

In vivo requirement for the paired domain and homeodomain of the *paired* segmentation gene product

Claudio Bertuccioli, Laurent Fasano[†], Susie Jun, Sheng Wang, Guojun Sheng and Claude Desplan*

Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10021, USA

[†]Present address: LGPD, CNRS Case 907, Luminy, 13288 Marseille Cedex 9, France

*Author for correspondence

SUMMARY

The *Drosophila* pair-rule gene *paired* is required for the correct expression of the segment polarity genes *wingless*, *engrailed* and *gooseberry*. It encodes a protein containing three conserved motifs: a homeodomain (HD), a paired domain (PD) and a PRD (His/Pro) repeat. We use a rescue assay in which *paired* (or a mutated version of *paired* in which the functions of the conserved motifs have been altered) is expressed under the control of its own promoter, in the absence of endogenous *paired*, to dissect the Paired protein in vivo. We show that both the HD and the N-terminal subdomain of the PD (PAI domain) are absolutely required within the same molecule for normal *paired* function. In contrast, the conserved C-terminal subdomain of the PD (RED domain) appears to be dispensable. Fur-

thermore, although a mutation abolishing the ability of the homeodomain to dimerize results in an impaired Paired molecule, this molecule is nonetheless able to mediate a high degree of rescue. Finally, a *paired* transgene lacking the PRD repeat is functionally impaired, but still able to rescue to viability. We conclude that, while Prd can use its DNA-binding domains combinatorially in order to achieve different DNA-binding specificities, its principal binding mode requires a cooperative interaction between the PAI domain and the homeodomain.

Key words: *paired*, paired domain, homeodomain, segmentation, *Drosophila*, Pax genes

INTRODUCTION

The subdivision of the developing *Drosophila* embryo into segments is a process governed by a hierarchy of interactions between four classes of segmentation genes: maternal coordinate genes, zygotic gap genes, pair-rule genes and segment polarity genes (Ingham, 1988; Nüsslein-Volhard and Wieschaus, 1980; StJohnston and Nüsslein-Volhard, 1992). Genes within this hierarchy are expressed in increasingly refined domains reflecting the progressive subdivision of the body plan. Many of these genes encode transcription factors that function by regulating the expression of downstream genes. These transcription factors often share functionally important conserved domains that are required for DNA-binding and/or dimerization (reviewed in Nelson, 1995). In addition, they usually also contain 'activation domains', which are believed to mediate interactions with other components of the transcription machinery, but which in general are less conserved structurally and at the sequence level (reviewed in Triezenberg, 1995). The helix-turn-helix (HTH), and its variants such as the homeodomain, the paired domain and the POU domain, is one of the most common and best-studied types of DNA-binding domain and is found in the product of a variety of developmental regulatory genes.

We have been analysing the function of the *paired* (*prd*) gene, a member of the pair-rule family of segmentation genes

(Nüsslein-Volhard and Wieschaus, 1980). In a *prd* mutant, the posterior part of even-numbered parasegments and the anterior part of odd-numbered parasegments are deleted. Around the time of cellularisation, *prd* is expressed in 14 approximately 2-cell-wide stripes spanning each parasegmental boundary (Gutjahr et al., 1994). In the stripes that span the even-odd parasegmental boundary, *prd* is required to activate the expression of the segment polarity genes *wingless* (*wg*), *engrailed* (*en*) and *gooseberry* (*gsb*) (Bopp et al., 1989; DiNardo and O'Farrell, 1987; Ingham et al., 1988). For convenience, these stripes will be referred to henceforth as '*prd*-dependent'. The regulation of these genes by *prd* occurs through interactions between *prd* and other pair-rule genes (Benedyk et al., 1994; DiNardo and O'Farrell, 1987; Ingham et al., 1988; Morrissey et al., 1991; Mullen and DiNardo, 1995; Weir et al., 1988). In addition to its role in patterning the *prd*-dependent stripes of *wg*, *en* and *gsb*, *prd* appears to also play a lesser role in patterning the alternate stripes of *wg* and *gsb*, since these stripes show significantly reduced expression levels in a *prd* mutant (Fig. 2) (Benedyk et al., 1994). However, this role for *prd* must be non-essential, since these portions of the embryo nonetheless develop normally in a *prd* mutant.

The *prd* gene encodes several conserved domains, which are found in a variety of *Drosophila* developmental regulatory genes: a Prd-class homeodomain (HD), a paired domain (PD) and a PRD (His-Pro sequence) repeat (Bopp et al., 1986;

Frigerio et al., 1986). These domains have also been found in regulatory genes of other organisms, from nematodes to vertebrates (reviewed in Chalepakis et al., 1993; Gruss and Walther, 1992; Halder et al., 1995; Noll, 1993). The PD and the *prd*-class HD both contain helix-turn-helix DNA-binding motifs (Wilson et al., 1995; Xu et al., 1995) and can bind in a sequence-specific manner to distinct DNA-binding sites independently of each other (Treisman et al., 1991; Wilson et al., 1993). The Paired (Prd) protein thus contains two independent DNA-binding domains, suggesting that it is a multifunctional transcriptional regulatory protein.

Previous studies have shown that Prd-class HDs are able to cooperatively dimerize on DNA and that a single amino acid substitution in the Prd HD (SQ50) can alter the spacing of the half-sites constituting the palindrome to which the HD binds, as well as the level of cooperativity between the two HDs (Wilson et al., 1993, 1995).

The PD (Bopp et al., 1986) is a sequence-specific DNA-binding domain (Treisman et al., 1991) that characterizes a growing family of developmental regulators known as the *Pax* genes. Many *Pax* genes are associated with mutant phenotypes in mouse, human and flies (for a review see (Strachan and Read, 1994)). A crystal structure of the *paired* PD bound to its target site reveals that the PD binds as a monomer and consists of two similar subdomains, each containing a HTH motif (Xu et al., 1995). These subdomains of the PD will henceforth be referred to as the PAI domain (N-terminal subdomain) and RED domain (C-terminal subdomain) (PAI+RED=Paired Domain) (Jun and Desplan, 1996). Although both subdomains are conserved, the RED domain does not contact the DNA, consistent with *in vitro* studies that showed it is dispensable for DNA-binding (Treisman et al., 1991). This result is also consistent with overexpression studies that indicate that the RED domain of *prd* is dispensable for inducing ectopic expression of target genes (Cai et al., 1994).

Relatively little is known about the third conserved motif of Prd, the PRD repeat. The PRD repeat is an approximately 20 amino acid-coding domain, rich in His-Pro, that has been found in several *Drosophila* genes (Frigerio et al., 1986). Although its function is not clearly understood, previous experiments have suggested that this motif might form part of a required activation domain (Cai et al., 1994; Han et al., 1989).

Since *paired* is the *Pax* gene that has been most extensively characterised molecularly (Treisman et al., 1989; Treisman et al., 1991; Wilson et al., 1993) and structurally (Wilson et al., 1995; Xu et al., 1995), and for which genetic targets are known (Bopp et al., 1989; DiNardo and O'Farrell, 1987; Ingham et al., 1988), it provides an ideal paradigm for studying the functions of conserved domains in *Pax* genes.

We have therefore designed a transgenic *paired* rescue assay, allowing us to express the *paired* gene (or mutated versions of *paired*) under the control of its own promoter and in a background lacking endogenous *paired* function. In this assay, a transgene containing a wild-type *prd* coding sequence is able to rescue all strong *prd* mutants to adulthood. We show that both the homeodomain and the PAI domain of Prd are absolutely required for normal function, and that both DNA-binding domains must be present within the same molecule. Although the PAI domain is absolutely required, the conserved RED domain is dispensable, consistent with crystallographic data that indicate that it does not bind to DNA

(Xu et al., 1995) and in contrast to other *Pax* genes where this domain has been shown to function as an independent DNA-binding domain (Epstein et al., 1994). In addition, we show that mutations in the homeodomain that abolish its ability to dimerize or alter the preferred spacing of the half-sites in the palindromic binding site recognized by the homeodomain dimer, result in Prd molecules which, although impaired, are nonetheless able to rescue well. Finally, we show that a truncated Prd molecule lacking the PRD repeat is subfunctional, but still able to rescue to viability, indicating that it only lacks a relatively minor part of its activation domain. We conclude that Prd contains multiple HTH DNA-binding domains, which it can use combinatorially in order to achieve different functional specificities. Its principal binding mode requires a cooperative interaction between its PAI domain and its homeodomain.

MATERIALS AND METHODS

Cloning of *prd* genomic sequence and construction of rescuing *prd* transgene

A 3 kb *EcoRI-HindIII* fragment corresponding to the 3' end of the *prd* promoter (Fig. 1), derived from the *prd*gen plasmid (Gift of Paul MacDonald, Stanford University), was used to probe an EMBL3 *Drosophila* genomic library (Frischauf et al., 1987) (provided by John Tamkun University of Arizona, Tucson). Six positive phage were isolated and purified to homogeneity. From one phage, approximately 15 kb of *prd* genomic DNA were isolated and divided into two adjoining fragments: a 12 kb 5' *HindIII* fragment (*prd* promoter) and a 3 kb 3' *HindIII* fragment (*prd* coding cassette), which were separately subcloned into pGEM9Zf(-) (Promega). The '*prd* promoter' fragment was excised from pGEM9Zf(-) as a *Sall-SpeI* fragment and subcloned into the pCaSpeR4 P-element vector (Thummel and Pirrotta, 1992), digested with *XhoI* and *SpeI* (Fig. 1). The same *SpeI* site in the pCaSpeR4 vector containing the *prd* promoter was used to subsequently insert different *prd* coding cassettes which were excised as *XbaI-XbaI* fragments from the pGEM9Zf(-) vector using *XbaI* sites in the pGEM polylinker and in the 3' untranslated region of *prd*. This *XbaI-XbaI* *prd* coding fragment contains approximately 0.9 kb of 3' untranslated DNA. The final construct contains an extra 6 bp of pGEM polylinker-derived DNA at the *HindIII* site separating the promoter and coding fragments. This extra 6 bp DNA is inserted in the 5' untranslated end of the *prd* transcript, 32 bp upstream of the Start codon. Orientation of fragments was verified by digestion and the junction between promoter and coding fragments was verified by dideoxy sequencing (Sanger et al., 1977). Labeling of probe and screening of library were performed as described (Maniatis et al., 1989).

Mutagenesis of *prd* coding cassettes

Specific point mutations were introduced into the *prd* coding cassette in pGEM9 using site directed mutagenesis (for *prdGS*, *prdNQ*, *prdAR* and *prdSQ*), or by swapping an *NcoI* fragment spanning a 1.65 kb fragment containing the mutation, derived from *prd* cDNA constructs generously donated by M. Weir (for *prdΔPBC*, *prdΔPR* and *prdΔPRT*) (Cai et al., 1994). Site-directed mutagenesis was performed as described (Maniatis et al., 1989), using T7 DNA polymerase (NEB) and the following oligos:

prdGS: 5' TGGGCAAAGGACGACTGTTGATGAAAACCTC 3'
prdSQ: 5'GCTGCTTGCGGAGACGGCCCGCCGGTTTGGAA-
 CCACACCTGGATGCG 3'
prdNQ: 5' ACGAGCACGCCGCTGGCTGAACCACAC 3'
prdAR: 5' CCACACCTGGATGCGTCTCTCCGTGAGATTGGT 3'

In all cases, mutations were verified by dideoxy sequencing (Sanger et al., 1977) both in pGEM and once inserted in the P-element vector.

Generation of transgenic lines, genetic crosses and analysis of lines

The host strain used for injection of pCaSpeR4 constructs, and which was used as a wild-type strain in experiments is *D(1)6 w⁻y⁻* (gift of S. DiNardo). DNA was injected at a concentration of 300 µg/ml in TE, along with 75 µg/ml of helper construct (Rubin and Spradling, 1982). Remobilisation of transgenes was achieved by crossing the transgene with a line carrying the 'jumpstarter' chromosome ($\Delta 2-3$). The cross was as follows: *w⁻; +/+; P/P* females were crossed with *w⁺; Sp/CyO; $\Delta 2-3$. Sb/TM2* males (gift of W. Engels) to produce dysgenic males: *w⁻; Sp or CyO/+; P/ $\Delta 2-3$. Sb*. These males were mated to *D(1)6 w⁻y⁻* females and their progeny was scored for remobilisation events.

Insertions on the X or III chromosome were homozygosed for the insertion (or balanced if homozygous lethal/sterile) and 'double balanced' on chromosome II to generate either *P/P; Sco/SM6Beve-lacZ*, or *Sco/SM6Beve-lacZ; P/P* stocks. These 'double balanced' P-element stocks were crossed with 'double balanced' *prd* stocks: *prd/SM6Beve-lacZ; MkrS/TM2* and the progeny was homozygosed to give a *prd/SM6Beve-lacZ; P/P* final stock (or equivalent for X insertion).

Between two and fifteen independent insertions of each transgene were tested for adult rescue. In order to test for adult rescue for a particular line, two *prd/SM6Beve-lacZ; P/P* stocks containing the *prd^{f32.12}* and *prd^{2.45}* alleles, were then crossed. Rescue of *prd^{f32.12}/prd^{2.45}* transheterozygotes was judged by the loss of *SM6*, which carries the *Roi* adult marker.

Subsequently, several healthy *prd^{f32.12}/SM6Beve-lacZ; P/P* stocks were amplified in order to analyse cuticle phenotypes. Individual cuticles were given scores reflecting the number of correctly formed denticle belts. Between 200 and 400 cuticles for each line were then analysed statistically in order to obtain a measure of the range of rescue mediated by each transgenic line. The range of phenotypes observed between the different lines of a specific transgene in all cases formed a continuous distribution, implying that any variability observed between these lines was probably the result of insertional position effects.

Based on cuticle phenotypes, the strongest line was then analysed for rescue of target gene expression in *prd* mutant embryos, which were identified by loss of the *eve-lacZ* marker. All lines analysed for target gene expression were tested (in a *prd^{f32.12}* mutant (no protein expression) background) to insure that transgenic Prd expression levels were normal. Although there is significant variability in the intensity of Prd antibody staining observed between individual embryos in a given batch (e.g. wild type), it was possible to estimate levels of transgenic Prd expression by comparing levels (within the same batch) in several embryos expressing only transgenic Prd and several embryos expressing both transgenic and endogenous Prd. All strong lines were thus judged to express normal levels of transgenic Prd. Furthermore, the weaker, alternate stripes of transgenic Prd (see Fig. 1) are clearly visible in all these strong lines, whereas they disappear in weaker lines (not discussed with respect to results). This provides an independent, 'internal' measure of transgenic expression levels, and we therefore concluded by both criteria that transgenic Prd expression was robust in all strong lines tested.

In order to generate a stock that was homozygous for both a *prdGS* transgene (on III) and a *prdNQ* transgene (on the X), the following cross was made:

Females *w⁻.NQ/ w⁻.NQ; prd^{f32.12}/SM6Beve-lacZ* were crossed with males *w⁻; prd^{f32.12}/SM6Beve-lacZ; prdGS/prdGS* to generate males *w⁻.NQ/Y; prd^{f32.12}/SM6Beve-lacZ; prdGS/+*. These males were backcrossed to the parental females to generate *w⁻.NQ/ w⁻.NQ; prd^{f32.12}/SM6Beve-lacZ; prdGS/+* progeny which were homozygosed. All *prd* alleles and the *prd* deficiency used were obtained from the

Tübingen stock center. These alleles are *prd^{f32.12}* (strong), *prd^{2.45.17}* (strong), *prdIIB* (weak), *prdIIN*(weak), *prdIIW*(weak), *prd^{6L}*(strong), *prd^{X3}*(strong) and *Df(2L)prd^{1.7}*. The *SM6Beve-lacZ* chromosome, which carries an *eve-lacZ* marker insertion, was a generous gift of M. Weir.

Raising antibodies to Prd and affinity purification of the antibodies

We raised antibodies to the C-terminal non-conserved portion of the Prd protein, in order to minimise antibody cross-reactivity to other proteins; we expressed a GST-fusion protein containing a peptide that spans the C-terminal half of Prd but in which the PR has been deleted. This peptide was derived by PCR amplifying a plasmid that contains a 57 bp PR deletion (pAR *prd rpt Δ* , gift of E. Harris), using two primers that introduce an artificial *Bam*HI site at the 5' end and an artificial *Eco*RI at the 3' end. This amplified fragment was subcloned into the pGEX-2T vector (Pharmacia). The segment of DNA amplified encodes a peptide, which spans the last 258 residues of Prd, with the PR deleted (residues 355-552 Δ 572-613). This GST fusion protein was expressed and purified on Glutathione sepharose (Pharmacia) and used to inoculate rabbits directly, along with Hunter's Titermax adjuvant (CytRx). Inoculations were performed by Hazleton Research products (Denver, PA).

Serum from the final bleed was then affinity purified on GST-sepharose to which the same peptide that had been used for immunisation had been crosslinked, as described (Harlow and Lane, 1988): Bound antibodies were eluted off the column using 100 mM glycine pH2.5, and neutralized with 1/10 volume of 1M Tris pH 8. They were then dialyzed 3 \times against 2 litres 1 \times PBS; 0.02% NaN₃, and frozen in small aliquots in 10% glycerol, 1% BSA at -20°C. These antibodies have been used at 1/10,000 dilution on embryos and do not require preadsorption before use.

In situ hybridisation and antibody staining of embryos

0-4 hour embryos were collected, washed, fixed and stained, both for in situ hybridisation and staining with Prd antibodies, as described (Simpson-Brose et al., 1994), omitting proteinase K treatment and refixation prior to hybridisation, since this dramatically decreases Prd staining.

The following plasmids were transcribed to make DIG RNA probes: pKS-gsb (gift of E. Harris), linearised with *Sal*I, transcribed with T7 RNA polymerase (NEB); pKS-lacZ (gift of M. Simpson) (Simpson-Brose et al., 1994), linearised with *Sac*I, transcribed with T3 RNA polymerase; pKSII(+)-wg (gift of S. Dougan), linearised with *Bam*HI, transcribed with T7; pKSII(+)-en (gift of J. Mullen), linearised with *Hind*III, transcribed with T7. Anti-Prd antibodies were used at 1:5000 dilution. Embryos were mounted in methyl salicylate: Canada balsam (1:2) and photographed with Nomarski optics using Ektachrome 160 slide film.

Cuticle preparations

Embryos were collected for 2 hours and allowed to age 20-22 hours. They were then washed and dechorionated in bleach. After dechorionation, the embryos were washed and transferred to an Eppendorf containing heptane/methanol. The embryos were then vortexed for 10 seconds to devitellinise them. Heptane and interface were removed and the embryos were washed 3 \times in methanol before being washed 3 \times in PBS/0.1% Tween. The embryos were transferred to a slide and most of the PBS/0.1% Tween was removed with a tissue. 80 µl of Hoyer's:lactate (1:1) was added to the embryos on the slide and they were covered with a coverslip and incubated at 70°C to clear the cuticles. Cuticles were examined using dark-field optics and photographed using Ektachrome 160 slide film.

Expression of Δ PBC peptide and gel mobility shift assay

Using a template containing *prd* Δ PBC genomic DNA, a fragment spanning the PD and HD was amplified by PCR (using 5'PrdPD and

3'PrdHDRI oligos, see 'PCR' below). The PCR product was digested and cloned into the *NdeI* site of pET14b (Novagen). Peptide was expressed in BL21 cells as described (Treisman et al., 1989). Crude bacterial extracts were prepared and used for gel mobility assay.

Oligos were gel purified and labeled with Klenow and $\alpha^{32}\text{P}$ -dATP. In 20 μl , the protein was diluted with GS buffer (15 mM Tris pH 7.5; 6.5% glycerol; 90 mM KCl; 0.7 mM EDTA; 0.2 mM DTT; 0.5 mg/ml BSA; 50 ng/ μl poly dIdC; 0.5% NP40) and mixed with 100 pg of labeled probes for 20 minutes at RT. The mixture was loaded onto an 8% non-denaturing acrylamide (29:1, polyacrylamide: bis-acrylamide) gel buffered in 0.25 \times TBE and electrophoresed at 15 V for 2 hours. The gel was fixed in a 10% HAc, 10% EtOH solution for 15 minutes, dried and exposed to a phosphorimager screen (Molecular Dynamics).

PD-binding site oligo: 5' GATCTTAGTCGTCACGGTTCGTTA-GCTAG 3'

HD-binding site oligo: 5' GATCCTGAGTCTAATCGATTACTG-TACAGCTAG 3'

PCR amplification of genomic DNA derived from flies or embryos

DNA was extracted from flies or embryos as described (Ashburner, 1989). In order to identify the mutations in *prd* alleles, homozygous mutant embryos were identified by their phenotype at the end of embryogenesis and collected. The DNA extracted from 20 homozygous *prd* embryos was PCR amplified using two primers that span the entire *prd* gene, (EMS5' and EMS3'). The product of this PCR was digested with *HindIII* (which cuts at an internal site close to the 5' end of the *prd* gene) and *EcoRI* (which cuts within the EMS3' primer) and subcloned into pGEM9 (Promega). Mutations were identified by dideoxy sequencing (Sanger et al., 1977) of the cloned DNA.

Similarly, in order to determine that the *prd* Δ PBC transgene indeed contains a mutation, DNA was extracted from a single transgenic fly and subjected to PCR amplification using two primers that span a 550 bp region of the *prd* gene from within the PD (*mutseq3*) (but 5' to the Δ PBC deletion) to the 3' end of the HD (3'PrdHDRI). The product of this PCR was run on a 1% Agarose TBE gel in order to analyse the products.

For PCR amplification of *prd* genomic sequences, 35 cycles were performed as follows: 94°C 1 minute; 59°C 1 minute; 72°C 5 minutes. For amplification of *prd* Δ PBC coding sequence for subcloning, 30 cycles were performed as follows: 94°C 1 minute; 50°C 1 minute; 72°C 1 minute. In all cases, PCR amplification was performed in 100 μl 20 mM Tris pH 7.5; 50 mM KCl; 2 mM MgCl₂; 0.2 mM dNTPs, 1 unit Taq polymerase (Perkin Elmer Cetus) and 500 ng of each primer.

Oligos:

5'PrdPD: 5' CAATAGAATTTTCATATGGGCCAGGGGCGCGTCAA 3'

3'PrdHDRI: 5' CTATGAATTCTCATGAGACCGAGGTGTGCTGCTTG 3'

mutseq3: 5' AGACTGGCTCCATTAGAC 3'

EMS5': 5' GATATGGCCGACGTCTGG 3'

EMS3': 5' CCAGAATTCTCAGCGGTTTCTCAAACG 3'

RESULTS

An in vivo assay for *prd* function

In order to dissect *paired* function in vivo, we generated germ-line transformant lines carrying a transgene in which the endogenous *prd* 5' regulatory sequences (referred to henceforth as the *prd* 'promoter') are used to drive expression of a *prd* coding 'cassette' into which mutations of interest can be introduced. By crossing this transgene into a genetic background

that lacks endogenous *prd*, the function of the transgene can be assayed in terms of its ability to rescue missing *prd* function. This function is assayed at three levels: rescue of *prd* mutants to adulthood, rescue of the *prd* mutant cuticular phenotype and rescue of *prd* target gene expression. With respect to the rescue of target gene expression, it should be noted that these targets are defined genetically and need not be direct targets of the Prd protein (Mullen and DiNardo, 1995). Because its expression is driven by the *prd* promoter, transgenic *prd* is expressed only in the cells where endogenous *prd* is usually expressed, and at levels similar to those of endogenous *prd*. Such assay systems, which have previously been used to dissect other HD proteins (Furukubo-Tokunaga et al., 1993; Heberlein et al., 1994), thus provide an extremely sensitive, physiologically accurate, way to analyse gene function in the developing *Drosophila* embryo.

Since 10 kb of *prd* promoter is sufficient to drive expression of a *lacZ* reporter gene RNA in a striped pattern that closely resembles that of *prd* (Gutjahr et al., 1994), we decided to use this 5' DNA to drive the expression of the *prd* transgene. We thus generated a P-element construct in which a 12 kb promoter fragment drives the expression of a 3 kb coding 'cassette' which contains the *prd* ORF and approximately 1 kb of 3' sequence (Fig. 1). The *prd* coding cassette allows mutations of interest to be introduced into *prd*.

Rescue by a wild-type *prd* transgene

In a first round of experiments, we tested the ability of a wild-type *prd* transgene to rescue *prd* function. We first examined its ability to rescue seven different *prd* alleles (four strong, three weak; see Materials and methods) over a *prd* deficiency, and every heteroallelic combination of strong *prd* alleles. In all cases, a wild-type *prd* transgene was able to rescue 90-95% of the *prd* mutant embryos to adulthood, when present in two copies, and the rescued adults appeared phenotypically normal. Rescue could also occur with a single copy of the transgene, but the proportion of mutants rescued was reduced to approximately 25% and many of the rescued flies showed partial segment fusions, especially in the abdomen. This suggests a dosage requirement for the rescuing transgene.

We then analysed the cuticles derived from a balanced *prd*^{32.12} stock carrying two copies of a wild-type *prd* transgene. *prd*^{32.12} was chosen as the genetic background for all our analyses of cuticle phenotypes and *prd* target gene expression because, of the *prd* alleles at hand, it was the only healthy, strong *prd* allele that did not express Prd protein (as detectable with our antibodies to the Prd C terminus). As such it provided an ideal background in which to analyse not only the rescue of cuticle phenotype and target gene expression but also the expression of transgenic Prd protein.

Whereas 1/4 of cuticles derived from a balanced *prd*^{32.12} stock looked completely mutant (Fig. 2A), almost all cuticles derived from the same stock carrying two copies of a wild-type *prd* transgene looked completely wild type (Fig. 2B), with the remaining few showing only partial segmentation defects. These results indicate that the mutant cuticles were rescued by the wild-type *prd* transgene and looked essentially phenotypically wild type. However, although rescued homoallelic *prd*^{32.12} mutants completed embryogenesis to form normal larvae, they did not reach adulthood. This was presumably due to other lethal mutations that had accumulated on the *prd*^{32.12} chromosome since the stock was established.

We then examined *prd* target gene expression in a *prd*^{32.12} stock carrying two copies of the strongest rescuing wild-type transgene (based on cuticle phenotypes; see Materials and Methods): Fig. 2B shows that a wild-type transgene could rescue the expression of all three *prd* targets analysed. Although the onset of expression of *prd*-dependent *en* stripes appeared slightly delayed relative to wild type (Fig. 2C), and levels of early expression of *wg* and *gsb* were slightly weaker (but less dramatically affected than *en*) than wild type in all stripes, this initial deficit was rapidly compensated for so that, by stage 9, all rescued expression was indistinguishable from wild type. A similar delay in the onset of *en* expression has been observed in embryos heterozygous for endogenous *prd* (S. DiNardo, personal communication). These results suggest that two copies of a wild-type *prd* transgene may function approximately as well as a single copy of endogenous *prd*. However, although a wild-type *prd* transgene may be somewhat subfunctional relative to endogenous *prd*, it could nonetheless mediate very good rescue of *prd* target gene expression. In addition, these results show that the wild-type transgene could rescue both odd- and even-numbered stripes of both *wg* and *gsb* expression (compare levels of expression in mutant embryos, Fig. 2A), again consistent with a role for *prd* in all segments.

We could thus obtain good rescue of *prd* function at all three levels examined our assay. This assay therefore allows us to test the effects of specific mutations in *prd* in vivo.

Sterility of rescued males reveals a requirement for *prd* in accessory gland formation

Interestingly, the rescued males in all our experiments were all sterile, although the females were fertile. Since *prd* rescue has been achieved with a different transgene (Gutjahr et al., 1994) containing approximately 10 kb of 5' and 5 kb of 3' DNA and these rescued flies were not male-sterile, it seems likely that our transgene (which contains 4 kb less 3' DNA) lacks a 3' enhancer required for fertility. Testes and genital tracts from wild-type and rescued males were stained with an antibody to Prd protein and compared; while wild-type animals showed strong Prd staining in the accessory glands (Fig. 2D), the accessory glands in rescued males were missing (P. Gönczy, personal communication). These results indicate a requirement for *prd* in accessory gland formation. However, since *prd* mutants die at the end of embryogenesis, it has never previously been possible to identify this late requirement for *prd* function and this fact underscores the unique perspective which this transgenic rescue system can provide.

The effect of specific mutations abolishing the ability of either the PD or the HD to bind to DNA

Since Prd contains two (or possibly three) independent DNA-binding domains, the HD and the PAI/RED domain (Treisman et al., 1991), we wished to test the relative contribution each domain makes to Prd function in vivo. For example, since in vitro data (Treisman et al., 1991) suggest that mutations abolishing DNA binding via one domain do not abolish DNA binding via the remaining domain, we wanted to test whether equivalent mutations would have an effect in vivo. In addition, we wanted to determine whether the activation of all the different targets of *prd* requires DNA-binding via sites bound by the same combination of DNA-binding domains, or whether distinct targets are in fact activated through sites for distinct DNA-binding domains. Finally, we wished to test whether both the PD and the HD are required within the same Prd molecule, or whether two Prd molecules with a single functional HD and PD each, are able to transcomplement.

In order to examine these issues, we generated two point mutants of *prd*: *prdGS* and *prdNQ*. We decided to use point mutants rather than deletions in order to cause the least possible disruption of the protein overall. The mutation in *prdGS* is a single amino acid change at position 15 of the PD, which renders the PAI domain of Prd unable to bind to DNA (Treisman et al., 1991). This mutation was first identified in the *undulated* allele of mouse *Pax-1* as the basis for the phenotype (Chalepakos et al., 1991; Treisman et al., 1991). In the crystal structure of the Prd PD bound to DNA (Xu et al., 1995), Gly 15 makes contacts with two different bases in the minor groove of the binding site, and it is likely that spacing constraints would not permit a Ser residue to fit functionally at

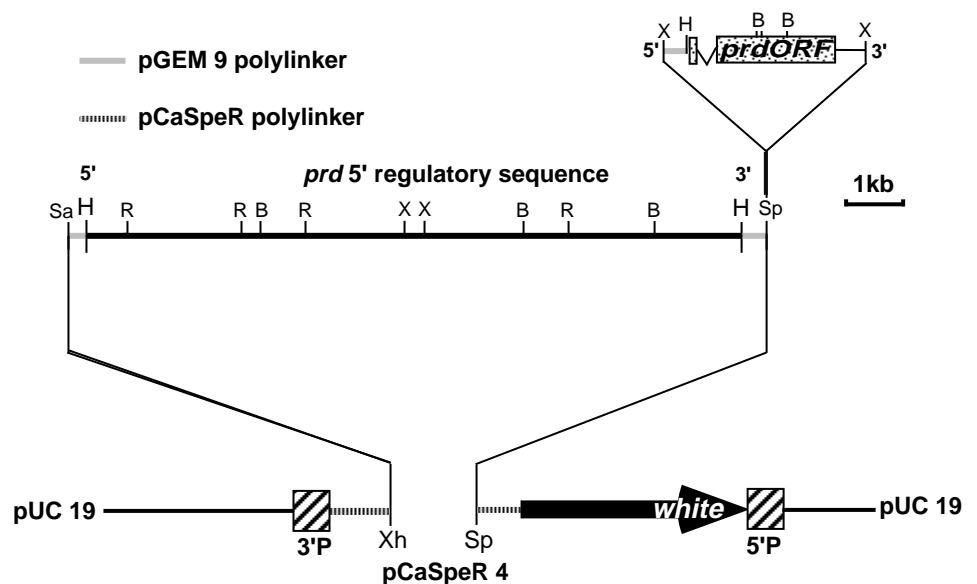


Fig. 1. Construction of rescuing *prd* transgene. Approximately 15 kb of *prd* genomic DNA derived from an EMBL3 *Drosophila* library were divided into two adjoining fragments: a 12 kb 5' *Hind*III fragment ('*prd* 5' regulatory sequence', shown as bold line) and a 3 kb 3' *Hind*III-*Xba*I genomic fragment consisting of a coding cassette containing the *prd* ORF ('*prd* ORF', shown as hatched box) and approximately 0.9 kb of 3' sequences. 6 extra bp (relative to endogenous *prd*) were introduced at the *Hind*III site separating the two fragments, which lies 32 bp upstream of the Start codon, within the untranslated 5' end of the *prd* transcript. Construct is drawn to scale (scale bar), showing the Casper4 vector linearised. Total size of rescue construct is 23 kb. Restriction sites are abbreviated as follows: H, *Hind*III; R, *Eco*RI; X, *Xba*I; B, *Bam*HI; Sa, *Sa*II; Sp, *Spe*I; Xh, *Xho*I.

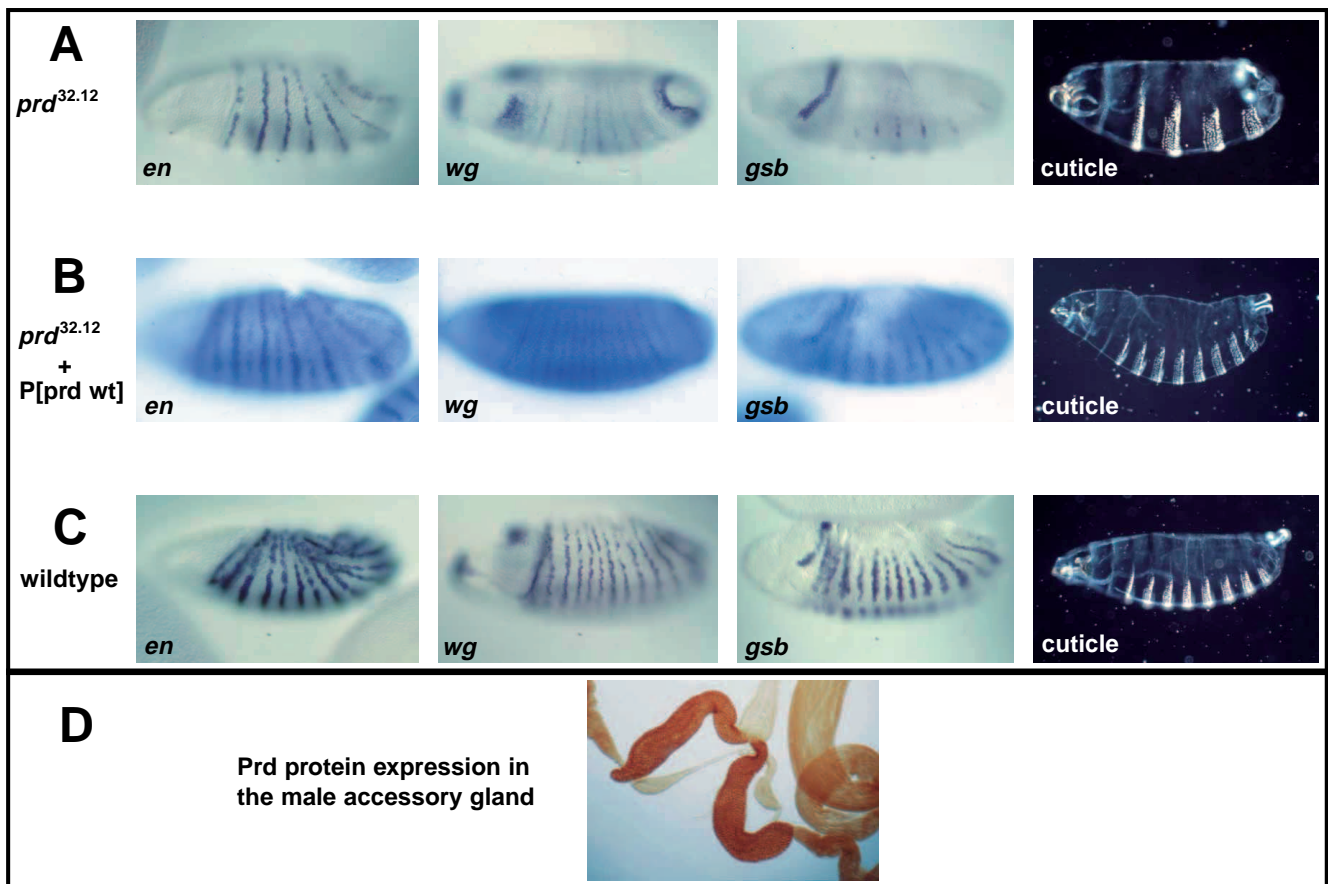


Fig. 2. Rescue of expression of *prd* target genes *en*, *wg* and *gsb* and resulting cuticle mediated by a wild-type *prd* transgene. Embryos were probed with DIG RNA probes to the *prd* target genes as well as *lacZ*, allowing *prd*^{32.12} mutant embryos to be identified by loss of *eve-lacZ* staining. (A) *prd*^{32.12} mutant embryos. (B) *prd*^{32.12} mutant embryos rescued by a wild-type *prd* transgene. Onset of rescued *en* expression is slightly delayed relative to wild type and overall levels of *wg* and *gsb* in all stripes are somewhat weak (but not as dramatically affected as *en*). Rescued cuticle is phenotypically wild type. (C) Wild-type embryos. (D) Expression of Prd protein in wild-type accessory glands. Anti-Prd antibody staining was performed on wild-type male testes to reveal Prd protein expression in the accessory glands.

this position. In addition, it is likely that the overall structure of the β -turn containing the G15 residue would be disrupted by the GS15 mutation.

Similarly, the mutation in *prdNQ* is a single amino acid change at the absolutely conserved position 51 of the HD, which abolishes the ability of the HD to bind to DNA (Hanes and Brent, 1991). The N51 residue of the Prd HD has been shown to mediate a crucial base contact with the core of the HD-binding site (Wilson et al., 1995).

In all cases tested, neither *prdNQ* nor *prdGS* were able to rescue *prd* function in an adult rescue assay. Consistent with this, 1/4 of the cuticles derived from a *prd*^{32.12} stock carrying either two copies of a *prdNQ* transgene, or two copies of a *prdGS* transgene look mutant (Fig. 3A,B).

In order to determine whether *prdNQ* and *prdGS* are able to transcomplement, we analysed cuticles derived from stocks carrying both two copies of *prdNQ* and two copies of *prdGS*. These showed essentially the same phenotypic distribution as the *prdGS* lines alone, indicating that no transcomplementation between *prdGS* and *prdNQ* can be observed in cuticles (Fig. 3C).

Analysis of *prd* target expression in these lines showed that neither *prdNQ* nor *prdGS*, nor both together were able to

rescue target gene expression, except for a very weak transient rescue of *wg* expression, which occurred in all cases at the same time as the onset of wild-type *wg* expression, but faded shortly after stage 7 (Fig. 3).

Although neither *prdGS* nor *prdNQ* was able to rescue *prd* function, staining of these lines for Prd protein, in a *prd*^{32.12} (protein null) background, indicated that the transgenes were indeed all expressed at normal levels (data not shown; see Materials and methods). Furthermore, when crossed with a transgenic line containing a *lacZ* reporter driven by PD-binding sites, a *prdNQ* transgene is able to direct expression of the reporter in a *prd*-like pattern, indicating that it remains biologically active (G. S. and C. D., unpublished data). Similarly, *prdGS* is also biologically active, since it is able to poison endogenous *prd* function, resulting in a weakening of both *en* and *gsb* stripes and giving rise to larvae showing a partially penetrant, weak, *prd* phenocopy (Fig. 3D).

Rescue of *prd* mutants by a *prd* transgene carrying a deletion of the C-terminal subdomain of its PD (Δ PBC)

Since the results obtained for the *prdGS* transgene indicated that a functional PAI domain is required for *prd* function in

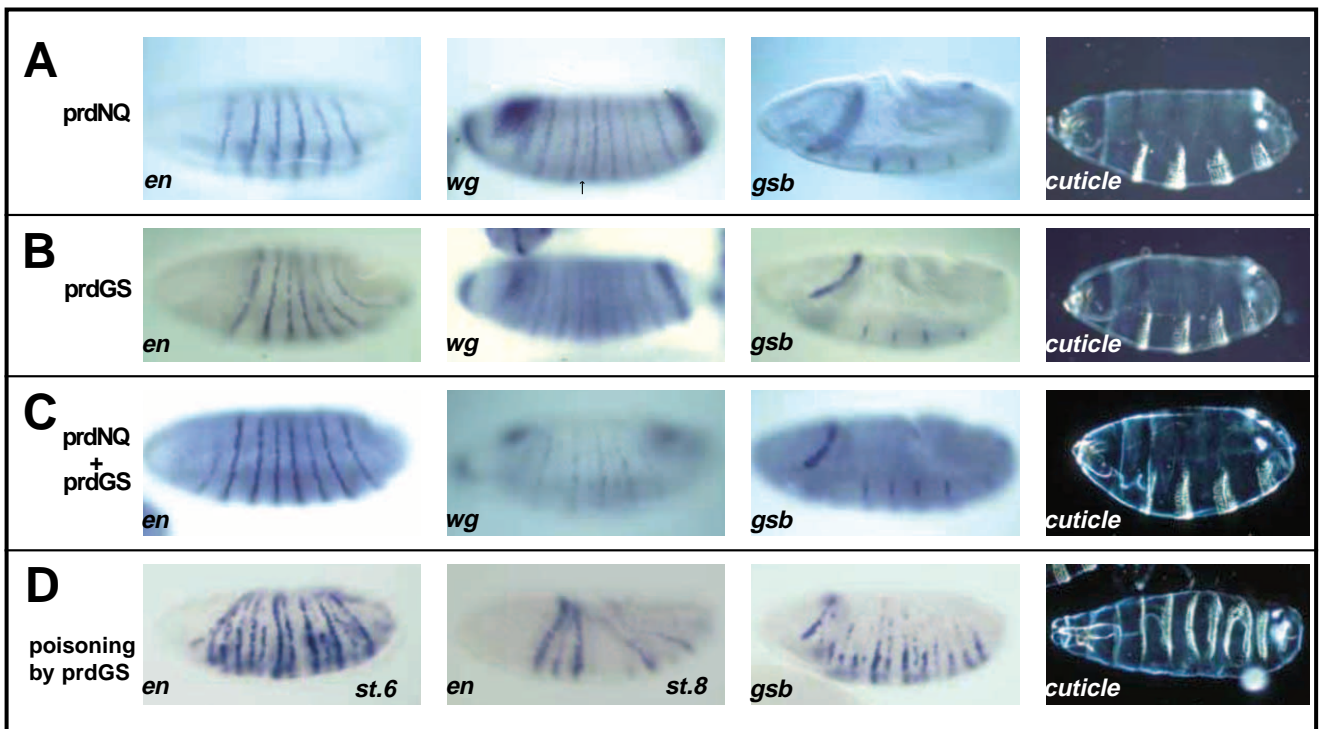


Fig. 3. Rescue of expression of *prd* target genes *en*, *wg* and *gsb* and resulting cuticle mediated by *prdGS* and *prdNQ* transgenes, in *prd*^{32.12} mutant embryos. Embryos were probed with DIG RNA probes to the *prd* target genes and *lacZ*, allowing *prd*^{32.12} mutant embryos to be identified by loss of *eve-lacZ* marker. (A) *prd*^{32.12} embryos carrying 2 copies of a *prdNQ* transgene. No rescue of target gene expression is observed, other than weak, transient rescue of *wg* expression (marked with arrow). (B) *prd*^{32.12} embryos carrying 2 copies of a *prdGS* transgene. Again, only weak, transient rescue of *wg* expression is observed. (C) *prd*^{32.12} embryos carrying 2 copies of both *prdNQ* and *prdGS* transgenes. No transcomplementation between the two transgenes is observed. Again, only weak, transient rescue of *wg* expression is observed. (D) In embryos with at least one copy of endogenous *prd* (*lacZ* positive), presence of one or two copies (genotype cannot be identified) of a *prdGS* transgene causes poisoning of *en* (2 embryos, stages 6 and 8) and *gsb* expression. Poisoned cuticle is shown.

vivo, we wished to determine the relative contribution of the RED domain to *prd* function in vivo. As presented in the Introduction, the RED domain may be dispensable for *prd* function.

In order to test this idea, we introduced into our assay the same deletion of the RED domain (*prd*ΔPBC) as was tested in a *hs-prd* assay (Cai et al., 1994) (generously provided by M. Weir). The *prd*ΔPBC transgene contains a deletion of amino acids 75-125 in the PD. Based on the crystal structure, this deletion should not affect DNA binding via the PAI domain. As expected from our previous in vitro studies, we demonstrated that a PrdΔPBC (PD-HD) peptide binds in vitro to the same consensus site as selected with the entire PD (Fig. 4A), indicating that the DNA-binding functions of the PrdΔPBC PD are the same as for the wild-type PD.

In an adult rescue assay, a *prd*ΔPBC transgene was able to rescue *prd* function almost as well as a wild-type *prd* transgene. However, more than half of the *prd*ΔPBC lines tested were only able to rescue to adulthood when the transgene was present in two copies, indicating that *prd*ΔPBC function may be somewhat reduced relative to a wild-type transgene. Analysis of cuticles and *prd* target gene expression in the strongest (see Materials and Methods) *prd*ΔPBC line, shows that it is able to rescue *prd* function to a degree indistinguishable from the rescue obtained with a wild-type *prd* transgene (Fig. 5A).

Since we were surprised by the degree of rescue mediated

by the *prd*ΔPBC transgene, we wanted to confirm that this transgenic line indeed contained a deletion. We thus PCR amplified genomic DNA derived from flies carrying the transgene (both from a stock and from a rescued fly) and were able to demonstrate that the ΔPBC deletion is indeed present (Fig. 4B) in the transgenic flies.

An alternative interpretation is that the endogenous mutant protein may serve as a landing platform for the exogenous protein lacking the RED domain. There are several reasons to believe that it is not the case. First, the results are completely consistent with our molecular and structural analyses indicating that the RED domain is dispensable. The same RED domain deletion also exhibits the same normal function when tested by a heat-shock overexpression assay in vivo (Cai et al., 1995). Second, a wild-type construct achieved qualitatively identical rescue to adulthood for *prd*^{32.12}/Df *prd* and *prd*^{2.45}/Df *prd*, a real null combination (or for any combination of strong alleles). Third, as no protein is detectable in *prd*^{32.12} mutants using an antibody against the C-terminal end of the molecule, the longest possible molecule would completely lack an activation domain. *prd*^{K3} (see above) contains a stop codon just after the homeodomain and therefore lacks the entire activation domain (Fig. 6), just like the putative best possible product of *prd*^{32.12}. It exhibits a dramatic 'poisoning' dominant negative effect which is never seen with *prd*^{32.12}, arguing that *prd*^{32.12} is truly biologically inactive.

Rescue by *prd* transgenes affected in HD dimerization

The *prd* HD binds *in vitro* as a dimer (Wilson et al., 1993). In order to dissect the role of HD dimerization *in vivo*, we analysed two different mutant transgenes in which HD dimerization is affected.

(1) *prdAR43*

The AR43 mutation affects a conserved residue in Prd class HDs. *In vitro* data indicate that the AR43 mutation causes the Prd HD to completely lose its ability to cooperatively dimerize and the crystal structure of a Prd HD dimer on DNA shows that the two A43 residues are positioned close to each other in the HD dimer (Wilson et al., 1995). As such, the AR43 mutation changes a small (permissive) residue to a large (non-permissive) residue, which should be unable to fit functionally in the dimer. Nonetheless, the PrdAR43 HD retains its ability to bind as a wild-type monomer (Wilson et al., 1995).

The majority of *prdAR43* lines were able to mediate adult rescue. However, rescue usually required two copies of the transgene and we observed a wide range of quality of adult rescue between lines, which correlated with different degrees of cuticle rescue.

Analysis of *prd* target gene expression in the strongest (see Materials and Methods) *prdAR43* line showed that, although quite good rescue occurred, the onset of target gene expression driven by the *prdAR43* transgene was significantly delayed and weak compared to that driven by a wild-type transgene (Fig. 5B). Odd-numbered *en* stripes were delayed and weak and, to a lesser extent, *wg* and *gsb* were weak in all stripes. The greatest deficit was seen with *en*, which showed an alternate strong/weak expression pattern that persisted until the end of germband extension. However, although target gene expression mediated by the *prdAR43* transgene showed deficits, the cuticle produced at the end of embryogenesis looked wild type (Fig. 5B).

These results indicate that HD-mediated cooperative dimerization provides a relatively important, but non-essential, component of the overall specificity of Prd.

(2) *prdSQ50*

The SQ50 mutation is a single amino acid substitution, which has two effects. First, it converts the monomeric DNA-binding specificity of the Prd HD *in vitro* from its normal TAATCG site to that of Ftz, which prefers to bind to TAATTG (Treisman et al., 1989). In addition, residue 50 has been shown to define DNA-binding specificity *in vivo* in other HD proteins (Capovilla et al., 1994; Schier and Gehring, 1992; Sun et al., 1995). Second, whereas a wild-type Prd HD cooperatively dimerizes on two different palindromic binding sites, P2 and P3, with approximately the same degree of cooperativity (50 \times), the SQ50 homeodomain cooperatively dimerizes only on P3 sites and with 200 \times cooperativity (Wilson et al., 1993). Finally, S50 is always found in Prd class HDs associated with a PD. This association suggests that S50 might be involved in HD-PD interactions.

Most of the *prdSQ50* lines tested were able to mediate rescue to adulthood, although some lines could only rescue when two copies of the transgene were present. They were able to mediate a range of cuticle rescue from good to full rescue.

Analysis of *prd* target gene expression and cuticles in the

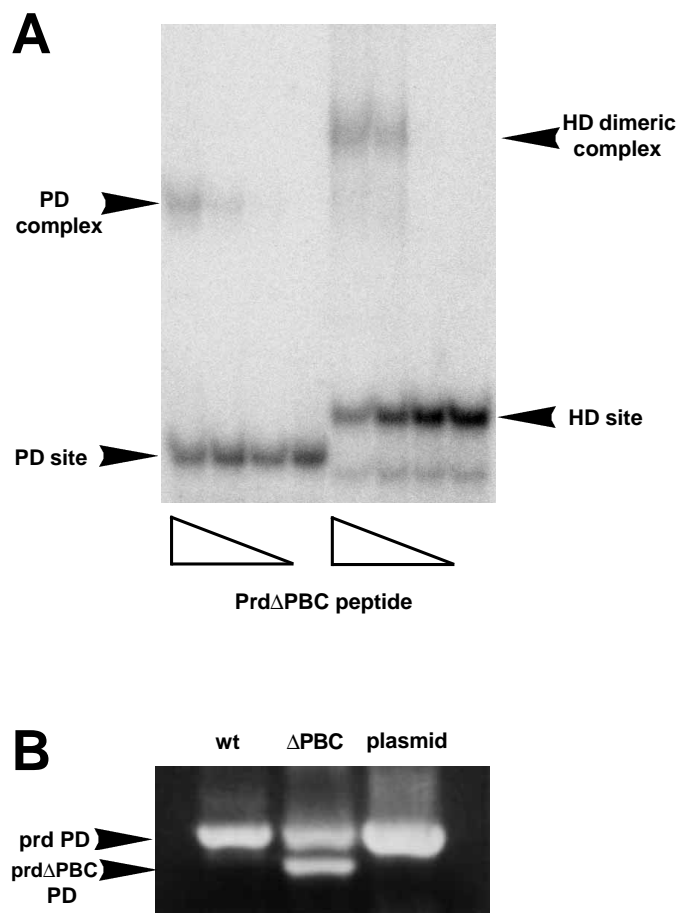


Fig. 4. (A) Prd Δ PBC is able to bind *in vitro* to the same site as selected with the entire PD. Gel shift assay of Prd Δ PBC peptide on PD- and HD-binding sites. Prd Δ PBC peptide corresponds to the region of Prd spanning the PD and HD and containing the Δ PBC deletion in the PD. Ten-fold serial dilutions of the Prd Δ PBC peptide are able to shift the consensus PD-binding site (left half of gel) and the consensus HD-binding site (right half of gel). 4th and 8th lanes contain free probe alone. The Prd Δ PBC peptide binds as a cooperative dimer to the HD consensus site. (B) the transgene in *prd* Δ PBC contains the Δ PBC deletion in its PD. PCR performed, using a set of primers spanning the region from within the 5' end of the PD to the 3' end of the HD, on genomic DNA derived from wild-type flies (lane 1), or from a *prd* 32.12/SM6; *prd* Δ PBC/*prd* Δ PBC stock (lane 2), or on a plasmid containing wild-type genomic *prd* DNA (lane 3). PCR on wild-type flies produces a single 550 bp band, identical to that produced from a plasmid containing wild-type genomic *prd* DNA. In contrast, PCR performed on the transgenic *prd* Δ PBC1 stock produces a doublet. The lower band (400 bp) corresponds to the Δ PBC deletion, indicating that the *prd* Δ PBC1 transgenic line indeed contains the deletion. The same doublet was seen with DNA derived from rescued *prd* Δ PBC1 flies (data not shown).

strongest (see Materials and Methods) *prdSQ50* line (Fig. 5C) indicated that *prdSQ50* was able to mediate approximately the same degree of rescue as a wild-type *prd* transgene (compare to Fig. 2B).

Thus, although the distribution of adult and cuticle rescue observed with different *prdSQ50* lines suggests that the *prdSQ50* transgene does not function at completely wild-type

levels, a strong *prd*SQ50 line was able to function essentially as a wild-type transgene.

These results indicate that, although the component of HD specificity mediated by S50 may play a role in the overall specificity of *prd*, this role may be relatively minor compared to the sum of the specificities mediated by its different DNA-binding domains together.

Rescue by a *prd* transgene containing a deletion of the PRD (His-Pro) repeat

In order to attempt to define the in vivo activation domain of Prd, we tested a deletion mutant that had been tested in a *hs-prd* assay: *prd*ΔPRT (gift of M. Weir) (Cai et al., 1994). This transgene contains a Stop codon that truncates the C-terminal 74 amino acids (539-613), including the PRD repeat (PR) and approximately 50 residues C-terminal to it. In a *hs-prd* assay, *hs-prd*ΔPRT loses its ability to drive ectopic target gene expression, although it retains some activity in cell culture (Cai et al., 1994). These results suggested that the deletion in *prd*ΔPRT deletes a required activation domain, consistent with previous results that had shown that a deletion of the C-terminal 121 amino acids of Prd resulted in loss of activation of an artificial *en* reporter in cell culture (Han et al., 1989). Furthermore, amino acids 539-613 of Prd have been shown to be capable of functioning as a heterologous activation domain (Cai et al., 1994), indicating that the region contained in these last 74 amino acids of Prd constitutes an activation domain, of which the PR forms a part.

Although the majority of *prd*ΔPRT lines were unable to rescue to adulthood, one line was able to mediate adult rescue when present in one or two copies. The degree of cuticle rescue mediated by the different *prd*ΔPRT lines formed a continuous distribution from fully mutant to fully wild type, with individual lines showing cuticle rescue in a range of phenotypes, which was consistent with their ability or inability to cause adult rescue.

Analysis of *prd* target gene expression and cuticles showed that the strongest *prd*ΔPRT line was able to rescue as well as a wild-type transgene (Fig. 5D). It is thus clear that, although *prd*ΔPRT transgenes may be subfunctional in general, the rescue mediated by a strong *prd*ΔPRT is as good as that of a wild-type transgene.

These results suggest that the ΔPRT deletion does not eliminate an essential activation domain of Prd, which remains to be identified.

In order to further address the location of this activation domain, we have determined the molecular basis of the mutations in two different alleles of *prd*: *prd*X3 and *prd*IIN (Fig. 6). The two alleles shed further light on the Prd activation domain. *prd*X3 contains a mutation truncating the last 288 amino acids of Prd. This mutation gives rise to a molecule that is completely non-functional by genetic criteria (strong allele) and which cannot transactivate a reporter gene driven by HD-binding sites in cell culture (L. Aguilar and C. B., unpublished data). In contrast, *prd*IIN contains two substitutions relative to wild type: TA406 (which may be a permissive polymorphism) and QStop487, which truncates the last 126 amino acids of Prd. Although *prd*IIN can only weakly transactivate a HD reporter in cell culture (L. Aguilar and C. B., unpublished), it displays a weak, temperature-sensitive, phenotype (Tearle and Nusslein-Volhard, 1987), indicating the PrdIIN molecule is at

least partially functional in vivo. By comparison to the non-functional PrdX3 molecule, this in turn suggests that the activation domain of Prd extends to between the beginning of the truncation in PrdIIN (residue 487) and the beginning of the truncation in PrdX3 (residue 325) (Fig. 6). By these criteria, the activation domain of Prd extends at least some 50 residues further N terminally than the activation domain defined by the ΔPRT mutation.

DISCUSSION

A wild-type *prd* transgene can rescue *prd* mutant embryos to adulthood

We have established a rescue assay to dissect *prd* function in vivo. In this assay, a transgene expressing *prd* (or mutated versions of *prd*, summarized in Fig. 7) under the control of the *prd* promoter is expressed in embryos lacking endogenous *prd* function, in order to test its ability to rescue missing *prd* function. Using two copies of a wild-type *prd* transgene, we have rescued 90-95% of *prd* mutants to adulthood for every transheteroallelic combination of 4 strong *prd* alleles and a *prd* deficiency. These rescued adults look phenotypically wild type. A single copy of a wild-type *prd* transgene can also rescue to adulthood, although only 25% of *prd* mutants are rescued and many of the rescued adults show partial segment fusions consistent with the notion that the quality of rescue observed is dependent on the dose of the rescuing *prd* transgene. This dosage dependence for *prd* function is reminiscent of that of mammalian *Pax* genes, which are usually haploinsufficient (for a review see (Strachan and Read, 1994)), and may reveal mechanistic commonalities.

Consistent with the high degree of phenotypic rescue observed in adults, both the cuticles and *prd* target gene expression rescued by a wild-type *prd* transgene show a high degree of rescue (Fig. 2). Although the onset of *en* expression is slightly delayed relative to wild type, and although *wg* and *gsb* seem to be expressed at slightly weaker levels than normal (but less affected than *en*), these slight deficits are rapidly compensated for, so that by stage 9 rescued target gene expression is indistinguishable from wild type. This compensation is likely to occur through interregulation between the target genes (DiNardo et al., 1988; Heemskerk et al., 1991), and may reflect a functional redundancy between *prd* and its target *gsb*, which encode functionally interchangeable proteins, which are expressed at overlapping times in the same cells (Li and Noll, 1994). As such, it is possible that a subfunctional Prd molecule may function sufficiently to establish *gsb* expression, which could then functionally substitute for *prd*. However, the initial expression of *en*, *wg* and *gsb* RNA driven by a *prd* transgene cannot depend on *gsb*, since the onset of expression of these genes precedes any potential *gsb* function.

Both the HD and PD of Prd are required, within the same molecule, for normal *prd* function

Results with *prd*GS and *prd*NQ transgenes (Fig. 3) indicate that, if a Prd molecule contains either a mutation abolishing the ability of its PAI domain (*prd*GS) or HD (*prd*NQ) to bind to DNA, it is no longer able to rescue *prd* function. Although very weak, transient expression of *wg* is observed with both *prd*GS and *prd*NQ, neither *en* nor *gsb* shows any rescue with

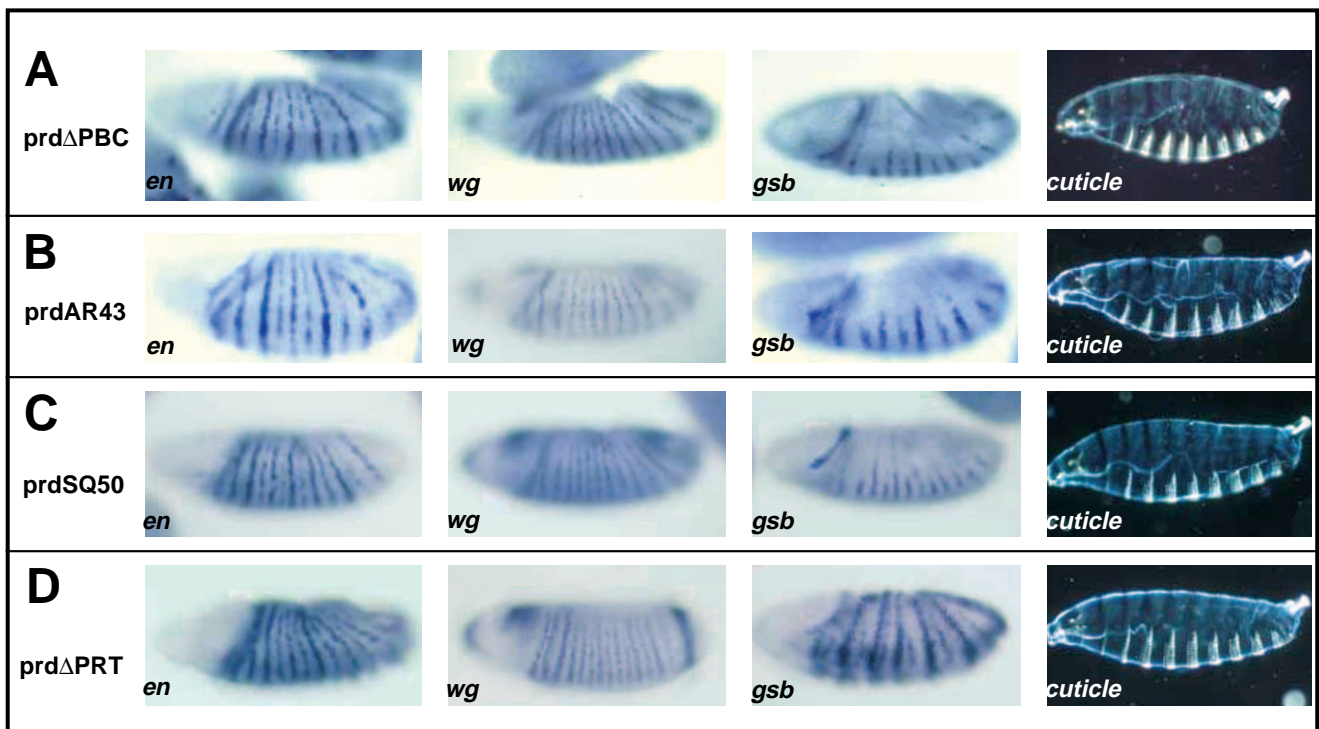


Fig. 5. Rescue of expression of *prd* target genes *en*, *wg* and *gsb* and resulting cuticle mediated by *prd* Δ PBC, *prd* AR43, *prd*SQ50 and *prd* Δ PRT transgenes, in *prd*^{32.12} mutant embryos. Embryos were probed with DIG RNA probes to the *prd* target genes and *lacZ*, allowing *prd*^{32.12} mutant embryos to be identified by loss of *eve-lacZ* marker. (A) *prd*^{32.12} embryos carrying 2 copies of a *prd* Δ PBC transgene. Rescued target gene expression is indistinguishable from that mediated by a wild-type transgene. (B) *prd*^{32.12} embryos carrying 2 copies of a *prd*AR43 transgene. Target genes are rescued but there is a significant delay in the onset of rescued *en* expression and rescued *wg* and *gsb* expression are weaker than with a wild-type transgene. (C) *prd*^{32.12} embryos carrying 2 copies of a *prd*SQ50 transgene. Rescued target gene expression is indistinguishable from that mediated by a wild-type transgene. (D) *prd*^{32.12} embryos carrying 2 copies of a *prd* Δ PRT transgene. Rescued target gene expression is indistinguishable from that mediated by a wild-type transgene.

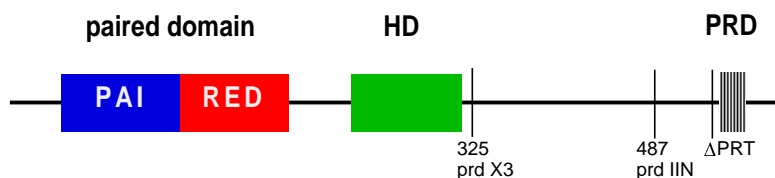


Fig. 6. Schematic representation of the truncations identified in the *prd* X3 and *prd* IIN alleles, compared to the truncation in the *prd* Δ PRT transgene.

either of the transgenes. Since *en* is required to maintain *wg* expression once established (Heemskerk et al., 1991), it is likely that *wg* fades shortly after stage 7 because it is not maintained. As such, *wg* cannot execute its function to pattern the developing epidermis; the resulting cuticles look completely mutant and no adult rescue occurs.

In addition, we determined that *prd*GS and *prd*NQ are unable to transcomplement each other to rescue *prd* function, when both are expressed in the same embryo (Fig. 3C). Again, very weak, transient expression of *wg* is observed, but no cuticle or adult rescue occurs. This result contrasts with results obtained from the *hs-prd* system, where a weak degree of transcomplementation has been observed (M. Weir, personal communication). Since cooperativity between the HD and PD of Prd has been demonstrated for DNA-binding in vitro, even when the two domains are in separate molecules (S. J. and C. D., unpublished data), it is likely that transcomplementation occurs at levels too weak to rescue *prd* function at physiolog-

ical levels of transgene expression, but that upon ectopic expression such transcomplementation may become visible. Alternately, it is possible that an intramolecular interaction between the mutated and functional domains within either *prd*GS or *prd*NQ might result in steric hindrance of intermolecular transcomplementation.

Finally, although an interdependence between the HD and PD has been shown for DNA binding of Pax-3, the closest mammalian relative of Prd (Underhill et al., 1995), all present data for Prd suggest that its PD and HD can function as truly independent DNA-binding domains and that mutations abolishing the ability of either the PD or the HD to bind to DNA do not affect the DNA-binding functions of the other domain, both in vitro (Treisman et al., 1991) or in cell culture (M. Weir, personal communication; G. S., unpublished observations). In addition, our results indicate that, at least for *prd*NQ, the mutation does not affect both DNA-binding domains, since *prd*NQ is able to drive expression of a PD-reporter in vivo (G.

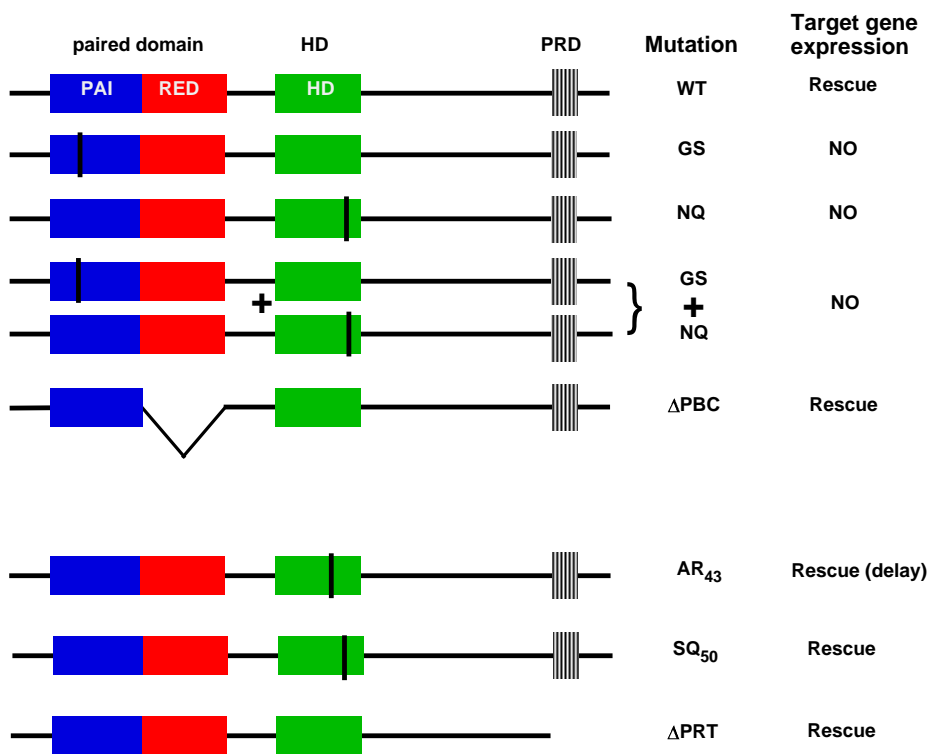
Rescue of *prd*⁻ mutant by transgenes

Fig. 7. Schematic representation of the rescue of target gene expression observed with different *prd* transgenes. The different transgenes tested in our in vivo assay are represented schematically. The conserved domains of Prd are highlighted, along with the specific mutations tested. The ability of these transgenes to rescue *prd* target gene expression is listed.

Sheng, C. B. and C. D., unpublished data). Furthermore, since the same dominant negative 'poisoning' effect has been described for both *prd*^{GS} and for *prd*^{ΔPB} (deletion of the entire PD) in a *hs-prd* assay (M. Weir, personal communication), this suggests that these two molecules function in similar ways independently of the presence or absence of a non-functional PD.

Our results indicate that both the HD and the PAI domain are required, within the same molecule, for normal *prd* function. The dramatic loss of function in *prd*^{GS} and *prd*^{NQ} mutants confirms that Prd binding to DNA via both its HD and PAI domain simultaneously (and probably cooperatively) constitutes a critical component of its overall functional specificity as a transcription factor. Consistent with this, a binding site containing a PD halfsite followed by a monomeric HD halfsite, which could mediate this mode of binding, has been identified in the *eve* promoter and has been shown to be required for *prd*-dependent refinement of late *eve* expression (Fujioka et al., 1996). A very similar binding site (PH0) has been defined in vitro as an optimal site for PD-HD cooperative binding (Jun and Desplan, 1996). However, although binding via its PAI domain and HD simultaneously appears to be a critical mode of function for Prd, these results do not exclude that other modes of binding may be used in specific circumstances.

The RED domain is dispensable for Prd function in vivo

The results for the *prd*^{ΔPBC} transgene indicate that the RED domain of Prd is dispensable for its function in vivo (Fig. 5A). Rescued expression of *prd* target genes is indistinguishable from that conferred by a wild-type transgene (Fig. 2B) and both

rescued cuticles and adults look phenotypically wild type when 2 copies of the transgene are present. This is consistent with in vitro (Treisman et al., 1991; Jun and Desplan, 1996) and structural data (Xu et al., 1995) that indicate that the RED domain is not required for normal binding of the Prd PD to DNA (Fig. 4), although it does appear to make a small contribution to the overall binding affinity of the Prd PD (Jun and Desplan, 1996). In addition, they are consistent with results obtained with a *hs-prd* assay system, which also indicated that the RED domain of Prd is not required to cause ectopic expression of *en* and *gsb* (Cai et al., 1994).

Since the RED domain and PAI domain of Prd fold into a very similar HTH structures (and indeed similar also to the HD), this suggests that, although the RED domain does not contact DNA, it could in principle do so. Indeed, this is the case in other *Pax* genes such as *Pax-5* (Czerny et al., 1993) and *Pax-6* (Epstein et al., 1994). In addition, a newly described *Pax* gene, *lune*, contains a HD and only a RED domain (Jun and Desplan, 1996).

These results suggest that the RED domain may indeed constitute an independent DNA-binding domain or that it may function as a DNA-binding domain in association with either a PAI domain, or with a HD. Since Prd thus contains three HTH DNA-binding domains, two in the PD and one in the HD, and these HTH domains appear to be used in different combinations in different *Pax* genes, it may be useful to think of them as modular in nature, perhaps similar to individual fingers in Zn finger proteins (Miller et al., 1985), which can be used in different combinations to achieve different DNA-binding specificities. The second HTH module in Prd thus appears to be dispensable for *prd* function in regulating segment polarity

gene expression. It may simply be redundant, or it may in fact have other functions that are either obscured by a redundancy between *prd* and *gsb*, or may be involved in other, as yet undefined, roles for *prd* later in development.

Cooperative dimerization of the HD constitutes an important, but non-essential component of the in vivo specificity of Prd

A mutation abolishing cooperative dimerization of the HD results in a subfunctional Prd molecule, arguing that cooperative dimerization of the HD provides a relatively important, but not critical, component of the overall specificity of the Prd protein. Indeed, some *prd* function may be mediated by cooperative dimerization of the HD, since a site foot-printed by Prd in a *prd*-responsive promoter element in the *gsb* promoter (Li and Noll, 1994) is similar to a previously defined optimal binding site for a *prd*-class HD dimer (P3) (Wilson et al., 1993).

In contrast, the SQ50 mutation, which changes the specificity of the HD from both P2 and P3 palindromic binding sites in wild-type Prd, to only P3 sites in the SQ50 mutant HD (Wilson et al., 1995), appears not to have a dramatic effect on *prd* function, although the range of cuticle rescue obtained between different lines suggests that the *prd*SQ transgene may be slightly impaired relative to a wild-type transgene.

Residue 50 of the HD had previously been shown to be important for monomeric DNA binding of HDs (Treisman et al., 1989). However, the relatively small contribution that this residue makes to the overall specificity of the Prd HD is not surprising in light of the fact that the SQ50 mutation makes only a relatively subtle difference to the dimeric binding of the Prd HD to P3 sites in vitro (Wilson et al., 1995) and the fact that we have shown that HD dimerization is not absolutely required for *prd* function.

Combined with results from the *prd*GS and *prd*NQ transgenes, these results indicate that binding via the PAI domain and HD simultaneously, provides a critical component of the overall functional specificity of Prd in vivo, with HD dimerization also being relatively important. At present it is not clear whether both modes of DNA binding could function simultaneously on the same binding sites in vivo. While binding sites selected for a PD-HD peptide in vitro have occasionally contained PH0-binding sites whose HD site is palindromic: ACTCAAGCGTGAC TAATTGATTA (e.g. PD-HD-HD) (Jun and Desplan, 1996), no such binding sites have been identified in vivo.

The in vivo activation domain of Prd appears to extend further N-terminally than the previously defined activation domain deleted in the *prd*ΔPRT deletion mutant

The *prd*ΔPRT lines as a whole appear to be significantly impaired compared to a wild-type transgene. The very wide range of phenotypic rescue, both between lines and within lines, is consistent with the idea that the *prd*ΔPRT transgene may be significantly subfunctional and hence particularly sensitive to position effects and minor variations in transgene expression levels between embryos. However, although the domain deleted in *prd*ΔPRT is able to function as a heterologous activation domain (Cai et al., 1994), it is non-essential to mediate good control of *prd* target genes in our assay system.

This in turn suggests that a region of the C-terminal half of Prd which is not deleted in *prd*ΔPRT is able to function as an activation domain. Analysis of *prd* mutant alleles suggests that the Prd activation domain extends between residues 325 and 487, or 50 residues further N-terminally than the domain deleted in *prd*ΔPRT (Fig. 6).

We are grateful to Michael Weir and Tad Goto for their insightful comments on this manuscript. We also thank Steve DiNardo for repeated input, and all the members of the DiNardo and Desplan labs for intellectually stimulating discussions and encouragement, on this work. We are particularly indebted to David Wilson for conceptual collaboration on many of the HD mutants; Michael Weir for a productive intellectual collaboration and for sharing both reagents and results prior to publication; Luc Aguilar for assistance in cloning *prd* alleles; and Mark Benedyk for the genomic library and help with cloning. We also acknowledge the excellent technical assistance of Yan Gu and Jason Ein and, in particular, Kathy Snook-Branda and Terry Turner for invaluable help with the injections and crosses.

REFERENCES

- Ashburner, M. (1989). *Drosophila: A Laboratory Manual*. 108-109. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
- Benedyk, M. J., Mullen, J. R. and DiNardo, S. (1994). *odd-paired*: a zinc finger pair-rule protein required for the timely activation of *engrailed* and *wingless* in *Drosophila* embryos. *Genes Dev.* **8**, 105-117.
- 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 in *Drosophila*. *Cell* **47**, 1033-1049.
- 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.
- 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.
- Capovilla, M., Brandt, M. and Botas, J. (1994). Direct regulation of *decapentaplegic* by *Ultrabithorax* and its role in *Drosophila* midgut morphogenesis. *Cell* **76**, 461-475.
- Chalepakis, G., Fritsch, R., Fickenscher, H., Deutsch, U., Goulding, M. and Gruss, P. (1991). The molecular basis of the *undulated/Pax-1* mutation. *Cell* **66**, 873-84.
- Chalepakis, G., Stoykova, A., Wijnholds, J., Tremblay, P. and Gruss, P. (1993). *Pax*: gene regulators in the developing nervous system. [Review]. *J. Neurobiol.* **24**, 1367-84.
- Czerny, T., Schaffner, G. and Busslinger, M. (1993). DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev.* **7**, 2048-2061.
- DiNardo, S. and O'Farrell, P. H. (1987). Establishment and refinement of segmental pattern in the *Drosophila* embryo: spatial control of *engrailed* expression by pair rule genes. *Genes Dev.* **1**, 1212-1225.
- 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.
- Epstein, J. A., Glaser, T., Cai, J., Jepeal, L., Walton, D. S. and Maas, R. L. (1994). Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. *Genes Dev.* **8**, 2022-2034.
- Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M. (1986). Structure of the segmentation of gene *paired* and the *Drosophila* PRD gene set as part of a gene network. *Cell* **47**, 735-746.
- Frischauf, A.-M., Lehrach, H., Poustka, A. and Murray, N. (1987). Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* **170**, 827-842.
- Fujioka, M., Miskiewicz, P., Raj, L., Gullledge, A. A., Weir, M. and Goto, T. (1996). *Drosophila* Paired regulates late *even-skipped* expression through a composite binding site for the paired domain and the homeodomain. *Development* **122**, 2697-2707.
- Furukubo-Tokunaga, K., Flister, S. and Gehring, W. J. (1993). Functional specificity of the Antennapedia homeodomain. *Proc. Natl. Acad. Sci. USA* **90**, 6360-6364.

- Gruss, P. and Walther, C. (1992). Pax in Development. *Cell* **69**, 719-722.
- 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.
- Halder, G., Callaerts, P. and Gehring, W. J. (1995). Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. *Science* **267**, 1788-1792.
- Han, K., Levine, M. S. and Manley, A. J. L. (1989). Synergic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56**, 573-583.
- Hanes, S. D. and Brent, R. (1991). A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* **251**, 426-430.
- Harlow, E. and Lane, D. (1988). *Antibodies: A Laboratory Manual*. Chapter 8, 313; Chapter 13, 519-552. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
- Heberlein, U., Penton, A., Falsafi, S., Hackett, D. and Rubin, G. M. (1994). The C-terminus of the homeodomain is required for functional specificity of the *Drosophila* *rough* gene. *Mech. Dev.* **48**, 35-49.
- Heemskerk, J., DiNardo, S., Kostriken, R. and O'Farrell, P. H. (1991). Multiple modes of *engrailed* regulation in the progression towards cell fate determination. *Nature* **352**, 404-410.
- Ingham, P. W. (1988). The molecular genetics of embryonic pattern formation in *Drosophila*. *Nature* **335**, 25-34.
- Ingham, P. W., Baker, N. E. and Martinez-Arias, A. A. (1988). Regulation of segment polarity genes in the *Drosophila* blastoderm by *fushi tarazu* and *even skipped*. *Nature* **331**, 73-75.
- Jun, S. and Desplan, C. (1996). Cooperative interactions between paired domain and homeodomain. *Development* **122**, 2639-2650.
- Li, X. and Noll, M. (1994). Evolution of distinct developmental functions of three *Drosophila* genes by acquisition of different cis-regulatory regions. *Nature* **367**, 83-7.
- Maniatis, T., Fritsch, E. F. and Sambrook, J. (1989). Analysis and cloning of eukaryotic genomic DNA. In *Molecular cloning: a Laboratory Manual*, Second Edition, vol. 2, pp. 9.2-9.62 Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.
- Miller, J., McLachlan, A. D. and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609-14.
- 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.
- Mullen, J. R. and DiNardo, S. (1995). Establishing parasegments in *Drosophila* embryos: Roles of the *odd-skipped* and *naked* genes. *Dev. Biol.* **169**, 295-308.
- Nelson, H. C. M. (1995). Structure and function of DNA-binding proteins. *Curr. Opin. Genet. Dev.* **5**, 180-189.
- Noll, M. (1993). Evolution and role of Pax genes. *Current Opin. Genetics Dev.* **3**, 595-605.
- Nusslein-Volhard, C. and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. *Nature* **287**, 795-801.
- Rubin, G. M. and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348-353.
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
- Schier, F. S. and Gehring, W. J. (1992). Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. *Nature* **356**, 804-807.
- Simpson-Brose, M., Treisman, J. and Desplan, C. (1994). Synergy between the Hunchback and Bicoid morphogens is required for anterior patterning in *Drosophila*. *Cell* **78**, 855-865.
- StJohnston, D. and Nusslein-Volhard, C. (1992). The Origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201-219.
- Strachan, T. and Read, A. P. (1994). Pax genes. *Current Opinion in Genet. Dev.* **4**, 427-438.
- Sun, B., Hursh, D. A., Jackson, D. and Beachy, P. A. (1995). Ultrabithorax protein is necessary but not sufficient for full activation of *decapentaplegic* expression in the visceral mesoderm. *EMBO J.* **14**, 520-535.
- Tearle, R. and Nusslein-Volhard, C. (1987). Tubingen mutants and stocklist. *Dros. Inform. Serv.* **66**, 209-269.
- Thummel, C. S. and Pirrotta, V. (1992). New pCaSpeR P-element vectors. *Dros. Inform. Serv.* **71**, 150.
- Treisman, J., Gonczy, P., Vashishtha, M., Harris, E. and Desplan, C. (1989). A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**, 553-562.
- Treisman, J., Harris, E. and Desplan, C. (1991). The Paired Box encodes a second DNA-binding domain in the Paired homeodomain protein. *Genes Dev.* **5**, 594-604.
- Triebenberg, S. J. (1995). Structure and function of transcriptional activation domains. *Curr. Opin. Genet. Dev.* **5**, 190-196.
- Underhill, D. A., Vogan, K. J. and Gros, P. (1995). Analysis of the mouse *Spotch-delayed* mutation indicates that the Pax-3 paired domain can influence homeodomain DNA-binding activity. *Proc. Natl. Acad. Sci. USA* **92**, 3692-3696.
- Weir, M. P., Edgar, B. A., Kornberg, T. and Schubiger, G. (1988). Spatial regulation of *engrailed* expression in the *Drosophila* embryo. *Genes Dev.* **2**, 1194-1203.
- Wilson, D., Sheng, G., Lecuit, T., Dostatni, N. and Desplan, C. (1993). Cooperative dimerization of paired class homeodomains on DNA. *Genes Dev.* **7**, 2120-34.
- 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**, 709-719.
- Xu, W., Rould, M., Jun, S., Desplan, C. and Pabo, C. (1995). Crystal structure of a paired domain-DNA complex at 2.5Å resolution reveals structural basis for Pax developmental mutations. *Cell* **80**, 639-650.