Nucleic Acids Research, 1995, Vol. 23, No. 17 3481–3487

Deformed expression in the *Drosophila* central nervous system is controlled by an autoactivated intronic enhancer

Lily Lou, Clare Bergson[1] and William McGinnis*

Departments of Pediatrics and [1]Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

Received May 15, 1995; Accepted July 18, 1995

EMBL accession no. X87111

ABSTRACT

***Deformed* (*Dfd*) is a *Drosophila* homeotic selector gene required for normal development of maxillary segment morphology in the larval and adult head. Consistent with this function, *Dfd* transcripts are restricted to epidermal, mesodermal and neural cells in the embyonic mandibular and maxillary primordia. Previous studies have identified a far upstream element in *Dfd* sequences which functions as an epidermal-specific autoregulatory enhancer. In a search through 35 kb of *Dfd* sequences for additional transcriptional control elements, we have identified a 3.2 kb DNA fragment containing an enhancer that mimics the expression of *Dfd* in the subesophageal ganglion of the embryonic central nervous system. This Neural autoregulatory enhancer (NAE) maps in the large *Dfd* intron just upstream of the homeobox exon and requires Dfd protein function for its full activity. A 608 bp NAE subfragment retains regulatory function that is principally localized in the subesophageal ganglion. This small region of the *Drosophila melanogaster* genome contains numerous blocks of sequence conservation with a comparable region from the *Dfd* locus of *D.hydei*. A pair of conserved blocks of NAE sequence match a Dfd protein binding site in the epidermal autoregulatory element, while another conserved sequence motif is repeated multiple times within the 608 bp subelement.**

INTRODUCTION

The generation of diverse embryonic structures from the morphologically homogeneous *Drosophila* egg is controlled by the sequential functioning of several classes of patterning genes. Once the overall polarities and segmental fields of the embryo have been established by the maternal effect, gap, pair-rule and segment-polarity genes (1,2), the homeotic selector or HOM genes assign different identities to the newly defined segments and parasegments (3,4). The *Drosophila* HOM genes are a branch of the HOM/Hox gene family. Many lines of evidence indicate this gene family has an evolutionarily conserved role in patterning the anterior-posterior axis of animal embryos (5–7). In *Drosophila* embryos, HOM gene expression patterns are established in spatially restricted domains in response to coordinate, gap and pair-rule proteins, but just as important to the function of the HOM genes are the post-establishment expression patterns during which HOM gene expression is maintained and modulated (8). The post-establishment phases of homeotic gene expression are controlled by the repressive functions exerted by genes of the *Polycomb* group, by the activating functions exerted by genes such as *trithorax* and *brahma* and by auto- and cross-regulatory effects among the HOM genes themselves (9–11).

Autoregulation appears to be particularly important for maintaining the expression pattern of *Deformed* (*Dfd*), a HOM gene which assigns identities in the posterior head region of *Drosophila* (12). *Deformed* null mutants die at the end of embryogenesis and are missing most cuticular specializations of the maxillary and mandibular segments. The idea that *Dfd* autoactivation is important for the assignment of identity in maxillary epidermal cells rests on two types of evidence. First, *Dfd* transcription is normally initiated in *Dfd* mutant embryos carrying nonsense codons in the *Dfd* open reading frame, but normal transcript levels are not maintained in either epidermal or neural cells of these embryos (13). Secondly, ectopic expression of Dfd protein under the control of a heat shock promoter can activate the endogenous *Dfd* transcription unit in some cells of the thoracic and abdominal epidermis (14). This ectopic *Dfd* transcription is maintained for much of embryogenesis and induces the development of ectopic maxillary cuticular structures in the labial and thoracic segments. The DNA sequences from the *Dfd* locus that mediate this aspect of the autoactivation circuit map within a large upstream epidermal-specific enhancer (13,15).

Deformed is also expressed persistently in internal head mesodermal and neural cells, where its functions are less well understood than in the epidermis. Merrill *et al.* (16) noted that some *Dfd* hypomorphic mutants die soon after reaching adulthood, apparently as a result of sensory and/or motor defects that are likely to originate in neural or muscle tissue. Using hypomorphic and temperature sensitive mutations of *Dfd*, Restifo and Merrill (17) have recently shown that the gene is required for the formation of a normal cervical connective between the thoracic and subesophageal ganglia. The subesophageal ganglion, which innervates the gnathal region including the maxillary and mandibular segments,

*To whom correspondence should be addressed at present address: Department of Biology, Room 3100A Pacific Hall, 9500 Gilman Drive, University of California, San Diego, La Jolla, CA 92093-0357, USA

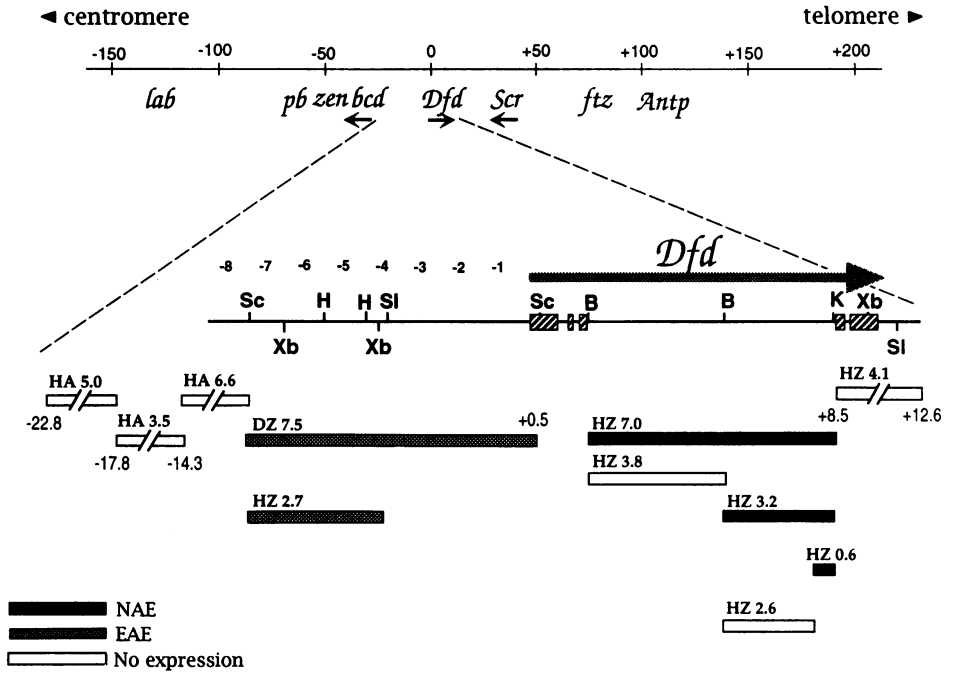

Figure 1. Regulatory constructs from the *Deformed* region. Approximately 35 kb was included in a series of reporter gene constructs which tested the depicted fragments for their ability to direct β-galactosidase expression in developing *Drosophila* embryos. The putative regulatory elements, denoted by bars, were combined with promoter and polyadenylation signals in the P-element reporter test vectors. DZ indicates use of the *Dfd* promoter attached to *lacZ* reporter gene. HA and HZ constructs contained the *lacZ* gene flanked by the *hsp70* promoter and polyA site. The gray bars denote regions which direct reporter gene expression in the epidermal portion of the wild type *Dfd* pattern (EAE) and the black bars specify the regions capable of directing expression primarily in the neural part of the *Dfd* pattern (NAE). The open bars represent regions which did not activate reporter gene expression in our studies.

includes the CNS cells that normally express *Dfd* transcripts and proteins during embryonic, larval and pupal stages (18–22).

In an effort to learn more about the sequences and factors that allow persistent, localized expression of HOM genes in different tissues, we have screened for additional regulatory elements which provide other aspects of the *Dfd* expression pattern. Our screen identified an enhancer that maps in the *Dfd* intron that directs expression in the central nervous system portion of the *Dfd* expression domain. We also present evidence that persistent CNS *Dfd* expression is under positive autoregulatory control. Portions of the CNS enhancer are conserved between divergent fly species and show partial functional conservation when tested in *D.melanogaster* embryos. Sequence analysis of these regions reveal two categories of repeated motifs, one of which shows similarity to previously identified Dfd protein binding sites.

MATERIALS AND METHODS

Reporter constructs

All of the reporter constructs shown in Figure 1 were made by subcloning fragments from the *Dfd* locus into restriction sites 5′ of the *hsp70* basal promoter in the enhancer test vector HZ50 (23). HA 5.0 includes a 5.0 kb *Sac*I fragment containing sequences from –22.8 to –17.8 kb relative to the *Dfd* transcription start, HA 3.5 a *Sac*I fragment containing sequences from –17.8 to –14.3, HA 6.6 a *Sac*I fragment containing sequences from –14.3 to –7.5, HZ 7.0 a *Bam*HI/*Kpn*I fragment from 1.5 to 8.5, HZ 3.8 a *Bam*HI fragment from 1.5 to 5.3, HZ 3.2 a *Bam*HI/*Kpn*I fragment from 5.3 to 8.5, HZ 2.6 a *Bam*HI/*Pst*I fragment from 5.3 to 7.9, HZ 0.6 a 608 bp *Pst*I/*Kpn*I fragment from 7.9 to 8.5 and HZ 4.1 a *Kpn*I/*Not*I fragment from 8.5 to 12.6. The construction of DZ 7.5, HZ 2.7 are described in Bergson and McGinnis (15). The H designation indicates the use of the *hsp70* promoter in the reporter construct, while D indicates that the endogenous *Dfd* promoter was used. All *D.melanogaster* DNA fragments were isolated from the genomic phage clones λA87, λ99 and λ100 (24,25).

HZ [Dh-600] was made by PCR amplification of a 677 bp *D.hydei* fragment homologous to the *D.melanogaster* HZ 0.6 insert. Primers included restriction sites to allow insertion of the amplified sequence into pBluescript DNA (Stratagene). *D.hydei* phage clone λDhDfd3, which contains 12.2 kb of sequence from the *hydei Dfd* gene (unpublished results), served as the template for amplification. The *D.hydei* genomic library was a generous gift from V. Herndon and S. Artavanis-Tsakonas. After sequencing/verification, the *D.hydei* 677 bp insert was subcloned into HZ50 upstream of the *hsp70* promoter.

The enhancer test constructs described above were co-injected with transposase-encoding helper plasmid pπ25.7wc (26) into ry^{506} embryos. DNA was prepared using double CsCl gradients or QIAGEN columns for purification. The standard procedure for P-element transformation was followed (27) with some preinjection modifications: for example embryos were dechorionated mechanically, using double-sided tape and green food-coloring and ampicillin were added to the injection mixture just prior to needle loading.

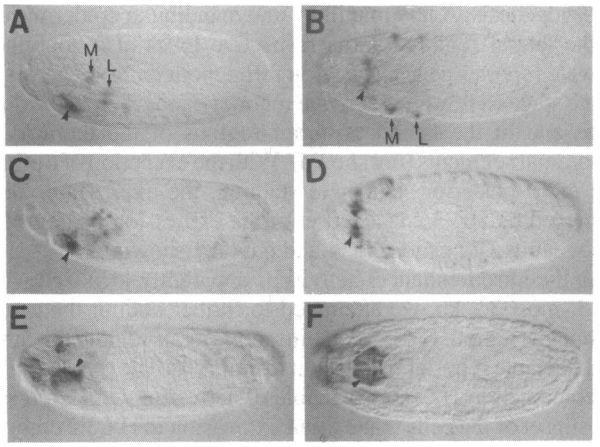

Figure 2. Expression pattern directed by the *Dfd* intronic enhancer. β-galactosidase expression directed by the HZ 3.2 construct is detected by immunolocalization. Panels (**A**), (**C**) and (**E**) show embryos in sagittal view; (**B**), (**D**) and (**F**) in horizontal views. Early Stage 11 embryos are shown in (A) and (B) and late Stage 11 in (C) and (D). Stage 16 embryos are shown in (E) and (F). Reporter gene expression is first detected early in Stage 11 (~5 h after egg lay) and is present throughout embryogenesis. The pattern comprises a subset of the *Dfd* expression pattern, including maxillary epidermal cells (M) and vental mandidular cells that will contribute to the subesophageal ganglion (arrowheads). Ectopic expression is seen in the labial lobe epidermis (L) and in a few cells in the posterior lateral regions of each segment. The pattern directed by the larger HZ 7.0 construct is indistinguishable from that of HZ 3.2 (data not shown).

Standard genetic crosses were used to determine chromosomal location of inserts and to establish homozygous or balanced lines. Multiple single insert lines were tested for each construct. These were either independently injected or derived from induced transposition by crossing single-insert lines with strains containing a genomic source of P transposase.

Immunohistochemistry

For whole mount *in situ* staining with antibodies, embryos were collected, fixed and devitellinized as previously described (28). β-galactosidase protein was detected with anti-β-gal polyclonal (Cappel) or monoclonal (Promega) serum, used at 1:1000 dilution after immunoadsorption against *Drosophila melanogaster* embryos. Biotinylated goat anti-rabbit IgG (Jackson Immunoresearch) secondary antibody was used at 1:500 dilution. Color was developed using Vectastain 4000 ABC reagent and diaminobenzidine reaction.

DNA sequencing

DNA sequence was determined on both strands using the dideoxy chain termination method (29), using Sequenase version 2.0, standard T7 and T3 primers and synthesized internal primers. Sequences were read and compared using the MacVector program (IBI).

RESULTS

Neural regulatory element within the *Dfd* locus

Previous work has shown that a region mapping between –3.9 and –6.6 kb relative to the *Dfd* transcription start can direct reporter gene expression in the maxillary and mandibular epidermis of Stage 11–17 *Drosophila* embryos (13,15,30). To search for additional enhancers that direct other aspects of *Dfd* expression in embryos, a number of restriction fragments mapping between –22.8 and +12.6 were tested for their ability to activate *hsp70* promoter-*lacZ* reporter gene expression in *Dfd*-like patterns in transformed embryos (Fig. 1; for a more detailed molecular description of the *Dfd* transcription unit, see ref. 25). All fragments

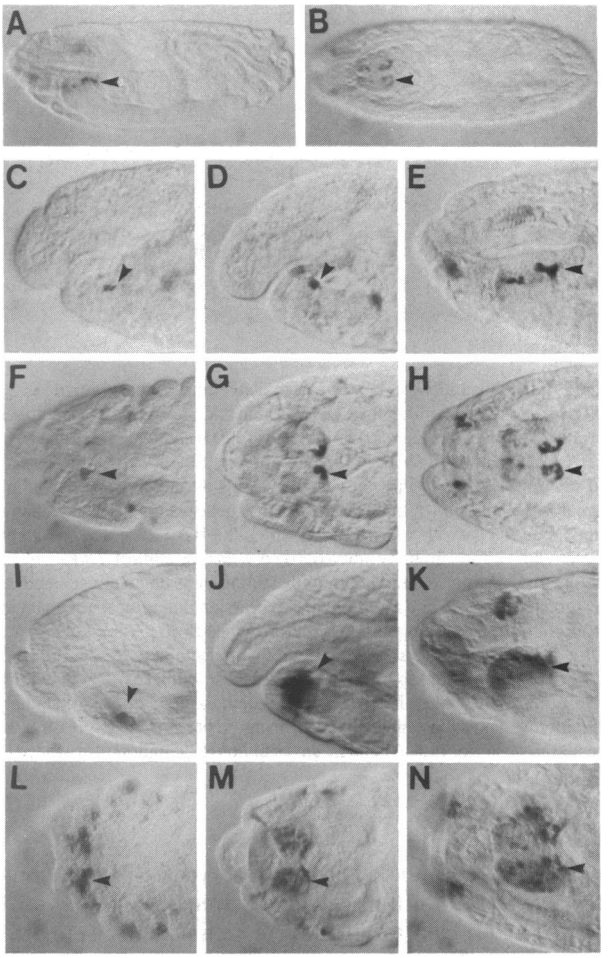

Figure 3. Reporter gene expression pattern of HZ 0.6, compared to HZ 3.2. β-galactosidase expression in all embryos visualized by antibody staining. The dorsal edge of the mandibular region of the subesophageal ganglion is marked by an arrowhead in all panels. (**A**) and (**B**) show sagittal and horizontal views of stage 15 HZ 0.6 embryos. (**C**), (**D**) and (**E**) show sagittal views of the head regions of HZ 0.6 embryos at Stage 11 (C), Stage 12 (D) and Stage 15 (E). Compare with sagittal views of HZ 3.2 expression patterns at comparable stages in panels (**I**), (**J**) and (**K**). Panels (**F**), (**G**) and (**H**) show horizontal views of HZ 0.6 embryos at Stage 11 (F), Stage 12 (G) and Stage 15 (H). Compare with horizontal views of HZ 3.2 expression patterns at comparable stages in panels (**L**), (**M**) and (**N**). Although fewer cells stain in the HZ 0.6 embryos than in HZ 3.2 lines, cells in the dorsal-medial region of the subesophageal ganglion, which stained most intensely in HZ 3.2 embryos, consistently exhibit the most abundant HZ 0.6 activity. Note that as the subesophageal ganglia develop and gradually rise to near vertical orientations, cells that were originally in dorsal positions appear to be posterior. There is also epidermal staining in a few maxillary and more posterior epidermal cells, most of which are out of the plane of focus in these photographs. This pattern of epidermal staining, some of which is in *Dfd* expressing cells, is seen in both HZ 0.6 and HZ 3.2 embryos.

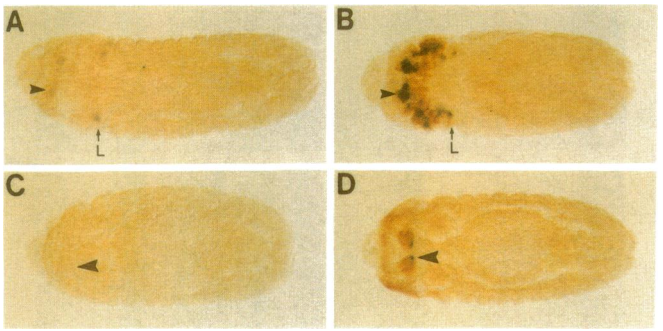

Figure 4. Double antibody stain of *Dfd* mutant embryos that contain HZ 3.2 or HZ 0.6. The brown staining marks Dfd protein and blue marks β-galactosidase. (**A**) and (**B**) show HZ 3.2 embryos and (**C**) and (**D**) show HZ 0.6 embryos. (A) The ventral aspect of a Dfd^{Rx1} mutant embryo at Stage 12; with no detectable Dfd staining. HZ 3.2 is active at very low levels in only a few cells within the normal *Dfd* expression domain. Some reporter gene staining is also seen in cells in the labial (L) and other segments in which the HZ 3.2 element is expressed in both wild type and mutant embryos. (B) A wild type control (Dfd^{Rx1}/balancer heterozygotes) which exhibits both Dfd staining in its characteristic epidermal plus proneural pattern and reporter gene expression that is most abundant in the primordia for the mandibular ganglion (arrowhead). (C) A horizontal view of a Dfd^{Rx1} mutant embryo at Stage 14; with no detectable Dfd staining. HZ 0.6 is not detectably expressed in the CNS of these embryos. (D) A wild type control (Dfd^{Rx1}/Balancer heterozygotes) which exhibits both Dfd staining in the epidermis and subesophageal ganglion and reporter gene expression that is detected in posterior dorsal cells of the mandibular ganglion (arrowhead).

were cloned into the enhancer test vector HZ50, multiple independent lines of transgenic flies were established for each construct and β-galactosidase expression patterns were detected by antibody staining of embryos.

Sequences in the large *Dfd* intron contained in HZ 7.0 activate reporter gene expression in a pattern that overlaps with the normal *Dfd* expression domain in the central nervous system. Constructs HZ 3.8 and HZ 3.2 independently test the regulatory activity of two subfragments within the HZ 7.0 construct, HZ 3.2 containing the right or 3′ portion of the intron. Embryos carrying the HZ 3.2 transgenic construct express β-galactosidase in a pattern indistinguishable from that directed by the parental HZ 7.0 (Fig. 3I and J). Embryos carrying HZ 3.8 exhibited no β-galactosidase expression (not shown). We also tested DNA extending from −7.5 to −22.8 kb in the three regulatory constructs, HA 5.0, HA 3.5 and HA 6.6. No β-galactosidase expression could be detected in transformants carrying these reporter constructs. Sequences from +8.5 to +12.6 were tested in HZ 4.1, which includes the 4th and 5th exons of the *Dfd* transcription unit as well as ~2 kb of downstream sequence. This construct also displayed no detectable regulatory activity in embryos.

Expression directed by sequences HZ 3.2 or HZ 7.0, is first detected in Stage 10 ventral mandibular cells (Fig. 2). Most of these cells lie within the mandibular neurogenic region and as judged by their size and shape many correspond to neuroblasts (30). At the same time a few cells in lateral regions of the maxillary and labial lobes also weakly activate reporter gene expression. In 8–10 h embryos, expression from HZ 3.2 is detected in the anteriormost cells of the developing subesophageal ganglion, which presumably derive from the original ventral mandibular expression domain. This subesophageal CNS expression is detected continuously to late stages (Stage 17) of embryogenesis. A few maxillary and mandibular epidermal cells in the lateral regions of the respective lobes also continue to activate reporter gene expression throughout embryogenesis and at late stages a few cells activate reporter expression in the dorsal ridge and in the lateral posterior regions of the thoracic and abdominal segments (not shown). With the exception of the labial and more posterior segmental staining, the expression pattern conferred by HZ 3.2 is a subset of the expression pattern of the endogenous *Dfd* gene (Fig. 4 and data not shown; 20).

As the intron enhancer activity is apparently localized within the 3′-most 3.2 kb, we attempted to further delimit the activity within this region. HZ 2.6 contains most of this region, but directs no detectable reporter expression in embryos. The remaining 608 bp contained in HZ 0.6 is capable of activating *lacZ* expression in a subset of the cells that exhibit expression in HZ 3.2 embryos. Figure 3 compares the expression patterns provided by HZ 3.2 and HZ 0.6. The highest reproducible levels of expression from HZ 3.2 are located in medial dorsal regions of the subesophageal ganglion (Fig. 3M and N). This region is virtually the only CNS domain where HZ 0.6 is active in late stage embryos (Fig. 3G and H), although a few cells in more ventral regions of the subesophageal ganglion exhibit weak expression at late stages. HZ 0.6 embryos also have a similar, but weaker expression pattern in the epidermis that mimics that seen for the HZ 7.0 and HZ 3.2 constructs.

The CNS enhancer is autoregulated

As the previously described epidermal enhancer (HZ 2.7) functions in a *Dfd* dependent manner (15), we tested whether the intronic CNS enhancer was dependent on *Dfd* function. In the first set of experiments, both HZ 3.2 and HZ 0.6 were tested in Dfd^{Rx1} mutant embryos. The Dfd^{Rx1} chromosome produces no detectable Dfd protein. These embryos and wild type controls, were double stained with both anti-Dfd and anti-β-galactosidase antiserum. As shown in Figure 4B and D, the expression patterns directed by the HZ 3.2 and HZ 0.6 reporter gene test constructs map within the normal Dfd protein expression domain, with the exception of a few cells that weakly express β-galactosidase in the labial and more posterior segments. In *Dfd* mutant embryos, the activity of both regulatory constructs is severely diminished or abolished in the central nervous system and in maxillary epidermal cells. The HZ 3.2 construct still retains weak ventral mandibular/CNS activity in *Dfd* mutants, with β-galactosidase staining barely detectable and in fewer cells (Fig. 4A). As expected, the regulatory activity of the constructs in cells which are outside the normal *Dfd* expression domain was not affected in the *Dfd* mutants (Fig. 4A). Since the 3.2 kb fragment and the 0.6 kb sub-fragment both direct expression principally in the central nervous system in a *Dfd*-dependent manner, we will refer to the regulatory elements that they contain as *Dfd* neural autoregulatory elements (*Dfd*-NAEs).

We also tested whether the activity of the HZ 3.2 construct could be ectopically activated in embryos in which the Dfd protein is ectopically expressed under the control of a *hsp70* promoter (*hsDfd* embryos). A single heat shock, applied at a variety of early embryonic stages ranging from 2.5 to 7 h, did not detectably induce HZ 3.2 in new cells. In *hsDfd* embryos receiving multiple heat shocks, some weak, ectopic expression of β-galactosidase was observed in the CNS that was not present in heat-shocked control embryos lacking the *hsDfd* transgene (data

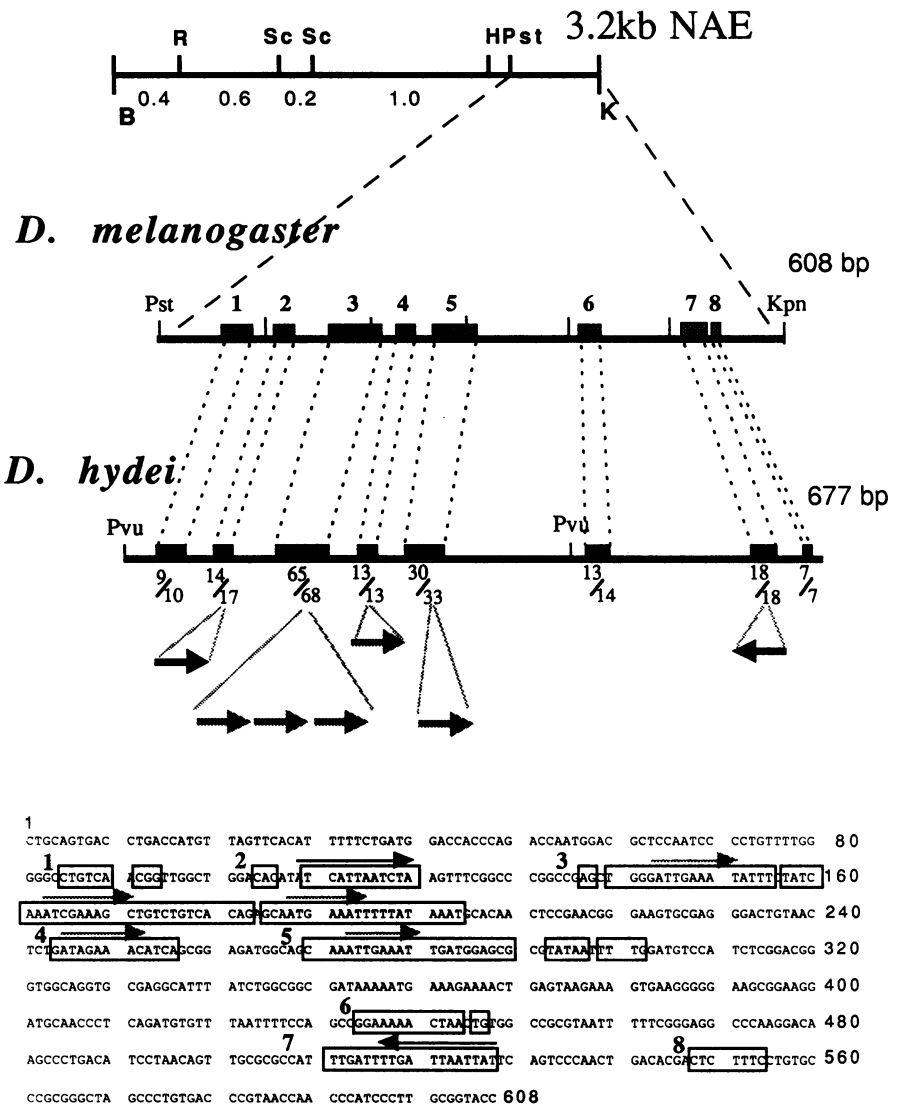

Figure 5. Sequence similarities between 0.6 NAE regions of *D.melanogaster* and *D.hydei*. The top of the figure shows a restriction map of the 3.2 NAE, including the location of the 608 bp *PstI–KpnI* fragment that is sufficient to supply some of the *Dfd*-NAE activity. B, *Bam*HI; R, *Eco*RI; Sc, *Sac*I; H, *Hin*dIII; K, *Kpn*I. The two lines below the restriction map represent the DNA sequences of the *D.melanogaster* and *D.hydei* NAE regions. The numbered shaded boxes depict the eight blocks of interspecifically conserved sequence, with the degree of identity given below the *D.hydei* blocks. Within these regions, an 8 bp motif (consensus: ATTGAAAT) is found five times, as indicated by the shaded arrows. The black arrows indicate homology with the Dfd binding site D from the upstream epidermal autoregulatory element. The bottom shows the DNA sequence of the 608 bp *PstI–KpnI* fragment from *D.melanogaster*. The boxes enclose the blocks of nucleotides that are identical in *D.hydei*. The repeated motifs of ATTGAAAT are overlined by shaded arrows, the homology blocks with Dfd binding site D are overlined with black arrows.

not shown). However, the penetrance of this effect was low and we conclude that ectopic Dfd protein expression is not sufficient to induce the robust activation of the HZ 3.2 regulatory element.

The *Dfd* NAE is evolutionarily conserved

The 608 bp NAE fragment within HZ 0.6 is the smallest element so far defined that is sufficient to supply *Dfd*-dependent expression in the mandibular ganglion. This fragment maps just upstream of the homeobox-containing exon in the *Drosophila melanogaster Dfd* locus. A 1.9 kb *SalI* fragment from the *D.hydei* genome cross-hybridizes with the 608 bp fragment of *D.melanogaster* under reduced stringency conditions (unpublished results) and maps just upstream of the homeobox-containing exon of the *Dfd* gene of that species. *D.melanogaster* and *D.hydei* have been evolutionarily separated for ~60 million years (31). The DNA sequence of the 608 bp fragment from *melanogaster* was obtained and compared to the sequence of a homologous 677 bp region derived from the *hydei* 1.9 kb fragment. There are eight blocks of highly conserved sequence arrayed in the same order and orientation between *melanogaster* and *hydei*, ranging in size from 7 to 68 bp, separated by divergent stretches with no extensive sequence similarity (Fig. 5).

Some interesting patterns emerge on analysis of the conserved sequences. Within homology blocks 3, 4 and 5, an 8 bp sequence motif is repeated 5-fold. The consensus sequence is ATTGAAAT and none of the copies differs in >2 positions. All of the copies are perfectly conserved between *D.melanogaster* and *D.hydei* and arranged in the same orientation. Also highly conserved between

the species in homology blocks 2 and 7 are two copies of a sequence similar to a region within the *Dfd* epidermal autoregulatory element (EAE) that includes a high affinity Dfd protein binding site. This sequence (site D in 13,32) in the *Dfd*-EAE is ATCATTAAGC, which is matched in NAE homology block 2 at 9/10 positions with ATCATTAATC and in homology block 7 (on the bottom strand) at 8/10 positions with ATAATTAATC. In addition, there are a few additional nucleotide matches between blocks 2 and 7 that are diagrammed Figure 5.

A *D.hydei* regulatory element exhibits conserved function

We next asked whether the *D.hydei* 677 bp homologous region was sufficient to confer an NAE-like regulatory function in *D.melanogaster* embryos. Transgenic flies were constructed which carried 677 bp of the *D.hydei* region in the HZ50PL enhancer test vector. These HZ[Dh-677] lines show reporter gene expression in the mandibular, maxillary and labial segments, similar to that directed by the *D.melanogaster* HZ 0.6 element. Both epidermal and CNS components of the pattern are present; however, the HZ[Dh-677] pattern includes many more epidermal cells than HZ 0.6. The CNS expression directed by HZ[Dh-677] becomes localized in the manibular region of the subesophageal ganglion but is weaker and includes fewer cells than HZ 0.6.

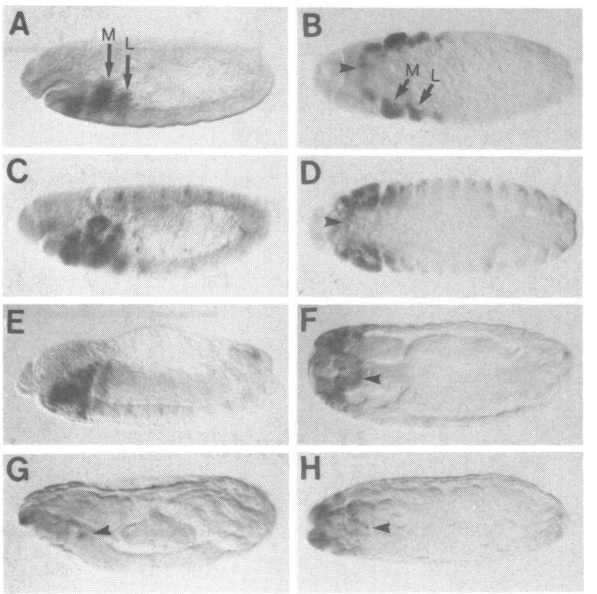

Figure 6. Expression pattern directed by the HZ[Dh677] element of *D.hydei* in *D.melanogaster* embryos. β-galactosidase expression in all embryos is visualised by antibody staining. (**A**), (**C**), (**E**) and (**G**) show sagittal sections and (**B**), (**D**), (**F**) and (**H**) show horizontal section s. (A) and (B) are late stage 10, (C) and (D) are stage 11, (E) and (F) are stage 13–14 and (G) and (H) are stage 16. Note the more extensive expression pattern in the epidermis of gnathal segments than that provided by HZ 0.6 (see Fig. 4) from *melanogaster* (M-maxillary lobe, L-labial lobe). The expression pattern provided by HZ[Dh677] in other body segments also includes more cells than HZ0.6. The arrowheads in (B) and (D) points out the low level expression in the primordia of the mandibular ganglion and the arrowheads in (F), (G) and (H) point out the dorsal edge of the mandibular region of the subesophageal ganglion.

DISCUSSION

In order to achieve its function in assigning segmental identity in the posterior head of *Drosophila*, *Dfd* gene expression must be restricted to a precise, persistent pattern in a variety of tissues, including the central nervous system (14,16,17). In these experiments, we have characterized a transcriptional regulatory element which directs expression in the mandibular ganglion of the developing *D.melanogaster* nervous system. This *Dfd*-NAE, which maps upstream of the homeobox exon in the large intron of *Dfd*, requires Dfd protein function for its full activity and therefore is a candidate for a Dfd protein target element. The smallest identified sub-element with autonomous activity in mandibular ganglion cells is 608 bp. Within this sub-element, eight blocks of conserved sequence are shared by both *D.melanogaster* and *D.hydei* and are likely to comprise an important part of the NAE regulatory information since sub-elements from both species are expressed in the subesophageal ganglion and gnathal epidermis.

Though the highest levels of expression directed by the *Dfd*-NAE are in *Dfd* expressing cells of the central nervous system, the element is also active in epidermal cells in the maxillary and mandibular segments and in a few more posterior epidermal cells. Many of the maxillary and mandibular epidermal cells that activate the NAE also express Dfd protein, but those in the labial and more posterior segments do not (20,21). Since the endogenous *Dfd* transcription unit is never activated posterior to the maxillary segment (18,19), the removal of the NAE from its normal context in the large *Dfd* intron has apparently removed it from repressive influences that would prevent its activation in more posterior segments. One possibility we considered is that combining another large enhancer element at *Dfd*, the 2.7 kb *Dfd*-EAE (see Fig. 1), along with the 7.0 kb NAE, might completely restrict the activity of the NAE to *Dfd* expressing cells. However, in transgenic embryos carrying both of these regulatory elements in tandem attached to a *Dfd* promoter/*lacZ* reporter, the NAE/EAE combination still provided ectopic reporter gene expression in a few epidermal cells of the labial and more posterior segments (C. Zeng and W.M, unpublished results).

Somewhat to our surprise, we found no enhancers in our 35 kb regulatory walk that direct a pattern similar to the initial, blastoderm stripe of *Dfd* transcription. *Dfd* transcripts first accumulate in a band six or seven cells wide in cellular blastoderm embryos and the activation of this earliest phase of *Dfd* expression requires the function of the *bicoid*, *hunchback* and *even-skipped* genes (18–20,33). None of the *Dfd* locus fragments that we tested supply this pattern, nor does the combination of the NAE, EAE and *Dfd* promoter attached to a *lacZ* reporter gene (C.Zeng and WM, unpublished results). Perhaps the stripe element for *Dfd* maps outside the region we tested, since it is known that *Dfd* transcription can be influenced by enhancers that map as much as 28 kb upstream (34). It is also possible that the blastoderm stripe element maps in a fragment we tested, but it is not capable of directing expression with the *hsp70* promoter.

The identification of the NAE brings to four the number of known enhancers that are regulated by *Dfd* in the maxillary and mandibular segments. The NAE and EAE contribute to the maintenance of expression of *Dfd* in the CNS and epidermis, respectively, and strong evidence exists that at least one module of the EAE is directly regulated by Dfd protein (13). The other two regulatory elements controlled by Dfd reside in *cis*-acting sequences of the *1.28* and *Distalless* genes (35,36). Though all

require Dfd protein function for their activity in embryos, no two elements supply the same pattern of expression, suggesting a rich combinatorial diversity in the regulatory interactions accomplished by Dfd. Comparison of sequences that are currently available from this diverse group of *Dfd*-regulated elements reveals no extensive blocks of sequence identity larger than the ATCATTAATC consensus that is present in both NAE and EAE (Fig. 5). In the EAE, this sequence comprises a high affinity Dfd protein binding site. The other notable sequence motif in the NAE is a repeat of ATTGAAAT, which is conserved in both the *melanogaster* and *hydei* versions of the element. The functional significance of these sequence motifs in the NAE awaits more detailed genetic characterization of the 608 bp subelement.

Many other homeotic genes exhibit autoregulatory functions in a variety of tissues. Notably, sequences in a 2.3 kb fragment upstream of the P2 promoter of the *Antennapedia* (*Antp*) gene supply reporter gene expression in thoracic regions (parasegments 3–5) of the central nervous system (37,38). This expression pattern in the CNS is repressed in posterior neuromeres by more posteriorly expressed homeotic proteins and within the thoracic ganglia, Antp protein function is required for the activity of the P2 regulatory element (38). It is interesting that the anterior CNS expression domain for one of the *Dfd*-like genes of the mouse, *Hoxb-4*, requires Hoxb-4 protein function (39). Regulatory elements from *Hoxb-4* that provide reporter gene expression in this region of the hindbrain have also been described (40). Since some evidence indicates that homeotic enhancers conserve position-specific regulatory functions whether tested in mouse or *Drosophila* embryos (41,42), more detailed functional dissections of the *Antp*, *Hoxb-4* and *Dfd* regulatory elements may reveal common strategies and factors in their CNS specificity and homeotic regulatory response.

ACKNOWLEDGEMENTS

For grant support we thank the NIH (RO1-HD28315 to WM) and a Physician Scientist Research Award to LJL (K11-HD00976). We are grateful to Virginia Herndon and Spyros Artavanis-Tsakonas for supplying the *D.hydei* genomic library. For much fly handling, cloning and sequencing assistance, we are very grateful to Nadine McGinnis. For advice on the manuscript, we thank Gabriel Gellon.

REFERENCES

1 Akam, M. E. (1987) *Development*, **101**, 1–22.
2 St. Johnston, D. and Nüsslein-Volhard, C. (1992) *Cell*, **68**, 201–219.
3 Lewis, E. B. (1978) *Nature*, **276**, 565–570.
4 Kaufman, T. C., Seeger, M. A. and Olsen, G. (1990) *Adv. Genet.*, **27**, 309–362.
5 McGinnis, W. and Krumlauf, R. (1992) *Cell*, **68**, 283–302.
6 Kenyon, C. (1994) *Cell*, **78**, 175–180.
7 Scott, M.P. (1995) *Cell*, **79** 1121–1124.
8 Dessain, S. and McGinnis, W. (1991) *Current Biol.* **1**, 275–282.
9 Paro, R. (1990) *Trends Genet.*, **6**, 416–421.
10 Kennison, J. A. (1993) *Trends Genet.*, **9**, 75–79.
11 Botas, J. (1993) *Current Opin. Cell Biol.*, **5**, 1015–1022.
12 McGinnis, W., Jack, T., Chadwick, R., Regulski, M., Bergson, C., McGinnis, N. and Kuziora, M. A. (1990) *Adv. Genet.*, **27**, 363–402.
13 Zeng, C.,Pinsonneault, J.,Gellon, G.,McGinnis, N. and McGinnis, W. (1994) *EMBO J.*, **13**, 2362–2377.
14 Kuziora, M. A. and McGinnis, W. (1988) *Cell*, **55**, 477–485.
15 Bergson, C. and McGinnis, W. (1990) *EMBO J.*, **9**, 4287–4297.
16 Merrill, V. K. L.,Turner, F. R. and Kaufman, T. C. (1987) *Dev. Biol.*, **122**, 379–395.
17 Restifo, L. L. and Merrill, V. K. L. (1994) *Dev. Biol.*, **162**, 465–485.
18 Chadwick, R. and McGinnis, W. (1987) *EMBO J.*, **6**, 779–789.
19 Martinez-Arias, A.,Ingham, P. W.,Scott, M. P. and Akam, M. E. (1987) *Development*, **100**, 673–68.
20 Jack, T.,Regulski, M. and McGinnis, W. (1988) *Genes Dev.*, **2**, 635–651
21 Mahaffey, J. W., Diederich, R. and Kaufman, T. C. (1989) *Development*, **105**, 167–174.
22 Diederich, R.J., Pattatucci, A.M. and Kaufman, T.C. (1991) *Development*, **113**, 273–281.
23 Hiromi, Y. and Gehring, W. J. (1987) *Cell*, **50**, 963–974.
24 Scott, M. P.,Weiner, A. J.,Hazelrigg, T. I.,Polisky, B. A.,Pirotta, V.,Scalenghe, F. and Kaufman, T. C. (1983) *Cell*, **35**, 763–776.
25 Regulski, M.,McGinnis, N.,Chadwick, R. and McGinnis, W. (1987) *EMBO J.*, **6**, 767–777.
26 Karess, R. E. and Rubin, G. M. (1984) *Cell*, **38**, 135–146.
27 Rubin, G. M. and Spradling, A. C. (1982) *Science*, **218**, 348–353.
28 Karr, T. L. and Kornberg, T. B. (1989) *Development*, **105**, 95–103.
29 Sanger, F.,Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
30 Campos-Ortega, J. A. and Hartenstein, V. (1985). *The Embryonic Development of Drosophila melanogaster.* Springer-Verlag, Berlin.
31 Beverly, S. M. and Wilson, A. C. (1984) *J. Mol. Evol.*, **21**, 1–13.
32 Regulski, M.,Dessain, S.,McGinnis, N. and McGinnis, W. (1991) *Genes Dev.*, **5**, 278–286.
33 Jack, T. and McGinnis, W. (1990) *EMBO J.*, **9**, 1187–1198.
34 Chouinard, S. and Kaufman, T. C. (1991) *Development*, **113**, 1267–1280.
35 Mahaffey, J. W., Jones, D. F.,Hickel, J. and Griswold, C. M. (1993) *Development*, **118**, 203–214.
36 O'Hara, E.,Cohen, B.,Cohen, S. M. and McGinnis, W. (1993) *Development*, **117**, 847–856
37 Boulet, A. M. and Scott, M. P. (1988) *Genes Dev.*, **2**, 1600–1614.
38 Appel, B. and Sakonju, S. (1993) *EMBO J.*, **12**, 1099–1109.
39 Ramirez-Solis, R.,Zheng, H.,Whiting, J.,Krumlauf, R. and Bradley, A. (1993) *Cell*, **73**, 279–294.
40 Whiting, J., Marshall, H., Cook M., Krumlauf, R., Rigby, P.W.J., Stott, D. and Allemann, R.K. (1991) *Genes Dev.*, **5**, 2048–2059.
41 Awgulewitsch, A. and Jacobs, D. (1992) *Nature*, **358**, 341–344.
42 Malicki, J. C.,L. Peschle, C. and McGinnis, W. (1992) *Nature*, **358**, 345–347.