The EMBO Journal vol.15 no.10 pp.2476–2487, 1996

An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif

Siu-Kwong Chan, Heike Pöpperl[1], Robb Krumlauf[1] and Richard S.Mann[2]

Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, NY 10032, USA and [1]MRC Laboratory of Developmental Neurobiology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

[2]Corresponding author

HOX homeoproteins control cell identities during animal development by differentially regulating target genes. The homeoprotein encoded by the *extradenticle* (*exd*) gene can selectively modify HOX DNA binding, suggesting that it contributes to HOX specificity *in vivo*. HOX–EXD interactions are in part mediated by a conserved stretch of amino acids termed the hexapeptide found in many HOX proteins. Here, we demonstrate that a 20 bp oligonucleotide from the 5′ region of the mouse *Hoxb-1* gene, a homolog of *Drosophila labial* (*lab*), is sufficient to direct an expression pattern in *Drosophila* that is very similar to endogenous *lab*. *In vivo*, this expression requires *lab* and *exd* and, *in vitro*, LAB requires EXD to bind this oligonucleotide. In contrast, LAB proteins with mutations in the hexapeptide bind DNA even in the absence of EXD. Moreover, a hexapeptide mutant of LAB has an increased ability to activate transcription *in vivo*. Partial proteolysis experiments suggest that EXD can induce a conformational change in LAB. These data are consistent with a mechanism whereby the LAB hexapeptide inhibits LAB function by inhibiting DNA binding and that an EXD-induced conformational change in LAB relieves this inhibition, promoting highly specific interactions with biologically relevant binding sites.

Keywords: cooperative binding/homeodomain/homeotic/*labial*/PBX

Introduction

The Homeotic Complex (HOX) family of homeoproteins are required for specifying cell fates and morphological differences along the anterior–posterior axes of most animals (Lewis, 1978; Wakimoto and Kaufman, 1981; reviewed in McGinnis and Krumlauf, 1992). However, despite their very different functions *in vivo*, they have very similar homeodomains and therefore bind similar DNA sequences *in vitro* (Desplan *et al.*, 1988; Hoey and Levine, 1988; Muller *et al.*, 1988; Kalionis and O'Farrell, 1993; Ekker *et al.*, 1994). These *in vivo* functional differences probably reflect differences in the target genes regulated by HOX proteins (Garcia-Bellido, 1977; Andrew and Scott, 1992; Botas, 1993). Thus, an unresolved problem is how HOX proteins, and homeoproteins in general, select and regulate the correct sets of target genes *in vivo*. Outside of their homeodomains, HOX proteins have no sequence similarities except for two short peptides: 'MXSYF', at their N-termini, and the hexapeptide (also called the 'YPWM' motif or pentapeptide) N-terminal to their homeodomains (Burglin, 1994). Paradoxically, despite extensive differences outside their homeodomains, many of their specific *in vivo* functions are determined by the relatively small differences in their homeodomains and in nearby residues (Kuziora and McGinnis, 1989; Gibson *et al.*, 1990; Mann and Hogness, 1990; Kuziora and McGinnis, 1991; Lin and McGinnis, 1992; Chan and Mann, 1993; Zeng *et al.*, 1993).

The EXD/PBX family of homeoproteins (Kamps *et al.*, 1990; Nourse *et al.*, 1990; Monica *et al.*, 1991; Flegel *et al.*, 1993; Rauskolb *et al.*, 1993) has provided an important clue for how HOX proteins achieve DNA binding specificity *in vivo*. Originally identified genetically as the *extradenticle* (*exd*) gene in *Drosophila melanogaster* (Jürgens *et al.*, 1984), EXD/PBX proteins can specifically modulate both the *in vivo* functions (Peifer and Wieschaus, 1990; Rauskolb and Wieschaus, 1994) and DNA binding specificities (Chan *et al.*, 1994; van Dijk and Murre, 1994; Chang *et al.*, 1995; Lu *et al.*, 1995; Phelan *et al.*, 1995; Pöpperl *et al.*, 1995) of HOX gene products. Therefore, EXD/PBX proteins behave as specificity cofactors for HOX proteins (reviewed in Mann, 1995).

In HOX proteins, several regions have been implicated in mediating an interaction with EXD/PBX (Chan *et al.*, 1994; Chang *et al.*, 1995; Johnson *et al.*, 1995; Phelan *et al.*, 1995). One of these regions, the hexapeptide, appears to be critical for mediating an interaction between HOX proteins and EXD/PBX upon binding to consensus DNA binding sites (Chang *et al.*, 1995; Knoepfler and Kamps, 1995; Neuteboom *et al.*, 1995; Phelan *et al.*, 1995). These binding sites show little or no preference for different HOX proteins; nearly all hexapeptide-containing HOX proteins interact with EXD/PBX upon binding these consensus sites.

Using the yeast two-hybrid assay, interactions between HOX proteins and EXD/PBX have also been observed (Chan *et al.*, 1994; Chang *et al.*, 1995; Johnson *et al.*, 1995). In one study (Johnson *et al.*, 1995), the hexapeptide in the *Drosophila* HOX protein Ultrabithorax (UBX) was shown to be important, but not sufficient, to mediate an interaction with EXD. Further, the UBX homeodomain and carboxy-terminal tail were also shown to contribute to this interaction. These data are consistent with DNA binding studies that also implicated sequences within and C-terminal to the UBX homeodomain in the interaction with EXD (Chan *et al.*, 1994). Taken together, the available data suggest that the hexapeptide is critical for HOX–EXD/PBX interactions, but that additional sequences in HOX proteins modulate and/or contribute to this interaction.

Although important for understanding how HOX specificity is achieved, very few *in vivo* HOX-regulated enhancers have been defined. Here, we have characterized the *in vitro* and *in vivo* properties of an autoregulatory enhancer from the mouse *Hoxb-1* gene (Pöpperl *et al.*, 1995). We show that both Hoxb-1 and labial (LAB) (the *Drosophila* homolog of *Hoxb-1*) require EXD to bind to a 20 bp element derived from this enhancer. Remarkably, this element is sufficient to generate a *labial*-like expression pattern in *Drosophila* embryos. LAB proteins that are deleted for or mutated in the conserved hexapeptide bind this element even in the absence of EXD and have increased transcriptional activation function *in vivo*. These results suggest that the hexapeptide inhibits LAB DNA binding *in cis* and that EXD acts to remove this inhibition, perhaps by inducing a conformational change in LAB.

Results

A 20 bp autoregulatory element from Hoxb-1 generates a labial-like expression pattern in flies

Three related DNA elements (referred to as repeat 1, repeat 2 and repeat 3) derived from the mouse *Hoxb-1* gene were previously shown to be necessary and sufficient for autoregulation of *Hoxb-1* in rhombomere 4 of the mouse hindbrain (Pöpperl *et al.*, 1995). This previous study also demonstrated that a 900 bp fragment containing all three of these repeats was sufficient to generate a partial *labial*-like expression pattern in *Drosophila* embryos (Pöpperl *et al.*, 1995). Here, we have characterized the expression pattern in *Drosophila* dictated by one of these sequences (repeat 3). A *lacZ* reporter gene with three copies of the 20 bp repeat 3 oligonucleotide was placed upstream of a minimal promoter and introduced into flies (this reporter gene is referred to as *3Xrpt3-lacZ*). *lacZ* expression appeared identical in four independent transformant lines. β-Galactosidase (β-gal) and LAB expression were monitored in parallel during embryogenesis by immunohistochemistry. With the exception of staining in the gastric caeca primordia (gc), *lacZ* expression was very similar to the endogenous *lab* gene (Figure 1) (Diederich *et al.*, 1989; Chouinard and Kaufman, 1991). β-gal was first detected in the invaginating posterior midgut (pmg, arrows in Figure 1A, B, D and E) at the end of germband extension (stage 10) but, unlike LAB, was absent from the head at this stage (Figure 1A and D). Double-label immunofluorescence experiments that detected LAB and β-gal indicated that the pmg *lacZ* expression was coincident with *lab* expression (Figure 1G–I). Weak *lacZ* expression in the head was first observed at approximately stage 12 (Figure 1B and E). Following germband retraction (stage 14), β-gal and LAB were observed in the endoderm (en) next to PS7 of the visceral mesoderm (vm), in epidermal (ep) and presumptive neural (pnr) cells of the head, and in the dorsal ridge (dr) (Figure 1C and F) (Diederich *et al.*, 1989; Chouinard and Kaufman, 1991). All head and dorsal ridge *lacZ* expression appeared to be coincident with *lab* expression (Figure 1J–L). The endodermal expression was further characterized by double-label experiments that detected UBX and β-gal. Like LAB (Immergluck *et al.*, 1990; Reuter *et al.*, 1990), β-gal was detected in endodermal cells that are adjacent to UBX-expressing cells in the visceral mesoderm (Figure 1M and N). In addition, unlike LAB, β-gal was detected further posteriorly. Because this expression is not observed in *lab*$^-$ embryos (see below, Figure 2), it is most likely due to residual β-gal present in pmg cells prior to fusion of the midgut (Figure 1M). All *3Xrpt3-lacZ* expression required intact repeat 3 sequences because no *lacZ* expression was detected in multiple transformant lines containing a reporter gene with a mutated version of repeat 3 (data not shown; see Materials and methods).

3Xrpt3 expression requires lab and exd

The expression of *lab* is controlled in part by autoregulation (Chouinard and Kaufman, 1991; Tremml and Bienz, 1992). To determine whether the *lab*-like pattern of *3Xrpt3-lacZ* was due to regulation by *lab*, this reporter gene was introduced into a *lab*$^-$ background. As with the endogenous *lab* gene, the head and midgut expression in germband retracted embryos was absent in *lab*$^-$ embryos (Figure 2E). Thus, the *lab*-like expression of *3Xrpt3-lacZ* requires the endogenous *lab* gene. In contrast, expression in the gastric caeca primordia, where *lab* is not expressed, was still present in *lab*$^-$ embryos (Figure 2E).

In addition to its role in co-regulating downstream genes with HOX proteins (Peifer and Wieschaus, 1990; Chan *et al.*, 1994; Rauskolb and Wieschaus, 1994), the homeoprotein extradenticle (EXD) has also been implicated as a cofactor in *lab* autoregulation (Pöpperl *et al.*, 1995). To assess whether expression from *3Xrpt3-lacZ* is dependent on *exd*, *lacZ* expression was assayed in an *exd*$^-$ background. Embryos in which both maternal and zygotic *exd* functions were eliminated had no detectable expression (Figure 2C), whereas removing only zygotic *exd* function did not eliminate expression (data not shown). We also tested whether expression of the endogenous *lab* gene depends on *exd*. Most *lab* expression was eliminated in embryos in which both maternal and zygotic *exd* functions were absent (Figure 2D). These data demonstrate that *exd* has a role in regulating the expression of *lab* and perhaps other HOX genes. Because *exd* is required for *lab* expression, loss of *3Xrpt3-lacZ* expression in an *exd* mutant background may in part be due to loss of *lab*. However, this cannot be true for loss of expression in the gastric caeca primordia because *3Xrpt3-lacZ* expression in these cells does not require *lab* (Figure 2).

LAB requires EXD to bind DNA with high affinity

The results from these *in vivo* studies demonstrate that a multimerized 20 bp element from the mouse *Hoxb-1* gene can dictate a *lab*-like expression pattern in flies that is both *lab*- and *exd*-dependent. EXD protein has been shown to modify the DNA binding properties of HOX proteins by cooperative interactions (Chan *et al.*, 1994; van Dijk and Murre, 1994; Chang *et al.*, 1995; Lu *et al.*, 1995; Phelan *et al.*, 1995; Pöpperl *et al.*, 1995). To examine whether the requirement for *exd* function for *3Xrpt3-lacZ* expression could be due to cooperative binding between EXD and LAB, we used the Electrophoretic Mobility Shift Assay (EMSA) to characterize the binding of these homeoproteins to repeat 3. Unless otherwise stated, the proteins used in all DNA binding experiments were histidine tagged at their N-terminus and purified after their overexpression in *Escherichia coli* by Ni^{2+} affinity chromatography (Smith *et al.*, 1988). The EXD protein

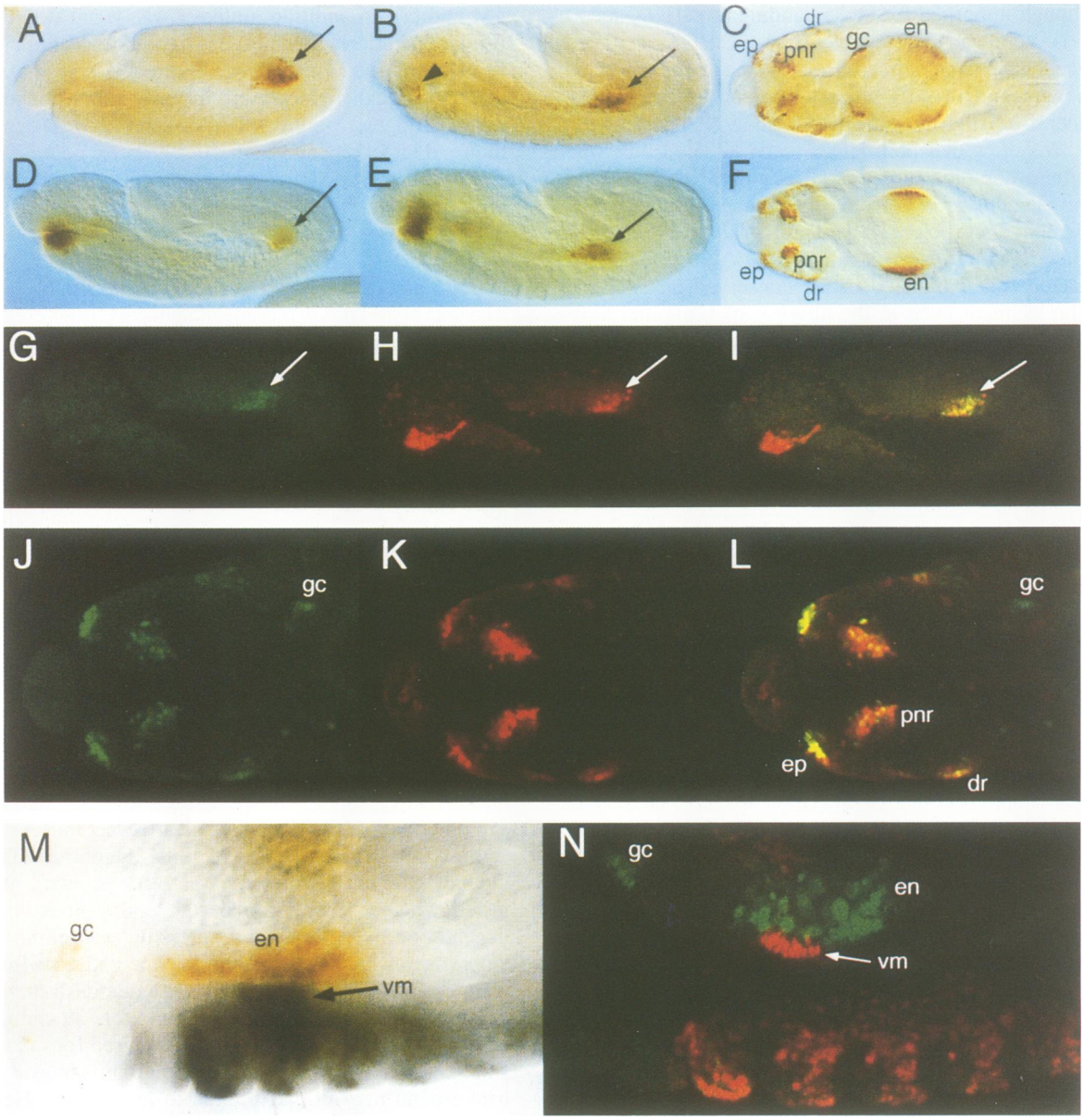


Fig. 1. A 20 bp element from *Hoxb-1* generates a *lab*-like expression pattern in *Drosophila*. Expression of *3Xrpt3-lacZ* (transformant line 3-5) (**A–C**) was compared with the endogenous *lab* gene (**D–F**). In stage 10 embryos (A and D), β-gal and LAB are both present in cells of the posterior midgut (pmg; arrows). In stage 12 embryos (B and E), expression in the pmg remains (arrows) and β-gal is first observed in the head (arrowhead). In stage 14 embryos (C and F), LAB and β-gal are detected in epidermal (ep) and presumptive neural (pnr) cells in the head, in the dorsal ridge (dr) and in the endoderm (en). In addition, β-gal is detected in the gastric caeca primordia (gc). (**G–L**) are confocal images of stage 10 (G–I) or stage 14 (J–L, head only) embryos doubly stained for β-gal (green) and LAB (red). In (G) and (J), only *lacZ* expression is shown, in (H) and (K) only *lab* expression is shown, and in (I) and (L) the double images are shown. The arrows point to pmg staining. (**M**) is a higher magnification view of the lateral side of a stage 13 embryo doubly stained for UBX (blue/black) and β-gal (brown) during fusion of the posterior and anterior midgut primordia. In the visceral mesoderm (vm), UBX is present only in PS7, whereas in the endoderm (en) β-gal is present more extensively along the anterior–posterior axis. (**N**) is a confocal view of a slightly older (stage 14) embryo doubly stained for UBX (red) and β-gal (green), showing that β-gal is present in endodermal cells immediately adjacent to UBX-expressing cells in the visceral mesoderm.

used in these experiments was a previously described 74 amino acid homeodomain-containing fragment (Chan *et al.*, 1994; Pöpperl *et al.*, 1995). Hoxb-1 was expressed as a previously described fusion with glutathione-*S*-transferase (GST) (Pöpperl *et al.*, 1995). The LAB proteins used in these experiments varied in size and are referred to by the N- and C-terminal LAB residues they contain (Figure 4B). The purity of these proteins was assessed by SDS–PAGE and ranged from ~50% to >95% (Figure 4B).

EXD weakly bound repeat 3 in the absence of LAB (Figure 3A, lanes 2 and 8). However, in the absence of EXD, no LAB(158–635) or GST–Hoxb-1 protein–DNA complexes were detected (Figure 3A, lanes 3 and 5). In contrast, in the presence of EXD, LAB(158–635) or GST–Hoxb-1 generated a slow mobility complex with repeat 3 (Figure 3A, lanes 4 and 6). Formation of the LAB–EXD complex was sequence specific because it was competed by a 50-fold excess of unlabeled repeat 3 oligo, but was not competed by 50-fold excess of mutant repeat 3 oligo (repeat 3*; Figure 3A, lanes 10–18).

The presence of both LAB and EXD in this complex was demonstrated by antibody supershift experiments

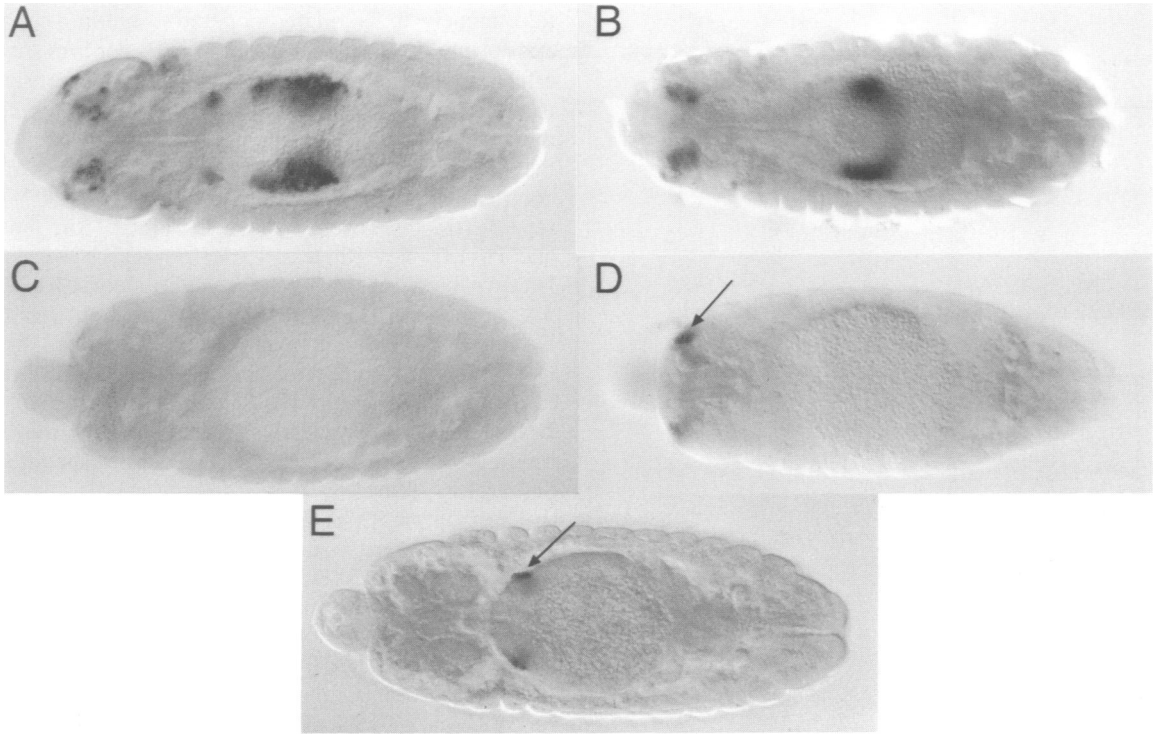


Fig. 2. Expression of *3Xrpt3-lacZ* requires *exd* and *lab*. β-gal protein expressed from *3Xrpt3-lacZ* (**A**, **C** and **E**) or RNA from the endogenous *lab* gene (**B** and **D**) were detected in wild-type (A and B), *exd*$^-$ (C and D) and *lab*$^-$ (E) stage 14 embryos. In *exd*$^-$ embryos (C and D), no *3Xrpt3-lacZ* expression was detected, whereas a small amount of *lab* expression in the head remains (D, arrow). In *lab*$^-$ embryos (E), all head and endodermal expression of *3Xrpt3-lacZ* is absent. In contrast, the gastric caeca expression remains (E, arrow) and confirms the presence of the reporter construct in these embryos.

using T7 epitope-tagged versions of these proteins. Upon the addition of an anti-T7 antibody, a significant fraction of the slow mobility complex could be supershifted in reactions with T7-EXD plus untagged LAB or with T7-LAB plus untagged EXD (Figure 3B, lanes 4, 5, 8 and 9). Addition of the anti-T7 antibody had no effect on the mobility of the free probe (Figure 3B, lane 1) nor did it induce a complex with T7-LAB in the absence of EXD (Figure 3B, lane 7).

To determine the stoichiometry of LAB molecules in the complex, binding reactions were performed (in the presence of EXD) with mixtures of truncated and full-length forms of LAB. If complexes could be generated with one truncated and one full-length form of LAB, a complex of intermediate mobility would be expected (Hope and Struhl, 1987). However, an intermediate complex was not observed (Figure 3B, lanes 10–12), suggesting that heterodimers between these two versions of LAB were not formed. Assuming that this truncated form of LAB was not impaired in its ability to dimerize, these data suggest that LAB binds this element as a monomer.

These data indicate that the LAB family of HOX proteins do not bind DNA on their own with high affinity despite the presence of a homeodomain. EXD, therefore, appears to be a cofactor that enables LAB to bind DNA with high affinity.

LAB truncations suggest an inhibitory role for the hexapeptide

In addition to 80% sequence identity in their homeodomains, LAB and Hoxb-1, which both bind cooperatively with EXD to repeat 3 (Figure 3A), share a 20 amino acid region (50% identity, 60% similarity) encompassing the hexapeptide (Figure 4A). Apart from the hexapeptide region and the homeodomain, there is little similarity between Hoxb-1 and LAB. Together with previous work demonstrating the importance of the hexapeptide in mediating an interaction between HOX and EXD/PBX proteins (Chang *et al.*, 1995; Knoepfler and Kamps, 1995; Neuteboom *et al.*, 1995; Phelan *et al.*, 1995), the sequence similarity between LAB and Hoxb-1 suggested that the hexapeptide may also contribute to the interaction between LAB/Hoxb-1 and EXD upon binding repeat 3. To test the role of the hexapeptide, binding reactions with truncated forms of LAB were performed in the absence and presence of EXD. Sequences N-terminal to the hexapeptide and C-terminal to the homeodomain were not required for cooperative binding with EXD (Figure 4B and C, lane 6). Deletion of the conserved hexapeptide greatly reduced cooperative binding with EXD (Figure 4B and C, lanes 7–10). In addition, these experiments revealed a novel role for the hexapeptide: strikingly, all three hexapeptide-deleted proteins no longer required EXD to form a protein–DNA complex (lanes 7, 9 and 11). These results suggest that the presence of the hexapeptide inhibits LAB DNA binding.

LAB derivatives LAB(436–567) and LAB(494–635), which do not contain the hexapeptide, were still stimulated to bind repeat 3 by EXD (quantitation of these data revealed a 3- to 5-fold stimulation; Figure 4C, lanes 7–10). These data indicate that, even in the absence of the hexapeptide, LAB may still interact with EXD. In contrast, a LAB derivative consisting essentially of only the homeodomain, LAB(500–567), bound equally well in the absence or presence of EXD (Figure 4C, lanes 11–12). Hence, residues

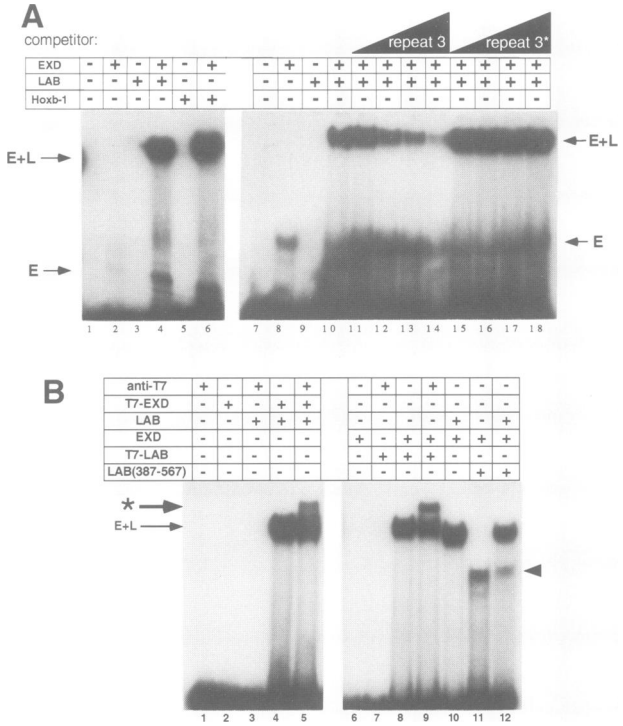


Fig. 3. LAB requires EXD to bind repeat 3 with high affinity. (**A**) EMSA experiments demonstrating that LAB and Hoxb-1 require EXD to bind with high affinity to repeat 3. In this panel, LAB refers to LAB(158–635) (see Figure 4B). EXD (50 ng) binds weakly on its own (lanes 2 and 8, 'E') and no Hoxb-1 or LAB binding was observed in the absence of EXD (lanes 3, 5 and 9). The amount of LAB used in each binding reaction was 80 ng. E+L indicates the EXD- and Hoxb-1- or LAB-dependent complex. In the presence of increasing amounts of unlabeled repeat 3 (6×, lane 11; 12×, lane 12; 25×, lane 13; 50×, lane 14), complex formation is gradually competed. In contrast, in the presence of similar amounts of an unlabeled mutant form of repeat 3 (repeat 3*), no competition was observed (lanes 15–18). (**B**) Both EXD and LAB are present in the complex. In this panel, LAB refers to LAB(158–635) and T7-LAB refers to a T7 epitope-tagged version of LAB(158–635) (see Figure 4B and Materials and methods). Lanes 1–5 demonstrate the presence of EXD and lanes 6–9 demonstrate the presence of LAB. A slower mobility complex (supershift, ★) is observed only in the presence of the T7 antibody ('anti-T7'), the T7-tagged protein (T7-EXD, lane 5 or T7-LAB, lane 9) and the untagged partner. In the absence of LAB, no supershift was observed with T7-EXD plus the anti-T7 antibody (not shown). Lanes 10–12: When two forms of LAB are used [LAB(158–635) and LAB(387–567) (see Figure 4B)], complexes with different mobilities are observed (lanes 10 and 11, the arrowhead indicates the LAB(387–567) complex). In combination (lane 12), only these two complexes are formed, suggesting that heterodimers are not formed.

N-terminal or C-terminal to the LAB homeodomain are required for this hexapeptide-independent interaction with EXD.

Apparently, the hexapeptide can function at variable distances from the homeodomain. This is illustrated by the similar binding behavior of Hoxb-1 and LAB(135–635) which have 18 and 110 residues separating the conserved methionine of the hexapeptide and the start of the homeodomain, respectively (Figures 3 and 4A) (Pöpperl *et al.*, 1995).

A hexapeptide mutant of LAB binds DNA in the absence of EXD and is hyperactive in vivo

The DNA binding behavior of the LAB truncations described here suggests a dual role for the LAB hexapeptide, namely to facilitate cooperative binding with EXD and to inhibit DNA binding *in cis*. To confirm that these effects are due to the hexapeptide and not another sequence in LAB, the DNA binding behavior of a hexapeptide point mutant was analyzed. In this mutant, LAB(158–635)^{AAA}, the sequence 'YKWM' was changed to 'AAAM'. In contrast to LAB(158–635), LAB(158–635)^{AAA} bound repeat 3 in the absence of EXD (Figure 5A, lanes 5–7 and 11–13). Nevertheless, LAB(158–635)^{AAA} was stimulated to bind repeat 3 by the addition of EXD (Figure 5A, lanes 11–16). Thus, the DNA binding behavior of LAB(158–635)^{AAA} is qualitatively identical to LAB truncations in which the hexapeptide was deleted (e.g. LAB(436–567); Figure 4). In addition, as judged by the DNase I footprinting and methylation techniques, DNA binding by LAB(158–635)^{AAA} is sequence specific (Chan and Mann, 1996).

The data presented in Figure 5A were quantitated and comparisons were made between the amount of the slow mobility (LAB±EXD) complex formed (see Materials and methods). In the absence of EXD, LAB(158–635)^{AAA} bound repeat 3 at least 40-fold better than LAB(158–635). EXD stimulated LAB(158–635) to bind repeat 3 by at least 500-fold and stimulated LAB(158–635)^{AAA} to bind repeat 3 by ~3-fold. These data are consistent with the suggestion that the hexapeptide of LAB inhibits DNA binding to repeat 3 *in cis* and contributes to, but does not entirely account for, the interaction with EXD.

The ability of the LAB hexapeptide to inhibit DNA binding *in cis* suggested the possibility that this motif could also inhibit LAB's ability to activate target genes *in vivo*. To address this question, transgenic fly stocks were generated that express full-length LAB proteins under the control of the heat shock promoter that are either wild-type (*HS:lab*) or mutated (*HS:lab*^{AAA}) in the hexapeptide. The hexapeptide mutation in *HS:lab*^{AAA} was identical to the 'YKWM' to 'AAAM' mutation in LAB(158–635)^{AAA}. The *in vivo* activities of these proteins were assayed by monitoring the phenotypes of the first instar larval cuticles and the expression of *3Xrpt3-lacZ*. For both heat shock constructs, two independent transformant lines were characterized and produced similar results. In addition, using an anti-LAB antibody, similar levels of LAB protein were detected in heat-shocked embryos from all four lines; these embryos were prepared identically to those stained for *lacZ* expression (data not shown; see Materials and methods). Therefore, *in vivo*, the stability and expression levels of LAB and LAB^{AAA} were similar.

When given two 20 min heat shocks between 3 and 6 h of embryogenesis (see Materials and methods), both *HS:lab* and *HS:lab*^{AAA} generated very similar transformations of the first instar larval cuticle. The most apparent effect of ectopic LAB or LAB^{AAA} on the first instar cuticle was a dramatic reduction in the cephalopharyngeal skeleton and a partial block of head involution (data not shown). These phenotypes were indistinguishable from those generated by a previously described *HS:lab* transgene (Heuer and Kaufman, 1992). Despite these similar cuticle transformations, ubiquitous expression of LAB^{AAA} during embryogenesis resulted in ~50% greater lethality than did equivalent expression of LAB (data not shown). Thus, the biological activities of LAB and LAB^{AAA} are not equivalent.

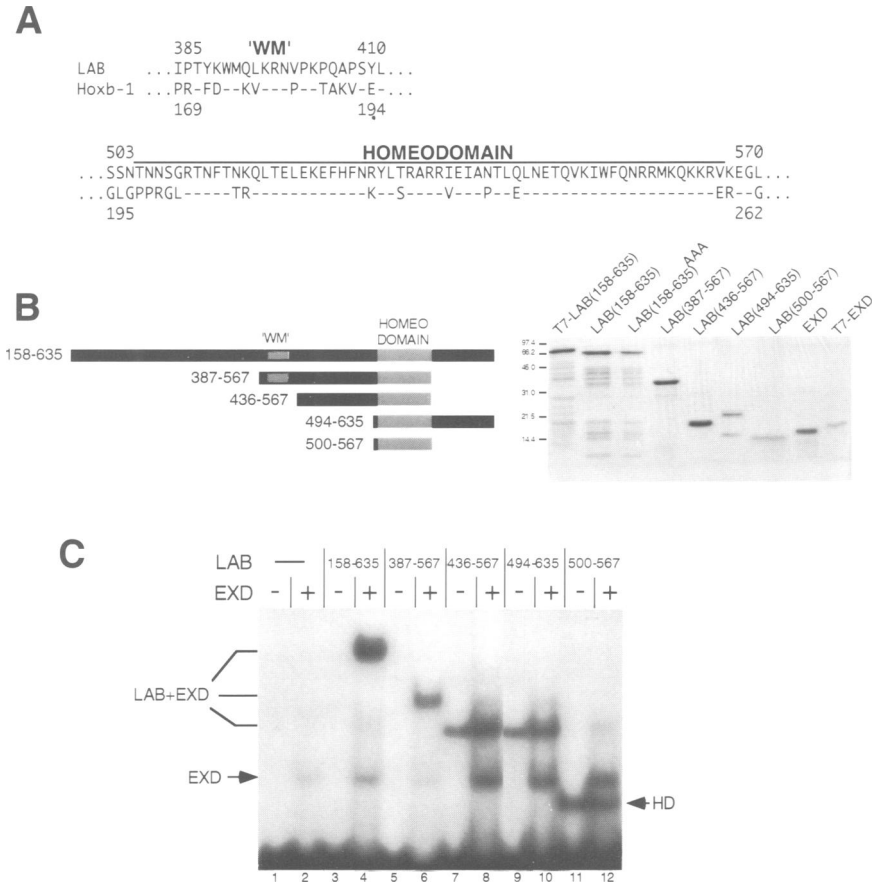


Fig. 4. The role of the hexapeptide in LAB. (**A**) Two regions, the hexapeptide (indicated by the 'WM') and the homeodomain (indicated by an overline), are conserved between LAB (top line) and Hoxb-1 (bottom line). Identical residues are shown as a '-'. In LAB and Hoxb-1, the hexapeptide and the homeodomain are separated by 110 and 18 amino acids, respectively. (**B**) The left panel illustrates various LAB truncations. The numbers refer to the wild-type LAB amino acids that are present in these proteins. The hexapeptide region is represented as a dark gray box (WM) and the homeodomain as a dark gray box. The right panel is a Coomassie-stained SDS–polyacrylamide gel showing the various proteins used in this work. The purity of the full-length proteins was estimated to be ~50% to >95%. The numbers to the left indicate molecular weight markers in kilodaltons. (**C**) DNA binding to repeat 3 by the various LAB truncations in the absence or presence of EXD was analyzed using EMSA. The version of LAB used is indicated by the amino acids present in the protein [see (B), above]. LAB(158–635) and LAB(387–567), which both contain the hexapeptide, bind repeat 3 only in the presence of EXD (lanes 3–6). LAB(436–567) and LAB(494–635), which do not contain the hexapeptide, bind DNA in the absence of EXD (lanes 7 and 9). However, their binding is still increased in the presence of EXD by 3- to 5-fold (lanes 8 and 10). LAB(500–567), which is essentially only the homeodomain, binds equally well in the absence or presence of EXD (lanes 11 and 12).

In addition to an increase in lethality, ectopic expression of LAB^{AAA} resulted in more extensive induction of the *3Xrpt3-lacZ* reporter gene than did ectopic expression of LAB (Figure 5B–D). Two 20 min heat shocks of *3Xrpt3-lacZ*; *HS:lab* embryos resulted in weak ectopic *lacZ* expression that was primarily limited to the clypeolabrum (Figure 5C). In contrast, two 20 min heat shocks of *3Xrpt3-lacZ*; *HS:lab*^{AAA} embryos resulted in extensive expression throughout the head, in two clusters of cells in the posterior abdomen, and in scattered cells throughout the embryo (Figure 5D). Expression of *lacZ*, however, was not uniform throughout the embryo, demonstrating that there are other factors that limit the expression of *3Xrpt3*. These results suggest that the hexapeptide normally functions to inhibit the transcriptional activation potential of LAB *in vivo*. When heat shock-induced expression of LAB^{AAA} was carried out in the absence of all *exd* function, no expression of *3Xrpt3-lacZ* was observed (data not shown), demonstrating that EXD is still required for LAB^{AAA} to activate this reporter gene *in vivo*.

Induction of a conformational change in LAB by EXD

The data presented above indicate that the hexapeptide region acts *in cis* to inhibit the interaction between LAB and repeat 3. The presence of EXD could induce a conformational change in LAB that releases this inhibition. To detect such a conformational change, a partial proteolysis experiment was performed (Shuman *et al.*, 1990; Roberts and Green, 1994). In the absence of EXD, brief digestion of T7-LAB(158–635) with trypsin primarily resulted in one N-terminal fragment (labeled 'A') of ~33 kDa (Figure 6A, lanes 3–5). Cleavage site A maps to Arg436, 26 residues C-terminal to the hexapeptide (Figure 6B). In the presence of increasing concentrations of EXD, a novel, faster migrating species (labeled 'B') of ~9 kDa was observed (Figure 6A, lanes 6–9). Cleavage at site A remained constant. The appearance of cleavage product B depended upon the addition of both trypsin and EXD (Figure 6A, compare lanes 2 and 8). Moreover, incubation with a similar concentration of other homeo-

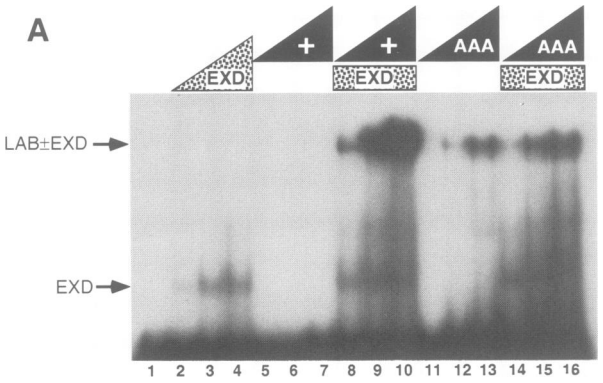
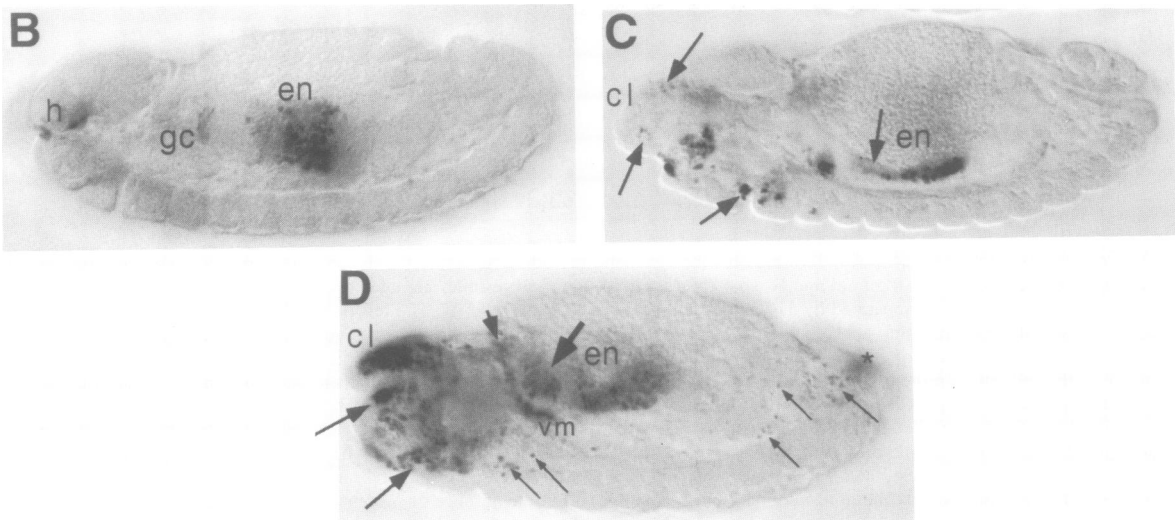


Fig. 5. A LAB hexapeptide point mutant binds repeat 3 without EXD and is hyperactive *in vivo*. (**A**) An EMSA experiment comparing the DNA binding behavior of LAB(158–635) (lanes 5–10; '+') with LAB(158–635)^{AAA} (lanes 11–16; 'AAA') in the absence (lanes 5–7 and 11–13) and presence (lanes 8–10 and 14–16) of 50 ng EXD. Lane 1 is the free probe and lanes 2–4 show a titration of EXD, alone (lane 2, 25 ng; lane 3, 50 ng; lane 4, 100 ng). The positions of the EXD and LAB±EXD complexes are indicated. The small size of the EXD protein used in these experiments (74 amino acids) may account for the similar mobilities of the LAB(158–635)^{AAA} and LAB(158–635)+EXD complexes. (**B–D**) β-gal was detected in *3Xrpt3-lacZ* (B), *3Xrpt3-lacZ*; *HS:lab* (C) and *3Xrpt3-lacZ*; *HS:lab*^{AAA} stage 14 heat-shocked embryos. All three photographs are lateral views. In (B), wild-type *lacZ* expression in the endoderm (en), gastric caeca (gc) and head (h) is indicated. (C) Ubiquitous expression of LAB resulted in weak ectopic *lacZ* expression (indicated by arrows) in the clypeolabrum (cl), in the first thoracic segment and in anterior cells of the endoderm. This embryo is slightly rotated relative to those in (B) and (D), resulting in a 'thinner' appearing endoderm. (D) Ubiquitous expression of LAB^{AAA} resulted in strong ectopic *lacZ* expression throughout the head (arrows) with the highest levels in the clypeolabrum (cl). In addition, ectopic *lacZ* expression was observed in the anterior portion of the endoderm (arrow), in visceral mesoderm (vm) cells adjacent to the gastric caeca primordia, in two clusters of cells in the posterior abdomen (asterisk; only one cluster is visible in this photograph) and in scattered cells throughout the trunk (thin arrows).



proteins (ANTP, PHOX-1) did not induce a change in the trypsin cleavage pattern (Figure 6A, lanes 3–4). The simplest interpretation of these data is that, in the absence of DNA, EXD can interact with and induce a conformational change in LAB.

Discussion

A previous study demonstrated that a multimerized 20 bp fragment derived from the mouse *Hoxb-1* gene was sufficient for Hoxb-1-dependent expression in rhombomere 4 of the mouse hindbrain (Pöpperl *et al.*, 1995). We show here that, when multimerized, this 20 bp fragment is also sufficient to generate a pattern of expression in *Drosophila* that is very similar to the endogenous *lab* gene, a *Hoxb-1* homolog. The small size of this HOX-dependent enhancer element is remarkable. Moreover, its activity requires two genes, *lab* and *exd*, which both encode homeoproteins that bind in a highly cooperative manner to this enhancer. Therefore, this small enhancer element is a valuable tool for studying HOX specificity *in vitro* and *in vivo*.

The hexapeptide inhibits LAB DNA binding

Previous experiments demonstrated a role for the hexapeptide in the interaction with EXD/PBX proteins (Chang *et al.*, 1995; Knoepfler and Kamps, 1995; Neuteboom *et al.*, 1995; Phelan *et al.*, 1995) and the experiments presented here provide an additional example of the importance of this motif. In addition, however, these experiments suggest a novel function for the hexapeptide, namely to inhibit or modify DNA binding *in cis*. The dual role for the hexapeptide is illustrated in Figure 7. We suggest that, in the absence of any cofactor, LAB exists in a conformation in which the hexapeptide prevents a high-affinity interaction between its homeodomain and DNA. Mutating or deleting this peptide uncovers LAB's ability to bind DNA. *In vitro*, EXD appears to be able to alleviate the inhibition of DNA binding by the hexapeptide because wild-type LAB binds strongly to repeat 3 in the

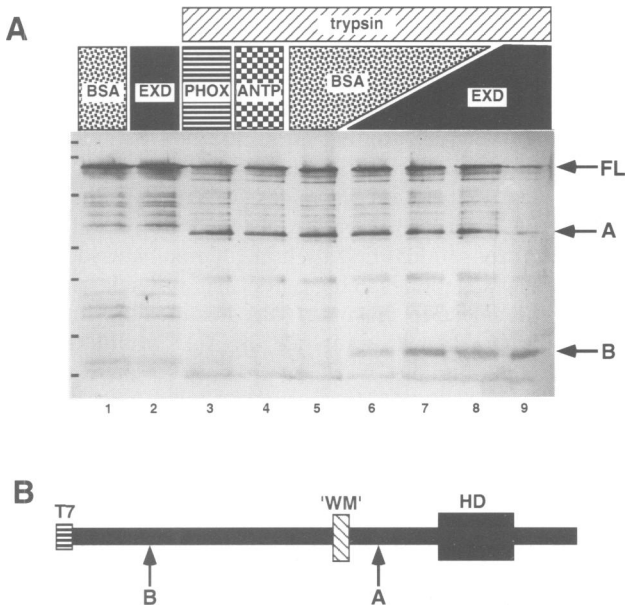


Fig. 6. An EXD-induced conformational change in LAB. (**A**) Immunoblot analysis of T7-LAB(158–635) partially digested with trypsin in the presence of increasing amounts of EXD. Five hundred nanograms T7-LAB(158–635) were incubated with 4 μg BSA (lane 1), 4 μg EXD (lane 2), 8 μg GST–PHOX1 (Gruenberg *et al.*, 1994) (lane 3), 4 μg ANTP (lane 4), 5 μg BSA (lane 5) and 1, 2, 4 or 5 μg EXD (lanes 6–9, respectively). In addition, lanes 6–9 had compensating amounts of BSA to keep the total amount of protein constant. Only the proteins in lanes 3–9 were treated with trypsin. In the absence of EXD, trypsin generated two major products that are detected with the T7 antibody: full-length T7-LAB(158–635) (FL) and product 'A'. Increasing amounts of EXD generated increasing amounts of an additional cleavage product, 'B'. The small bars on the left indicate the positions of the following molecular weight markers: 97.4, 66.0, 45.0, 31.0, 21.5, 14.4 and 6.5 kDa. (**B**) Approximate positions in T7-LAB(158–635) of the cleavage sites generating fragments 'A' and 'B'. The T7 tag is indicated at the left (N-terminal) end. Based on the specificity of trypsin cleavage and the sizes of these fragments, we deduce that cleavage site A is at Arg436 and site B is at Lys234.

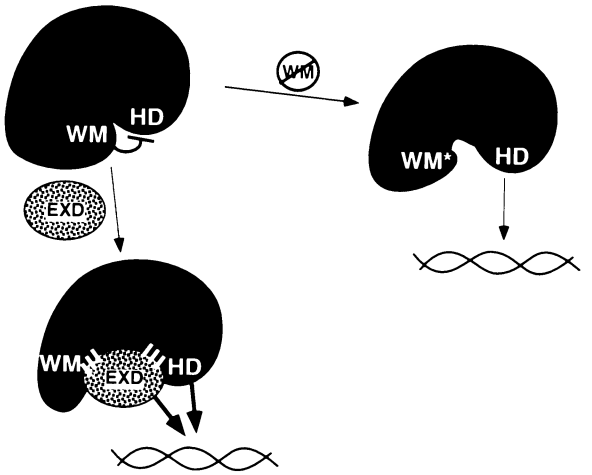


Fig. 7. Dual function of the LAB hexapeptide. In the conformation shown in the upper left, the LAB hexapeptide (indicated by the conserved dipeptide WM) inhibits the homeodomain (HD) from interacting with DNA, possibly by a direct interaction. When this motif is mutated or deleted (WM*), LAB is able to bind DNA. Alternatively, EXD promotes strong cooperative DNA binding by interacting with the hexapeptide and causing a conformational change, allowing both homeodomains to interact with DNA. Contacts between LAB and EXD (indicated by short white lines) are proposed to exist both within and outside the LAB hexapeptide.

presence of this cofactor. Thus, the model proposes that EXD interacts with the hexapeptide and causes a conformational change in LAB, a suggestion that is supported by the EXD-induced change in LAB protease sensitivity (Figure 6). In this altered conformation, the LAB–EXD complex binds DNA with high specificity and affinity. The model predicts that the precise arrangement of EXD and LAB binding sites in the DNA is critical to cooperative binding and this prediction has been confirmed by the biochemical characterization of this complex (Chan and Mann, 1996). Further, the model accounts for the observation that the LAB class of HOX proteins do not appear to bind DNA with high affinity (Phelan *et al.*, 1994; Pöpperl *et al.*, 1995; this work).

LAB proteins that lack or are mutated in the hexapeptide are still stimulated by EXD to bind repeat 3 DNA (Figures 4 and 5). These results suggest that sequences outside the hexapeptide also contribute to the interaction with EXD. Similarly, UBX proteins deleted for the hexapeptide were also stimulated to bind DNA by EXD (Chan *et al.*, 1994). In these experiments, sequences C-terminal to the UBX homeodomain were found to be important. Taken together, these results suggest that the hexapeptide is not the only sequence in HOX proteins that is important for interacting with EXD. Consistent with this suggestion, putative protein–protein interactions between UBX and EXD detected by the yeast two-hybrid assay were stronger when the UBX hexapeptide was present (Johnson *et al.*, 1995). However, in these experiments, the UBX homeodomain was necessary and the N-terminal (hexapeptide-containing) domain was not sufficient for an interaction with EXD (Johnson *et al.*, 1995), again suggesting that the homeodomain of UBX contributes to the interaction with EXD.

Additional support for the suggestion that the hexapeptide is not the sole EXD interaction domain comes from the observation that, *in vivo*, LAB^{AAA} is hyperactive, not less active, in activating *3Xrpt3-lacZ*. Furthermore, LAB^{AAA} and LAB induce similar homeotic transformations. If the hexapeptide were the only EXD interaction domain in LAB, then a strong hexapeptide mutation like LAB^{AAA} would be expected to have reduced activity *in vivo* due to its inability to interact with a required cofactor.

The inhibition of DNA binding by the hexapeptide of LAB is comparable with that of the LIM domains, which are present in the N-terminal portions of the LIM family of homeoproteins (Way and Chalfie, 1988; Freyd *et al.*, 1990; Karlsson *et al.*, 1990). Analogous to the results presented here, the LIM domain inhibits DNA binding *in cis* (Sánchez-García *et al.*, 1993) and blocks the *in vivo* functions of these homeoproteins (Taira *et al.*, 1994). However, to date, no cofactor analogous to EXD has been identified that reverses these inhibitory effects.

Inhibition of LAB function in vivo by the hexapeptide

The experiments that compare the effects of ectopic expression of LAB and LAB^{AAA} (Figure 5) suggest that the hexapeptide is also inhibitory *in vivo*. However, these experiments indicate that the LAB–EXD interaction has

additional complexity. These results are more easily discussed by defining the following four cell types: (i) cells that never appear able to express the reporter gene, e.g. most ectodermal cells of the trunk segments; (ii) cells that express the reporter gene only in response to LAB^{AAA} expression; (iii) cells that express the reporter gene in response to LAB or LAB^{AAA}; (iv) cells that express the reporter gene in the absence of ectopic LAB expression. These differences are unlikely to be due to differences in EXD expression because EXD protein is probably uniformly present up to stage 10 of embryogenesis due to maternal expression (Rauskolb *et al.*, 1993; S.Gonzalez-Crespos and G.Morata, personal communication; M.Abu-Shaar and R.S.Mann, unpublished observations). However, as discussed below, post-translational modification of EXD or additional factors present *in vivo* could account for these differences.

We suggest that the properties of cell type (i) are due to the presence of a repressor or the absence of a required activator. *In vivo*, this additional factor is predicted to bind repeat 3 in addition to LAB and EXD. Cell types (iii) and (iv) appear to have all the required activators; therefore, expression of the reporter gene in these cells is only limited by the presence of LAB. Cell type (ii), which activates the reporter gene in response to LAB^{AAA}, but not in response to LAB, appears to be missing a factor that is not required when the hexapeptide is mutated. The *in vitro* experiments which show that LAB^{AAA} no longer requires EXD to bind repeat 3 suggest that this factor is EXD. However, if EXD expression is uniform, it must be a modified form of EXD that is the relevant difference between these cell types. Alternatively, there may be another factor, present only in cell types (iii) and (iv), that works together with EXD *in vivo* to overcome inhibition by the hexapeptide.

Interestingly, previous studies of a visceral mesoderm enhancer derived from the *dpp* gene also suggested the possibility of a post-translational modification of EXD (Sun *et al.*, 1995). In those experiments, UBX binding to the *dpp* enhancer was found to be necessary but not sufficient to activate transcription. Additional factors or post-translational modification of EXD, which is also a required cofactor for this enhancer, were suggested to account for its restricted activation.

Finally, we note that *exd* function is still required for LAB^{AAA} to activate *3Xrpt3-lacZ* expression. This requirement for *exd* may be indirect, i.e. EXD could be necessary for the synthesis of another required factor. Alternatively, EXD may still have to bind *3Xrpt3-lacZ* to activate transcription, perhaps to facilitate the recruitment of additional transcription factors.

Could the hexapeptide be inhibitory in other HOX proteins?

There are many studies demonstrating that full-length (hexapeptide-containing) forms of HOX proteins other than LAB bind DNA with high affinity (Beachy *et al.*, 1988, 1993; Muller *et al.*, 1988; Regulski *et al.*, 1991; Capovilla *et al.*, 1994). Thus, it is clear that the hexapeptide does not completely block DNA binding by most HOX proteins. However, direct comparisons of hexapeptide-containing and hexapeptide-mutated versions of these other HOX proteins have not been carried out, leaving open the possibility that the hexapeptides in other HOX proteins partially inhibit DNA binding. Alternatively, the hexapeptide regions in other HOX proteins may modify their DNA binding properties in more subtle ways. For instance, by altering conformation, the hexapeptide may alter the sequence specificities of HOX proteins. According to this idea, optimal binding sites identified by homeodomains in the absence of the hexapeptide region may differ from those identified by homeodomains associated with their hexapeptides. This model has implications for target gene selection *in vivo*. If the hexapeptide inhibits DNA binding, then the presence of EXD would simply facilitate the binding of HOX proteins to their target genes. Alternatively, if the presence of the hexapeptide modifies binding specificity, the presence or absence of EXD could change the subset of target genes that are bound and regulated by HOX proteins.

We also note that although many HOX proteins contain a hexapeptide, sequence comparisons reveal that the hexapeptides of orthologs (HOX gene homologs from different species, e.g. *lab* and *Hoxb-1*) have more amino acids in common than do HOX genes from the same HOX cluster (e.g. *lab* and *Ubx*) (Burglin, 1994; Mann, 1995). In other words, the hexapeptide appears to have co-evolved with its associated homeodomain. Thus, while all hexapeptides may share the ability to interact with EXD/PBX, this co-evolution suggests that for many HOX proteins the homeodomain and hexapeptide function together, perhaps in a manner that is analogous to the interaction described here for LAB.

In vivo HOX binding sites can be low-affinity sites in vitro

A significant amount of work, including the determination of three-dimensional structures of four homeodomain–DNA complexes (Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; Billeter *et al.*, 1993; Hirsch and Aggarwal, 1995), has provided a detailed view of how HOX proteins, by themselves, recognize and bind to high-affinity DNA binding sites (Laughon, 1991; Gehring *et al.*, 1994). In several cases, high-affinity binding sites have been shown to be important for the function of an enhancer element *in vivo* (Regulski *et al.*, 1991; Vachon *et al.*, 1992; Capovilla *et al.*, 1994; Zeng *et al.*, 1994; Sun *et al.*, 1995). The work presented here demonstrates that HOX proteins, by themselves, bind some *in vivo* binding sites very poorly. Specifically, LAB binding to repeat 3 sequences is undetectable in standard EMSA experiments. Only in conjunction with its cofactor, EXD, does LAB bind with high affinity. A similar scenario to this has been observed in yeast, where the homeoprotein MAT-**a**1 binds to its target sequence only with its cofactor MAT-α2 (Goutte and Johnson, 1988, 1993; Dranginis, 1990). Thus, DNA binding by HOX proteins, themselves, cannot be used as the sole criterion for defining *in vivo* binding sites.

Implications of an inhibitory domain in HOX proteins

The LAB–EXD interaction described here serves at least two functions: to remove the inhibitory effect of the hexapeptide and to enhance DNA binding to relevant target sequences by cooperative interactions. The inhibitory function of the LAB hexapeptide may help prevent LAB

from binding and activating inappropriate target genes *in vivo*. Further, such a mechanism that requires two proteins for strong and productive DNA binding may be useful for several additional reasons. First, it provides the possibility for independently regulating the activity of the two proteins, perhaps by post-translational modification. Second, it raises the possibility that other cofactors direct the same HOX protein to different target sequences. For instance, the yeast homeoprotein MAT-α2 has two cofactors, MAT-**a**1 and MCM1, that promote MAT-α2 binding to different DNA sequences (Johnson, 1993). Third, in the absence of any cofactor, HOX proteins may have yet another DNA binding specificity. All of these possibilities illustrate that HOX–EXD/PBX interactions result in the potential for additional flexibility in the control of HOX target gene expression.

Materials and methods

Reporter genes and expression analysis

lacZ reporter genes were constructed by blunt-end cloning three wild-type (repeat 3) [GGGGTGATGGATGGGCGCTG] or mutated (repeat 3*) [GGGGTGTCGACTGGGCGCTG] oligos in tandem in the P-element vector CPLZN which expresses nuclear-localized β-gal (Bier *et al.*, 1989). The number and orientation of oligo sequences were confirmed by sequencing. The reporter constructs were introduced into the germline of the w^{1118} mutant stock (Lindsley and Zimm, 1992) of *D.melanogaster* by standard procedures. At least two independent transformants for each reporter gene showed similar *lacZ* expression patterns. Line 3-5, which has a homozygous viable insertion of *3Xrpt3-lacZ* on the second chromosome, was used for all experiments with the wild-type repeat 3. This element was crossed into a lab^{11} background (Lindsley and Zimm, 1992) that was balanced by TM6B 22UZ (Irvine *et al.*, 1991) (containing a *Ubx-lacZ* gene) to enable identification of lab^{11} homozygous embryos. *lacZ* expression was detected using a rabbit anti-β-gal antibody (Cappell). The secondary antibody was a donkey anti-rabbit directly conjugated to horseradish peroxidase (HRP) (Jackson Labs). Expression of *lab* was detected by *in situ* hybridization (Tautz and Pfeifle, 1989) using a probe derived from a *lab* cDNA kindly provided by W.Gehring (Mlodzik *et al.*, 1988) or by using a rabbit anti-LAB antibody generously provided by T.Kaufman. The double immunofluorescence stains used a mouse anti-β-gal (Promega) plus rabbit anti-LAB or rabbit anti-β-gal plus the anti-UBX antibody, mAbFP3.38; the secondary antibodies were a goat anti-rabbit–rhodamine and a goat anti-mouse–fluorescein (Boehringer). For Figure 1M, both UBX and β-gal were detected using HRP, but the UBX detection was in the presence of Ni^{2+} as described previously (Patel, 1994). exd^- embryos, devoid of both maternal and zygotic *exd* functions, were generated and identified by their segmentation phenotype as previously described (Chan *et al.*, 1994) and exd^-; *3Xrpt3-lacZ* embryos were obtained by crossing females with exd^- germlines to 3-5 homozygous males.

EMSA

GST–Hoxb-1 protein has been described elsewhere (Pöpperl *et al.*, 1995). All LAB proteins were expressed and purified as HIS-tagged fusion proteins (Smith *et al.*, 1988). T7-tagged versions were constructed by inserting an oligonucleotide encoding the T7 epitope [MASMTG-GQQMG] into the *Nde*I site of the pET-HIS14b expression vector (Novagen) and confirmed by DNA sequencing. Probes for EMSA were ^{32}P labeled by end-filling a single base overhang at the 3′ end. Amounts used in the binding reactions (unless otherwise stated) were: T7 antibody (Novagen), 1 μg; LAB, ~1.5 pmol; HIS-tagged EXD-homeodomain peptide (Chan *et al.*, 1994), 50 ng. Five hundred nanograms of poly(dI/dC) were included in all reactions in a total volume of 25 μl. For all LAB derivatives, ~1.5 pmol were used in each EMSA binding reaction. Compensating amounts of bovine serum albumin (BSA) were included to maintain constant protein levels.

The data in Figure 5A were quantified using a phosphorimager. The amount of the slow mobility complex in lane 3 (no LAB protein) was 0.03% of the total c.p.m. present in the lane (= background); in lane 6 [40 ng LAB(158–635)], 0.02%; in lane 9 [40 ng LAB(158–635) + 50 ng EXD], 10.8%; in lane 12 (40 ng LAB(158–635)^{AAA}], 1.2%; and in lane 15 [40 ng LAB(158–635)^{AAA} + 50 ng EXD], 3.6%.

LAB expression plasmids

LAB(158–635) was constructed by blunt-end cloning the *Apa*LI–*Xba*I fragment from the LAB cDNA (Mlodzik *et al.*, 1988) into the *Nde*I site of pET-HIS14b. LAB(387–567) was generated by polymerase chain reaction (PCR) with primer lab7 [AATCATATGAGCAGCATCCCC-ACC] and lab6 [GTAGATCTACTCCTTCACGCGCTTCTT] using the *lab* cDNA as the template. lab7 introduced a *Nde*I site (underlined) five residues N-terminal to the hexapeptide. lab6 created a stop codon at the end of homeodomain followed by a *Bgl*II site (underlined). The *Nde*I–*Bgl*II fragment was cloned into pET-HIS14b. LAB(436–567) was generated by deleting the *Sac*II–*Nde*I fragment from LAB(387–567) and a *Sac*II linker was inserted to restore the reading frame. LAB(494–635) was generated by deleting the *Nde*I–*Bam*HI (partial) fragment from LAB(158–635) and LAB(500–567) was generated by an *Nde*I–*Pst*I (partial) digestion of LAB(387–567). After blunting with T4 polymerase and re-ligation, the reading frame was restored. LAB(158–635)^{AAA} was constructed by a three-part ligation. Part 1: pET-HIS14b digested with *Nde*I and *Xho*I. Part 2: a PCR product using oligos T7 and sk163 [CATCGCGGCCGCGGTGGGGATGCTGCTGGA] and pETHIS14b-(158–635) as the template was digested with *Not*I (underlined) and *Xba*I, cloned into Bluescript, sequenced, and reisolated as a *Not*I–*Nde*I fragment. Part 3: a PCR product using oligos T3 and sk162 [ACCGCGGCCGCGATGCAACTCAAGAGGAAT] and the *labial* cDNA plasmid, pLabialSspI (Chouinard and Kaufman, 1991) (provided by T.Kaufman) as the template was digested with *Not*I (underlined) and *Xho*I, cloned into Bluescript, sequenced, and reisolated as a *Not*I–*Xba*I fragment.

Heat shocks

CaSpeR-*HS:lab* was constructed by inserting an *Eco*RI (blunt)–*Kpn*I fragment from pLabialSspI (Chouinard and Kaufman, 1991) into in a CaSpeR-based vector containing the *hsp70* promoter and the *Ubx* polyadenylation signal (phs10). CaSpeR-*HS:lab*^{AAA} was constructed by replacing the *Nsi*I–*Sph*I fragment of CaSpeR-*HS:lab* with the equivalent fragment from pETHIS14b-(158–635)^{AAA}. Multiple transformants for each transgene were examined and generated equivalent results. Heat shocks: 3 h collections of embryos from the cross 3-5×*HS:lab* or the cross 3-5×*HS:lab*^{AAA} were aged 3 h, heat shocked 20 min at 37°C, aged 2 h at 25°C, heat shocked 20 min at 37°C and aged 8 h before fixation. In an exd^- background: females with exd^- mosaic germlines were crossed to 3-5; *HS:lab*^{AAA} males; the embryo collections and heat shocks were as described above.

Partial proteolysis assay

Five hundred nanograms of purified T7-LAB(158–635) were incubated with 0, 1, 2, 9 or 4 μg EXD, 8 μg GST–PHOX1 (Grueneberg *et al.*, 1992), 4 μg ANTP (Chan *et al.*, 1994) or 4 μg BSA in 25 mM Tris–HCl (pH 7.9), 50 mM KCl, 0.5 mM dithiothreitol (DTT) and 0.5 mM EDTA in a volume of 20 μl. In the EXD reactions, the total protein concentration was held constant (at 4 μg) by the addition of BSA. After incubation on ice for 30 min, 5 ng of freshly diluted trypsin (Boehringer) were added and the samples incubated at 25°C for 1 min. After stopping the reaction with 5× loading buffer, the samples were boiled for 2 min, resolved by 16% SDS–PAGE and transferred to nitrocellulose for immunoblotting with a T7 antibody (Novagen) as described by the manufacturer.

Acknowledgements

We thank M.Bienz and the Indiana Stock Center for fly stocks, W.Gehring and T.Kaufman for DNAs, T.Kaufman for the LAB antibody, and A.Aggarwal, R.Axel, P.Feinstein, S.Goff, T.Jessell, B.Konforti, G.Struhl, D.Thanos and A.Tomlinson for helpful comments on this manuscript and for valuable discussions. We also thank P.Smith for P element injections and fly maintenance. This work was supported by grants from the NIH and the Searle Foundation awarded to R.S.M. H.P. was supported by a HFSP postdoctoral fellowship.

References

Andrew,D.J. and Scott,M.P. (1992) Downstream of the homeotic genes. *New Biol.*, **4**, 5–15.

Beachy,P.A., Krasnow,M.A., Gavis,E.R. and Hogness,D.S. (1988) An *Ultrabithorax* protein binds sequences near its own and the *Antennapedia* P1 promoters. *Cell*, **55**, 1069–1081.

Beachy,P.A., Varkey,J., Young,K.E., von Kessler,D.P., Sun,B.I. and Ekker,S.C. (1993) Cooperative binding of an Ultrabithorax homeodomain protein to nearby and distant DNA sites. *Mol. Cell. Biol.*, **13**, 6941–6956.

Bier,E. *et al.* (1989) Searching for pattern and mutation in the *Drosophila* genome with a P-*lacZ* vector. *Genes Dev.*, **3**, 1273–1287.

Billeter,M., Qian,Y.Q., Otting,G., Muller,M., Gehring,W. and Wuthrich,K. (1993) Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *J. Mol. Biol.*, **234**, 1084–1093.

Botas,J. (1993) Control of morphogenesis and differentiation by HOM/Hox genes. *Curr. Opin. Cell Biol.*, **5**, 1015–1022.

Burglin,T. (1994) A comprehensive classification of homeobox genes. In Duboule,D. (ed.), *Guidebook to the Homeobox Genes*. Oxford University Press, Oxford, UK, pp. 26–71.

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.

Chan,S.-K. and Mann,R.S. (1993) The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes Dev.*, **7**, 796–811.

Chan,S.-K. and Mann,R.S. (1996) A structural model for a HOX-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc. Natl Acad. Sci. USA*, in press.

Chan,S.-K., Jaffe,L., Capovilla,M., Botas,J. and Mann,R.S. (1994) The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with Extradenticle, another homeoprotein. *Cell*, **78**, 603–615.

Chang,C.-P., Shen,W.-F., Rozenfeld,S., Lawrence,H.J., Largman,C. and Cleary,M. (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox Proteins. *Genes Dev.*, **9**, 663–674.

Chouinard,S. and Kaufman,T.C. (1991) Control of expression of the homeotic *labial* (*lab*) locus of *Drosophila melanogaster*: evidence for both positive and negative autogenous regulation. *Development*, **113**, 1267–1280.

Desplan,C., Theis,J. and O'Farrell,P.H. (1988) The sequence specificity of homeodomain-DNA interaction. *Cell*, **54**, 1081–1090.

Diederich,R.J., Merrill,V.K., Pultz,M.A. and Kaufman,T.C. (1989) Isolation, structure, and expression of *labial*, a homeotic gene of the Antennapedia Complex involved in *Drosophila* head development. *Genes Dev.*, **3**, 399–414.

Dranginis,A.M. (1990) Binding of yeast a1 and α2 as a heterdimer to the operator DNA of a haploid-specific gene. *Nature*, **347**, 682–685.

Ekker,S.C., Jackson,D.G., von Kessler,D.P., Sun,B.I., Young,K.E. and Beachy,P.A. (1994) The degree of variation in DNA sequence recognition among four *Drosophila* homeotic proteins. *EMBO J.*, **13**, 3551–3560.

Flegel,W.A., Singson,A.W., Margolis,J.S., Bang,A.G., Posakony,J.W. and Murre,C. (1993) Dpbx, a new homeobox gene closely related to the human proto-oncogene pbx1 molecular structure and developmental expression. *Mech. Dev.*, **41**, 155–161.

Freyd,G., Kim,S.K. and Horvitz,H.R. (1990) Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene lin-11. *Nature*, **344**, 876–879.

Garcia-Bellido,A. (1977) Homeotic and atavic mutations in insects. *Am. Zool.*, **17**, 613–630.

Gehring,W.J., Qian,Y.Q., Billeter,M., Furukubo-Tokunaga,K., Schier, A.F., Resendez-Perez,D., Affolter,M., Otting,G. and Wuthrich,K. (1994) Homeodomain-DNA recognition. *Cell*, **78**, 211–223.

Gibson,G., Schier,A., LeMotte,P. and Gehring,W.J. (1990) The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell*, **62**, 1087–1103.

Goutte,C. and Johnson,A. (1988) a1 protein alters the DNA binding specificity of α2 repressor. *Cell*, **52**, 875–882.

Goutte,C. and Johnson,A.D. (1993) Yeast a1 and alpha 2 homeodomain proteins form a DNA-binding activity with properties distinct from those of either protein. *J. Mol. Biol.*, **233**, 359–371.

Grueneberg,D.A., Natesan,S., Alexandre,C. and Gilman,M.Z. (1992) Human and *Drosophila* homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science*, **257**, 1089–1095.

Heuer,J.G. and Kaufman,T.C. (1992) Homeotic genes have specific functional roles in the establishment of the *Drosophila* embryonic peripheral nervous system. *Development*, **115**, 35–47.

Hirsch,J. and Aggarwal,A. (1995) Structure of the Even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity. *EMBO J.*, **14**, 6280–6291.

Hoey,T. and Levine,M. (1988) Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. *Nature*, **332**, 858–861.

Hope,I.A. and Struhl,K. (1987) GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. *EMBO J.*, **6**, 2781–2784.

Immergluck,K., Lawrence,P.A. and Bienz,M. (1990) Induction across germ layers in *Drosophila* mediated by a genetic cascade. *Cell*, **62**, 261–268.

Irvine,K.D., Helfand,S.L. and Hogness,D.S. (1991) The large upstream control region of the *Drosophila* homeotic gene *Ultrabithorax*. *Development*, **111**, 407–424.

Johnson,A. (1993) A combinatorial regulatory circuit in budding yeast. In Yamamoto,K. and McKnight,S. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 975–1006.

Johnson,F.B., Parker,E. and Krasnow,M. (1995) Extradenticle protein is a selective cofactor for the *Drosophila* homeotics: Role of the homeodomain and YPWM amino acid motif in the interaction. *Proc. Natl Acad. Sci. USA*, **92**, 739–743.

Jürgens,G., Wieschaus,E., Nüsslein-Volhard,C. and Kluding,H. (1984) Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* II. Zygotic loci on the third chromosome. *Roux's Arch. Dev. Biol.*, **193**, 283–295.

Kalionis,B. and O'Farrell,P.H. (1993) A universal target sequence is bound *in vitro* by diverse homeodomains. *Mech. Dev.*, **43**, 57–70.

Kamps,M.P., Murre,C., Sun,X.-h. and Baltimore,D. (1990) A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell*, **60**, 547–555.

Karlsson,O., Thor,S., Norberg,T., Ohlsson,H. and Edlund,T. (1990) Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. *Nature*, **344**, 879–882.

Kissinger,C.R., Liu,B., Martin-Blanco,E., Kornberg,T.B. and Pabo,C.O. (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. *Cell*, **63**, 579–590.

Knoepfler,P. and Kamps,M. (1995) The pentapeptide motif of Hox proteins is required for cooperative DNA binding with Pbx1, physically contacts Pbx1, and enhances DNA binding by Pbx1. *Mol. Cell. Biol.*, **15**, 5811–5819.

Kuziora,M.A. and McGinnis,W. (1989) A homeodomain substitution changes the regulatory specificity of the deformed protein in *Drosophila* embryos. *Cell*, **59**, 563–571.

Kuziora,M.A. and McGinnis,W. (1991) Altering the regulatory targets of the Deformed protein in *Drosophila* embryos by substituting the Abdominal-B homeodomain. *Mech. Dev.*, **33**, 83–93.

Laughon,A. (1991) DNA binding specificity of homeodomains. *Biochemistry*, **30**, 11357–11367.

Lewis,E.B. (1978) A gene complex controlling segmentation in *Drosophila*. *Nature*, **276**, 565–570.

Lin,L. and McGinnis,W. (1992) Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev.*, **6**, 1071–1081.

Lindsley,D.L. and Zimm,G.G. (1992) *The Genome of Drosophila melanogaster*. Academic Press, San Diego, CA.

Lu,Q., Knoepfler,P., Scheele,J., Wright,D. and Kamps,M. (1995) Both Pbx1 and E2A-Pbx1 bind the DNA motif ATCAATCAA cooperatively with the products of multiple murine *Hox* genes, some of which are themselves oncogenes. *Mol. Cell. Biol.*, **15**, 3786–3795.

Mann,R.S. (1995) The specificity of homeotic gene function. *BioEssays*, **17**, 855–863.

Mann,R.S. and Hogness,D.S. (1990) Functional dissection of Ultrabithorax proteins in *D. melanogaster*. *Cell*, **60**, 597–610.

McGinnis,W. and Krumlauf,R. (1992) Homeobox genes and axial patterning. *Cell*, **68**, 283–302.

Mlodzik,M., Fjose,A. and Gehring,W.J. (1988) Molecular structure and spatial expression of a homeobox gene from the labial region of the Antennapedia-complex. *EMBO J.*, **7**, 2569–2578.

Monica,K., Galili,N., Nourse,J., Saltman,D. and Cleary,M.L. (1991) PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-oncogene PBX1. *Mol. Cell. Biol.*, **11**, 6149–6157.

Muller,M., Affolter,M., Leupin,W., Otting,G., Wuthrich,K. and Gehring,W.J. (1988) Isolation and sequence-specific DNA binding of the *Antennapedia* homeodomain. *EMBO J.*, **7**, 4299–4304.

Neuteboom,S., Peltenburg,L., van Dijk,M. and Murre,C. (1995) The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins. *Proc. Natl Acad. Sci. USA*, **92**, 9166–9170.

Nourse,J., Mellentin,J., Galili,N., Wilkinson,J., Stanbridge,E., Smith,S. and Cleary,M. (1990) Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell*, **60**, 535–545.

Patel,N.H. (1994) Imaging neuronal subsets and other cell types in whole mount *Drosophila* embryos and larvae using antibody probes. In Goldstein,L.S.B. and Fyrberg,E. (eds), *Drosophila melanogaster: Practical Uses in Cell Biology*. Academic Press, New York, pp. 445–487.

Peifer,M. and Wieschaus,E. (1990) Mutations in the *Drosophila* gene *extradenticle* affect the way specific homeo domain proteins regulate segmental identity. *Genes Dev.*, **4**, 1209–1223.

Phelan,M.L., Sadoul,R. and Featherstone,M.S. (1994) Functional differences between HOX proteins conferred by two residues in the homeodomain N-terminal arm. *Mol. Cell. Biol.*, **14**, 5066–5075.

Phelan,M.L., Rambaldi,I. and Featherstone,M. (1995) Cooperative interactions between HOX and PBX proteins mediated by a conserved peptide motif. *Mol. Cell. Biol.*, **15**, 3989–3997.

Pöpperl,H., Bienz,M., Studer,M., Chan,S.-K., Aparicio,S., Brenner,S., Mann,R. and Krumlauf,R. (1995) Segmental expression of *Hoxb-1* is controlled by a highly conserved autoregulatory loop dependent upon *exd/Pbx*. *Cell*, **81**, 1031–1042.

Rauskolb,C. and Wieschaus,E. (1994) Coordinate regulation of downstream genes by extradenticle and the homeotic selector proteins. *EMBO J.*, **13**, 3561–3569.

Rauskolb,C., Peifer,M. and Wieschaus,E. (1993) *extradenticle*, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene *pbx1*. *Cell*, **74**, 1–20.

Regulski,M., Dessain,S., McGinnis,N. and McGinnis,W. (1991) High-affinity binding sites for the Deformed protein are required for the function of an autoregulatory enhancer of the *Deformed* gene. *Genes Dev.*, **5**, 278–286.

Reuter,R., Panganiban,G.E., Hoffmann,F.M. and Scott,M.P. (1990) Homeotic genes regulate the spatial expression of putative growth factors in the visceral mesoderm of *Drosophila* embryos. *Development*, **110**, 1031–1040.

Roberts,S. and Green,M. (1994) Activator-induced conformational change in general transcription factor TFIIB. *Nature*, **371**, 717–720.

Sánchez-García,I., Osada,H., Forster,A. and Rabbitts,T.H. (1993) The cysteine-rich LIM domains inhibit DNA binding by the associated homeodomain in Isl-1. *EMBO J.*, **12**, 4243–4250.

Shuman,J., Vinson,C. and McKnight,S. (1990) Evidence of changes in protease sensitivity and subunit exchange rate on DNA binding by C/EBP. *Science*, **249**, 771–774.

Smith,M.C., Furman,T.C., Ingolia,T.D. and Pidgeon,C. (1988) Chelating peptide-immobilized metal ion affinity chromatography. A new concept in affinity chromatography for recombinant proteins. *J. Biol. Chem*, **263**, 7211–7215.

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.

Taira,M., Otani,H., Saint-Jeannet,J.P. and Dawid,I.B. (1994) Role of the LIM class homeodomain protein Xlim-1 in neural and muscle induction by the Spemann organizer in Xenopus. *Nature*, **372**, 677–679.

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.

Tremml,G. and Bienz,M. (1992) Induction of *labial* expression in the *Drosophila* endoderm: response elements for dpp signalling and for autoregulation. *Development*, **116**, 447–456.

Vachon,G., Cohen,B., Pfeifle,C., McGuffin,M.E., Botas,J. and Cohen,S. (1992) Homeotic genes of the Bithorax complex repress limb development in the abdomen of the *Drosophila* embryo through the target gene *distal-less*. *Cell*, **71**, 437–450.

van Dijk,M. and Murre,C. (1994) *extradenticle* raises the DNA binding specificity of homeotic selector gene products. *Cell*, **78**, 617–624.

Wakimoto,B.T. and Kaufman,T.C. (1981) Analysis of larval segmentation in lethal genotypes associated with the antennapedia gene complex in *Drosophila melanogaster*. *Dev. Biol.*, **81**, 51–64.

Way,J.C. and Chalfie,M. (1988) mec-3, a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans*. *Cell*, **54**, 5–16.

Wolberger,C., Vershon,A.K., Liu,B., Johnson,A.D. and Pabo,C.O. (1991) Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*, **67**, 517–528.

Zeng,W., Andrew,D.J., Mathies,L.D., Horner,M.A. and Scott,M.P. (1993) Ectopic expression and function of the *Antp* and *Scr* homeotic genes: the N terminus of the homeodomian is critical to functional specificity. *Development*, **118**, 339–352.

Zeng,C., Pinsonneault,J., Gellon,G., McGinnis,N. and McGinnis,W. (1994) Deformed protein binding sites and cofactor binding sites are required for the function of a small segment-specific regulatory element in *Drosophila* embryos. *EMBO J.*, **13**, 2362–2377.

Received on November 23, 1995; revised on January 9, 1996