Recognition of distinct target sites by a unique Labial/Extradenticle/Homothorax complex

Andreas Ebner*, Clemens Cabernard, Markus Affolter‡ and Samir Merabet*,†,‡

Biozentrum der Universität Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

*These authors contributed equally to this work
†Present address: IBDM/LGPD, Case907, Parc Scientifique de Luminy, 13288, Marseille Cedex 09, France
‡Authors for correspondence (markus.affolter@unibas.ch and merabet@ibdm.univ-mrs.fr)

Accepted 31 January 2005

Development 132, 1591-1600
Published by The Company of Biologists 2005
doi:10.1242/dev.01721

Summary

Hox genes encode evolutionarily conserved transcriptional regulators, which define regional identities along the anteroposterior axis of multicellular animals. In *Drosophila*, Hox proteins bind to target DNA sequences in association with the Extradenticle (Exd) and Homothorax (Hth) co-factors. The current model of Hox-binding selectivity proposes that the nucleotide sequence identity defines the Hox protein engaged in the trimeric complex, implying that distinct Hox/Exd/Hth complexes select different binding sites and that a given Hox/Exd/Hth complex recognizes a consensus DNA sequence. Here, we report that the regulation of a newly identified Lab target gene does not rely on the previously established consensus Lab/Exd/Hth-binding site, but on a strongly divergent sequence. Thus Lab, and most probably other Hox proteins, selects different DNA sequences in regulating downstream target genes. These observations have implications with regard to the current model of Hox-binding selectivity.

Key words: Transcription, Enhancer, Hox proteins, Binding selectivity, Gene regulation

Development

Introduction

Hox genes are involved in the specification of segmental identity along the anteroposterior axis of multicellular animals (McGinnis and Krumlauf, 1992). They encode transcription factors containing two characteristic features, a 60 amino acid (aa) DNA-binding motif, the homeodomain (HD), as well as a short motif upstream of the HD, the hexapeptide (HX). The mode of HD/DNA interaction is highly conserved (Gehring et al., 1994). Consequently, the DNA-binding properties of Hox proteins are similar, raising the paradox that proteins bearing equivalent biochemical properties reach in vivo distinct regulatory effects, to ultimately initiate different developmental programs (Hayashi and Scott, 1990; Mann, 1995).

This paradox has been partially solved by the finding that Hox proteins bind to DNA in association with TALE (three amino acid extension) homeodomain proteins: Extradenticle (Exd; Pbx proteins in vertebrates) and Homothorax (Hth; Meis/Prep proteins in vertebrates) (Pai et al., 1998; Peifer and Wieschaus, 1990; Rauskolb et al., 1993; Rieckhof et al., 1997). In *Drosophila*, the formation of the trimeric complex relies on Hox/Exd and on Exd/Hth interactions. These protein contacts raise DNA-binding affinity of Hox proteins. In addition, as each partner of the complex contacts DNA, the nucleotide site is larger than the Hox monomer site, leading to an increased specificity in target site recognition (Chan et al., 1994; Chan and Mann, 1996; Ryoo and Mann, 1999).

One of the best characterized *cis*-acting elements regulated by a Hox protein is the autoregulatory enhancer of the homeotic gene *labial* (*lab*) (Grieder et al., 1997; Marty et al., 2001; Tremml and Bienz, 1992). The full-length enhancer recapitulates endodermal *lab* expression from stage 12 onwards and integrates signalling and Hox inputs through different modules. The Hox responsive element (HRE) has been narrowed down to a 47 bp sequence. It is composed of a unique binding site for Lab and its co-factors Exd and Hth (Marty et al., 2001; Ryoo et al., 1999), and drives, on its own, expression in a subset of *lab*-expressing cells.

An important characteristic of the *lab* enhancer is its functional conservation in worms, flies and mouse. Trans-species analysis of the autoregulatory enhancers of the mouse *Hoxb1* or of the worm *ceh-13* genes have shown that these elements behaved like the *lab* gene in *Drosophila* midgut endoderm; suggesting that molecular mechanisms underlying *lab* class gene autoregulation are evolutionary conserved (Popperl et al., 1995; Streit et al., 2002). This conclusion is further supported by the findings that all these enhancers harbour the same Lab/Exd-binding site composed of the consensus sequence TGATGGAT(T/G)G.

The *lab* HRE served as a paradigm to establish the Hox-binding selectivity model, a model that explains how distinct Hox proteins in a complex with the same co-factor, Exd, reach distinct DNA-binding properties. Analysis of protein-DNA contacts indicated that Labial and Extradenticle bind the TGAT[GG]ATGG sequence in a head-to-tail orientation (Chan and Mann, 1996), with Exd and Lab contacting, respectively, the 5′ TGATGG and 3′GGATGG nucleotides; a prediction that was later confirmed by crystallographic studies (Piper et al., 1999). The first GG nucleotides of the Lab half site are contacted by the N-terminal arm of the HD, while the other

Development

nucleotides contact the third helix of the HD. Most importantly, it was shown that mutation of the two central nucleotides of this heterodimer site (enclosed inside brackets) from GG to TA was sufficient to change the Hox protein in the complex; from Lab to Ultrabithorax or to Deformed (Chan and Mann, 1996; Chan et al., 1997). These observations, together with similar experiments carried out on another enhancer (Ryoo and Mann, 1999), led to the proposal of the 'DNA-binding selectivity model', which defines a prototypical Hox/Exd-binding site composed of the TGAT[NN]ATNN sequence. In this model, the two central NN nucleotides are instructive with regard to the Hox protein that would associate with Exd: subtle differences in DNA sequences will thus result in selecting different Hox/Exd complexes.

According to the Hox DNA-binding selectivity model, we thought that target genes for Lab should be found in close vicinity to the consensus TGAT[GG]ATGG sequence, that should be further linked to a CTGTCA Hth-binding site. We thus screened the *Drosophila* genome for such sites, and tested neighbouring transcripts for expression and regulation by Lab. Surprisingly, the approach led to the identification of a single novel Lab target gene, whose regulation yielded unexpected findings with regard to the Hox DNA-binding selectivity model.

Materials and methods

In silico approach for research of consensus Lab/Exd/Hth binding sites

We first scanned the *Drosophila* genome for the consensus Lab/Exd/Hth-binding site with the following sequence: TGATGGAT(T/G)G (N)(40) CTGTCA, and searched for the presence of genes located no more than 10 kb away. We identified 30 putative binding sites and 40 potential target genes around them. Dig-RNA labelling probes corresponding to 16 of these potential target genes were synthesized in order to characterize their expression profiles by in situ hybridization: except *lab* and *CG11339*, none of them showed an expression pattern similar to *lab* (or to any other Hox gene) and were thus not further considered. When we recently repeated the in silico approach using the research tool available on the fly enhancer web site (http://flyenhancer.org), we found 51 sites closed to around 100 genes. The increase number of sites and associated targets might result from completion and novel annotations of the *Drosophila* genome databases. Therefore, we cannot exclude that, in addition to *lab* itself, other Lab target genes are effectively regulated via a consensus Lab/Exd/Hth-binding site.

In situ and antibody staining

Embryo collection, in situ hybridisation and immunodetection to whole embryos were performed according to standard procedures. Digoxigenin-labelled *lab* and *CG11339* antisense RNA probes were generated according to the manufacturer's protocol (Boehringer-Mannheim) with SP6 promoters. The anti-β-Gal antibody was produced in mouse (Promega), and the Lab antibody was produced in rabbit and affinity purified (U. Nussbaumer and M. Affolter, unpublished). Secondary antibodies were conjugated with horseradish peroxidase (AB kit) or FITC/RITC fluorochromes (Jackson). For Lab fluorescent staining, the signal was amplified with the aid of a Tyramid Signal Amplification kit (NEN life sciences).

Electrophoretic mobility shift assays (EMSAs)

Proteins for EMSA were produced with the TNT T7-coupled in vitro transcription/translation system (Promega). EMSAs were performed in 20 μl as described (Popperl et al., 1995). The Lab protein used in this study was either full-length (EST clone RE63854 obtained from the UK HGMP Resource Centre, cloned into PcDNA3: Fig. 4A, lanes 3-8, 11-13, 16-18, 24-26; and Fig. 5B) or from amino acid 158 to its C terminus (Fig. 4A, lanes 21-23, and Fig. 4B, lanes 3-5; (Chan and Mann, 1996). His-Exd was full length (a gift from Richard Mann: Fig. 4A and Fig. 5B) or from residue 1 to 323 (Fig. 4B) (Chan et al., 1997). The His-Hth construct (Ryoo et al., 1999) included amino acids 59 to the C terminus subcloned into pET14b (Novagen). The *proboscipedia* (Cribbs et al., 1992), *Sex combs reduced* (LeMotte et al., 1989), *Deformed* (Lin and McGinnis, 1992), *Antennapedia* (Schneuwly et al., 1987), *abdominalA* (Merabet et al., 2003) and *AbdominalB* (Celniker et al., 1989) cDNAs were full length cloned in PcDNA3-expressing vector (Invitrogen). *Ubx* cDNA was full length in T7pLink vector (Galant and Carroll, 2002). Each Hox cDNA was sequenced and analysed for protein expression by band shift experiments on control oligonucleotides containing a consensus Hox/Exd/Hth-binding site (Gebelein et al., 2002).

Band shift experiments with wild type or mutated *EVIII* enhancers in Fig. 4A were performed both with the truncated or full-length Lab and Exd proteins, and led to identical results. The oligonucleotides used for the mobility shift assays were annealed, and blunt-ended with [α-^{32}P]dATP nucleotides and Klenow polymerase. The sequence of the *lab48/95* control oligonucleotide containing the consensus Lab/Exd/Hth-binding site from *lab* enhancer is 5′tggtTGATGGATTGcccggcgccgaCTGTCAccgctccaagaac3′ (Passner et al., 1999). Sequence of the oligonucleotide containing the divergent Lab/Exd/Hth-binding site from the *EVIII* enhancer is 5′tttgtcgcatcgTGATCAATTAcagCTGACTgggttg3′. Mutated *EVIII* oligonucleotides (Fig. 4A) were made by replacing the core binding sites of Lab (ATTA), Exd (TGAT) or Hth (CTGA) by guanosine nucleotides, or were as indicated (Fig. 5A).

Fly stocks and transformants

Transformant lines were generated by standard procedures. All *Drosophila* reporter genes were generated by standard cloning procedures, and genomic inserts were cloned into the Asp718 site of the nuclear *lacZ* encoding P element vector pCβ (a gift of Konrad Basler). Original genomic inserts and mutations of the 150 bp fragment *EVIII* were generated using a PCR-based approach with 5′ and 3′ primers coupled to a Asp718 site. For each reporter construct, the *lacZ* expression pattern was determined for several independent transformant lines. In each case, the majority of transformants of a given construct showed identical expression patterns. For the analysis of expression in mutant backgrounds, the following alleles were used: *lab*^{vd1} (Diederich et al., 1989), *hth*^{P2} (kindly provided by R. Mann) and *exd*^{YO12} [which was used to generate female germline mosaics as described by Rauskolb et al. (Rauskolb et al., 1993)].

Results

The consensus Lab/Exd/Hth binding sequence is not sufficient to identify Lab target genes

The rationale of our approach was that any gene in the vicinity of a consensus Lab/Exd/Hth-binding site could represent a potential Lab target gene. We thus screened the *Drosophila* genome for sequences matching the Lab/Exd sequence TGATGGAT(T/G)G, associated with a Hth-binding site CTGTCA within 40 nucleotides. This led to the identification of 30 loci. For half of them, we analysed the expression of neighbouring transcriptional units (see Materials and methods). Only two transcripts were expressed in the endoderm in a pattern reminiscent of *lab*. The first one corresponded to the *lab* transcript itself. The second one, called *CG11339*, is located at chromosomal position 100B4-B5 (FlyBase report). We concluded from this in silico approach

Development

that the consensus Lab/Exd/Hth-binding sequence was not sufficient to efficiently identify Lab target genes.

CG11339 is regulated by Lab, Exd and Hth

Two cDNAs, *LP8211* and *CG11339*, which map in the vicinity of the TGATGGATGG(N)(14)CTGTCA sequence at 100B4-B5, were recovered from EST clones. Both correspond to the same transcription unit; *LP8211* containing additional 5′ end sequences that include the Lab/Exd/Hth-binding site (Fig. 2A). In situ hybridization with antisense probes derived from *LP8211* or *CG11339* cDNAs were identical (not shown).

The expression pattern of *CG11339* was compared to that of *lab*, which is expressed from stage 12, both in the endoderm primordia and in the head (Fig. 1A). Upon germ band retraction, strong expression is detected in the endodermal cells of the midgut, where *lab* transcription is induced to high levels by the Decapentaplegic (Dpp) signalling pathway (Fig. 1B). *Lab* expression is also detected in the developing brain, in a specific neuromere called the tritocerebrum (Hirth et al., 2001). Expression of *CG11339* was detected as early as *lab* in the endoderm (Fig. 1C). At later stages, *CG11339* was also weakly expressed in the entire visceral mesoderm and in parts of the head of the embryo (Fig. 1D). By confocal analyses of doubly stained embryos, we observed that expression of *CG11339* in the endoderm is confined to *lab*-expressing cells (Fig. 1E,F). We also noted that the most anterior *lab*-positive cells in the endoderm do not express *CG11339*.

We next investigated whether *CG11339* is indeed regulated by Lab and its co-factors Exd and Hth. In homozygous lab^{vd1} mutant embryo, the expression of *CG11339* was nearly completely abolished at stage 12 (Fig. 1G), and absent in endodermal cells at stage 14 (Fig. 1H). By contrast, expression of *CG11339* in the visceral mesoderm and in the head was still observed in *lab* mutants in later stages (not shown). In hth^{P2} mutant embryos, as well in embryos lacking *exd* maternal and zygotic contributions, we observed a strong reduction in the transcript level of *CG11339,* both at early and late stages of embryogenesis (Fig. 1I-L). Expression of *CG11339* in the visceral mesoderm and in the head was also lost, indicating that *CG11339* could be regulated by *hth* and *exd* in these domains. We thus concluded that *CG11339* expression in the endoderm is regulated by *lab* and its co-factor encoding genes *exd* and *hth*.

The genomic sequences responsible for *lab*-dependent expression of *CG11339* do not include the consensus Lab/Exd/Hth site identified by the in silico approach

To identify the enhancer responsible for *lab*-dependent expression of *CG11339*, we first tested a 2 kb genomic fragment containing the consensus Lab/Exd/Hth-binding site identified by our in silico approach (fragment A: Fig. 2A). This fragment overlaps the 3′ end of the first exon and the 5′ end of the first intron of the longest isoform of *CG11339* (Fig. 2A). It was cloned upstream of a *lacZ* reporter gene and transgenic lines for this construct were obtained by P-element mediated transformation. Surprisingly, this DNA element did not display any enhancer activity (Fig. 2B). Among several other genomic fragments tested (not shown), one was able to drive a *lacZ* expression profile similar to *CG11339*. This 2 kb genomic element maps 200 bp upstream of the most 5′ sequences of *CG11339* (fragment B: Fig. 2A and 2B). Analysis of three sub-elements (fragments C-E: Fig. 2A,B) allowed us to restrict the enhancer region of *CG11339* to a 769 bp fragment (fragment E).

The fragment E drives *lacZ* expression in midgut endoderm and in the head, but not in the visceral mesoderm (Fig. 2B). Moreover, activity of this enhancer displayed the same genetic requirements as expression of *CG11339*, as it was inactive in embryos homozygous mutant for *lab* and *hth* (not shown). To further narrow down the sequence responsible for expression of *CG11339* in the endoderm, eight overlapping sub-elements of the *E* enhancer were analysed (elements EI to EVIII: Fig. 2A). Three elements reproduced the *lacZ* expression profile of

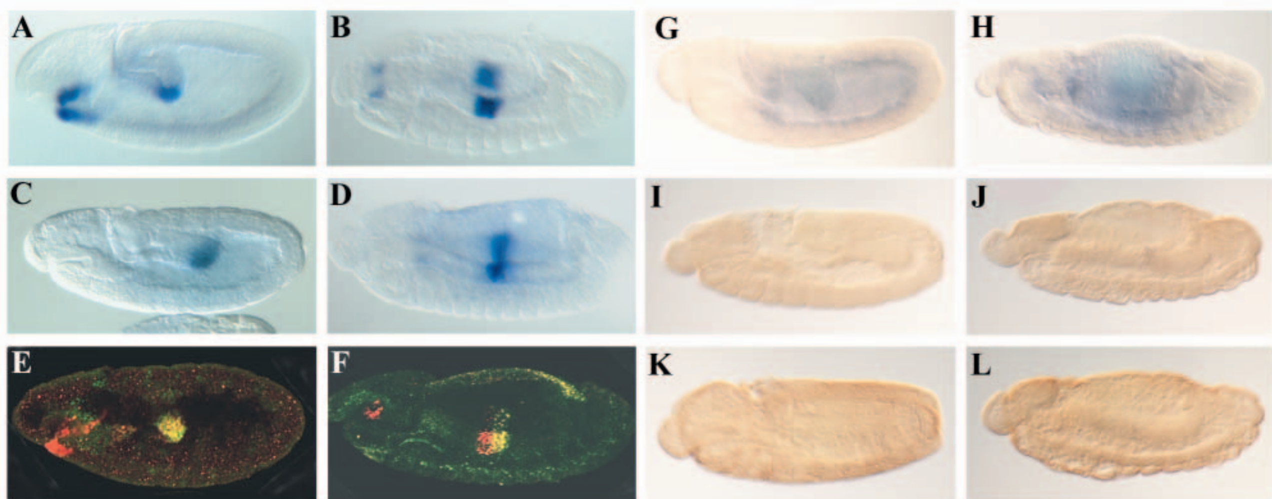


Fig. 1. Expression of *CG11339* in *Drosophila* midgut endoderm is regulated by *lab*, *exd* and *hth*. (A-L) All embryos are shown at stages 12 (A,C,E,G,I,K) and 14 (B,D,F,H,J,L). Anterior towards the left and posterior towards the right. (A,B) Expression of *lab* RNA. (C,D) Expression of *CG11339* RNA. (E,F) Confocal analysis of *CG11339* (in situ; green) and Lab (immunostaining; red). (G,H) Loss of *CG11339* RNA expression in lab^{vd1} homozygous mutant embryos. (I,J) Loss of *CG11339* RNA expression in hth^{P2} homozygous mutant embryos. (K,L) Loss of *CG11339* RNA expression in maternal and zygotic exd^{Y012} mutant embryos.

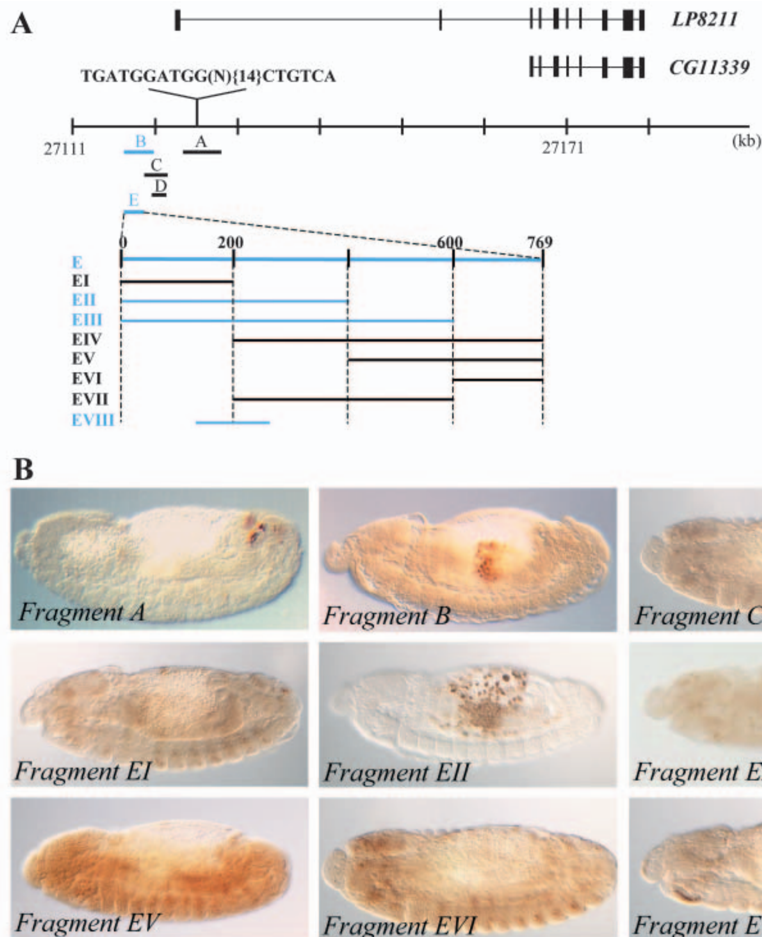


Fig. 2. Enhancer characterisation of the genomic region responsible for *CG11339* expression in endoderm. (A) Scheme of the 100B4-B5 genomic region of *CG11339*. The two isoforms are indicated, as well as the sequence of the consensus Lab/Exd/Hth-binding site found by in silico approach. Genomic fragments tested for enhancer activities are shown: fragments testing positive are shown in blue. (B) β-Gal staining in lines carrying the *lacZ* transgene under the control of the different genomic regions, corresponding to fragments A-E and sub-fragments EI-EVIII derived from fragment E. All embryos are shown for stage 14 of embryogenesis.

the *E* enhancer (fragments EII, EIII and EVIII; Fig. 2B), although these enhancers drive a somewhat broader expression in the endoderm, and in scattered cells in the amnioserosa. As positive enhancers *EII* and *EIII* include *EVIII*, we concluded that the 150 bp enhancer *EVIII* (and not the original 2 kb enhancer that included the consensus Lab/Exd/Hth site identified by in silico approach) bears the sequences responsible for *lab*-dependent expression of *CG11339*.

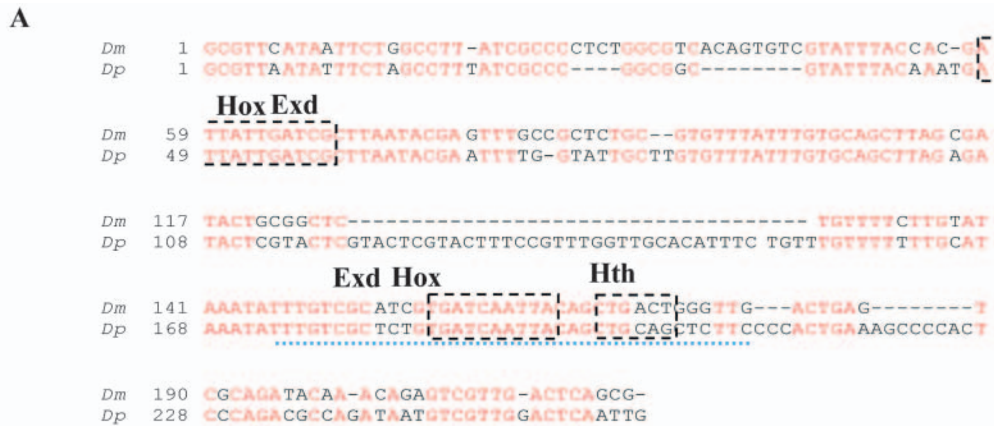
A sequence divergent from the consensus Lab/Exd/Hth-binding site is responsible for activity of the *EVIII* enhancer in the endoderm

Examination of sequence of the *EVIII* enhancer failed to identify a consensus Lab/Exd/Hth-binding site. We thus considered that the Lab/Exd/Hth complex might recognise a divergent sequence, and searched for presence of a general Hox/Exd TGAT[NN]ATNN motif, with the central NN nucleotides distinct from the ones that have so far been shown to confer Lab recruitment. One such site was identified (TGAT[CA]ATTA; Fig. 3A). This site is separated by three nucleotides from a sequence CTGACT, which differs by two nucleotides (underlined) from the Hth consensus binding site CTGTCA. We also searched for Exd and Hox half sites that would be arranged in a non conventional orientation. This led to the identification of a second Hox/Exd-like binding sequence, ATTATTGATCG, with the 5′ end of the Exd-binding site overlapping the 5′ end of the Hox-binding site (Fig. 3A). This atypical Hox/Exd sequence is not flanked by a putative Hth-binding site. Alignment of *Drosophila melanogaster* and *Drosophila pseudoobscura* sequences further showed that both sites are located within two blocks of conserved sequence, suggesting that they might be functionally important for *CG11339* regulation.

We next analysed the functional importance of these two putative binding sequences for in vivo activity of the *EVIII* enhancer (for comparison, wild-type expression is shown in Fig. 3B,C). We started by mutating the complete Hox/Exd sequence of the two putative Hox/Exd-binding sites, and observed that the mutated enhancer loses its activity in the endoderm (Fig. 3D,E). The same observation was made upon mutation of the Hox/Exd-binding site located most 3′ in the *EVIII* enhancer (Fig. 3F,G). By contrast, mutation of the atypically oriented Hox/Exd-binding site, located more 5′, did not abolish endoderm activity of the enhancer: it resulted in a somewhat general reduction of enhancer activity, including the expression in the head. (Fig. 3H,I). We conclude that *lab*-dependent activity of the *EVIII* enhancer relies on the most 3′ Hox/Exd sequences that strongly diverge from the consensus Lab/Exd sequence.

The Lab/Exd/Hth complex binds the *EVIII* enhancer to a sequence divergent from the consensus binding site

To test for direct binding of Lab and its co-factors Exd and Hth



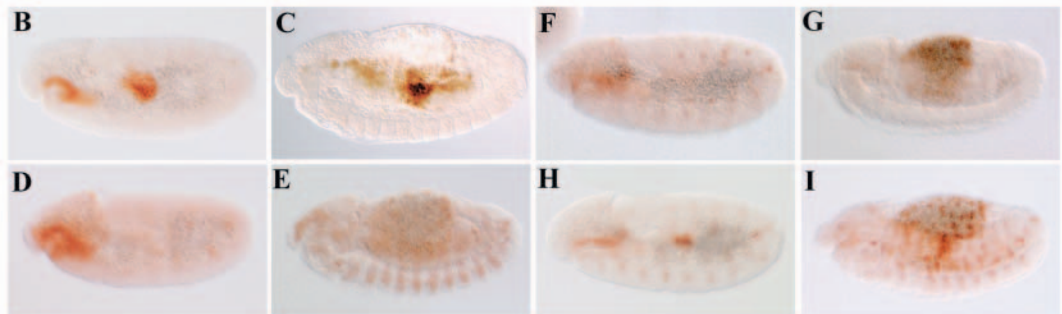


Fig. 3. The divergent Lab/Exd/Hth-binding site is crucial for in vivo activity of *EVIII* enhancer of *CG11339*. (A) Sequence alignment of the *Drosophila melanogaster* (*Dm*) 150 bp fragment EVIII with the corresponding genomic region of *Drosophila pseudoobscura* (*Dp*). Identical nucleotides are outlined in red. The putative Lab/Exd- and Hth-binding sites are indicated (broken boxes), as well as the oligonucleotide used for gel mobility experiments in Fig. 4 (broken blue lines). (B-I) All embryos are shown at stages 12 (B,D,F,H) and 14 (C,E,G,I). (B,C) Expression of β-Gal protein under the control of wild-type enhancer *EVIII*. This expression was lost when fragment EVIII carried mutations in the two putative Lab/Exd-binding sites (D,E) or in the divergent Lab/Exd/Hth-binding site (F,G); and was diminished with mutations in the atypically oriented Hox/Exd-like binding site (H,I).

on the *EVIII* enhancer, the conserved element bearing the functional divergent Lab/Exd/Hth-binding site was used as oligonucleotide for band shift experiments (underlined by blue broken line in Fig. 3A). We found that Lab alone (Fig. 4A, lane 3), with Exd (Fig. 4A, lane 4), or with Exd and Hth (Fig. 4A, lane 5), binds on *EVIII*. Binding specificity was confirmed by the addition of a polyclonal Lab antibody, that inhibited Lab binding (Fig. 4A, lane 6), but also Lab/Exd (Fig. 4A, lane 7) and Lab/Exd/Hth (Fig. 4B, lane 8) complex formation.

To verify that the putative Hox, Exd and Hth sites identified within enhancer *EVIII* are responsible for assembling a Lab/Exd/Hth complex, we mutated independently the core sequences expected to bind each of these factors. Mutation of the Hox half site resulted in the loss of Lab (Fig. 4A, lane 11), as well as Lab/Exd (Fig. 4A, lane 12) and Lab/Exd/Hth (Fig. 4A, lane 13) binding. Mutation of the Exd half site did not impair Lab monomer binding (Fig. 4A, lane 16), but abolished the formation of Lab/Exd (Fig. 4A, lane 17) and Lab/Exd/Hth (Fig. 4A, lane 18) complexes. Finally, mutation of the Hth core site resulted in the selective loss of the Lab/Exd/Hth triple complex, while not affecting binding of Lab and Lab/Exd (not shown).

A comparison between the *EVIII* and the *lab48/95* or *repeat3* enhancers highlights important differences in the Lab DNA-binding properties. First, Lab binds as a monomer to *EVIII*, while it can not do so on *repeat3* (Chan and Mann, 1996) or *lab48/95* (Grieder et al., 1997). Second, although Lab and Exd cooperatively bind on these two enhancers, no cooperative binding was observed on *EVIII*. The previous reports that characterized Lab DNA-binding properties used truncated Lab and Exd proteins. This however does not account for the observed differences, as full-length Lab still does not bind as a monomer (compare lanes 21 and 24 in Fig. 4A), and synergises with Exd for complex formation on *lab48/95* (compare lanes 22 and 25 in Fig. 4A). Further support for the idea that the distinct binding behaviour of Lab relies on intrinsic properties of the *EVIII* site comes from the observation that truncated Lab also binds *EVIII* as a monomer (Fig. 4B, lane 3) and does not cooperate with truncated Exd for DNA binding (Fig. 4B, lane 4).

To assess to what extent *EVIII* is selective in assembling Hox/Exd or Hox/Exd/Hth complexes, band shift experiments with all other *Drosophila* Hox proteins were performed. The results demonstrate a high degree of selectivity: dimeric Hox/Exd or trimeric Hox/Exd/Hth complexes only form with Lab (Fig. 4B, lanes 4 and 5), while the other Hox proteins only bind to *EVIII* as monomers and with distinct binding affinities (Fig. 4B, lane 6-26). We concluded from these experiments that the Lab/Exd/Hth triple complex specifically recognizes the TGAT[CA]ATTACAGCTGACT sequence of *EVIII*, which strongly diverges from the established consensus TGAT[GG]AT(T/G)G (N)(1-40) CTGTCA sequence.

Sequence requirements for Lab, Lab/Exd and Lab/Exd/Hth binding to *EVIII*

As the Lab binding site of *EVIII* was strongly different from the *repeat3* and *lab48/95* sequences, we investigated its nucleotide requirements for the Lab/Exd assembly. For this purpose, we analyzed the effects of several mutations (listed in Fig. 5A) by band shift experiments. In the canonical Hox/Exd-binding site, the proteins contact nucleotide from the same DNA strand in a given orientation. Mutations *rev1* (Fig. 5B,

Development

Fig. 4. The site of crucial importance for in vivo activity of enhancer *EVIII* is specifically recognized by Lab and its two co-factors Exd and Hth. (A) Band-shift experiments showing that Lab (lane 3), Lab/Exd (lane 4) and Lab/Exd/Hth (lane 5) specifically bind to the oligonucleotide *EVIII*. Specificity of the binding was verified by adding a polyclonal anti-Lab antibody to the reactions (lanes 6-8), or by mutating in a G stretch the putative core binding sites of Lab (lanes 11-13), Exd (lanes 16-18) and Hth (not shown). We noted that the Exd-binding site mutation affected Lab monomere binding in presence of the co-factor (lanes 17-18). No different DNA-binding activities were observed between the truncated (158-629) and the full-length form of Lab on the consensus Lab/Exd/Hth-binding site of the *lab48/95* enhancer (lanes 19-26). Lanes 1, 9, 14 and 19 are probes alone. Lanes 2, 10, 15 and 20 are probes in presence of the rabbit reticulocyte lysate. For each gel, the following lanes correspond to the Lab protein alone (3, 6, 11, 16, 21 and 24), with Exd (4, 7, 12, 17, 22 and 25), or with Exd and Hth proteins (5, 8, 13, 18, 23 and 26). Lab proteins were full length, except in lanes 21-23, where the 158-629 form (see Materials and methods) described in previous studies was used. (B) Gel mobility experiments with oligonucleotide *EVIII*, in presence of the Hox proteins Lab (lanes 3-5), Proboscipedia (Pb; lanes 6-8), Sex combs reduced (Scr; lanes 9-11), Deformed (Dfd; lanes 12-14), Antennapedia (Antp; lanes 15-17), Ultrabithorax (Ubx; lanes 18-20), AbdominalA (AbdA; lanes 21-23) and AbdominalB (AbdB; lanes 24-26). Hox proteins are alone (lanes 3, 6, 9, 12, 15, 18, 21, 24), with Exd (4, 7, 10, 13, 16, 19, 22, 25) or with Exd and Hth (5, 8, 11, 14, 17, 20, 23, 26). Exd and Hth proteins correspond to the truncated isoforms described in previous studies (see Materials and methods). Lanes 1 and 2 correspond to probe alone or with rabbit reticulocyte lysate, respectively. Hox/Exd or Hox/Exd/Hth complexes are only observed with Lab (lanes 4 and 5). Lab proteins were full length, except in lanes 3-5. Lab binding, alone or in combination with the Exd and Hth co-factors, is indicated by arrowheads.

lanes 8-10) and *rev2* (not shown), that respectively alter the orientation or strand location of the Lab-binding site abolish Lab binding, alone or as part of a dimeric or trimeric complexes. We next generated mutations altering the two central nucleotides CA of the TGAT[CA]ATTA Lab/Exd-binding site present in *EVIII*. We observed that the effect of the mutation was dependent of the identity switch. When these nucleotides are changed to TT (oligonucleotide $C_5A_6{>}T_5T_6$), no binding of Lab (Fig. 5B, lane 13), Lab/Exd (Fig. 5B, lane 14) or Lab/Exd/Hth is observed (Fig. 5B, lane 15). By contrast, changes to GG (oligonucleotide $C_5A_6{>}G_5G_6$), i.e. the identity found in *repeat3* and *lab* cis-regulatory sequences, decrease Lab monomer binding (Fig. 5B, lane 18), whereas Lab/Exd (Fig. 5B, lane 19) and Lab/Exd/Hth (Fig. 5B, lane 20) complexes form more efficiently. Finally introducing a T in place of a C at position 5 (oligonucleotide $C_5{>}T_5$) does not impair Lab monomer binding (Fig. 5B, lane 23), but strongly affects the Lab/Exd (Fig. 5B, lane 24) and Lab/Exd/Hth (Fig. 5B, lane 25) complex formation. We thus conclude that, despite important differences with binding to *repeat3* or *lab48/95*, the formation of the Lab/Exd complex on *EVIII* also crucially depends of the identity of the two central [NN] nucleotides, and requires similar strand location and orientation of the Lab half site.

A peculiarity of the *EVIII* target sequence is that Lab binds to it as a monomer. Drastic mutation of the core Lab binding site (oligonucleotide *EVIII(labmut)* in Fig. 4A, and oligonucleotide $A_7T_8{>}G_7G_8$ in Fig. 5B, lanes 28-30), as well as more discrete alterations at the 3′ end of this site (oligonucleotides $A_{10}{>}G_{10}$, $T_9{>}G_9$ and $T_9A_{10}{>}G_9G_{10}$: lanes 33-35, 38-40 and 43-45 in Fig. 5B, respectively) strongly reduce or alleviate Lab monomer binding. As strong reduction of Lab monomer binding also resulted from mutations at the 5′ end of the Lab half site (corresponding to the positions 5 and 6), it suggests that Lab binding to *EVIII* both requires nucleotides at the 5′ and 3′ ends of the Hox core site. Most surprisingly, while

Development

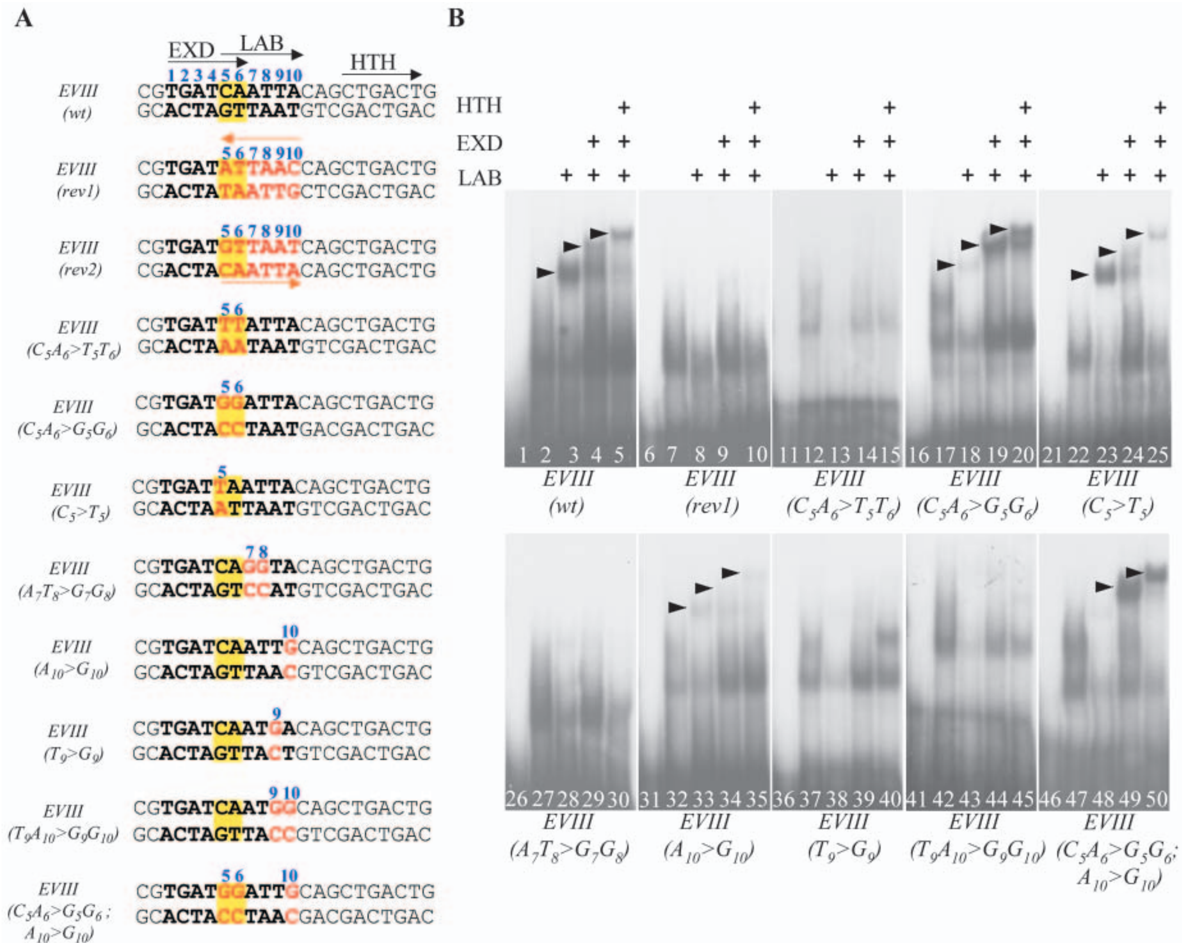


Fig. 5. Sequence requirements for Lab, Lab/Exd and Lab/exd/Hth binding to *EVIII*. (A) Sequences of the mutated *EVIII* oligonucleotides tested for band-shift experiments with Lab and the co-factors Exd and Hth. Positions of each nucleotide within the Lab/Exd sequence are numbered 1 to 10, and orientations of each binding site are indicated (arrowheads). The two central nucleotides of the Lab/Exd-binding site are highlighted in yellow, while mutated nucleotides are in red. (B) Band shift experiments with the control wild-type and the mutated versions of the *EVIII* oligonucleotide listed in A. The identity of the oligonucleotide used is indicated below the gel. For each gel, the first two lanes correspond to the probe, alone (first lane) or with the rabbit reticulocyte lysate (second lane). The next three lanes are in the same order: Lab alone (lanes 3, 8, 13, 18, 23, 28, 33, 38, 43 and 48), Lab with Exd (4, 9, 14, 19, 24, 29, 34, 39, 44 and 49) and Lab with Exd and Hth (lanes 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50).

a mutated oligonucleotide at the 3′ end of the Lab half site forbad Lab monomer, Lab/Exd and Lab/Exd/Hth binding (oligonucleotide A_{10}>G_{10}: Fig. 5B, lanes 33-35), combining this change with mutations at the 5′ end of the site (oligonucleotide C_5A_6>G_5G_6;A_{10}>G_{10}: Fig. 5A) can result in a compensatory effect, leading to Lab/Exd (Fig. 5B, lane 49) and Lab/Exd/Hth (Fig. 5B, lane 50) complex formation. On such an altered site, Lab still does not bind as a monomer (Fig. 5B, lane 48), but the efficiency of Lab/Exd and Lab/Exd/Hth complexes formation is highly increased.

Different Lab-responsive enhancers drive spatially and temporally distinct expression patterns

As different Lab-responsive enhancers are available, we compared their spatial and temporal characteristics. With the exception of the enhancer identified in this study, enhancers are either from the Drosophila *lab* gene [enhancers *lab550* and *lab48/95* (Grieder et al., 1997; Marty et al., 2001)] or from the mouse *Hoxb1* gene [enhancer *repeat3*) (Popperl et al., 1995)]. Sequence characteristics of the Lab/Exd or Lab/Exd/Hth binding sites are shown in Fig. 6A.

Double-labelling experiments monitoring enhancer activity with regard to Lab expression were performed. The *lab550* enhancer, which associates a HRE and a Dpp Response element, reproduces most accurately the spatial characteristics of *lab* expression, but does not provide early expression at stage 12 (Fig. 6B,C). The *lab48/95* enhancer, which is the HRE part of *lab550*, displays the same temporal characteristics as *lab550*, but its activity is weaker and confined to the most posterior *lab*-expressing cells (Fig. 6D,E). The *repeat3* enhancer activity differs from *lab550* and *lab48/95* (Fig. 6F,G): it is active earlier, with an onset of activity detectable from stage 13, and in cells posterior to the *lab* expressing cells at stage 14, cells that were Lab positive at earlier stages (Immergluck et al., 1990). Finally, *EVIII* is the only enhancer that is active as early as Lab, with expression detected in endoderm at stage 12 (Fig. 6H). Although we observed some *EVIII* activity in Lab-negative cells, more sensitive enzymatic

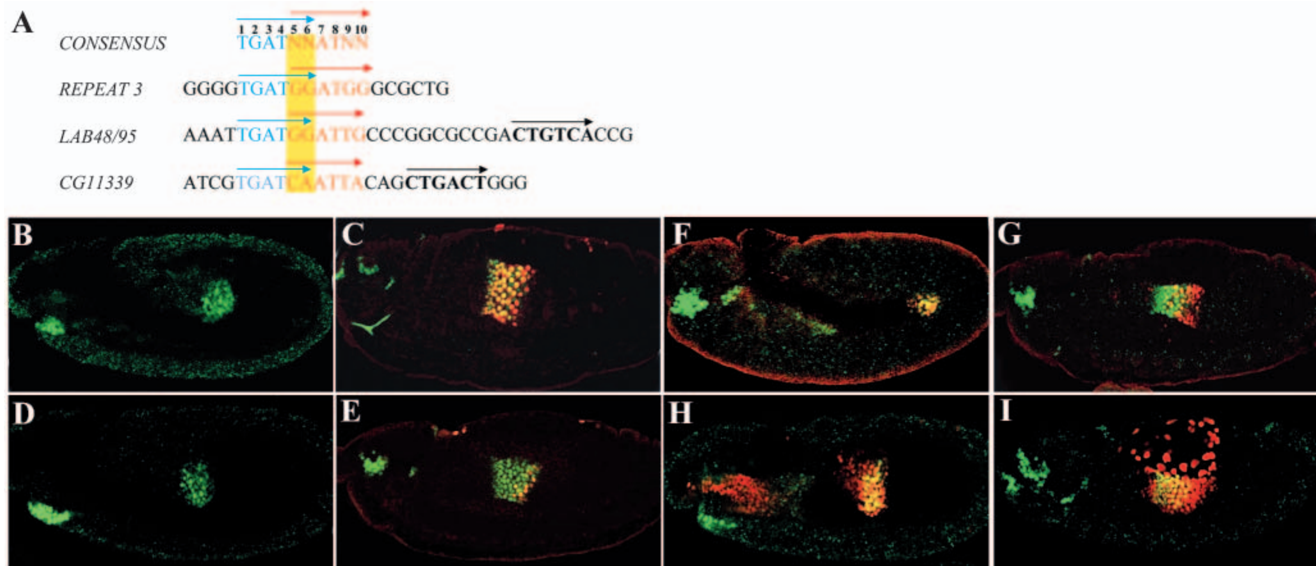


Fig. 6. Midgut expression comparisons between several enhancers regulated by Lab. (A) Sequences of the Lab/Exd-binding sites of enhancers thought to be regulated by Lab (red) and its co-factors Exd (blue) and Hth (bold). The two central nucleotides of the Lab/Exd-binding site are highlighted in yellow. These enhancers are derived from the mouse *Hoxb1* (*repeat3*), or the *Drosophila* Lab (*lab550* and *lab48/95*) genes. The sequence of the divergent Lab/Exd/Hth-binding site of the *EVIII* enhancer of *CG11339* is also shown for comparison. (B-I) All embryos are shown at stages 12 and 14. Confocal analysis of Lab (immunostaining; green) and β-gal (immunostaining; red) expressions in lines carrying the *lacZ* transgene under the control of enhancers *lab550* (B,C), *lab48/95* (D,E), *repeat3* (F,G) and *EVIII* of *CG11339* (H,I).

staining shows that Lab is present in these cells, a conclusion further supported by the loss of enhancer activity in all endodermal cells in *lab* mutant embryos (not shown). At stage 14, *EVIII* activity identifies all cells that have seen Lab at early stages (Immergluck et al., 1990), and therefore defines a pattern broader than Lab-expressing cells at stage 14 (Fig. 6I). These experiments show that the same Lab/Exd complex acts on distinct enhancer sequences to drive spatially and/or temporally distinct expression patterns.

Discussion

Understanding how Hox proteins trigger diversified morphogenesis requires the identification of the mechanisms underlying appropriate target gene selection as well as appropriate target gene regulation, which relies on controlling Hox transregulatory properties (Galant and Carroll, 2002; Merabet et al., 2003; Ronshaugen et al., 2002). At present, most studies have focused on how Hox proteins cooperate with two classes of co-factors, Exd/Pbx and Hth/Meis, to reach DNA-binding selectivity.

Although not valid for the regulation of all Hox target genes (Capovilla et al., 1994; Galant et al., 2002; Grienenberger et al., 2003; Li et al., 1999), the Hox-binding selectivity model is a useful conceptual framework for understanding how Hox proteins, which as monomers display similar DNA-binding properties, reach specificity in target site recognition by interacting with a single co-factor, Exd. This model implies that distinct Hox/Exd complexes select different binding sites, which has been well documented. First, in vitro studies have shown that the prototypical TGAT[NN]ATNN Hox/Exd site recruits Lab or Ubx, depending on the identity of the two central NN nucleotides: GG selects Lab/Exd, while TT or TA recruits a Ubx/Exd complex (Chan and Mann, 1996). Second, the *Distalless* regulatory element that mediates repression by Ubx contains a Hox/Exd site where the two central nucleotides are TT (Gebelein et al., 2002). Third, switching the identity of these two central nucleotides from GG to TA, within the context of *repeat3*, leads to the recruitment of a Dfd/Exd complex instead of Lab/Exd, and in vivo transformed the Lab-responsive enhancer into a Dfd-responsive enhancer (Chan et al., 1997). Similar DNA binding preferences were also observed with the vertebrate Hox and Pbx homologues (Chang et al., 1996).

We used an in silico approach based on the Hox DNA-binding selectivity model to find novel Lab target genes. Although the approach identified 40 putative target sequences for the Lab/Exd/Hth complex, expression analysis of half of them only identified a single novel Lab target, *CG11339*. This suggests that sequences mediating Lab regulatory function in vivo are insufficiently well defined, which is further supported by the finding that the regulation of *CG11339* did not rely on the consensus Lab/Exd/Hth-binding site used for the in silico approach, but on a strongly divergent sequence. These results have implications both with regard to the mode of Lab DNA-binding and more generally to the Hox-binding selectivity model.

Previous work proposed that Lab is very peculiar among all other Hox proteins, in the sense that it does not bind DNA as a monomer, but do so in association with the co-factor Exd. Mutation of the HX motif confers to Lab the capacity to bind DNA in absence of Exd. Accordingly, it was proposed that the HX exerts an inhibitory effect on Lab DNA binding, which is neutralized when interaction occurs with Exd (Chan et al., 1996). This conclusion was reached by studying the DNA-binding properties of Lab on the mouse *repeat3* enhancer.

Here, we observed that this conclusion does not hold on another target sequence, the *EVIII* enhancer of *CG11339*, indicating that the previous conclusion could reflect a specialisation of Lab activity with regard to its autoregulation, rather than a general feature that distinguish the mode of Lab DNA binding from that of other Hox proteins.

The Hox-binding selectivity model also implies that a given Hox/Exd complex should recognize a consensus nucleotide sequence in downstream target genes, which, owing to the lack of well characterised Hox target sequences, still remains to be experimentally validated. We found that the sequence responsible for Lab-mediated regulation of *CG11339* is TGAT[CA]ATTA, which diverges from the TGAT[GG]ATTG site mediating *lab* autoregulation, at the two central positions that are predicted to define the choice of the Hox protein recruited with Exd. The fact that Lab can recognize target sequence differing at the central NN nucleotide is also observed upon mutation of these nucleotides from GG to TA in the *lab550* autoregulatory enhancer (Grieder et al., 1997). Thus, Lab can form a complex with Exd and activates transcription in vivo on at least three sequences that differ with regard to the identity of the central NN nucleotides: GG in *repeat3*, TA in the mutated *lab* enhancer and CA in *CG11339*.

As altering the GG identity of the central NN nucleotides in *repeat3* to TA or TT alleviates Lab/Exd complex assembling, the readout of the nucleotide identity at the central NN positions most probably depends upon neighbouring nucleotides that are different in *repeat3*, *lab48/95* and *CG11339*. Examination of the three sites shows that the Exd half sites are conserved, while the Hox half site differs at the most 3′ end. In support for a role of nucleotides lying in the Hox half site in the readout of the identity of the central NN nucleotides, we found that loss of Lab/Exd complex assembly following mutations at the 3′ end of the Hox half site can be reversed by modifying the two central positions (Fig. 5). This compensatory effect might result from subtle changes in contacting helix 3 of the HD, which in turn might modify the sequence requirement at the central NN position for efficient Lab/Exd recruitment. The importance of the Hox half site 3′ end sequences is further supported by the observation that Scr and Dfd both bind in vitro and act in vivo on a prototypical Hox/Exd site that shares a TA at the central NN position, but differs in the identity of nucleotides at the 3′ end of the Hox half site: GA for Dfd (Chan et al., 1997) and CT for Scr (Ryoo and Mann, 1999).

Variability in the sequence and spacing of the Hth-binding site might also influence the choice of the Hox protein that will preferentially form a complex with Exd and Hth. In any case, our study clearly shows that one Hox/Exd complex can recognize divergent sequences in two different regulated target genes. Although the two central nucleotides play a crucial role in assembling a specific Hox/Exd complex (Chan and Mann, 1996; Chan et al., 1997; Ryoo and Mann, 1999), added complexity to the Hox-binding selectivity model need to be considered, and the nature of these two base pairs will not necessarily predict which Hox protein will selectively bind with the co-factor Exd.

Finally, our data might also open perspectives on the mechanisms underlying the establishment of complex and distinct transcriptional patterns downstream of Hox genes. Hox transcription factors are usually expressed in broad domains, yet downstream target genes are often activated or repressed only in part of the Hox expression domain. It has previously been shown that regulatory regions of downstream target genes integrate signalling inputs (Grienenberger et al., 2003; Marty et al., 2001), which provides additional positional information to restrict downstream target gene activation. These observations highlight the importance of the environment of the Hox/Exd-binding sequence in mediating transcriptionally distinct outputs. Here, we show that Lab responsive enhancers that bear Lab/Exd-binding sites drive distinct expression patterns, both with regard to spatial and temporal characteristics. It suggests that in addition to environmental cues, the identity of the Hox/Exd sites might also be instructive.

We are grateful to Yacine Graba for discussions and helpful comments of the manuscript. We thank Richard Mann, Paul Baumgartner, Urs Kloter, Frederic Prince, Martin Müller, Frank Hirth, François Karch, Nicole deButts, William McGinnis, Sean Carroll and the Bloomington stock center for plasmids, flies and antibodies. This work was supported by the Suiss National Science Foundation, and the Kantons Basel-Stadt and Basel-Land. S.M. was supported by l'Association pour la Recherche contre le Cancer (ARC) and an EMBO postdoctoral fellowships.

References

Capovilla, M., Brandt, M. and Botas, J. (1994). Direct regulation of *decapentaplegic* by *Ultrabithorax* and its role in midugt morphogenesis. *Cell* **76**, 461-475.

Celniker, S. E., Keelan, D. J. and Lewis, E. B. (1989). The molecular genetics of the bithorax complex of *Drosophila:*: characterization of the products of the *Abdominal-B* domain. *Genes Dev.* **3**, 1424-1436.

Chan, S.-K. and Mann, R. S. (1996). A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer. *Proc. Natl. Acad. Sci. USA* **93**, 5223-5228.

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.

Chan, S.-K., Pöpperl, H., Krumlauf, R. and Mann, R. S. (1996). An extradenticle-induced conformational change in a HOX protein overcomes an inhibitory function of the conserved hexapeptide motif. *EMBO J.* **15**, 2476-2487.

Chan, S. K., Ryoo, H. D., Gould, A., Krumlauf, R. and Mann, R. S. (1997). Switching the in vivo specificity of a minimal Hox-responsive element. *Development* **124**, 2007-2014.

Chang, C. P., Brocchieri, L., Shen, W. F., Largman, C. and Cleary, M. L. (1996). Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus. *Mol. Cell. Biol.* **16**, 1734-1745.

Cribbs, D. L., Pultz, M. A., Johnson, D., Mazzulla, M. and Kaufman, T. C. (1992). Structural complexity and evolutionary conservation of the Drosophila homeotic gene proboscipedia. *EMBO J.* **11**, 1437-1449.

Diederich, R. J., Merrill, V. K. L., 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.

Galant, R. and Carroll, S. B. (2002). Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**, 910-913.

Galant, R., Walsh, C. M. and Carroll, S. B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development* **129**, 3115-3126.

Gebelein, B., Culi, J., Ryoo, H. D., Zhang, W. and Mann, R. S. (2002). Specificity of distalless repression and limb primordia development by abdominal hox proteins. *Dev. Cell* **3**, 487-498.

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

Grieder, N. C., Marty, T., Ryoo, H. D., Mann, R. S. and Affolter, M. (1997).

Development

Synergistic activation of a *Drosophila* enhancer by HOM/EXD and DPP signaling. *EMBO J.* **16**, 7402-7410.
Grienenberger, A., Merabet, S., Manak, J., Iltis, I., Fabre, A., Berenger, H., Scott, M. P., Pradel, J. and Graba, Y. (2003). Tgf{beta} signaling acts on a Hox response element to confer specificity and diversity to Hox protein function. *Development* **130**, 5445-5455.
Hayashi, S. and Scott, M. P. (1990). What determines the specificity of action of Drosophila homeodomain proteins? *Cell* **63**, 883-894.
Hirth, F., Loop, T., Egger, B., Miller, D. F., Kaufman, T. C. and Reichert, H. (2001). Functional equivalence of Hox gene products in the specification of the tritocerebrum during embryonic brain development of Drosophila. *Development* **128**, 4781-4788.
Immergluck, K., Lawrence, P. A. and Bienz, M. (1990). Induction across germ layers in Drosophila mediated by a genetic cascade. *Cell* **62**, 261-268.
LeMotte, P. K., Kuroiwa, A., Fessler, L. I. and Gehring, W. J. (1989). The homeotic gene *Sex Combs Reduced* of Drosophila: gene structure and embryonic expression. *EMBO J.* **8**, 219-227.
Li, X., Veraksa, A. and McGinnis, W. (1999). A sequence motif distinct from Hox binding sites controls the specificity of a Hox response element. *Development* **126**, 5581-5589.
Lin, L. and McGinnis, W. (1992). Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev.* **6**, 1071-1081.
Mann, R. S. (1995). The specificity of homeotic gene function. *BioEssays* **17**, 855-863.
Marty, T., Vigano, M. A., Ribeiro, C., Nussbaumer, U., Grieder, N. C. and Affolter, M. (2001). A HOX complex, a repressor element and a 50 bp sequence confer regional specificity to a DPP-responsive enhancer. *Development* **128**, 2833-2845.
McGinnis, W. and Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283-302.
Merabet, S., Kambris, Z., Capovilla, M., Berenger, H., Pradel, J. and Graba, Y. (2003). The hexapeptide and linker regions of the AbdA Hox protein regulate its activating and repressive functions. *Dev. Cell* **4**, 761-768.
Pai, C. Y., Kuo, T. S., Jaw, T. J., Kurant, E., Chen, C. T., Bessarab, D. A., Salzberg, A. and Sun, Y. H. (1998). The Homothorax homeoprotein activates the nuclear localization of another homeoprotein, extradenticle, and suppresses eye development in Drosophila. *Genes Dev.* **12**, 435-446.
Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S. and Aggarwal, A. K. (1999). Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex. *Nature* **397**, 714-719.
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.
Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L. and Wolberger, C. (1999). Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation. *Cell* **96**, 587-597.
Popperl, H., Bienz, M., Studer, M., Chan, S. K., Aparicio, S., Brenner, S., Mann, R. S. 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., 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**, 1101-1112.
Rieckhof, G. E., Casares, F., Ryoo, H. D., Abu-Shaar, M. and Mann, R. S. (1997). Nuclear translocation of extradenticle requires homothorax, which encodes an extradenticle-related homeodomain protein. *Cell* **91**, 171-183.
Ronshaugen, M., McGinnis, N. and McGinnis, W. (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature* **415**, 914-917.
Ryoo, H. D. and Mann, R. S. (1999). The control of trunk Hox specificity and activity by Extradenticle. *Genes Dev.* **13**, 1704-1716.
Ryoo, H. D., Marty, T., Casares, F., Affolter, M. and Mann, R. S. (1999). Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex. *Development* **126**, 5137-5148.
Schneuwly, S., Kuroiwa, A. and Gehring, W. J. (1987). Molecular analysis of the dominant homeotic *Antennapedia* phenotype. *EMBO J.* **6**, 201-206.
Streit, A., Kohler, R., Marty, T., Belfiore, M., Takacs-Vellai, K., Vigano, M. A., Schnabel, R., Affolter, M. and Muller, F. (2002). Conserved regulation of the Caenorhabditis elegans labial/Hox1 gene ceh-13. *Dev. Biol.* **242**, 96-108.
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.

Development