The EMBO Journal vol.15 no.8 pp.1961–1970, 1996

DEAF-1, a novel protein that binds an essential region in a Deformed response element

Cornelius T.Gross and William McGinnis[1,2]

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114, USA

[1]Present address: Department of Biology, University of California, San Diego, La Jolla, CA 92093-0357, USA

[2]Corresponding author

A 120 bp homeotic response element that is regulated specifically by *Deformed* in *Drosophila* embryos contains a single binding site for Deformed protein. However, a 24 bp sub-element containing this site does not constitute a Deformed response element. Specific activation requires a second region in the 120 bp element, which presumably contains one or more binding sites for Deformed cofactors. We have isolated a novel protein from *Drosophila* nuclear extracts which binds specifically to a site in this second region. This protein, which we call DEAF-1 (Deformed epidermal autoregulatory factor-1), contains three conserved domains. One of these includes a cysteine repeat motif that is similar to a motif found in *Drosophila* Nervy and the human MTG8 proto-oncoprotein, and another matches a region of *Drosophila* Trithorax. Mutations in the response element designed to improve binding to DEAF-1 *in vitro* resulted in increased embryonic expression. Conversely, small mutations designed to diminish binding to DEAF-1 resulted in reduced expression of the element. Thus, DEAF-1 is likely to contribute to the functional activity, and perhaps to the homeotic specificity, of this response element. Consistent with this hypothesis, we have discovered DEAF-1 binding sites in other Deformed response elements.

Keywords: DEAF-1/*Deformed*/DNA binding protein/*Drosophila*/homeotic/*HOXD4*

Introduction

The proteins encoded in the HOM-C–Hox gene complexes are responsible for specifying different cell fates along the developing anterior–posterior body axis of metazoans (McGinnis and Krumlauf, 1992; Kenyon, 1994). They are believed to accomplish this morphogenetic function by the differential transcriptional regulation of a variety of distinct effector genes. However, the DNA binding function in the Hox proteins is contained in highly conserved homeodomains which share up to 93% amino acid identity. Consistent with this, many of these homeodomain proteins have similar DNA binding specificities *in vitro* and *in vivo* (Desplan *et al.*, 1988; Hoey and Levine, 1988; Affolter *et al.*, 1990; Florence *et al.*, 1991; Dessain *et al.*, 1992; Ekker *et al.*, 1994; Walter and Biggin, 1994). With such similar DNA binding specificities, how do these proteins achieve the regulatory specificity necessary to direct different anterior/posterior cell fates during development?

A few models exist in other systems for how homeodomain proteins acquire functional specificity. In yeast, the MATα2 homeodomain protein requires the ubiquitously expressed MCM1 protein to recognize its target with high affinity (reviewed by Johnson, 1992). MCM1 modulates MATα2 target recognition by requiring a particular binding site spacing to achieve cooperative binding interactions (Smith and Johnson, 1992). A related mechanism is used by the human Oct-1 homeodomain protein which will form complexes with the cofactor/activator protein VP16 in the presence of certain octamer DNA binding sites (Herr, 1992; Walker *et al.*, 1994). These models suggest that additional factors may be required to achieve homeodomain regulatory specificity for Hox proteins.

One possible Hox specificity cofactor is encoded by the *extradenticle* (*exd*) gene. Zygotic mutations that lower the levels of *exd* function, which is also supplied maternally, result in head defects and homeotic transformations in the thorax and anterior abdomen of *Drosophila* embryos (Peifer and Wieschaus, 1990). In these embryos, the expression patterns of Hox proteins are unchanged, suggesting that a reduction in Exd protein concentration alters the regulatory specificity of some Hox proteins. Biochemical studies have shown recently that the Exd protein, as well as its mammalian Pbx homologs, binds DNA cooperatively with some Hox proteins (Chan *et al.*, 1994; van Dijk and Murre, 1994; Chang *et al.*, 1995; Johnson *et al.*, 1995; Pöpperl *et al.*, 1995).

Indirect evidence for the existence of additional *Drosophila* homeodomain protein cofactors comes from experiments in which high affinity Engrailed or Fushi-tarazu homeodomain protein binding sites were unable to constitute *engrailed-* or *fushi-tarazu*-dependent response elements in embryos (Vincent *et al.*, 1990; Nelson and Laughon, 1993; Schier and Gehring, 1993). Apparently, oligonucleotides containing homeodomain binding sites and adjacent sequences rarely contain enough information to form specific homeodomain response elements because they lack binding sites for essential cofactors. Bicoid response elements are the only known exception to this generalization, although tandemly repeated Bicoid binding sites require certain spacings for *bicoid*-dependent regulatory activities in embryos (Hanes *et al.*, 1994).

We are using the *Drosophila Deformed* (*Dfd*) gene as a model system to understand the requirements for specific regulation by Hox proteins. Dfd protein is expressed in the maxillary and mandibular segments of the developing embryonic head, and is required for the formation of structures derived from these segments (Chadwick and McGinnis, 1987; Merrill *et al.*, 1987; Regulski *et al.*, 1987). The best characterized Dfd response element is a

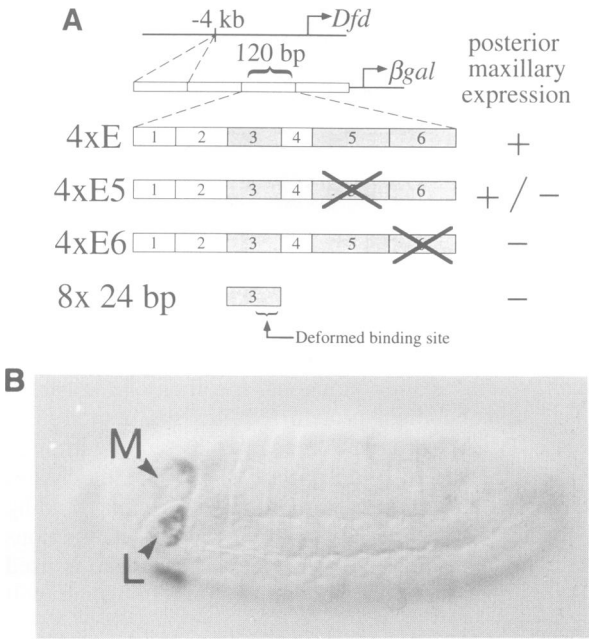

Fig. 1. Three regions of a 120 bp *Dfd* response element are required for its segment-specific activation in embryos. (**A**) A 120 bp fragment (module E) mapping 4 kb upstream of the *Dfd* transcription unit drives β-gal expression in the posterior maxillary segment when multimerized to four copies (Zeng *et al.*, 1994). Three regions of module E (3, 5 and 6, shaded boxes) are required for proper expression in the embryo. A 24 bp oligonucleotide including the Dfd binding site in region 3 is not sufficient to form a Dfd response element when multimerized to eight copies. (**B**) β-gal expression driven by the 8×24 bp construct shown in (A), showing absence of staining in the posterior lateral maxillary segment where module E is expressed and ectopic staining in the labial (L) and dorsal maxillary (M) segments.

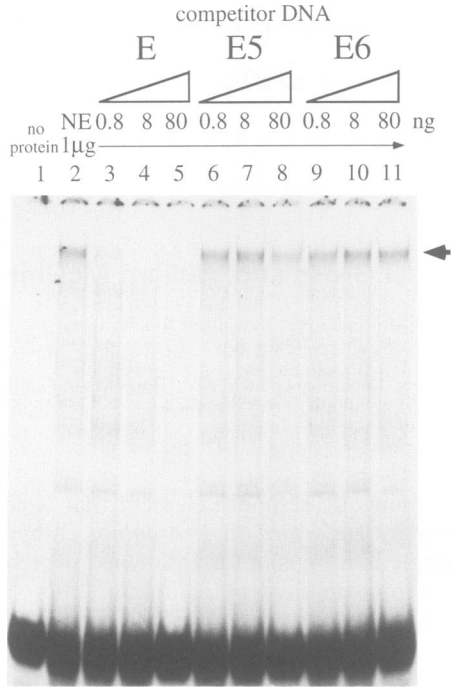

Fig. 2. Gel mobility shift assay of 0–12 h embryonic nuclear extracts with a region 5–6 oligonucleotide. A prominent low mobility complex (arrowhead) appears in the presence of 1 μg of nuclear extract and a 60 bp radiolabeled oligonucleotide-containing sequence from region 5–6 (lane 2; see Materials and methods for probe sequence). This complex is competed successfully by increasing amounts of identical unlabeled oligonucleotide (lanes 3–5). Region 5–6 oligonucleotides containing the same mutations which disrupt the function of module E in embryos are severely compromised in their ability to compete DEAF-1 binding (lanes 6–8 and 9–11). At high concentration, the E5 mutant begins to compete DEAF-1 binding while the E6 mutant fails to compete even at the highest concentrations used.

120 bp fragment of the *Dfd* epidermal autoregulatory element, called module E, which recapitulates *Dfd* expression in the posterior maxillary segment (Zeng *et al.*, 1994). Module E contains a single Dfd binding site that is required for the function of the element in embryos. Module E also contains another 51 bp region, called 5–6, that is required to generate a functional Dfd response element. Mutations in either the 5 or the 6 portions of this region reduce or abolish the activity of the element (Zeng *et al.*, 1994; summarized in Figure 1). Region 5–6, which contains a large inverted repeat sequence, presumably harbors binding sites for protein cofactors which help Dfd protein activate module E. By screening for embryonic proteins that specifically bind region 5–6, we have isolated a novel DNA binding protein which we call DEAF-1 (Deformed epidermal autoregulatory factor-1). We present evidence that DEAF-1 acts through module E in embryos, and that it may function on other Dfd response elements as well.

Results

A Dfd binding site is not sufficient to create a Dfd response element

When multimerized to four copies, module E activates embryonic reporter gene transcription in a *Dfd*-dependent pattern that is restricted to the posterior maxillary epidermis from stage 13 onward (Zeng *et al.*, 1994). However, a multimerized 24 bp oligonucleotide containing the Dfd binding site of module E is activated outside the Dfd expression domain in the labial segment and in a few dorsal maxillary cells (Figure 1). This result, in combination with previous data on the sequences required and sufficient for the function of module E (Zeng *et al.*, 1994; summarized in Figure 1), suggested that region 5–6 contains binding sites for factors that modulate the activity and segmental specificity of this homeotic response element. We undertook to identify proteins in *Drosophila* embryonic nuclear extracts which specifically bound this region.

Embryonic extracts contain a specific binding activity for region 5–6

Crude nuclear extracts from 0–12 h *Drosophila* embryos were incubated with a 56 bp radiolabeled oligonucleotide encompassing regions 5 and 6, and bound DNA complexes were resolved using a gel mobility shift assay (Figure 2). In order to distinguish specific from non-specific DNA binding activities, competition experiments were performed with increasing amounts of unlabeled binding site. Identical competition experiments done with unlabeled mutant binding sites define a prominent DNA binding activity which is dependent on specific sequences in both regions 5 and 6 (Figure 2). We have named this DNA binding activity DEAF-1. The binding site mutations in the E5 and E6 competitors that define the DEAF-1 activity

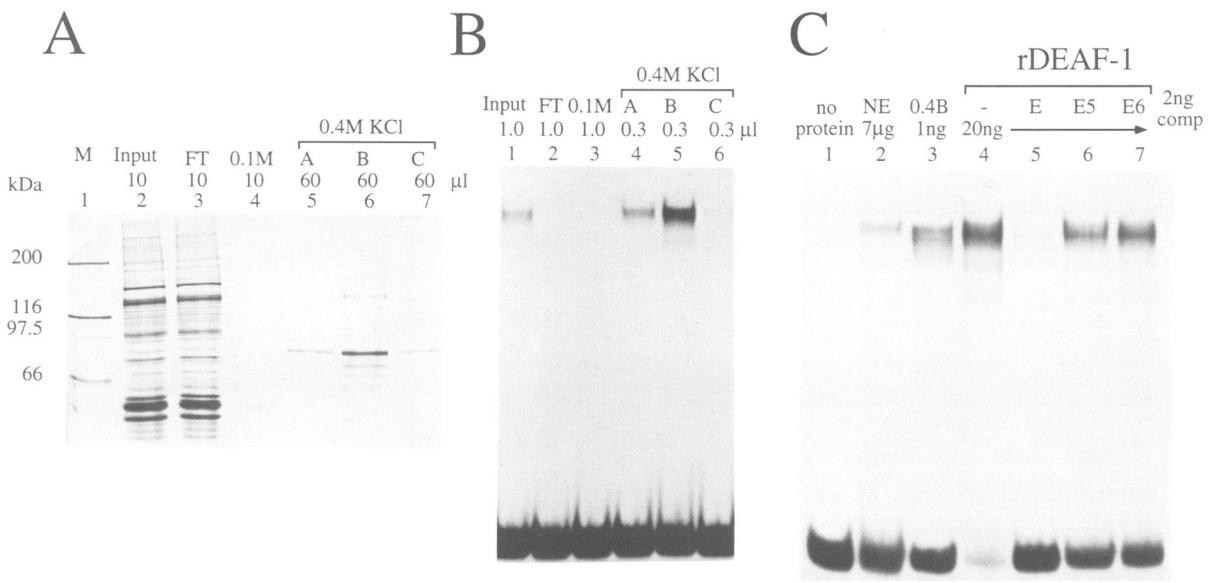

Fig. 3. Purification and reconstitution of the DEAF-1 DNA binding activity. (**A**) Silver-stained SDS protein gel of DNA affinity column fractions from the final step of the purification procedure. A small number of proteins, including a prominent 85 kDa band, are retained on the column and eluted at 0.4 M KCl (lane 6). (**B**) Gel mobility shift assay of the fractions shown in (A). DEAF-1 activity is retained on the column and eluted principally in the second 0.4 M KCl fraction (lane 5). The labeled probe is a 60 bp radiolabeled oligonucleotide-containing sequence from region 5–6 (see Materials and methods). (**C**) Gel mobility shift assay of bacterially expressed putative DEAF-1 protein, showing that it binds the same region 5–6 oligonucleotide, generating a complex of similar mobility to the embryonic DEAF-1 binding activity, and, like the embryonic activity, is competed specifically by wild-type but not region 5 or 6 mutant oligonucleotides. For comparison, the mobility shifts generated by crude embryonic extract (lane 2), and 0.4 M KCl fraction B purified embryonic DEAF-1 (lane 3), are shown.

are identical to those depicted in Figure 1 that inactivate the function of module E. Although other binding complexes may exist which are difficult to detect due to their low abundance or their failure to form stable complexes in mobility shift assays, we chose to pursue DEAF-1 as a candidate Dfd cofactor because of its prominence and its specificity for both regions 5 and 6.

Purification and cloning of the binding activity

Using the gel mobility shift assay, we were able to develop a chromatographic purification procedure for DEAF-1 (see Materials and methods). The final step involved purification over a DNA affinity column made by linking region 5–6 oligonucleotides to a Sepharose CL2B resin (Kadonaga and Tjian, 1986). Fractions from this column contained a prominent polypeptide band migrating at 85 kDa which co-eluted with DEAF-1 DNA binding activity (Figure 3A and B). UV cross-linking experiments demonstrated that a polypeptide migrating at 85 kDa could bind specifically to the region 5–6 site (data not shown), suggesting that the 85 kDa species in Figure 3A was indeed a component of the DEAF-1 activity. Five pmol of the 85 kDa protein were excised from a gel and submitted for peptide sequencing. Degenerate oligonucleotides complementary to two sequenced peptides were used to PCR amplify cDNA fragments from a 8–12 h *Drosophila* plasmid cDNA library (Brown and Kafatos, 1988). Radiolabeled fragments from this amplification were then used to isolate seven related cDNAs from the same library. The four largest (2.4 kb) cDNAs had identical restriction maps, and produced a polypeptide that migrated at 85 kDa after being transcribed and translated *in vitro* (data not shown). Furthermore, the *in vitro* translated protein bound specifically to region 5–6 in gel mobility shift assays, strongly suggesting that the isolated cDNA encoded DEAF-1 DNA binding activity (data not shown).

DEAF-1 protein contains three conserved domains

The DNA sequence of one of the large cDNAs revealed a single long open reading frame (ORF) that would encode a 576 amino acid protein (Figure 4). The sequence of both purified DEAF-1 tryptic peptide fragments matched amino acid sequences in this ORF. To confirm that the 576 amino acid ORF encoded the DEAF-1 binding activity first observed in crude embryonic extracts, the ORF was cloned into a polyhistidine-tagged bacterial expression vector. Full-length protein purified by means of the tag bound specifically to the region 5–6 oligonucleotide, and generated a mobility shift complex similar to that of the embryonic protein (Figure 3C), indicating that the embryonic DEAF-1 DNA binding activity is encoded in the 576 amino acid polypeptide, which we will hereafter refer to as DEAF-1.

The DEAF-1 protein contains three significant regions of sequence similarity to proteins in the NCBI database. In the middle of the protein there is a 113 amino acid region that has 54% identity to a predicted protein encoded by human cDNA fragments of unknown function (Figures 4 and 5). This region also shows more distant similarity to a pair of human nuclear phosphoproteins identified in interferon-induced cell lines (Kadereit *et al.*, 1993) and the putative *Caenorhabditis elegans* protein CEC44F1.2. We call this 113 amino acid region the KDWK domain after the conserved amino acid motif at its core. The KDWK domain also contains a number of conserved, periodically spaced, aliphatic residues and clusters of conserved basic residues (Figure 5). Towards the C-terminus of the protein there is a 32 amino acid region

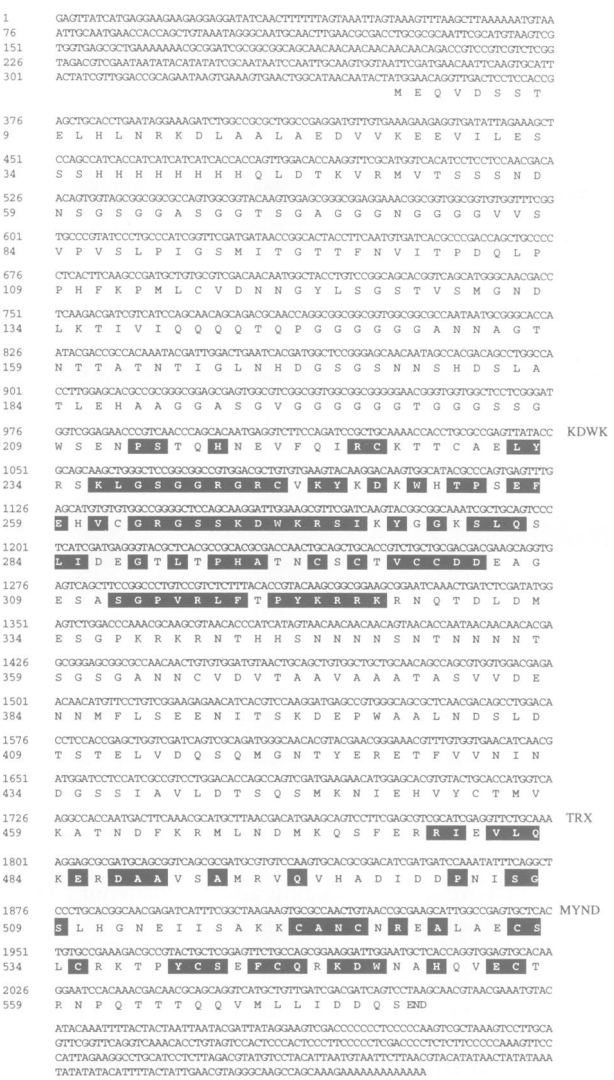

Fig. 4. Nucleotide and predicted amino acid sequence encoded in a DEAF-1 cDNA. Three domains of sequence identity to proteins in the database are denoted by reverse font. These include a 113 amino acid KDWK domain in the middle of the DEAF-1 protein, and two short domains near the C-terminus; TRX, with similarities to Trithorax (Mazo *et al.*, 1990; Stassen *et al.*, 1995), and MYND, with similarities to the Myeloid translocation protein (Miyoshi *et al.*, 1993; Erickson *et al.*, 1994), and Nervy (Feinstein *et al.*, 1995).

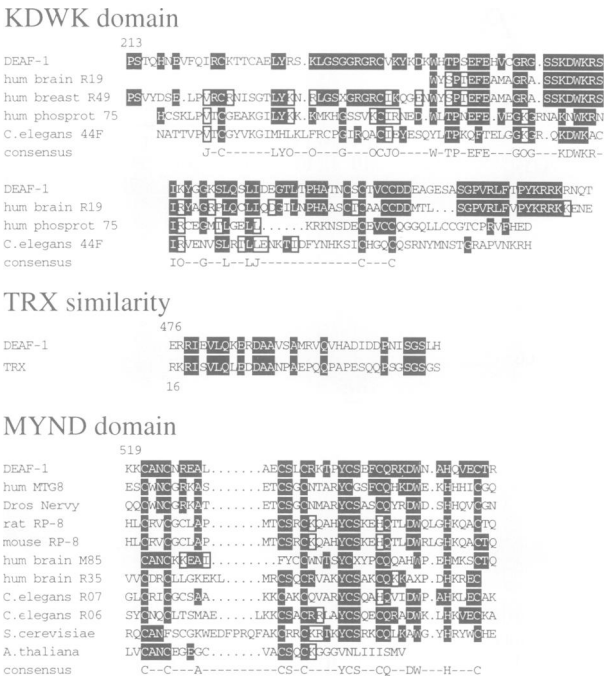

Fig. 5. Alignments of three domains of identity between DEAF-1 and proteins in the database. Reverse font indicates identities and open boxes indicate similarities. Periods show gaps which have been indroduced to maximize sequence identity. The consensus sequence is shown below (O = basic residue, J = aliphatic residue). The 113 amino acid central domain, named KDWK for its core homology, has similarities to a putative human brain protein (dbest R19688), a human breast protein (dbest R49909), a human nuclear phosphoprotein (Kadereit *et al.*, 1993) and *C.elegans* C44F1.2 protein (gp Z49067). The 32 amino acid region of DEAF-1 matches sequences near the N-terminus of the *Drosophila* TRX protein (Mazo *et al.*, 1990; Stassen *et al.*, 1995), at 15 positions. The MYND domain shows the conserved repeated pattern of cysteine and histidine residues that is reminiscent of zinc finger domains and is shared with human MTG8 (Miyoshi *et al.*, 1993), *Drosophila* Nervy (Feinstein *et al.*, 1995), rat, mouse and *C.elegans* RP-8 proteins (Owens *et al.*, 1991; gp U10903; pir S43602), putative *C.elegans* R06F6.4 protein (gp Z46794), *Saccharomyces cerevisiae* 6543.7 protein (gp Z49807), *Arabidopsis thaliana* protein (dbest T45013) and human brain proteins (dbest M85494 and R35199).

with 47% identity to the N-terminus of the *Drosophila* Trithorax protein (Figures 4 and 5; Mazo *et al.*, 1990; Stassen *et al.*, 1995). Finally, at the C-terminus there is a 37 amino acid region which shares a highly conserved cysteine repeat motif with a number of previously identified proteins (Figures 4 and 5). We call this motif the MYND domain after the three best characterized members of this family [Myeloid translocation protein 8 (MTG8) (Miyoshi *et al.*, 1993; Erickson *et al.*, 1994), *Drosophila* Nervy protein (Feinstein *et al.*, 1995) and DEAF-1]. The two C-X-X-C and two C/H-X-X-X-C structures of this domain are similar to the cysteine 'knuckle' described as the basic structural building block of many zinc finger proteins (Schwabe and Klug, 1994).

The gene encoding DEAF-1 maps to the 76C interval on the left arm of chromosome 3 (data not shown). This region is uncovered by the deficiency *Df(3L)kto*^{2} which contains at least two known homeotic regulatory genes, *absent, small or homeotic discs-1* (*ash-1*; Tripoulas *et al.*, 1994) and *kohtalo* (Kennison and Tamkun, 1988). DEAF-1 does not correspond to *ash-1* (Tripoulas *et al.*, 1994; C.T.Gross, unpublished results), and is unlikely to correspond to *kohtalo* as we have not detected changes in the codons of DEAF-1 in three different EMS-induced *kohtalo* mutants (A.Veraksa and W.McGinnis, unpublished results).

Embryonic expression of DEAF-1

Having isolated this response element binding protein, we set out to test for its involvement in *Dfd* autoregulation. One prediction is that the DEAF-1 protein should be expressed in maxillary epidermal cells during embryogenesis. Abundant DEAF-1 transcripts were detected in stage 3 embryos (Figure 6), indicating that DEAF-1 is maternally provided. Using antibodies raised against the bacterially expressed DEAF-1 protein, we observed ubiquitous nuclear staining throughout most of embryogenesis (Figure 6). After stage 15, DEAF-1 protein staining became more prominent in the central nervous system, consistent with the preferential accumulation of DEAF-1 transcripts in the CNS after stage 14 (data not shown).

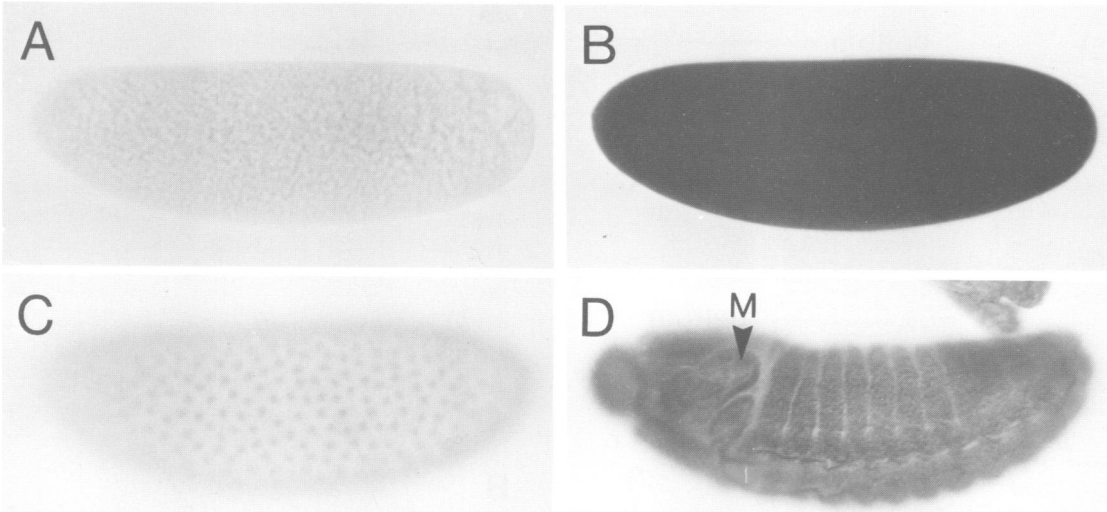

Fig. 6. Localization of DEAF-1 mRNA and protein in embryos. (**A**) *In situ* hybridization with full-length DEAF-1 RNA sense probe of a stage 3 embryo. (**B**) *In situ* hybridization with full-length DEAF-1 RNA antisense probe of a stage 3 embryo showing ubiquitous staining prior to the onset of zygotic transcription. (**C**) Anti-DEAF-1 antiserum staining of a stage 2 embryo showing nuclear localization. (**D**) Anti-DEAF-1 antiserum staining of a stage 14 embryo showing ubiquitous staining, including maxillary segment cells (M) in which module E is activated at this embryonic stage.

This expression pattern is consistent with DEAF-1 being a transcriptional regulator of the module E Dfd response element.

DEAF-1 binding interactions with region 5–6

Another prediction is that different base substitutions that increase or decrease the affinity of region 5–6 for DEAF-1 should result in parallel changes in the activity of these mutant response elements in the context of module E in embryos. In order to design such mutants, we set out to define nucleotides in region 5–6 that contribute to DEAF-1 binding interactions.

Footprinting of module E with recombinant DEAF-1 protein revealed a broad protected site extending the full length of region 5–6 and showing most extensive protection of the right end of region 6 (Figure 7A and B). Methylation interference experiments identified two Gs and a single A in this region whose methylation interfered with DEAF-1 binding, confirming that DEAF-1 makes most intimate contact with this portion of its site (Figure 7A and B). Binding interference by both methylated A and G residues suggests that DEAF-1 interacts with both major and minor groove bases in the core binding site, which contains a TTCG sequence (bottom strand of region 6; Figure 7B). The observation that DEAF-1 binds with higher affinity to region 6 than to region 5 is consistent with the ability of the E6 mutation to interfere most strongly with binding in competition assays (Figures 2 and 3C).

Mutations which decrease or increase DEAF-1 binding

Next, we designed a panel of small mutations in regions 5 and 6 predicted to reduce DEAF-1 *in vitro* binding. We first tested the relative affinity of these mutant binding sites for recombinant DEAF-1 protein *in vitro*, and then tested the function of some mutants in module E Dfd response elements in transgenic embryos. The relative affinity of mutated oligonucleotides was assayed in competition experiments similar to that used to identify DEAF-1 in crude extracts. Mutations with the strongest effect on DEAF-1 binding map to region 6 (Figure 8). E6.3, a 3 bp mutation in the right end of region 6, severely compromised the ability of the region 5–6 oligonucleotide to compete DEAF-1 binding. In region 5, on the other hand, no single mutation had a dramatic effect on DEAF-1 binding, suggesting that DEAF-1 affinity for region 5 sequences depends on multiple independent interactions.

We found that all mutations that reduced DEAF-1 affinity *in vitro* resulted in reduced activity of module E in embryos. For example, the E6.3 element activated moderate levels of expression in the maxillary segment intermediate between the E5 and wild-type module E, in parallel with its intermediate binding affinity *in vitro* (Figure 9). Differences in DEAF-1 affinity, however, did not predict the relatively dramatic reduction in embryonic activity caused by the E5.2 mutation, which had a negligible effect on DEAF-1 binding *in vitro*. The E5.3 and E6.4 elements also showed less embryonic activity than can be explained solely by their reduced *in vitro* affinity for DEAF-1 (Figure 8). If we assume that DEAF-1 is acting through these sequences, then either these mutations affect the function of DEAF-1 in a way that cannot be detected in our *in vitro* binding assays or additional activators require these sequences to regulate module E in embryos.

In order to test further whether DEAF-1 might act through module E in embryos, we took advantage of the partially symmetric nature of region 5–6 to design variants which might function as improved DEAF-1 binding sites. Region 5–6 oligonucleotides were synthesized in which the left side of the large imperfect repeat was mutated to exactly match the right half of the repeat, creating a large uninterrupted symmetric sequence (the E5pal mutant; Figure 10). As a consequence of copying sequences from region 6 into region 5, additional copies of the TTCG core binding motif normally found in region 6 were now repeated in region 5. A similar approach was used to

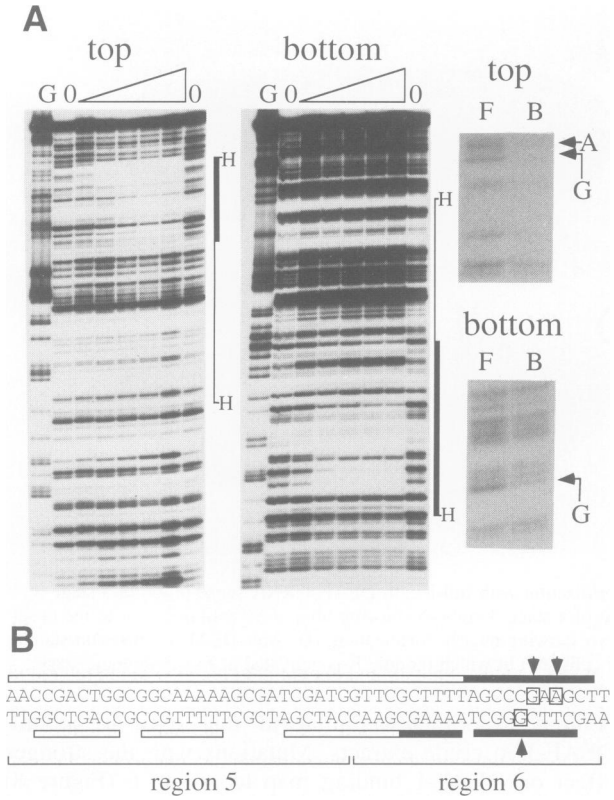

Fig. 7. Footprinting and methylation interference of module E with bacterially produced DEAF-1. (**A**) Increasing amounts of purified recombinant DEAF-1 (3, 6, 12, 24 and 48 ng) were incubated with end-labeled module E DNA as described in Materials and methods. The extent of 120 bp module E sequence is indicated with a bar at right. Region 5–6 within module E is shown as a thick bar. Top and bottom refer to the sequence as published in Zeng *et al.* (1994) (also shown in B). Panels at right show methylation interference of DEAF-1 binding in region 5–6. Lane F, free (unbound) region 5–6; lane B, bound region 5–6. At least two residues (an A and a G) are reduced in the bound lane of the top strand, and a single G residue is reduced in the bottom strand. (**B**) Sequence of 51 bp of region 5–6, summarizing the DNA binding data shown in (A). Footprinting data is indicated with a solid bar (strong binding) and an open bar (weak binding). Methylation interference data is indicated by arrow heads and boxes.

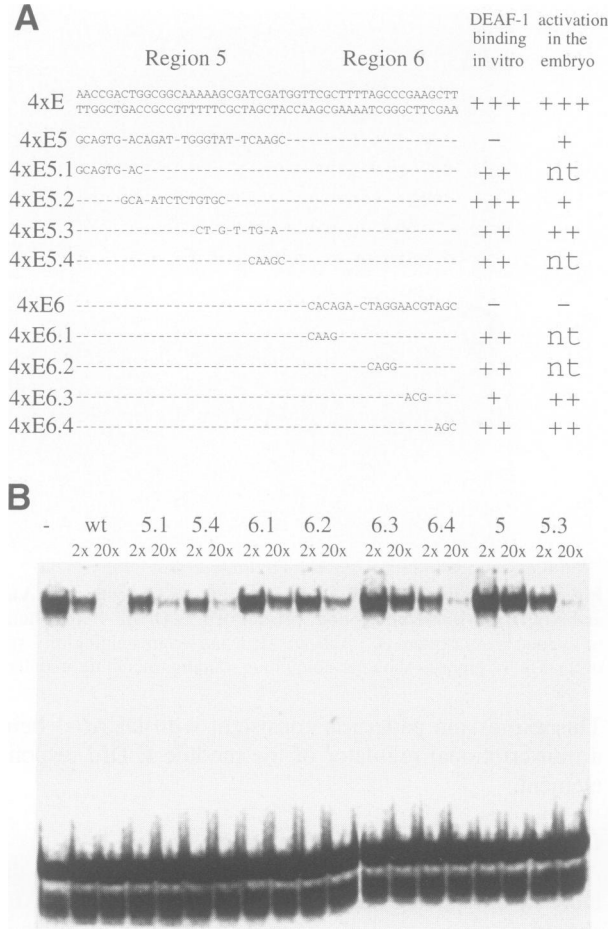

Fig. 8. Small mutations in region 5 and 6 disrupt DEAF-1 binding *in vitro* and compromise activity of module E in the embryo. (**A**) Sequences of the original region 5 (E5) and region 6 (E6) mutations shown in Figure 1 and the sequences of eight smaller substitution mutations. On the right is a summary of the *in vitro* DNA binding data shown in (B) and the *in vivo* expression data shown in Figure 9A. nt indicates the construct was not tested in embryos. (**B**) Gel mobility shift competition assay, similar to that shown in Figure 2, used to compare DEAF-1 binding affinity among oligonucleotides containing the mutations shown in (A). Binding by 1 µl of MonoQ-purified embryonic DEAF-1 to a 60 bp oligonucleotide containing region 5–6 was competed with a 2- or 20-fold molar excess of the indicated wild-type or mutant unlabeled oligonucleotides. Binding data for E6 and E5.2 are not shown.

perfect a smaller inverted repeat within region 6, which also had the consequence of duplicating the core binding motif normally found in the right end of region 6 (the E6L mutant; Figure 10).

To test relative binding affinities of the mutant elements, we measured the dissociation rates of DEAF-1 bound to region 5–6 oligonucleotides containing these mutations (Figure 10). Dissociation rate measurements, which do not depend on the absolute activity of DNA or protein, were used to minimize complications due to improper folding of these symmetric oligonucleotides and the possible introduction of new DEAF-1 binding sites. We found that DEAF-1 forms E5pal binding complexes that are 2-fold more stable, and E6L complexes that are 1.5-fold more stable, than wild-type (Figure 10).

When tested in embryos, both E5pal and E6L showed increased expression relative to wild-type module E. In these experiments, the mutants and control were tested as single copy regulatory constructs in order to make the *in vivo* assay more sensitive to improvements in activity (Zeng *et al.*, 1994). The good correlation between DEAF-1 binding affinity and *in vivo* activity for these mutants argues that the function of this element is dependent on its ability to bind DEAF-1 or possibly a related protein with very similar binding characteristics.

DEAF-1 binding to other Dfd response elements

To address the question of whether DEAF-1 may be required more generally for Dfd function, we asked whether two other known Dfd response elements bind DEAF-1. Zeng *et al.* (1994) described module F, a 460 bp *Dfd* epidermal autoregulatory element lying just downstream of module E. DEAF-1 footprinted four regions within module F. One of these binding regions is within a 77 bp fragment that deletion studies showed was crucial to the activity of module F, and that was itself sufficient to serve as a weak Dfd response element in embryos (Figure 11A; Zeng *et al.*, 1994). This 77 bp region also

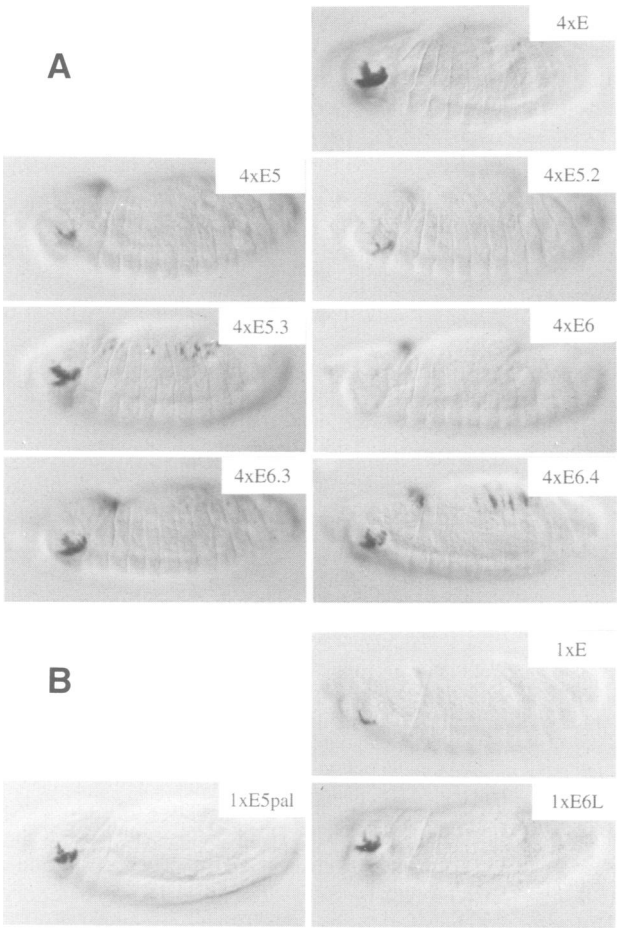

Fig. 9. β-gal staining of wild-type and mutant module E constructs. (**A**) Small DEAF-1 binding site mutations in region 5 and 6 (shown in Figure 8) result in less module E activity in the maxillary segment at stage 14. (**B**) Improved DEAF-1 binding variants of module E (Figure 10) exhibit increased module E activity in the maxillary segment at stage 14. In these latter embryos, module E variants were inserted as a single rather than four copy multimers in reporter constructs to provide increased sensitivity to changes in regulatory activity.

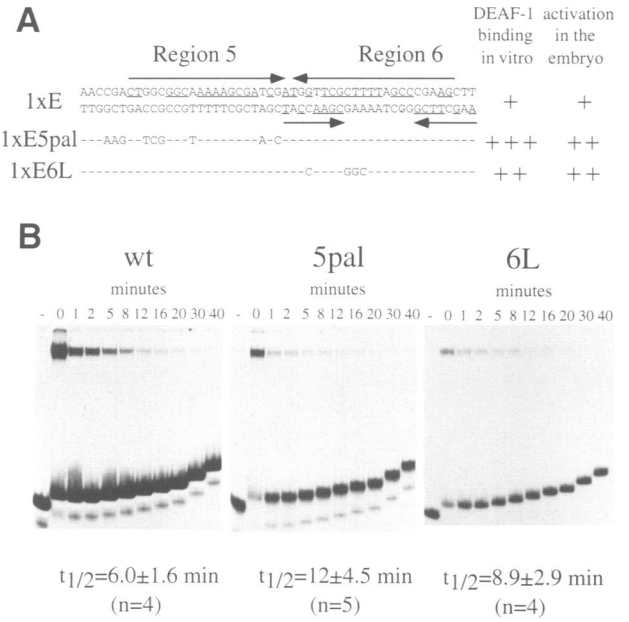

Fig. 10. Mutations which improve the imperfect inverted repeats found in region 5 and 6 lead to binding sites with increased affinity for DEAF-1 *in vitro* and increased activity in the embryo. (**A**) Sequences of wild-type module E and the symmetric mutants E5pal and E6L. Arrows and underlines indicate the imperfect inverted repeats used to design E5pal and E6L. On the right is a summary of the *in vitro* DEAF-1 binding data shown in (B) and the *in vivo* data shown in Figure 9B. (**B**) Representative gel mobility shift off-rate experiments used to measure differences in DEAF-1 binding affinity among oligonucleotides containing the mutations shown in (A). Dissociation rate constants are indicated as $t_{1/2}$ below each panel, along with the standard deviation and number of independent experiments.

contains a single Dfd binding site and thus may function in a manner similar to module E.

We have also footprinted a 0.5 kb regulatory element of the human *Dfd* homolog HOXD4 which drives hindbrain-specific expression in transgenic mice and functions as a maxillary-specific regulatory element in *Drosophila* embryos (Malicki *et al.*, 1992). Seven regions within this element were footprinted by DEAF-1, suggesting that DEAF-1 might contribute to its function in *Drosophila* embryos. In addition, we discovered a number of DEAF-1 footprint sites in a 0.6 kb fragment of the *1.28* gene that has been shown to serve as a maxillary-specific Dfd response element in the embryo (Mahaffey *et al.*, 1993; Mohler *et al.*, 1995; J.Pederson, C.Gross, W.McGinnis and J.Mahaffey, unpublished results). Comparison of the high affinity binding regions in all these elements revealed two common properties: conservation of at least one TTCG motif, and the nearby presence of one or more additional TCG motifs (Figure 11). We have not yet recognized any pattern in the spacing and orientation of the conserved TCG repeats.

Discussion

We have identified and purified a novel protein from *Drosophila* nuclear extracts by virtue of its binding specifically to an essential sequence in a 120 bp Dfd response element. The sequence of the gene encoding this DNA binding activity, which we call DEAF-1, reveals that it encodes a protein with extensive regions of structural similarity to previously studied *Drosophila* and mammalian proteins of unknown biochemical function (Mazo *et al.*, 1990; Kadereit *et al.*, 1993; Miyoshi *et al.*, 1993; Feinstein *et al.*, 1995; Stassen *et al.*, 1995), suggesting that some of these might also encode sequence-specific DNA binding functions.

A number of lines of evidence support the hypothesis that DEAF-1 functions directly via module E in embryos. First, mutations that interfere with DEAF-1 binding *in vitro* compromise the function of module E in the embryo (Figure 8). Both the large E5 and E6 mutations strongly reduce module E activity, with E6 mutations having more severe effects. This correlates with the observation that E6 mutations interfere more with DEAF-1 *in vitro* binding when compared with E5 mutations (Figures 2 and 7). More importantly, the small E6.3 mutation in the core of the DEAF-1 results in a module E element that has both reduced affinity for DEAF-1, and reduced maxillary activity in embryos (Figures 8 and 9).

Further evidence that DEAF-1 acts directly on module E in embryos comes from the increased expression of the E5pal and E6L constructs containing improved DEAF-1

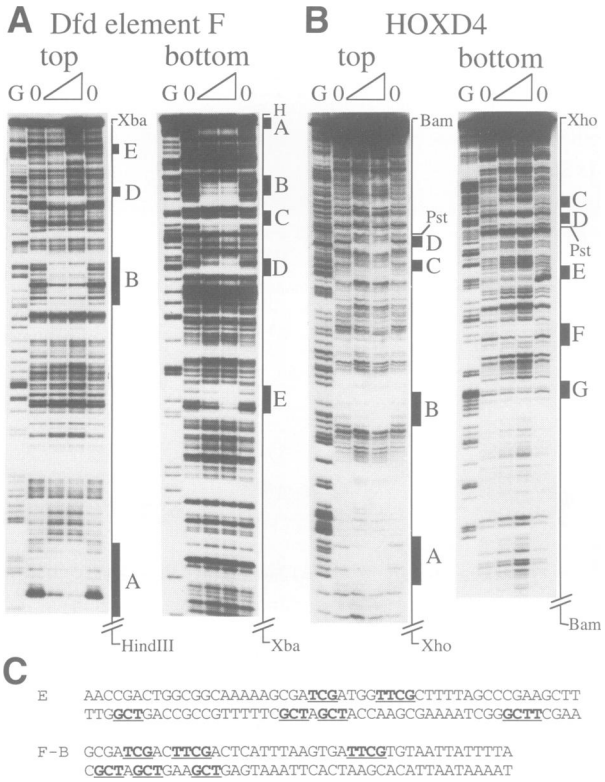

Fig. 11. Footprinting of other Dfd response elements with DEAF-1. Increasing amounts of bacterial DEAF-1 (12 and 48 ng) were incubated with end-labeled DNA as described in Materials and methods. Bars indicate protected regions in the sequence. Strong sites are those which show equivalent protection at both protein concentrations. (**A**) Footprinting of the *Dfd* epidermal autoregulatory module F, a 460 bp *Hin*dIII–*Xba*I fragment (see Zeng *et al.*, 1994). (**B**) Footprinting of the human HOXD4 hindbrain-specific regulatory element, a 0.5 kb *Bam*HI–*Xho*I fragment which functions as a maxillary-specific Dfd response element in *Drosophila* embryos (Malicki *et al.*, 1990). (**C**) Comparison of the high affinity DEAF-1 binding site in region 6 of module E, and the DEAF-1 binding site B in module F, which is contained within a 77 bp minimal Dfd response element (Zeng *et al.*, 1994). Both high affinity binding sites contain at least one TTCG core binding motif surrounded by additional TCG sequences.

binding sites (Figure 10). In these mutants, improved binding and embryonic activity is associated with the improved symmetry of sequences in region 5 and 6. However, the only common structural feature of both mutants is a duplication of the high affinity TTCG binding sequences found at the right end of region 6, suggesting that this explains their enhanced function (Figure 10). As anticipated, the complementary symmetric mutations, E6pal and E6R, in which the right side of the imperfect repeats were mutated to match the left, show reduced DEAF-1 binding *in vitro* (data not shown), and for E6pal, dramatically reduced expression in embryos (C.Zeng, unpublished data). Viewed together, these data suggest that DEAF-1 or a protein with very similar binding characteristics interacts directly with module E in embryos.

There may be other factors influencing module E activity in region 5–6. The E5.3 mutation results in a moderate reduction of activity in embryos, although its DEAF-1 binding affinity is only slightly reduced. The E5.2 and E6.4 mutations also have similar properties, presumably signaling the presence of other binding sites required for the full activity of region 5–6. Such additional input to module E would not be surprising, since a considerable amount of DNA sequence conservation exists between the *D.melanogaster* and *D.hydei* versions of module E that is not explicable by conservation of Dfd and DEAF-1 binding sites alone (Zeng *et al.*, 1994).

It has been proposed that the Cys-XX-Cys motif forms a 'knuckle' structure that is the basic building block of many zinc binding proteins (Schwabe and Klug, 1994). Many of these zinc binding motifs serve as DNA binding domains. However, we have found that a truncated form of the DEAF-1 protein containing just the C-terminal 84 amino acids comprising the MYND domain fails to bind DNA specifically in either a gel shift or footprint assay. Furthermore, a full-length DEAF-1 protein harboring Cys to Ser mutations in both the second and third cysteines of the MYND domain binds indistinguishably from the wild-type protein in a gel shift assay (C.Gross, unpublished results). While it is possible that the MYND domain contributes subtly to the DEAF-1 DNA binding function, we favor the hypothesis that this domain serves a non-DNA binding function. It is clear from studies of other proteins that such Cys/His repeat motifs can serve diverse functions, including forming active sites, stabilizing the tertiary structure of domains and possibly forming protein–protein interaction surfaces (Schwabe and Klug, 1994).

We find that all DEAF-1 footprinted sites contain TCG, and high affinity sites contain TTCG plus additional nearby TCG sequences (Figure 11). The module E, module F-B and *1.28* element binding sites fall into this category. The effect of bases flanking the TTCG core on DEAF-1 binding appears to be modest. The 1.5-fold increase in binding of E6L, where four bases flanking the TTCG motif in region 6 are changed, suggests that DEAF-1 discriminates subtly between bases at these positions (Figure 10). This trend is supported by the converse mutation, E6R, which shows a 1.5-fold decrease in DEAF-1 binding (data not shown). Considered together with the evidence for multiple independent interactions with region 5, it appears that DEAF-1 presents a large DNA binding surface, possibly formed by a DEAF-1 multimer, which achieves highest binding affinity when binding to sites containing multiple TCG sequences, but which can bind a single TCG-containing site with moderate affinity.

Is DEAF-1 involved in the selective action of Dfd on the module E regulatory element? If so, it apparently does so without binding directly to Dfd protein. In numerous attempts, we have failed to detect cooperative binding interactions between DEAF-1 and Dfd using the gel mobility shift assay and either crude or purified embryonic DEAF-1 protein (data not shown). Thus we postulate additional factor(s) which help DEAF-1 specifically recognize Dfd and, together with Dfd and DEAF-1, form an activating transcription complex.

The DEAF-1 protein expression pattern strongly suggests that DEAF-1 functions outside the maxillary segment, perhaps together with other homeotics. The DEAF-1 expression pattern is reminiscent of a number of trithorax group genes, including Brahma and Trithorax, which are also widely expressed during early embryogenesis and expressed at abundant levels in the late embryonic CNS

(Elfring *et al.*, 1994; Sedkov *et al.*, 1994). The trithorax group genes are required for the regulation and/or function of the *Drosophila* Hox genes (Kennison, 1993). The slight sequence similarity between DEAF-1 and Trithorax proteins (Figure 4), coupled with the evidence for DEAF-1 as an activator of *Dfd* expression, suggests that DEAF-1 may function within the ill-defined realm of the trithorax group activators. Our present model is that DEAF-1 binds directly to module E in the embryo and synergistically activates Dfd protein in combination with other, as yet unidentified, factors. The presence of DEAF-1 binding sites in a number of Dfd response elements suggests that a requirement for DEAF-1 activity may be a common feature of enhancers targeted by Dfd.

Materials and methods

Gel mobility shift assay

Drosophila embryonic nuclear extracts were prepared as described in Biggin and Tjian (1988). Gel mobility shift assays were performed as described in Dessain *et al.* (1992), with the following modifications. Herring testis competitor DNA was used at 10 μg/ml with crude extracts, and 1 μg/ml with purified protein, and NP-40 was added to 0.1%. The top strand of the 60 bp region 5–6 oligonucleotide used in gel mobility shift and competition assays is: 5′agctAACCGACTGGCGGCAAAAAGCGATCGATGGTTCGCTTTTAGCCCGAAGCTTcagct3′ (sequences not from region 5–6 are shown in lower case). In some cases, binding oligonucleotides lacked the leftmost four nucleotides, which were shown not to affect DEAF-1 binding affinity.

Protein purification and cloning

All purification steps were carried out at 4°C, and all steps except the DNA affinity column were executed using a Pharmacia FPLC apparatus. All buffers were based on HEMG [50 mM HEPES, 0.1 mM EDTA, 12 mM $MgCl_2$, 10% glycerol, 1 mM dithiothreitol (DTT), supplemented from the MonoQ step onward with 0.1% NP-40]. Crude nuclear extract (35 ml, ~15 mg/ml) was applied to a 30 ml heparin–Sepharose column and DEAF-1 activity was eluted with 0.4 M KCl. The eluate (18 ml) was passed through an 800 ml S300 HR sizing column (Pharmacia) and DEAF-1 activity appeared just following the void fractions. Pooled S300 fractions (35 ml) were further purified on an 8 ml MonoQ column, with peak DEAF-1 activity eluting near 0.18 M KCl. Final purification was achieved using a specific DNA affinity column (Kadonaga and Tjian, 1986) made using multimerized 56 bp oligonucleotides spanning region 5–6. Pooled MonoQ fractions diluted to 0.1 M KCl (10 ml total) and containing 33 μg/ml herring testis competitor DNA were applied to a 1 ml affinity column (35 μg DNA/ml resin) and DEAF-1 activity was eluted with 0.4 M KCl. Overall purification of DEAF-1 activity was ~6000-fold, with DEAF-1 protein estimated to be 0.02% of total protein in crude extract.

For peptide sequencing, the MonoQ step was skipped and S300 fractions were passed directly onto the affinity column (150 μg/ml herring testis competitor DNA). Gel slices containing 50 pmol of DEAF-1 were submitted to the Keck Foundation Biotechnology Resource Laboratory protein sequencing facility, where the sequences of two tryptic peptides were determined: Nterm, 252 WHTPSEFEHVCGR 264; Cterm, 495 VQVHADIDDPNISGSLHGNEIISAK 519. A set of complementary inosine-containing degenerate primers was made to each of these peptides and used to amplify DNA fragments from a *Drosophila* plasmid cDNA library (Brown and Kafatos, 1988). The single 0.8 kb product was radiolabeled and used to isolate eight cDNAs from the same library. Seven of these cDNAs were related and four of these seven appeared identical, containing inserts of 2.4 kb. A single cDNA was sequenced in both directions and has been submitted to EMBL GenBank under accession No. U46686.

To make recombinant DEAF-1 protein in bacteria, the complete ORF was subcloned into the vector pET15b kan^R (Novagen). Following sonication, the insoluble pellet was dissolved in 4 M guanidine–HCl, 20 mM Tris pH 7.9, 0.5 M NaCl, 10% glycerol, 0.1% NP-40, 10 μM $ZnCl_2$, 1 mM DTT, and dialyzed successively against 2 M, 1 M and 0 M guanidine–HCl. The soluble fraction retained ~50% of the total DEAF-1 protein and subsequently was purified over a nickel resin (Qiagen) and eluted with 1 M imidazole as described by the manufacturer. Final purification was achieved by passing this protein over the region 5–6-specific DNA affinity resin and eluting with 0.4 M KCl. The protein concentration was 12 μg/ml, as judged by Coomassie staining of an SDS gel.

For the production of guinea pig polyclonal antibodies, 0.6 mg of crude insoluble protein was separated on an SDS gel, eluted into SDS running buffer, and submitted to Pocono Rabbit Farms, Inc.

Footprinting and methylation interference

Binding reactions were carried out as for the gel mobility shift assay above in 25 μl. After 10 min binding, 25 μl of 10 mM $MgCl_2$, 5 mM $CaCl_2$ was added, followed by 3 μl of a 1/16 000 dilution of DNase (BM #776 785). Digestion was stopped after 1 min with an equal volume of stop solution (0.2 M NaCl, 20 mM EDTA, 1% SDS, 0.25 mg/ml tRNA), extracted with phenol–chloroform, and ethanol precipitated before separation by 6% PAGE.

Methylation interference was carried out following the protocol of Hatfull and Grindley (1986), with the following modifications. Dimethylsulfate-modified and twice ethanol precipitated DNA (~500 000 c.p.m.) was incubated with purified recombinant DEAF-1 protein and run on a gel as described for the gel shift mobility assay above. Bound and unbound DNA were excised and passively eluted into Q elution buffer (50 mM MOPS pH 7.0, 100 mM NaCl), followed by purification on Qiagen tip-5 columns and precipitation in the presence of 10 μg glycogen (BM). DNA was resuspended in 100 μl of 0.1 M NaOH, incubated at 90°C for 30 min and ethanol precipitated before separation by 6% PAGE.

Off-rate measurements

Gel mobility shift complexes were formed with radiolabeled DNAs and partially purified embryonic DEAF-1 (prepared as for peptide sequencing, see above). After 10 min binding, a 100-fold molar excess of unlabeled wild-type region 5–6 DNA was added to the reaction. At various times after the addition of competitor, small aliquots were loaded on a running gel (as described for the gel mobility shift assay above). The fraction of probe bound in each lane was quantitated using a Fuji BAS2000 imaging system, and off-rates were estimated by fitting plots of the ln(fraction bound) versus time curves with straight lines.

Mutant regulatory elements

Full-length 120 bp module E constructs were pieced together from 12 synthetic oligonucleotides covering regions 1–6, and cloned into the injection vector Casper CZIII β-gal as described in Zeng *et al.* (1994). Multiple strains carrying regulatory constructs inserted at different genomic locations were tested for each regulatory construct variant: 4×E (3), 4×E5 (3), 4×E6 (3), 4×E5.2 (3), 4×E5.3 (3), 4×E6.3 (2), 4×E6.4 (5), 1×E (5), 1×E5pal (5), 1×E6L (5). The oligonucleotides used to create the 8×24 bp construct were 5′-TCGAGAGCTAATGCGTGGCAATTAGGGG and 3′-CTCGATTACGCACCGTTAATCCCCAGCT.

Embryo staining

For RNA *in situ* staining, sense and antisense digoxigenin-labeled RNA probes were produced from the original cDNA library plasmid constructs. Staining was executed according to Tautz and Pfeifle (1989). Antibody staining for Dfd, β-gal and DEAF-1 were carried out essentially as described in Zeng *et al.* (1994), using affinity-purified guinea pig Dfd polyclonal antibody, mouse monoclonal β-gal antibody (Promega) and guinea pig polyclonal DEAF-1 antiserum (see above).

Acknowledgements

We thank Rick Austin for his patient expertise and for allowing us to test numerous fractions from his partially purified *Drosophila* embryonic extracts for DEAF-1 activity. We are very grateful to Chaoyang Zeng for allowing the use of many regulatory constructs. We thank Alyssa TenHarmsel for donating the nuclear extract used in Figure 1, and Mark Biggin for generous advice throughout the project and equally generous access to his FPLC, fly cages and other equipment. Jim Kennison helped enormously with unpublished mutations and useful advice, and Brian Florence provided enlightening criticism of the manuscript. We are also indebted to Kathy Stone at the Keck Foundation and Carl Wu, Peter Harte and Leslie Mounkes for providing antibodies used during the purification procedure. This work was supported by a grant from NIH (HD28315) to W.M. C.G. was partially supported by an NSF Graduate Fellowship.

References

Affolter,M., Percival-Smith,A., Müller,M., Leupin,W. and Gehring,W.J. (1990) DNA binding properties of the purified antennapedia homeodomain. *Proc. Natl Acad. Sci. USA*, **87**, 4093–4097.

Biggin,M.D. and Tjian,R. (1988) Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. *Cell*, **53**, 699–711.

Brown,N.H. and Kafatos,F.C. (1988) Functional cDNA libraries from *Drosophila* embryos. *J. Mol. Biol.*, **203**, 425–437.

Chadwick,R. and McGinnis,W. (1987) Temporal and spatial distribution of transcripts from the *Deformed* gene of *Drosophila*. *EMBO J.*, **6**, 779–789.

Chan,S.-K., Jaffe,L., Capovilla,M., Botas,J. and Mann,R. (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.L. (1995) Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev.*, **9**, 663–674.

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

Dessain,S., Gross,C.T., Kuziora,M. and McGinnis,W. (1992) Antp-type homeodomains have distinct DNA binding specificities that correlate with their target specificities in the embryo. *EMBO J.*, **11**, 991–1002.

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

Elfring,L., Deuring,R., McCallum,C., Peterson,C. and Tamkun,J.W. (1994) Identification and characterization of *Drosophila* relatives of the yeast transcriptional activator SNF2/SW12. *Mol. Cell. Biol.*, **14**, 2225–2234.

Erickson,P.F., Robinson,M., Owens,G. and Drabkin,H.A. (1994) The ETO portion of acute myeloid leukemia t(8;21) fusion transcript encodes a highly evolutionarily conserved, putative transcription factor. *Cancer Res.*, **54**, 1782–1786.

Feinstein,P.G., Kornfeld,K., Hogness,D.S. and Mann,R.S. (1995) Identification of homeotic target genes in *Drosophila melanogaster* including *nervy*, a proto-oncogene homologue. *Genetics*, **140**, 573–586.

Florence,B., Handrow,R. and Laughon,A. (1991) DNA-binding specificity of the fushi tarazu homeodomain. *Mol. Cell. Biol.*, **11**, 3613–3623.

Hanes,S.D., Riddihough,G., Ish-Horowicz,D. and Brent,R. (1994) Specific DNA recognition and intersite spacing are critical for action of the bicoid morphogen. *Mol. Cell. Biol.*, **14**, 3364–3375.

Hatfull,G. and Grindley,N. (1986) Analysis of γδ resolvase mutants *in vitro*: evidence for an interaction between serine-10 of resolvase and site I of *res*. *Proc. Natl Acad. Sci. USA*, **83**, 5429–5433.

Herr,W. (1992) Oct-1 and Oct-2: differential transcriptional regulation by proteins that bind to the same DNA sequence. In McKnight,S. and Yamamoto,K. (eds), *Transcriptional Regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 1103–1135.

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

Johnson,A. (1992) A combinatorial regulatory circuit in budding yeast. In McKnight,S. and Yamamoto,K. (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.

Kadereit,S., Gewert,D.R., Galabru,J., Hovanessian,A.G. and Meurs,E.F. (1993) Molecular cloning of two new interferon-induced, highly related nuclear phosphoproteins. *J. Biol. Chem.*, **268**, 24432–24441.

Kadonaga,J.T. and Tjian,R. (1986) Affinity purification of sequence-specific DNA binding proteins. *Proc. Natl Acad. Sci. USA*, **83**, 5889–5893.

Kennison,J.A. (1993) Transcriptional activation of *Drosophila* homeotic genes from distant regulatory elements. *Trends Genet.*, **9**, 75–79.

Kennison,J.A. and Tamkun,J.W. (1988) Dosage-dependent modifiers of *Polycomb* and *Antennapedia* mutations in *Drosophila*. *Proc. Natl Acad. Sci. USA*, **85**, 8136–8140.

Kenyon,C. (1994) If birds fly, why can't we? Homeotic genes and evolution. *Cell*, **78**, 175–180.

Mahaffey,J.W., Jones,D.F., Hickel,J. and Griswold,C.M. (1993) Identification and characterization of a gene activated by the deformed homeoprotein. *Development*, **118**, 203–214.

Malicki,J., Cianetti,L.C., Peschle,C. and McGinnis,W. (1992) A human *HOX4B* regulatory element provides head-specific expression in *Drosophila* embryos. *Nature*, **358**, 345–347.

Mazo,A.M., Huang,D., Mozer,B.A. and Dawid,I.B. (1990) The trithorax gene, a trans-acting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains. *Proc. Natl. Acad. Sci. USA*, **87**, 2112–2116.

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

Merrill,V.K.L., Turner,F.R. and Kaufman,T.C. (1987) A genetic and developmental analysis of mutations in the *Deformed* locus in *Drosophila melanogaster*. *Dev. Biol.*, **122**, 379–395.

Miyoshi,H., Kozu,T., Shimizu,K., Enomoto,K., Maseki,N., Kaneko,Y., Kamada,N. and Ohki,M. (1993) The t(8;21) translocation in acute myeloid leukemia results in production of an AML1–MTG8 fusion transcript. *EMBO J.*, **12**, 2715–2721.

Mohler,J., Mahaffey,J.W., Deutsch,E. and Vani,K. (1995) Control of *Drosophila* head segment identity by the bZIP homeotic gene *cnc*. *Development*, **121**, 237–247.

Nelson,H. and Laughon,A. (1993) *Drosophila* glial architecture and development: analysis using a collection of new cell-specific markers. *Devel. Biol.*, **202**, 341–354.

Owens,G.P., Hahn,W.E. and Cohen,J.J. (1991) Identification of mRNAs associated with programmed cell death in immature thymocytes. *Mol. Cell. Biol.*, **11**, 4177–4188.

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

Pöpperl,H., Bienz,M., Studer,M., Chan,S., 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 on exd/pbx, *Cell*, **81**, 1031–1042.

Regulski,M., McGinnis,N., Chadwick,R. and McGinnis,W. (1987) Developmental and molecular analysis of *Deformed*: a homeotic gene controlling *Drosophila* head development. *EMBO J.*, **6**, 767–777.

Schier,A.F. and Gehring,W.J. (1993) Analysis of a *fushi tarazu* autoregulatory element: multiple sequence elements contribute to enhancer activity. *EMBO J.*, **12**, 1111–1119.

Schwabe,J.W.R. and Klug,A. (1994) Zinc mining for protein domains. *Struct. Biol.*, **1**, 345–349.

Sedkov,Y., Tillib,S., Mizrokhi,L. and Mazo,A. (1994) The bithorax complex is regulated by *trithorax* earlier during *Drosophila* embryogenesis than is the Antennapedia complex, correlating with a bithorax-like expression pattern of early *trithorax* transcripts. *Development*, **120**, 1907–1917.

Smith,D.L. and Johnson,A.D. (1992) A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of the homeodomains of an α2 dimer. *Cell*, **68**, 133–142.

Stassen,M.J., Bailey,D., Fox,S., Chinwalla,V. and Harte,P.J. (1995) The *Drosophila* trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech. Dev.*, **52**, 1–15.

Tautz,D. and Pfeifle,C. (1989) A non-radioactive *in situ* hybridization method for localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma*, **98**, 81–85.

Tripoulas,N.A., Hersperger,E., La Jeunesse,D. and Shearn,A. (1994) Molecular genetic analysis of the *Drosophila melanogaster* gene *absent, small or homeotic discs1* (*ash1*). *Genetics*, **137**, 1027–1038.

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

Vincent,J.P., Kassis,J.A. and O'Farrell,P.H. (1990) A synthetic homeodomain binding site acts as a cell type specific, promoter specific enhancer in *Drosophila* embryos. *EMBO J.*, **9**, 2573–2578.

Walker,S., Hayes,S. and O'Hare,P. (1994) Site-specific conformational alteration of the Oct-1 POU domain–DNA complex as the basis for differential recognition by Vmw65(VP16). *Cell*, **79**, 841–852.

Walter,J. and Biggin,M.D. (1994) Two homeodomain proteins bind with similar specificity to a wide range of DNA sites in *Drosophila* embryos. *Genes Dev.*, **8**, 1678–1692.

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 6, 1995; revised on December 12, 1995