiments were carried out in the context of *LmLd* which is fully active as a late element.

The first mutation, PTE-pd, changes four nucleotides in the apparent core of the PD-binding sequence (Fig. 3). In the second mutation, PTE-hd, the putative HD-binding sequence is changed from **TAATTG** to **TCGAGG**, significantly altering the core TAAT sequence. The third mutation, PTE-phd, was designed to interfere with the binding of both the HD and the PD of Prd to the PTE. It changes the HD-binding sequence, **TAATTG**, to **CGACTG**, and the PD-binding sequence **ACCGTGGC** to **ACCGTGGT** (Fig. 3). This eliminates both the core TAAT sequence of the HD-binding site, and changes a single nucleotide in the PD-binding site that is expected to reduce the affinity of the PD, since a change in this position of the site-selected PD site significantly reduces the affinity of Prd (Jun and Desplan, 1996). As shown in Fig. 4, the affinity of Prd for PTE-pd was reduced more than 4-fold (lanes 6-10) when compared to wild-type PTE (lanes 1-5), and the affinity for PTE-hd was more than 16-fold lower (lanes 11-15). The affinity for PTE-phd was even lower (by at least 60-fold; data not shown). We note that PTE-pd retains limited PD DNA-binding activity (data not shown), consistent with the fact that only 4 of the 11 base pairs predicted to be contacted by the PD

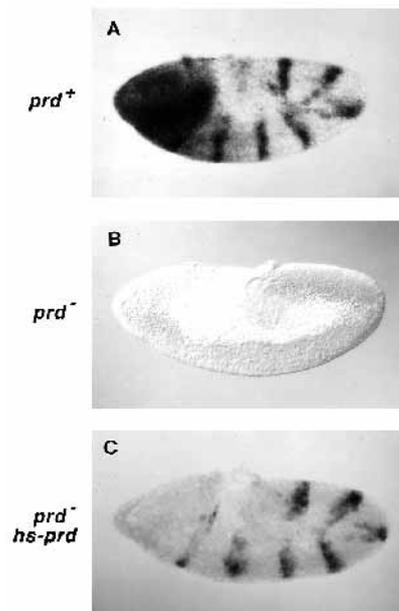


Fig. 6. Ectopic expression of Prd activates late *eve* expression. *L-lacZ* striped expression is observed in *prd*⁺ embryos (A) but not *prd*⁻ embryos (B) indicating that *prd* is required for activation and maintenance of *L-lacZ* stripes. The loss of striped expression is rescued in heat-treated *hs-prd prd*⁻ embryos (C). The rescued stripes are expanded posteriorly, just as observed in the *hs-prd prd*⁺ sibling embryos (not shown).

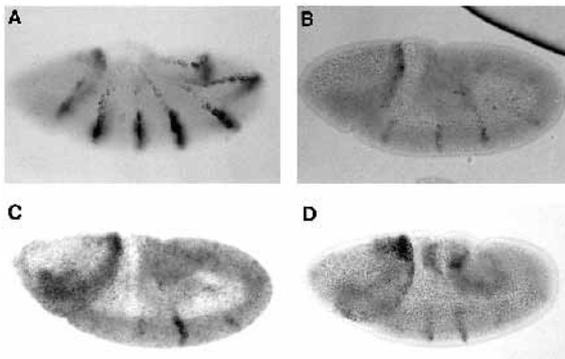


Fig. 7. Mutations in PTE that diminish Prd binding in vitro lead to decreased expression of late stripes in vivo. Expression of β -galactosidase mRNA in transformants of the LmLd-*lacZ* (A) or its mutagenized versions (B-D) is shown. Transgenes are all shown in a *prd*⁺ background. Mutagenized bases in each transgene are shown in Fig. 3. Gel shifts of these mutagenized PTEs with the Prd protein are shown in Fig. 4. (B) PTE-pd mutation; (C) PTE-hd mutation; (D) PTE-phd mutation. In the PTE-pd, PTE-hd and PTE-phd transformants (B-D), expression of late stripes is reduced to a level comparable to that of the complete deletion of the PTE (Fig. 2E). These observations suggest that intact HD- and PD-binding sites are both required for the in vivo activity of the PTE.

are mutated (Xu et al., 1995). When tested in vivo, the mutations PTE-pd, -hd and -phd all caused reduced activity of LmLd-*lacZ*. The effects of all three mutations on stripe activation and maintenance are roughly comparable (Fig. 7) and approach that of the deletion 12 construct (RV Δ 41S), which lacks the PTE entirely (Fig. 2E). Moreover, like deletion 12, all three mutant constructs show only limited in vivo rescue activity when placed upstream of *eve* coding sequences, in contrast to wild-type LmLd (results not shown; see Fig. 2 legend). Thus, both putative PD and HD-binding sites are required for the activity of the PTE. From this observation, in conjunction with the results of the deletion analysis and the fact that the PTE sequence closely matches the highest affinity site selected in vitro, we conclude that the PTE is a direct target site of Prd in vivo.

DISCUSSION

The early (E) and late (L) control elements of the *eve* promoter participate in distinct phases of the refinement of spatial information during segmentation of the early *Drosophila* embryo. Early elements direct expression in broad stripes, each of which spans the entire odd-numbered parasegmental primordium. This gives way, under the control of the late element, to expression in sharply defined narrow stripes in the anterior cell row of each primordium. Both phases are important for correct regulation of downstream genes in the cascade. The early expression is critical for defining expression boundaries of several pair-rule genes including *prd*, *runt*, *sloppy-paired* (*slp*) and *odd-skipped* (Fujioka et al., 1995; see below). Late *eve* expression regulates a later phase of expression of these same pair-rule genes, thereby providing for the initiation of strong, continuous *en* stripes in odd-numbered parasegments (Fujioka et al., 1995).

Prd participates in *eve* stripe refinement

We have analysed the *cis*-regulatory element for late *eve* expression and have found that this multistripe element (see Introduction) is regulated by the Prd protein through an evolutionarily conserved sequence (PTE). This observation supports our model for the formation of *eve* late stripes (Fig. 8; see also Fujioka et al., 1995). The model proposes that bell-shaped early *eve* stripes act as morphogenetic gradients that differentially regulate *prd*, *runt* and *slp*. Although all three genes are repressed by *eve*, it is postulated that *runt* and *slp* are sensitive to lower levels of Eve protein than is *prd*. This creates a row of cells at the anterior edge of early *eve* stripes that express *prd* but not *runt* or *slp*. Since *prd* is an activator of late *eve* (this study), and *runt* and *slp* are repressors (Cadigan et al., 1994a,b; Manoukian and Krause, 1993), activation of the late *eve* is restricted to this row of cells.

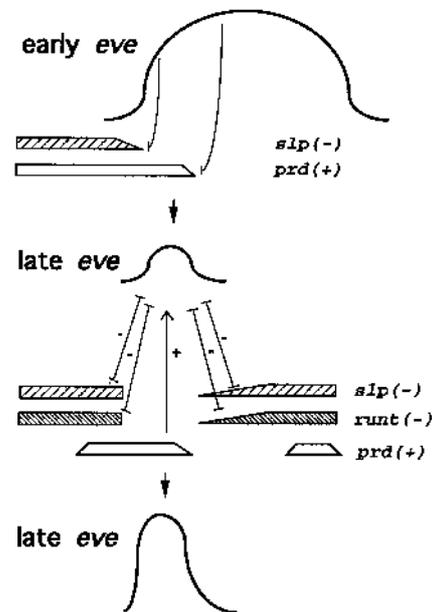


Fig. 8. A model of *eve* refinement. We propose that early *eve* stripes act as bell-shaped morphogenetic gradients (see Fujioka et al., 1995). Target genes are repressed by different concentrations of Eve in early stripes, leading to specific patterning of activators of late stripes (*prd*) and repressors (*slp*, *runt*). *prd* has two (early and late) phases of regulation, and only the early phase appears to be sensitive to *eve* repression. In contrast, *runt* appears to be sensitive to *eve* repression only at its late, 14 stripe stage (during gastrulation and thereafter), while *slp* is sensitive to *eve* at all stages (see Fujioka et al., 1995). Early *runt* stripes that overlap the posterior portion of early *eve* stripes are not shown. Late *eve* stripes are initially activated at the anterior portion of early stripes by *prd*, which also specifies the posterior border of *eve* late stripes. The initial anterior borders of late stripes are determined primarily by *slp*, and to a lesser degree by late *runt*, which appears later than *slp*. After cellularization, *prd* continues to enhance *eve* late stripes after becoming insensitive to repression by Eve. Cross-regulatory interactions (mutual repression) of late *eve* with *slp* and late *runt* result in sharpening of the anterior borders of late *eve* stripes (and of the juxtaposing borders of *slp* and late *runt* stripes). They may also help maintain the posterior borders of *eve* late stripes as illustrated. Since *slp* and *runt* appear to be direct repressors of *en*, this model also provides an explanation for the sharpness of the anterior borders of odd-numbered *en* stripes (see Fujioka et al., 1995).

In its regulation of late *eve*, the function of Prd may be augmented by a structural homolog of Prd, the Gooseberry (Gsb) protein, which also contains a paired-class HD and a PD. Recent studies (Li and Noll, 1994) have shown that, when appropriately expressed, the Prd protein can substitute for Gsb, suggesting that the two proteins may be functionally equivalent. Hence, since Gsb is turned on at about gastrulation in the same cell rows in which late *prd* is expressed, it is possible that Gsb contributes in part to later maintenance of late *eve* expression in collaboration with Prd protein, which continues to be expressed through much of germ band elongation. However, since *gsb*⁻ embryos have normal *en* expression which depends on late *eve* for full activation in odd-numbered parasegments (Hidalgo, 1991; Li and Noll, 1993; Fujioka et al., 1995), it appears that Prd protein alone is sufficient to activate late *eve* in the absence of *gsb* function. In contrast, in *prd*⁻ embryos in which odd-numbered *gsb* stripes are deleted (Bopp et al., 1989), there is no activation of late *eve* (this study), consistent with Prd and Gsb functions both being absent.

Central to our model (Fig. 8) is the idea that Prd functions combinatorially on the late element. Strong support for this comes from the effects of ectopic Prd on *L-lacZ* expression. Uniform expression of Prd does not cause uniform *L-lacZ* expression. Rather, *L-lacZ* stripes are expanded posteriorly by only one or two cells (Fig. 6), implying that other regulators act in combination with Prd. Indeed, ectopic Prd has the equivalent effect on *en* expression: the odd-parasegment *en* stripes, which coincide with *L-lacZ*, are expanded 1-2 cells posteriorly, suggesting that Prd normally specifies the posterior borders of these stripes (Morrissey et al., 1991). Like *L-lacZ*, the anterior borders of these stripes are thought to be specified by the repressors *slp* and *runt*. The question remains how these repressors might function in combination with Prd at the molecular level. With the identification of a functional Prd site within L, this question can now be addressed.

The PTE is an in vivo target of Prd

The PTE element is a conserved sequence that deletion analysis indicates is critical for function of the late element. Our present study shows that Prd can bind to the PTE in vitro and that Prd functions through this sequence in vivo. The significant loss of activity caused by mutations in the PTE (Figs 2, 7) suggests that the interaction of the site with Prd plays a critical role that cannot be substituted for by any other sequences within our minimal single-copy L element (LmLd). Nevertheless, although sequence analysis of LmLd does not reveal any conspicuous sequence similarities to previously identified Prd-binding sites, the possibility that there may be secondary Prd site(s) within L is suggested by our observation that several PTE deletion constructs have extremely weak stripes that appear to be maintained through germ band elongation (see Fig. 2 legend). This contrasts with the complete absence of L activity in *prd*⁻ embryos.

The e5 site (Fig. 3), located just upstream of the *eve* TATAA box, is a previously identified in vitro target of Prd that has been used to study the interaction of Prd, and its mammalian homologues, with DNA (Chalepakakis et al., 1994; Hoey and Levine, 1988; Treisman et al., 1991). However, the function of e5 in vivo has not been addressed. We have recently shown that removal of the e5-binding site, along with some adjacent

sequence, from the promoter-proximal sequence (-275 to +166) normally included in our reporter constructs, causes somewhat weakened expression of *L-lacZ* stripes (result not shown). A similar reduction was observed previously when the promoter-proximal region was replaced with a heat-shock promoter (Goto et al., 1989). This observation, together with the presence of a similar sequence in the *D. hydei* promoter (Fig. 3), suggests that the e5 site may function in vivo in collaboration with the PTE. However, our observation that PTE deletions have much more severe consequences in vivo than the e5 deletion, and that Prd has a much higher affinity for the PTE than for e5 (data not shown), suggests that the PTE plays a more critical role for Prd mediated activation of late *eve* expression. We should also note that the *eve* promoter contains another site, the e4 sequence, which is bound by Prd in vitro (Hoey and Levine, 1988; Treisman et al., 1991). However, unlike the PTE or e5, this sequence is not conserved in *D. hydei* (data not shown; see below).

The PTE has juxtaposed binding sites for the PD and the HD

The Prd protein, like a number of other Pax proteins, has both a HD and a PD (Walther et al., 1991). Although both have DNA-binding activities, the functional relationship between these two domains in vivo has remained elusive. PCR-selection experiments have revealed optimal target sequences for each of these domains when tested alone (Jun and Desplan, 1996; Wilson et al., 1993; Xu et al., 1995). These optimal sequences can be simultaneously aligned (see Fig. 3) with both the PTE and e5, forming adjacent half-sites. As was shown for e5 (Treisman et al., 1991), the PTE can be bound by Prd molecules lacking either one of the two binding activities (Miskiewicz et al., 1996). Furthermore, the PTE matches PH0, a PCR-selected site for simultaneous binding by the PD and HD (Jun and Desplan, 1996), at many of the bases thought to contact the Prd protein based on structural studies (Wilson et al., 1995; Xu et al., 1995). The order and relative orientation of the PD and HD sites in both the PTE and e5 is the same as in PH0. Moreover, the spacing between the two sites is the same in the PTE and PH0, but is one base closer in e5 (Fig. 3). In contrast, the order in the PTE is the opposite to that in e4, which is a relatively poor binding site for Prd (Treisman et al., 1991).

The PTE is the first identified in vivo target of Prd that contains both a PD and a HD target sequence. Previous analysis of the *gooseberry* promoter has identified a putative A/T-rich Prd-binding site (Li and Noll, 1994) which, based on comparison with PCR-selected sites, may be bound by the Prd HD (Wilson et al., 1995). Paired domain sites are probably also important for the regulation of *gooseberry*, since both PD- and HD-binding activities of the Prd protein are required for in vivo regulation of *gooseberry*, as they are for the regulation of *en* and *hedgehog* (Bertuccioli et al., 1996; Miskiewicz et al., 1996). Similarly, we have found that both the PD and HD target sites of the PTE are required for Prd regulation of *L-lacZ*. Mutation in each of these target sites leads to reduced binding of Prd to the PTE in vitro (Fig. 4), and to significantly weakened *L-lacZ* stripes in vivo (Fig. 7), consistent with the idea that Prd function requires both of its DNA-binding domains. Whether both binding activities need be present in the same Prd molecule for *L-lacZ* regulation has yet to be

determined. Ectopic expression studies of Prd regulation of segment-polarity genes have shown that the two binding activities can come from separate molecules: co-expression of a PD mutant protein with a HD mutant protein restores Prd function in embryos, whereas neither mutant alone is sufficient (Miskiewicz et al., 1996). However, it remains possible that the wild-type Prd protein uses both binding activities in the same molecule (see also Bertuccioli et al. 1996; Jun and Desplan 1996). With the identification of the PTE, it is now possible to investigate these questions, as well as structural constraints on the target sequence. For example, how do spacing and orientation of the HD and PD sites affect Prd binding and function in vivo?

Prd provides a model for Pax protein function

Prd is a member of the Pax family of transcription factors, all of which contain a PD, and some of which also contain a HD (Walther et al., 1991). Some of the mammalian members of this family have been associated with defined mutations and birth defects. For example, mutation of Pax-3, the closest mammalian homolog of Prd, which contains both a PD and a HD, is associated with Waardenburg Syndrome in humans, affecting neural crest cells and other derivatives of the dorsal neural tube (Baldwin et al., 1992; Burri et al., 1989; Tassabehji et al., 1992). Our observation that Prd functions in vivo through the PTE raises the possibility that Pax-3 and its homologs also function through composite binding sites, with adjacent sequences for HD and PD binding.

We thank Claude Desplan, Claudio Bertuccioli and Susie Jun for discussion of unpublished data and technical advice. We are grateful to Dieter Maier for providing us with the *D. hydei* library. We also thank James B. Jaynes for discussions and careful reading of the manuscript. This work was supported by NSF grant IBN-9507406 to T. G. and NIH grant GM42752 and ACS grant JFRA-430 to M. W.

REFERENCES

- Akam, M. (1987). The molecular basis for metameric pattern in the *Drosophila* embryo. *Development* **101**, 1-22.
- Ananthan, J., Baler, R., Morrissey, D., Zuo, J., Lan, Y., Weir, M. and Voellmy, R. (1993). Synergistic activation of transcription is mediated by the N-terminal domain of *Drosophila fushi tarazu* homeoprotein and can occur without DNA binding by the protein. *Mol. Cell. Biol.* **13**, 1599-1609.
- Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O. and Milunski, A. (1992). An exonic mutation in *HuP2* paired domain gene causes Waardenburg's syndrome. *Nature* **355**, 637-638.
- Baumgartner, S. and Noll, M. (1991). Network of interactions among pair-rule genes regulating *paired* expression during primordial segmentation of *Drosophila*. *Mech. Dev.* **33**, 1-18.
- Bertuccioli, C., Fasano, L., Jun, S., Wang, S., Sheng, G. and Desplan, C. (1996). In vivo requirement for the paired domain and homeodomain of the *paired* segmentation gene product. *Development* **122**, 2673-2685.
- Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986). Conservation of a large protein domain in the segmentation gene *paired* and in functionally related genes of *Drosophila*. *Cell* **47**, 1033-1040.
- Bopp, D., Jamet, E., Baumgartner, S., Burri, M., and Noll, M. (1989). Isolation of two tissue-specific *Drosophila* paired box genes, *pox meso* and *pox neuro*. *EMBO J.* **8**, 3447-3457.
- Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G. and Noll, M. (1989). Conservation of the paired domain in metazoans and its structure in three isolated human genes. *EMBO J.* **8**, 1183-1190.
- Cadigan, K. M., Grossniklaus, U. and Gehring, W. J. (1994a). Functional redundancy: the respective roles of the two *sloppy paired* genes in *Drosophila* segmentation. *Proc. Natl. Acad. Sci. USA* **91**, 6324-6328.
- Cadigan, K. M., Grossniklaus, U. and Gehring, W. J. (1994b). Localized expression of *sloppy paired* protein maintains the polarity of *Drosophila* parasegments. *Genes Dev.* **8**, 899-913.
- Cai, J., Lan, Y., Appel, L. F. and Weir, M. (1994). Dissection of the *Drosophila* paired protein: functional requirements for conserved motifs. *Mech. Dev.* **47**, 139-150.
- Carroll, S. B. and Scott, M. P. (1986). Zygotically active genes that affect the spatial expression of the *fushi tarazu* segmentation gene during early *Drosophila* embryogenesis. *Cell* **45**, 113-126.
- Chalepakis, G., Wijnholds, J. and Gruss, P. (1994). Pax-3-DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains. *Nucleic Acids Res.* **22**, 3131-3137.
- Coulter, D., Swaykus, E. A., Beran-Koehn, M. A., Goldberg, D., Wieschaus, E. and Schedl, P. (1990). Molecular Analysis of *odd-skipped*, a zinc finger encoding segmentation gene with a novel pair-rule expression pattern. *EMBO J.* **8**, 3795-3804.
- DiNardo, S., Sher, E., Heemskerk-Jongens, J., Kassis, J. A. and O'Farrell, P. H. (1988). Two-tiered regulation of spatially patterned *engrailed* gene expression during *Drosophila* embryogenesis. *Nature* **332**, 604-609.
- 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**, 363-367.
- Frasch, M. and Levine, M. (1987). Complementary patterns of *even-skipped* and *fushi tarazu* expression involve their differential regulation by a common set of segmentation genes in *Drosophila*. *Genes Dev.* **1**, 981-995.
- Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M. (1986). Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**, 735-746.
- Fujioka, M., Jaynes, J. B. and Goto, T. (1995). Early *even-skipped* stripes act as morphogenetic gradients at the single cell level to establish *engrailed* expression. *Development* **121**, 4371-4382.
- Goto, T., Macdonald, P. and Maniatis, T. (1989). Early and late periodic patterns of *even-skipped* expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell* **57**, 413-422.
- Grossniklaus, U., Pearson, R. K. and Gehring, W. J. (1992). The *Drosophila sloppy paired* locus encodes two proteins involved in segmentation that show homology to mammalian transcription factors. *Genes Dev.* **6**, 1030-1051.
- Gutjahr, T., Vanario-Alonso, C. E., Pick, L. and Noll, M. (1994). Multiple regulatory elements direct the complex expression pattern of the *Drosophila* segmentation gene *paired*. *Mech. Dev.* **48**, 119-128.
- Han, K., Levine, M. S. and Manley, J. L. (1989). Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56**, 573-583.
- Harding, K., Hoey, T., Warrior, R. and Levine, M. (1989). Autoregulatory and gap gene response elements of the *even-skipped* promoter of *Drosophila*. *EMBO J.* **8**, 1205-1212.
- Hidalgo, A. (1991). Interactions between segment polarity genes and the generation of the segmental pattern in *Drosophila*. *Mech. Dev.* **35**, 77-87.
- Hiroimi, Y., Kuroiwa, A. and Gehring, W. J. (1985). Control elements of the *Drosophila* segmentation gene *fushi tarazu*. *Cell* **43**, 603-613.
- Hoey, T. and Levine, M. (1988). Divergent homeobox proteins recognize similar DNA sequences in *Drosophila*. *Nature* **332**, 858-861.
- Howard, K., Ingham, P. and Rushlow, C. (1988). Region-specific alleles of the *Drosophila* segmentation gene *hairly*. *Genes Dev.* **2**, 1037-1046.
- Howard, K. R. and Struhl, G. (1990). Decoding positional information: regulation of the pair-rule gene *hairly*. *Development* **110**, 1223-1231.
- Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**, 25-34.
- Jiang, J., Hoey, T. and Levine, M. (1991). Autoregulation of a segmentation gene in *Drosophila*: combinatorial interaction of the *even-skipped* homeo box protein with a distal enhancer element. *Genes Dev.* **5**, 265-277.
- Jun, S. and Desplan, C. (1996). Cooperative interaction between paired domain and homeodomain. *Development* **122**, 2639-2650.
- Klingler, M. and Gergen, J. P. (1993). Regulation of *run* transcription by *Drosophila* segmentation genes. *Mech. Dev.* **43**, 3-19.
- Li, X., and Noll, M. (1993). Role of the *gooseberry* gene in *Drosophila* embryos: maintenance of *wingless* expression by a *wingless-gooseberry* autoregulatory loop. *EMBO J.* **12**, 4499-4509.
- Li, X. and Noll, M. (1994). Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different cis-regulatory regions. *Nature* **367**, 83-87.
- Maier, D., Preiss, A. and Powell, J. R. (1990). Regulation of the segmentation gene *fushi tarazu* has been functionally conserved in *Drosophila*. *EMBO J.* **9**, 3957-3966.

- Manoukian, A. S. and Krause, H. M.** (1992). Concentration-dependent activities of the *even-skipped* protein in *Drosophila* embryos. *Genes Dev.* **6**, 1740-1751.
- Manoukian, A. S. and Krause, H. M.** (1993). Control of segmental asymmetry in *Drosophila* embryos. *Development* **118**, 785-796.
- Miskiewicz, P., Morrissey, D., Lan, Y., Raj, L., Kessler, S., Fujioka, M., Goto, T. and Weir, M.** (1996). Both the paired domain and homeodomain are required for *in vivo* function of *Drosophila* Paired. *Development* **122**, 2709-2718.
- Morrissey, D., Askew, D., Raj, L. and Weir, M.** (1991). Functional dissection of the *paired* segmentation gene in *Drosophila* embryos. *Genes Dev.* **5**, 1684-1696.
- Nusslein-Volhard, C. and Wieschaus, E.** (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.
- Pankratz, M. J., Seifert, E., Gerwin, N., Billi, B., Nauber, U. and Jackle, H.** (1990). Gradients of *Krüppel* and *knirps* gene products direct pair-rule gene stripe patterning in the posterior region of the *Drosophila* embryo. *Cell* **61**, 309-317.
- Riddihough, G. and Ish-Horowicz, D.** (1991). Individual stripe regulatory elements in the *Drosophila hairy* promoter respond to maternal, gap, and pair-rule genes. *Genes Dev.* **5**, 840-854.
- Rubin, G. M. and Spradling, A. C.** (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348-353.
- Sackerson, C.** (1995). Pattern of conservation and divergence at the *even-skipped* locus of *Drosophila*. *Mech. Dev.* **51**, 199-215.
- Sambrook, J., Fritsch, E. F. and Maniatis, T.** (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Press.
- Small, S., Kraut, R., Hoey, T., Warrior, R. and Levine, M.** (1991). Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes Dev.* **5**, 827-839.
- Stanojevic, D., Hoey, T. and Levine, M.** (1989). Sequence-specific DNA-binding activities of the gap proteins encoded by *hunchback* and *Krüppel* in *Drosophila*. *Nature* **341**, 331-335.
- Tassabehji, M., Read, A. P., Newton, A. E., Harris, R., Balling, R., Gruss, P. and Strachan, T.** (1992). Waardenburg's syndrome patients have mutations in the human homologue of the *Pax-3* paired box gene. *Nature* **355**, 635-636.
- Tautz, D. and Pfeifle, C.** (1989). A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98**, 81-85.
- Thummel, C. S., Boulet, A. M. and Lipshitz, H. D.** (1988). Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74**, 445-456.
- Treisman, J., Harris, E. and Desplan, C.** (1991). The paired box encodes a second DNA-binding domain in the *paired* homeo domain protein. *Genes Dev.* **5**, 594-604.
- Walther, C., Guenet, J.-L., Simon, D., Deutsch, U., Jostes, B., Goulding, M. D., Plachov, D., Balling, R. and Gruss, P.** (1991). Pax: A murine multigene family of paired box-containing genes. *Genomics* **11**, 424-434.
- Warrior, R. and Levine, M.** (1990). Dose-dependent regulation of pair-rule stripes by gap proteins and the initiation of segment polarity. *Development* **110**, 759-767.
- Wilson, A. C., Ochman, H. and Prager, E. M.** (1987). Molecular time scale for evolution. *Trends In Genetics* **3**, 241-247.
- Wilson, D., Sheng, G., Lecuit, T., Dostatni, N. and Desplan, C.** (1993). Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev.* **7**, 2120-2134.
- Wilson, D. S., Guenther, B., Desplan, C. and Kuriyan, J.** (1995). High resolution crystal structure of a paired (Pax) class cooperative homeodomain dimer on DNA. *Cell* **82**, 709-720.
- Xu, W., Rould, M. A., Jun, S., Desplan, C. and Pabo, C. O.** (1995). Crystal structure of a paired domain-DNA complex at 2.5 Å resolution reveals structural basis for Pax developmental mutations. *Cell* **80**, 639-650.
- Yu, Y. and Pick, L.** (1995). Non-periodic cues generate seven *ftz* stripes in the *Drosophila* embryo. *Mech. Dev.* **50**, 163-175.

(Accepted 28 June 1996)