Antp-type homeodomains have distinct DNA binding specificities that correlate with their different regulatory functions in embryos

Scott Dessain¹, Cornelius T. Gross²,
Michael A. Kuziora^{2,3} and William McGinnis^{1,2}

Departments of ¹Biology and ²Molecular Biophysics and Biochemistry
Yale University, New Haven, CT 06511, USA

³Present address: Department of Biological Sciences, University of
Pittsburgh, Pittsburgh, PA 15260, USA

Communicated by W. Gehring

Much of the functional specificity of *Drosophila* homeotic selector proteins, in their ability to regulate specific genes and to assign specific segmental identities, appears to map within their different, but closely related homeodomains. For example, the *Drosophila* Dfd and human HOX4B (Hox 4.2) proteins, which have extensive structural similarity only in their respective homeodomains, both specifically activate the *Dfd* promoter. In contrast, a chimeric Dfd protein containing the Ubx homeodomain (Dfd/Ubx) specifically activates the *Antp* P1 promoter, which is normally targeted by Ubx. Using a variety of DNA binding assays, we find significant differences in DNA binding preferences between the Dfd, Dfd/Ubx and Ubx proteins when *Dfd* and *Antp* upstream regulatory sequences are used as binding substrates. No significant differences in DNA binding specificity were detected between the human HOX4B (Hox 4.2) and *Drosophila* Dfd proteins. All of these full-length proteins bound as monomers to high affinity DNA binding sites, and interference assays indicate that they interact with DNA in a way that is very similar to homeodomain polypeptides. These experiments indicate that the ninth amino acid of the recognition helix of the homeodomain, which is glutamine in all four of these Antp-type homeodomain proteins, is not sufficient to determine their DNA binding specificities. The good correlation between the *in vitro* DNA binding preferences of these four Antp-type homeodomain proteins and their ability to specifically regulate a Dfd enhancer element in the embryo, suggests that the modest binding differences that distinguish them make an important contribution to their unique regulatory specificities.

Key words: *Deformed* gene/DNA binding/homeodomain/human HOX4B (Hox 4.2)

Introduction

Homeotic selector genes encode a family of proteins that assign different identities to the *Drosophila* head and body segments during development (Garcia-Bellido, 1977; Lewis, 1978; Wakimoto and Kaufman, 1981; Struhl, 1982). The selector proteins all have closely related versions of the homeodomain, an ~60 amino acid domain which is also present in many other developmental control proteins (McGinnis *et al.*, 1984a,b; Scott and Weiner, 1984).

Homeotic selector proteins are a subset of the Antp-type (also known as Antp-related, A-type or Class I) homeodomain proteins and include the products of *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), *Antennapedia* (*Antp*), *Ultrabithorax* (*Ubx*), *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*) (Regulski *et al.*, 1985; Scott *et al.*, 1989). Antp-type homeodomains are ~60–90% identical to the Antp prototype and usually have an additional hexapeptide homology just upstream of the homeodomain (Regulski *et al.*, 1985; Mavilio *et al.*, 1986). They are distantly related ($\leq 50\%$ identity) to other classes that include the engrailed class, paired class, POU class, or bicoid homeodomains (reviewed in Scott, 1989). Homeodomains have a DNA binding function (Desplan *et al.*, 1985; Hoey and Levine, 1988; Müller *et al.*, 1988) and homeodomain proteins can act as transcriptional regulators (Jaynes and O'Farrell, 1988; Biggin and Tjian, 1989; Driever and Nüsslein-Volhard, 1989; Han *et al.*, 1989; Krasnow *et al.*, 1989; Samson *et al.*, 1989; Struhl *et al.*, 1989; Winslow *et al.*, 1989; Ohkuma *et al.*, 1990). Structural studies of the Antennapedia and engrailed homeodomains indicate that homeodomains fold into a helix–turn–helix (HTH) motif (Laughon and Scott, 1984; Otting *et al.*, 1988; Qian *et al.*, 1989). Some recent studies indicate that α -helices 2 and 3 of the homeodomain, though structurally related to the HTH motif of prokaryotic DNA binding proteins, interact in a different way with DNA binding sites (Treisman *et al.*, 1989; Otting *et al.*, 1990; Kissinger *et al.*, 1990; Hanes and Brent, 1991).

Embryonic studies using chimeric proteins indicate that the Antp-type homeodomains play a major role in determining the developmental and regulatory specificity of the entire protein (Kuziora and McGinnis, 1989, 1990; Mann and Hogness, 1990; Gibson *et al.*, 1990). For example, ectopic expression of Dfd protein in the embryo causes the ectopic transcriptional activation of the endogenous *Dfd* locus. In epidermal cells, this ectopic activation is mediated through the *Dfd* autoregulatory enhancer (Bergson and McGinnis, 1990; Regulski *et al.*, 1991). Replacement of the Dfd homeodomain with the Ubx homeodomain results in a hybrid protein that no longer regulates endogenous *Dfd*, but has gained the ability to regulate one of the normal targets of the Ubx protein, the *Antp* transcription unit (Kuziora and McGinnis, 1989; Figure 1a). Another chimeric Dfd protein containing the homeodomain of Abd-B exhibits a target specificity in the embryo that reflects the target specificity of *Abd-B* (Kuziora and McGinnis, 1990).

Additional evidence implicating the homeodomain in determining specific developmental regulatory functions comes from functional studies of two mammalian *Hox* genes, which function as attenuated versions of their *Drosophila* homologues when expressed in *Drosophila* embryos. Studies of the *Drosophila* homeotic selector gene *Deformed* and one of its human homologues, *HOX4B* (Hox 4.2) has shown that ectopic expression of either gene in *Drosophila* embryos and

larvae specifically activates the endogenous *Deformed* transcription unit and is accompanied by similar dominant homeotic transformations in adults (McGinnis *et al.*, 1990). A similar functional conservation exists for the mouse *Hox-2.2* gene, which is able to induce homeotic transformations in *Drosophila* embryos and adults that are characteristic of its *Drosophila* homolog, the *Antennapedia* gene (Malicki *et al.*, 1990). Since the homeodomain regions are the only parts of the mammalian proteins that are extensively conserved in comparison to their *Drosophila* homologs, either regions outside of the homeodomain play a minor role in developmental specificity or the regions outside the homeodomain which confer specificity are not highly conserved in amino acid sequence.

How do different homeodomains specify such different regulatory effects in the embryo? Presumably the functional specificity results from interactions with other cellular proteins and DNA sites with generic or specific affinities for different Antp-type homeodomains (Hayashi and Scott, 1990). Most previous DNA binding studies provide little support for the idea that Antp-type homeodomains have different DNA binding specificities *in vitro* or in yeast cells (Desplan *et al.*, 1988; Treisman *et al.*, 1989; Hanes and Brent, 1989), although recent studies of the Ubx and ftz homeodomains indicate that their preferred binding sites are different (Ekker *et al.*, 1991; Florence *et al.*, 1991). Between different classes of homeodomain proteins, an important contributor to binding specificity was found to be homeodomain residue #50, but this residue is identical in Dfd, Ubx and other Antp-type proteins. One drawback to many previous assays of homeodomain DNA binding specificity is that they have been done under conditions that are likely to be less discriminating than the conditions which exist in the embryo, where the stringent ionic conditions, the high concentration of competitor DNA sites and the low abundance of binding protein may increase binding selectivity (von Hippel *et al.*, 1974; Kao-Huang *et al.*, 1977; Ptashne, 1986). In addition, some homeodomain binding studies have been done with DNA sequences which are not known to be relevant target sites in embryos (Desplan *et al.*, 1988; Hanes and Brent, 1989; Treisman *et al.*, 1989; Vincent *et al.*, 1990). This might also increase the difficulty of discovering relevant differences in DNA binding specificity.

In this report we show that the Dfd, Dfd/Ubx and Ubx proteins exhibit distinct DNA binding specificities using binding sequences from a *Dfd* autoregulatory element and comparison sequences from *Antp* upstream elements. The distinct DNA binding specificity of the *Drosophila* Dfd homeodomain is conserved in its human homolog, HOX4B (Hox 4.2). The differences in DNA binding specificity in our *in vitro* assays show a good correlation with embryonic functional specificity and suggest that distinct DNA binding specificities contribute to the regulatory specificity of even closely related homeodomain proteins.

Results

The different embryonic regulatory specificities of the Dfd, Dfd/Ubx and human HOX4B (Hox 4.2) proteins are summarized in Figure 1a. Both the wild type Dfd and the human HOX4B (Hox 4.2) proteins activate the *Dfd* transcription unit in embryos, but have no detectable regulatory effect on *Antp* (Kuziora and McGinnis, 1989;

McGinnis *et al.*, 1990). When the Ubx homeodomain is precisely substituted for the Dfd homeodomain, the resulting Dfd/Ubx chimeric protein has lost the ability to regulate *Dfd* transcription, but has gained the ability to regulate a normal target of Ubx, the *Antp* transcription unit (principally acting on the P1 promoter). The Dfd and Dfd/Ubx proteins differ only in the 22 amino acid residues depicted in Figure 1b. The positions of the amino acid changes are superimposed on the predicted α -helical regions in the Dfd and Ubx homeodomains. Most of the changes are in the N- and C-terminal regions of the homeodomain. The third α -helix in the homeodomain is analogous to the recognition helix of the prokaryotic HTH motif (Qian *et al.*, 1989). The Dfd and Ubx homeodomains have no amino acid differences in this region; note particularly that both conserve Gln50 (at position 9 of helix 3). The Dfd and human HOX4B (Hox 4.2) proteins conserve only small blocks of amino acid sequence outside their respective homeodomains (McGinnis *et al.*, 1990), but are highly conserved in a 66 amino acid region encompassing the homeodomain, with only the seven differences shown in Figure 1b.

In order to test whether the DNA binding specificities of these proteins were different when tested in stringent *in vitro* assays, we produced Dfd, human HOX4B (Hox 4.2), Ubx and Dfd/Ubx chimeric proteins in *Escherichia coli*. After induction, all four proteins were purified from soluble lysates over the same sequence-specific DNA affinity column. The column contained ligated, double-stranded DNA oligonucleotides with a high affinity binding site for the Dfd protein derived from site A in the *Dfd* autoregulatory enhancer (Regulski *et al.*, 1991).

Binding specificities of the Dfd, HOX4B (Hox 4.2), Dfd/Ubx and Ubx proteins were first tested on large regions of upstream *Dfd* and *Antp* regulatory DNA in immunoprecipitation assays and filter elution assays, and then the four proteins were tested on the highest affinity fragments using quantitative nitrocellulose filter binding assays. In embryos, *Dfd* autoregulation in maxillary and mandibular epidermal cells is mediated through an upstream enhancer that maps between -3.9 and -6.6 kb relative to the *Dfd* transcription start (Figure 2b). A 900 bp *HindIII* fragment within this enhancer constitutes a minimal *Dfd* autoregulatory element and contains high affinity binding sites for Dfd protein. Site-directed mutagenesis of four high affinity Dfd protein binding sites within the element inactivates its embryonic regulatory function (Bergson and McGinnis, 1990; Regulski *et al.*, 1991). This 900 bp fragment thus contains the best candidates for target sites that directly interact with Dfd protein in the embryo.

The *Antp* P1 promoter is repressed in embryonic cells by the Ubx protein, and activated by the Dfd/Ubx protein (Bermingham *et al.*, 1990; Kuziora and McGinnis, 1989). *Antp* P1 regulatory sequences that span the interval from -9.4 to +7.8 kb relative to the P1 transcription start have been shown to direct Ubx-dependent expression in embryos (Zink *et al.*, 1991), but Ubx-dependent target sequences have not yet been sharply delimited. Within the *Antp* upstream region, there are DNA fragments that preferentially bind Ubx protein with high affinity (Beachy *et al.*, 1988; Figure 2b) and Ekker *et al.* (1991) have recently defined optimal binding sites for purified Ubx homeodomains *in vitro*. We have used these fragments and binding sites, as described below, in comparative binding assays with Dfd

Fig. 1. Structure and function of the *Dfd*, human HOX4B (Hox 4.2) and *Dfd/Ubx* proteins. (a) The heat shock induced expression of *Dfd* protein or the human HOX4B (Hox 4.2) protein activates the endogenous *Dfd* transcription unit at ectopic sites in *Drosophila* embryos. The substitution of the *Ubx* homeodomain for the *Dfd* homeodomain results in a protein that is unable to activate *Dfd* transcription and has gained the ability to activate *Antp* P1 transcription. (b) The *Dfd/Ubx* protein differs from *Dfd* at the 22 amino acid positions shown beneath the sequence of the *Dfd* homeodomain. HOX4B (Hox 4.2) differs from *Dfd* at the six positions shown, has very limited amino acid sequence similarity to *Dfd* outside the homeodomain and consists of only 255 residues in comparison to the 586 residues of *Dfd* and *Dfd/Ubx*. The three (or four) predicted α -helical regions of the homeodomains are indicated by the bars. The helix 4 region, which is distinct from helix 3 in the *Antp* but not the engrailed structure, is represented by an open bar (Qian *et al.*, 1989; Kissinger *et al.*, 1990). Also indicated is the position of the ninth amino acid of the third α -helix (Gln50) which strongly influences DNA binding specificity *in vitro* and in yeast cells (Treisman *et al.*, 1989; Hanes and Brent, 1989).

autoregulatory sequences to evaluate the *in vitro* binding specificities of *Dfd*, *Dfd/Ubx*, *Ubx* and human HOX4B (Hox 4.2) proteins.

Immunoprecipitation assays (McKay 1981) of *Dfd* and *Antp* P1 DNA fragments bound to *Dfd* and *Dfd/Ubx* proteins indicate that the two proteins have distinct DNA fragment binding specificities. In a binding reaction with *HindIII/EcoRI* digested pS8.0 (Figure 2b), which contains the upstream autoregulatory enhancer within an 8.0 kb

region of *Dfd* regulatory DNA, *Dfd* protein has highest affinity for the 900 bp *HindIII* fragment and lower affinity for a 1.4 kb fragment (Figure 2a). Similar amounts of *Dfd/Ubx* protein bind much less of the 900 bp *HindIII* fragment and do not detectably bind the 1.4 kb fragment. In a binding reaction with the *Antp* P1 region digested with *Hinf* I, *Dfd* protein binds a 740 bp and a 314 bp fragment, but both of these *Antp* P1 fragments appear to have much lower affinity for *Dfd* protein than either of the fragments

Fig. 2. Dfd, Dfd/Ubx and Ubx protein binding to *Antp* P1 and *Dfd* DNA fragments. **(a)** Binding reactions contained either no protein (0); 125 or 250 ng Dfd protein (D); or 125 or 250 ng Dfd/Ubx protein (D/U). The binding DNA consisted of radiolabeled digests of the *Antp* P1 (*Antp* P1) and *Dfd* (S8.0 *Dfd*) upstream regions. Bound complexes were precipitated with anti-Dfd antibodies bound to *S.aureus* cells and visualized by autoradiography after denaturing PAGE. S8.0 *Dfd* was digested with *HindIII/EcoRI* and *Antp* P1 was digested with *HinfI*. Input fragments are shown in the lanes marked (M). Fragment lengths are indicated in base pairs at right. **(b)** Maps of the clones used in this study. S8.0 *Dfd* is an 8.0 kb genomic *SacI* fragment extending from 7.5 kb upstream of the transcription initiation site to the beginning of the *Dfd* open reading frame, cloned into the *SacI* site of pBluescript+ (Stratagene). Protein binding fragments are marked above the map with shaded bars, including the 900 bp high affinity Dfd protein binding element. It is a subset of the 2.7 kb autoregulatory element (Bergson and McGinnis, 1990; Regulski *et al.*, 1991) which is indicated by the shaded bar below the map. *Antp* P1 contains a 7 kb *HindIII* fragment extending from 6.3 kb upstream of the transcription start site to 0.7 kb downstream of the transcription start site, cloned into the *HindIII* site of pEMBL8+ (J.Bermingham, personal communication). Protein binding regions are indicated above the map, including the 314 bp, 543 bp and 472 bp high affinity Ubx binding regions. This figure is based on Beachy *et al.* (1988) and Laughon *et al.* (1986).

bound in *Dfd* upstream DNA (Figure 2a). Similar amounts of Dfd/Ubx protein show higher affinity for two other *Antp* P1 DNA fragments of 472 bp and 543 bp (Figure 2b). Previous experiments have shown that the Ubx protein exhibits the same relative affinity for *Antp* P1 fragments (Beachy *et al.*, 1988) as we observe for the Dfd/Ubx protein, with highest affinity for the 472 bp *Hinf I* fragment and lower affinities for the 543 bp and 314 bp fragments.

To confirm and extend these results, we used the high affinity DNA fragments identified in the immunoprecipitation assays to quantitatively compare the DNA binding

specificities of Dfd and Dfd/Ubx proteins in filter binding assays. Radiolabelled DNA fragments, 900 bp *Dfd* and 472 bp *Antp* P1, were bound under high stringency conditions to increasing concentrations of either Dfd or Dfd/Ubx protein and filtered through nitrocellulose (Riggs *et al.*, 1970). The results are shown in Figure 3, with the binding curves representing the average of three binding assays. As was suggested by the immunoprecipitation assay data, the wild type Dfd protein exhibits a marked preference for the 900 bp *Dfd* fragment. At the highest concentration of protein tested, it binds 2.5 times more of the 900 bp *Dfd*

Fig. 3. Filter binding assays of Dfd, Dfd/Ubx and Ubx proteins. Ten femtomoles of each radiolabelled DNA fragment were incubated with increasing amounts of Dfd, Dfd/Ubx and Ubx proteins, and then filtered through nitrocellulose to separate bound from unbound DNA. The percentage of the total input DNA bound to the filter was measured with a scintillation counter and is shown as a function of the amount of protein in each binding reaction. Binding at each protein concentration was repeated three times and plotted individually; lines are drawn through the average percent DNA bound for each protein concentration. Solid triangles, 900 bp *Dfd* DNA fragment. Open squares, 472 bp *Antp* DNA fragment. (a) Dfd protein binding. (b) Dfd/Ubx protein binding. (c) Ubx protein binding.

fragment than of the 472 bp *Antp* fragment (Figure 3a). In contrast, the Dfd/Ubx protein shows a much higher affinity for the 472 bp *Antp* fragment. At the highest amount of protein tested, it binds 4.3-fold more *Antp* fragment than *Dfd* fragment (Figure 3b). At the highest levels tested, the molar concentration of protein in these assays closely approximates the concentration of the ftz protein in *Drosophila* blastoderm nuclei (Krause *et al.*, 1988).

If the Ubx homeodomain functions as an autonomous

Fig. 4. Dfd, human HOX4B (Hox 4.2) and Dfd/Ubx binding to DNA fragments in the *Dfd* upstream autoregulatory enhancer. Binding reactions contained either no protein (0); or indicated amounts (in nanograms) of purified Dfd, human HOX4B (Hox 4.2), or Dfd/Ubx proteins. The amount of HOX4B (Hox 4.2) used was half that of Dfd and Dfd/Ubx to obtain similar molar concentrations in the binding reactions. The binding DNA consisted of the same radiolabelled digests of the *Dfd* upstream and *Antp* P1 upstream DNAs as used in Figure 2a and depicted in Figure 2b. Bound complexes were trapped on nitrocellulose filters; bound DNA fragments were eluted from the filters as described in Nowock and Sippel (1982) and visualized by autoradiography after denaturing PAGE. Input fragments are shown in the lanes marked (M). Fragment lengths are indicated in base pairs at right.

DNA binding domain, the DNA binding specificity of the Dfd/Ubx protein should be the same as the wild type Ubx protein. To test this we compared the DNA binding affinities of Dfd/Ubx and Ubx proteins on the 900 bp *Dfd* and 472 bp *Antp* fragments. A filter binding assay with purified Ubx protein indicates that it has a 1.7-fold greater affinity for the *Antp* fragment than for the *Dfd* fragment at the highest amount shown (Figure 3c). Although the Ubx protein does have higher affinity for the *Antp* fragment, it does not discriminate between the fragments as much as the Dfd/Ubx protein. Therefore, protein sequences outside of the Ubx homeodomain can influence its DNA binding specificity. The influence of associated protein sequences on homeodomain binding activities has been noted by Hoey *et al.* (1988), using hybrid even-skipped and engrailed proteins.

We also compared the binding specificity of Dfd and HOX4B (Hox 4.2) proteins using filter elution and quantitative filter binding assays. In the first set of experiments, each protein was incubated with radiolabelled *HindIII/EcoRI* digested pS8.0 in a binding reaction. Bound complexes were trapped on nitrocellulose filters and bound DNA fragments were eluted and separated on polyacrylamide gels (Nowock and Sippel, 1982; Beachy *et al.*, 1988). Using this binding assay, the Dfd and HOX4B (Hox 4.2) proteins have indistinguishable binding preferences, with highest affinity for the 900 bp *HindIII* fragment and lower affinity for a 1.4 kb fragment (Figure 4a). Similar amounts of Dfd/Ubx protein bind much less of the 900 bp *HindIII* fragment and preferentially bind an 800 bp fragment in the *Dfd* upstream region. Significant

binding affinity between the Dfd/Ubx protein and the 800 bp fragment was not detected in immunoprecipitation assays, presumably due to inhibitory effects of the polyclonal anti-Dfd antiserum.

Quantitative filter binding assays also indicate that the DNA binding specificities of Dfd and HOX4B (Hox 4.2) proteins are very similar. Radiolabelled DNA fragments, 900 bp *Dfd* and 472 bp *Antp* P1, were bound under high stringency conditions to increasing concentrations of either Dfd or HOX4B (Hox 4.2) protein and filtered through nitrocellulose (Riggs *et al.*, 1970). The results are shown in Figure 5, with the binding curves representing the average of three binding assays. Both the Dfd and HOX4B (Hox 4.2) proteins preferentially bind the 900 bp *Dfd* fragment and have lower affinity for the 472 bp *Antp* P1 fragment.

Footprint assays indicate that the different relative binding affinities of the Dfd, Dfd/Ubx, Ubx and HOX4B (Hox 4.2) proteins for the two DNA fragments do not result from differences in the number of sites capable of being recognized by the three proteins. The Dfd, Dfd/Ubx and HOX4B (Hox 4.2) proteins all footprint sites on the 472 bp *Antp* fragment which were previously identified as Ubx binding sites by Beachy *et al.* (1988) (P.DeJager and S.Dessain, data not shown). The Dfd/Ubx, Ubx and HOX4B (Hox 4.2) proteins all recognize the four high affinity Dfd protein binding sites on the 900 bp *Dfd* fragment (Regulski *et al.*, 1991; P.DeJager and S.Dessain, data not shown). This agrees with previous studies which have indicated that Antp-type homeodomain proteins recognize similar sites (Treisman and Desplan, 1989; Affolter *et al.*, 1990; Percival-Smith, 1990) and indicates that the differences in DNA binding specificity that we observe do not simply result from whether the proteins can recognize additional sites within a fragment. Instead, they must result from differences in the relative affinities for the multiple binding sites within the fragments (see below).

Because we used DNA fragments with multiple protein binding sites for our binding assays, the observed differences in DNA binding specificity could result from homeodomain-mediated differences in affinities for single sites, or from cooperative protein-protein interactions, or from a combination of these. To test whether affinity for single DNA binding sites could account for the differences, we performed a series of experiments with single binding site oligonucleotides (see below) containing sequences from the *Dfd* 900 bp fragment (site D), from the *Antp* 472 bp fragment (site A-1) and from a Ubx optimal binding site (U_{op} ; Ekker *et al.*, 1991; Figure 6).

We first employed an electrophoretic mobility shift assay to test the relative affinity of the Dfd, Dfd/Ubx and HOX4B (Hox 4.2) proteins for the three oligonucleotides, using binding conditions comparable to those used in the previous fragment binding assays. In this assay the Dfd protein has highest affinity for the *Dfd* (D) site and significantly lower affinities for the *Antp* P1 (A-1) or U_{op} sites (Figure 6). The Dfd/Ubx protein has relatively low affinities for both the *Antp* P1 (A-1) and *Dfd* (D) sites, with higher affinity for the U_{op} binding site. This suggests that the preference in binding affinity of the Dfd/Ubx protein for the *Antp* P1 472 bp element, which contains multiple repeats of the *Antp* P1 (A-1) oligonucleotide sequence, may require cooperative protein-protein interactions (see Discussion). Like the *Drosophila* Dfd protein, the human HOX4B (Hox 4.2) protein binds with highest affinity to the *Dfd* (D) site,

Fig. 5. Filter binding assays of Dfd and human HOX4B (Hox 4.2) proteins. Ten femtomoles of each radiolabelled DNA fragment were incubated with increasing amounts of Dfd and human HOX4B (Hox 4.2) proteins, and then filtered through nitrocellulose to separate bound from unbound DNA. The percentage of the total input DNA bound to the filter was measured with a scintillation counter and is shown as a function of the amount of protein (in picomoles to compensate for size differences between the two proteins) in each binding reaction. Each protein concentration was tested three times and plotted individually; lines are drawn through the average percent DNA bound for each protein concentration. Solid triangles, 900 bp *Dfd* DNA fragment. Open squares, 472 bp *Antp* DNA fragment. (a) Dfd protein binding. (b) HOX4B (Hox 4.2) protein binding. Data points plotted for the Dfd protein correspond to 50, 100, 200 and 500 ng protein.

with lower affinity to the *Antp* P1 (A-1) and U_{op} sites (Figure 6). This indicates that, at least on these sites, the DNA binding preferences of the *Drosophila* and human proteins have also been conserved at the level of monomer site affinity. For unknown reasons, the Ubx protein exhibits no activity in gel shift assays using conditions as in Figure 6 (P.Beachy, personal communication and C.T.Gross, data not shown).

Biochemical and structural studies of protein-DNA complexes indicate that single homeodomains interact with a conserved TAAT core sequence (Affolter *et al.*, 1990; Otting *et al.*, 1990; Percival-Smith *et al.*, 1990; Kissinger *et al.*, 1990). The *Dfd* (D) and U_{op} oligos each contain only one TAAT motif, suggesting that they optimally interact with single protein molecules. To test this, we performed gel shift

Deformed D site

AATTCCTGCAGCCCCGAGAACATCATTAAAGCTTAATTTTGAGAGGGGATC
TTAAGGACGTCGGGCTCTTGTAGTAATTCGAATTAATACTCTCCCCTAG

Antennapedia A-1 site

GATCCCCCGATATTATTATTATTATTATTATCGATGGGCTGCAGGAATT
CTAGGGGGCTATAATAATAATAATAATAATAGCTACCCGACGTCCTTAA

Ubx optimal site

TCGAGGGATCCATAGCCATTAAAGATCTGAATTCCTGCAGCCC
AGCTCCCTAGGTATCGGTAATTTTCTAGACTTAAGGACGTCGGG

Fig. 6. Gel shift binding assays on single binding sites. Gel mobility shift assay of Dfd, Dfd/Ubx and human HOX4B (Hox 4.2) proteins bound to oligonucleotides. Lanes 1–3: Deformed D site (D), Antennapedia A-1 site (A) and Ubx optimal site oligonucleotides (U) without added protein. Lanes 4–6: the same oligonucleotides as in lanes 1–3 with 2 μ l Dfd protein. Lanes 7–9: the same oligonucleotides as in lanes 1–3 with 2 μ l Dfd/Ubx protein. Lanes 10–12: the same oligonucleotides as in lanes 1–3 with 0.5 μ l human HOX4B (Hox 4.2) protein. Double-stranded DNA sequences of the three oligonucleotides used are shown, with ATTA core binding sequences underlined. The *Antp* A-1 oligonucleotide has multiple overlapping ATTA motifs; only one is underlined. Control experiments show that identical protein preparations from lysates of *E. coli* BL21 strains carrying expression plasmids without Dfd or Dfd/Ubx open reading frames exhibit no binding activity under these conditions (data not shown).

Fig. 7. Interference binding studies. Ethyl nitrosourea (ENU) and potassium permanganate (MnO₄) modified DNAs were isolated from gel mobility shift gels, cleaved and resolved on 15% denaturing polyacrylamide gels. Free (F) and bound (B) samples are shown adjacent to G+A sequencing lanes (Maxam and Gilbert, 1977). To the left of each figure is the DNA sequence of the labelled strand, alongside bars which indicate levels of high (dark shading) and low (lighter shading) interference by ENU induced ethylation. On the right of each figure, triangles indicate thymines which, when modified by KMnO₄, interfere strongly (black triangles) or weakly (open triangles) with binding. 5' and 3' strand designations are taken from Figure 9. (a) Binding of the Dfd protein to the 5' strand of the Dfd (D) site DNA. (b) Dfd protein bound to the 3' strand of the Dfd (D) site DNA. (c) Dfd/Ubx protein bound to the 5' strand of the U_{op} site DNA. (d) Dfd/Ubx protein bound to the 3' strand of the U_{op} site DNA.

assays with DNA fragments that had been modified with either ethyl nitrosourea (ENU) or potassium permanganate (KMnO₄). ENU ethylates phosphate groups and MnO₄⁻ preferentially oxidizes thymine bases, allowing the identification of phosphate and thymine modifications that interfere with binding. Bound and unbound oligonucleotides were eluted from shift gels comparable to those in Figure 6,

then cleaved and resolved on denaturing polyacrylamide gels (Figure 7).

Comparison of the cleavage patterns of input and bound DNA molecules indicates that the Dfd protein interacts with a single region of the Dfd (D) oligonucleotide including the TAAT sequence (Figures 7a,b and 9a). The Dfd/Ubx protein interacts with an analogous domain of the U_{op}

oligonucleotide (Figures 7c,d and 9b). In both cases the interference pattern is similar in size to that seen with purified, monomeric Antp homeodomain (Affolter *et al.*, 1990), strong evidence that the shifted bands shown in Figure 6 represent single site binding complexes.

Although the Dfd/Ubx protein has a low affinity for the the 900 bp *Dfd* fragment and no detectable regulatory effect on *Dfd* transcription in the embryo, the Ubx protein had considerably higher affinity for the 900 bp fragment. This led us to ask whether the Ubx protein might have a regulatory effect on *Dfd*. Loss-of-function *Ubx* alleles exhibit no ectopic *Dfd* expression (M.Regulski, unpublished results), but since the expression patterns of *Ubx* and *Dfd* do not normally overlap, any cross-regulatory relationships similar to the *Ubx* repression of the overlapping expression of *Antp* would not have been observed. To test whether Ubx protein has a regulatory effect on *Dfd* when they are expressed in the same regions of the embryo, we examined *Dfd* expression in embryos of a *Drosophila* strain which is heterozygous for a *Ubx* cDNA under heat shock promoter control (*HS:Ubx-1a*, Mann and Hogness, 1990). Embryos were heat shocked at 37°C and probed with antibodies to Dfd protein after 5 h of recovery at 25°C. As shown in Figure 8, *Dfd* expression is partially or completely repressed in *hsp70-Ubx* embryos. Control embryos, which expressed a truncated, non-functional *Ubx* protein under heat shock promoter control (*HS:Ubx-IVaFS*, Mann and Hogness, 1990), were treated in a similar manner and showed normal *Dfd* expression levels. Thus, for the Ubx protein, a relatively high binding affinity to the 900 bp *Dfd* fragment correctly anticipated a regulatory effect on *Dfd* expression in the embryo.

Discussion

Our results indicate that closely related Antp-type homeodomains can have different DNA binding specificities. There are 22 amino acid differences between the Dfd and Ubx homeodomains (Figure 1b) and 15 amino acids in the Ubx homeodomain which differ from those conserved between the HOX4B (Hox 4.2) and Dfd homeodomains. Any or all of these might be responsible for the observed differences in DNA binding specificity. Since position 9 of α -helix 3 is occupied by a glutamine in all the proteins that we tested, it follows that although the amino acid residue occupying this position is an important contributor to DNA binding specificity (Hanes and Brent, 1989; Treisman *et al.*, 1989), significant differences in binding affinity can be conferred by amino acid variation in other regions of the homeodomain. Structural analysis of the Antp and engrailed homeodomain–DNA complexes suggest that base-specific contacts are likely to be made by residues in helix 3 and by residues that are N-terminal to helix 1 (Otting *et al.*, 1990; Kissinger *et al.*, 1990). A recent study of the highly divergent Antp and TTF-1 homeodomains, both of which have Gln50, has shown that the residues that account for their distinct binding preferences map in both N- and C-terminal regions of the homeodomain (Damante and DiLauro, 1991).

The different binding specificities that we observe in our *in vitro* assays for the Dfd, Dfd/Ubx, Ubx and human HOX4B (Hox 4.2) proteins on *Dfd* autoregulatory sequences show a good correlation with the ability to specifically regulate *Dfd* transcription in embryos. The Dfd and HOX4B

Fig. 8. Repression of *Dfd* by ectopic Ubx protein expression. Embryos of a *Drosophila* strain heterozygous for a functional *hsp 70-Ubx 1a* gene (strain 1a22) were collected for 3 h, heat-shocked at 37°C for 1 h and left to recover for 5 h at 25°C before collecting and staining with antibodies to Dfd protein. Lateral views of three stage 11–12 embryos (Campos-Ortega and Hartenstein, 1985). Two of the embryos (HSU) contain the *hsp70-Ubx* gene and show full or partial repression of *Dfd* expression. The third embryo (WT) is a wild type control subjected to the same heat shock, recovery and staining treatment, and shows normal *Dfd* expression in the maxillary and mandibular segments.

(Hox 4.2) proteins, which both activate *Dfd* transcription (Kuziora and McGinnis, 1989; McGinnis *et al.*, 1990), bind with higher relative affinity to a *Dfd* autoregulatory element than to *Antp* P1 fragments with high binding affinity for Ubx protein. Both proteins also preferentially bind a single high affinity site (site D, containing an ATCATTA consensus) from the *Dfd* autoregulatory element, with relatively lower affinity for a Ubx optimal site and for a Ubx binding site from the 472 bp *Antp* fragment. So although the human HOX4B (Hox 4.2) protein has evolved independently from Dfd for the past 600 million years, it has conserved a near identical binding specificity and affinity. The Dfd/Ubx protein, which fails to activate *Dfd* transcription (Kuziora and McGinnis, 1989), has a reduced relative affinity for the *Dfd* autoregulatory element as well as for the high affinity *Dfd* (D) site and a much higher affinity for fragments that are preferentially bound by Ubx protein. The Ubx protein, which represses *Dfd* expression when expressed at high levels in the embryo, binds with moderately high affinity to a *Dfd* autoregulatory element. These correlations suggest that a distinct and unique DNA binding specificity is required to distinguish Dfd regulatory function from the regulatory functions of other Antp-type homeodomain proteins in the embryo.

The view that a unique binding specificity is important for Dfd and other Antp-type proteins does not, of course, preclude protein modifications or stable protein–protein

Fig. 9. Summary of the interference binding data. Dfd protein interference pattern of Dfd protein with the *Dfd* (D) site oligonucleotide. Dfd/Ubx protein interference pattern of Dfd/Ubx protein with the U_{op} site oligonucleotide. Interference patterns are projected onto representations of the DNA double helix. The DNA sequence is shown alongside the helix, with bases that are shared by the two oligonucleotides in boldface. Ethylated phosphates which interfere with binding are shown as grey (weak interference) and black (strong interference) filled circles. Thymines which interfere with binding when modified are circled; strong and weak interference are indicated with dark and light shading, respectively. At the bottom of the figure, sequences of the bound regions are shown, with bars indicating regions of ethylation-induced interference (dark shading, strong interference; light shading, weak interference) and diamonds indicating regions of KMnO₄ induced interference at Ts (black, strong; open, weak).

interactions as other important influences on functional specificity. Genetic results suggest that much of the specificity of such interactions will also be controlled by homeodomain sequence. Based on analogy with other transcriptional regulators, it seems possible that a range of Dfd response elements exist in DNA, some with high affinity monomer binding sites for Dfd protein and others with lower affinity in which specific protein-protein interactions make a greater contribution to the Dfd specificity of the regulatory response (Bender and Sprague, 1987; Rauscher *et al.*, 1988; Nakabeppu *et al.*, 1988; Halazonetis *et al.*, 1988; Goutte and Johnson, 1988; Johnson and McKnight, 1989; Stern *et al.*, 1989; Himmelfarb *et al.*, 1990; Glass *et al.*, 1990; Kristie and Sharp, 1990). The hypothesis that modest differences in binding affinity are important in discriminating the regulatory function of closely related homeodomain proteins does tend to de-emphasize the relative importance of residue 50 of the homeodomain, which although required for the recognition of specific binding sites, is only one of

a number of important homeodomain residues contributing to binding specificity and to the regulation of specific response elements.

The inability of the Dfd/Ubx protein to bind preferentially the *Antp* P1 (A-1) oligonucleotide in gel shift assays (Figure 6) suggests that the high affinity of Dfd/Ubx for the 472 bp *Antp* P1 DNA fragment does not result from high affinity interactions with single site sequences, but from cooperative or other higher order interactions. The Dfd/Ubx protein does have preferential affinity for the *Antp* P1 (A-1) oligonucleotide in immunoprecipitation assays, binding 10-fold more strongly than to the Dfd (D) oligonucleotide (data not shown). Therefore, we hypothesize that in association with the matrix of polyclonal antibodies and *S. aureus* cells, Dfd/Ubx is able to form stable complexes of more than one protein with additional TAAT motifs in the *Antp* P1 (A-1) oligonucleotide. In support of this hypothesis, we have observed lower mobility complexes in gel shift assays performed under less stringent binding conditions,

suggesting that more than one protein can be accommodated on the *Antp* P1 (A-1) oligonucleotide. We have also shown that a version of the *Antp* P1 (A-1) site containing only four TAA repeats is not detectably bound and immunoprecipitated by Dfd/Ubx protein as a monomer, but is bound with high preferential affinity by the Dfd/Ubx protein if five or more of these sites are ligated together (C.T.Gross, unpublished results), further suggesting that high affinity binding to the *Antp* P1 (A-1) site requires the formation of a multi-protein complex.

Differences in the ethylation interference patterns of the Dfd and Dfd/Ubx proteins provide additional evidence for the distinct binding functions of the two proteins. Binding of the Dfd/Ubx protein is inhibited by ethylation of the three phosphates directly 5' of the TAAT motif, while Dfd binding appears unaffected by the modification of this region of the DNA backbone. Binding of the Dfd protein is disrupted by ethylation of two phosphates upstream of the ACATCATTA sequence, including a phosphate on the opposite DNA strand, which do not interfere with Dfd/Ubx binding in comparable positions of the U_{op} site (Figure 9).

In general, our interference experiments with full-length proteins show few differences with the basic features of DNA interactions previously elucidated using 61–73 residue homeodomain-containing peptides. The DNA models in Figure 9 are oriented as shown based on analogy with structural studies so that the centrally located major groove would be filled with helix 3 of the homeodomain. Common features between our data and that obtained from purified *Antp*, *fushi tarazu* (*ftz*) and engrailed homeodomains (Otting *et al.*, 1990; Affolter *et al.*, 1990; Percival-Smith *et al.*, 1991; Kissinger *et al.*, 1990) are shown in Figure 9. These include the strong backbone interferences at upper left, presumably due to required interactions with the loop between helix 1 and helix 2 and the strong interferences at lower right due to interactions with the residues that extend N-terminal to helix 1.

The repressive effect of Ubx on *Dfd* is consistent with the cross-regulatory relationships between other Antennapedia and Bithorax complex genes, where more posterior genes in the complexes repress the expression of more anterior genes. However, the repressive effect is probably dependent on abnormally high levels of Ubx protein expression. Previous experiments with an *hsp70-Ubx* gene indicated no effect of ectopic Ubx protein on *Dfd* expression (González-Reyes and Morata, 1990). However, the González-Reyes and Morata *hsp70-Ubx* construct was apparently less potent than that of Mann and Hogness, as evidenced by the *Dfd*-dependent structures which still formed in the cuticle of their heat shocked embryos (cf. González-Reyes and Morata, 1990, Figure 2; Mann and Hogness, 1990, Figure 1). Additional evidence that normal levels of Ubx protein do not necessarily result in global *Dfd* repression come from *hsp70-Dfd* experiments, where it is known that the autoregulated expression of the *Dfd* gene can occur in a subset of the wild type *Ubx* expression domain (Kuziora and McGinnis, 1988). Taken together, these observations suggest that high levels of *Ubx*—higher than the normal expression levels in parasegment 6—may be required to significantly repress *Dfd*.

Materials and methods

Production and purification of proteins

Dfd and Dfd/Ubx and HOX4B (Hox 4.2) proteins were produced in BL21 *E. coli* cells essentially as has been described (*Dfd/Ubx* was subcloned

into the pARDfd expression construct; Jack *et al.*, 1988), with some modifications suggested by Michael Carey and Mark Biggin (personal communication). Ubx (isoform 1b; O'Connor *et al.*, 1988; Kornfeld *et al.*, 1989) protein was produced in a nalidixic acid-inducible *E. coli* strain, provided by Phil Beachy, as described (Mott *et al.*, 1985; Beachy *et al.*, 1988). 250 ml 37°C 2XYT (500 µg/ml ampicillin) was inoculated with a fresh colony and grown to 0.5–0.7 OD₆₀₀ and induced with 57 mg IPTG for 1–2 h. Cells were washed with 60 ml 20 mM HEPES pH 7.6, 200 mM NaCl, 4°C and then resuspended in 6 ml lysis buffer, 4°C: 20 mM HEPES pH 7.6, 200 mM NaCl, 50 mM β-mercaptoethanol, 5 µg/ml leupeptin, 5 µg/ml pepstatin A, 2 µg/ml aprotinin, 1 mM phenylmethylsulfonyl fluoride, 1 mM sodium metabisulfite, 2 mM benzamidine, 10% w/v glycerol, 0.1 mM EDTA, 0.5 mg/ml lysozyme. Cells were sonicated in 1.5 ml aliquots three times for 5 s each with a micro-tip cell disruptor at the highest setting at 4°C (Branson Sonic Power Co. #350). Lysates were spun in an Eppendorf centrifuge for 30 min at 4°C and the soluble fractions were frozen at –80°C. For immunoprecipitation and filter binding assays, proteins were purified in one pass over a DNA affinity column following the protocol of Kadonaga and Tjian (1986), except that 2 ml of the lysis supernatant was thawed, diluted with 8 ml 0.1 M KCl Z buffer (0.1 Z) and loaded directly onto a 0.5 ml column. The column was washed with 7.5 ml 0.1 Z, 1 ml of 0.2 Z and the protein was collected in the second column vol of a 1.5 ml wash with 0.6 Z. The column contained ligated multimers of the oligonucleotide:

5'-GATCCCTATCATTAGATGA

GGATAGTAATCTACTCTAG-5'

which was taken from site A of the 900 bp Dfd regulatory element (Regulski *et al.*, 1991). Protein concentrations were assayed using the method of Bradford against a BSA standard (Bradford, 1976) and the proteins were visualized with 12–14% SDS–PAGE and Western blotting (Jack *et al.*, 1988).

Immunoprecipitation assays

Immunoprecipitation assays were based on the protocol of McKay (1981). Binding reactions were performed in 50 µl binding buffer (20 mM HEPES pH 7.6, 210 mM KCl, 6.25 mM MgCl₂, 0.25 mM EDTA, 1 mM dithiothreitol (DTT), 10% glycerol (w/v), 100 µg/ml sonicated herring sperm DNA, 0.05% NP-40) with protein and 1 fmol ³²P-labelled DNA digest. After 30 min on ice, 15 µl of *S. aureus* cells (Calbiochem) pre-incubated with polyclonal rabbit anti-Dfd antibodies were added and incubated another 30 min. The cell complexes were centrifuged and then washed 2 X for 5 min each with 200 µl binding buffer with 0.5% NP-40 and 10 µg/ml sonicated herring DNA. The bound complexes were centrifuged and resuspended in 200 µl TE (10 mM Tris–HCl, pH 8, 1 mM EDTA), extracted with phenol/chloroform 50/50, ethanol precipitated and resolved on a 5% denaturing polyacrylamide gel. To prepare the *S. aureus* cells, 100 µl of the formalin fixed, 10% solution was centrifuged and washed twice with binding buffer (no added DNA), incubated with 1 ml 50% anti-Dfd rabbit serum, 50% glycerol for 1 h at 4°C. The cells were centrifuged and washed with 500 µl binding buffer, 1 µg/ml herring sperm DNA and then resuspended in 200 µl binding buffer, 1 µg/ml herring sperm DNA.

Filter binding assays

Filter binding assays were based on the protocols of Riggs *et al.* (1970) and Seneor *et al.* (1986). Proteins were incubated for 30–60 min on ice with 10 fmol probe DNA [kinase labelled with ³²P on both ends, (DNA concentrations were determined using a Hoefer TKO fluorimeter)] in 40 µl binding buffer containing 20 mM HEPES pH 7.6, 200 mM KCl, 5 mM MgCl₂, 0.25 mM EDTA, 1 mM DTT, 5% DMSO, 100 µg/ml sonicated herring sperm DNA, 125 µg/ml acetylated bovine serum albumin (BSA) (Boehringer Mannheim). Binding reactions were suctioned through 25 mm nitrocellulose filters which had been pre-soaked overnight in binding buffer without DNA and BSA, at a rate of 15–25 s/0.5 ml. 0.5 ml binding buffer (4°C, 100 mM KCl, no DNA or BSA) was then suctioned through the filter at the same rate. Filters were air dried and retained probe DNA was measured with a scintillation counter. Background retention in the absence of protein was <1%; background counts were subtracted from experimental values before plotting.

In the experiments involving elution of DNA fragments from the nitrocellulose filters (Figure 5), binding reactions were performed as above with protein and 1 fmol ³²P-labelled DNA digest, filtered through nitrocellulose and then washed with 1 ml of binding buffer containing 200 mM KCl. Filters containing the bound protein and DNA fragments were damp-dried and then, following the method of Nowock and Sippel (1982), soaked in 0.5 ml 0.1% Sarkosyl, 10 mM EDTA, 20 µg/ml predigested pronase for 90 min at 37°C. Eluted DNAs were ethanol precipitated at 4°C to prevent precipitation of detergent and then visualized on a 4% denaturing polyacrylamide gel.

Gel mobility shift and interference binding assays

Gel mobility shift assays were based on the protocol of Fried and Crothers (1981). Binding was carried out in 25 μ l containing 200 mM KCl, 20 mM HEPES pH 7.4, 0.25 mM EDTA, 1 mM DTT, 10% glycerol, 100 μ g/ml sonicated herring sperm DNA and 8000 c.p.m. radiolabelled oligonucleotide (see below). The precise DNA concentration could not be conveniently determined, but control experiments confirmed that doubling or halving the DNA concentration did not affect the percentage of bound DNA (data not shown). Protein was added and binding was allowed to proceed for 25 min at room temperature (23°C), at which point the reactions were loaded onto a TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.025% NP40) 4% polyacrylamide gel (40:1 crosslinking) running at 8 V/cm. After 20 min the voltage was increased to 20 V/cm and left to run for 4 h. Gels were fixed and exposed overnight. For quantification of bound and free DNA, bands were cut out of the dried gel and counted.

Proteins for the gel mobility shift assays were prepared by 30% ammonium sulfate precipitation of soluble bacterial lysates from Dfd, Dfd/Ubx and human HOX4B (Hox 4.2) over-producing *E. coli* strains and were estimated to be >80% pure by Coomassie staining of SDS-polyacrylamide gels. Protein concentrations were measured using the method of Bradford (1976): Dfd 1.2 μ g/ μ l, Dfd/Ubx 1.2 μ g/ μ l, human HOX4B (Hox 4.2) 0.93 μ g/ μ l.

Synthetic oligonucleotides cloned into pBluescript+ (Stratagene) were excised from the polylinker with restriction endonucleases to give fragments of 44 bp (U_{ox}) and 49 bp [*Dfd* (D) and *Antp* P1 (A-1)] (Figure 6). Plasmids were cut and labelled at one end with either [α - 32 P]dNTPs and Sequenase 2.0 (US Biochemical) or [γ - 32 P]ATP and polynucleotide kinase (New England Biolabs), cut at the other end and if necessary, treated with cold dNTPs and Sequenase 2.0 to fill in remaining single-stranded regions and then purified on 3.5% polyacrylamide gels.

For the interference binding assays, purified, end-labelled DNAs were treated with either ethyl nitrosourea (ENU, Sigma) (Siebenlist and Gilbert, 1980; Shuey and Parker, 1986) or potassium permanganate (KMnO₄, Sigma) (Rubin and Schmid, 1980; Truss *et al.*, 1990). All manipulations except for the gel-shift binding reactions took place in siliconized tubes and were performed with siliconized tips. Reactions containing 1 600 000 c.p.m. of end-labelled fragment (15 μ l) were diluted with 185 μ l 50 mM sodium cacodylate, 1 mM EDTA, before adding 200 μ l ENU-saturated ethanol and incubating for 60 min at 50°C. Reactions were precipitated on dry ice with 20 μ g glycogen (molecular biology grade, Boehringer Mannheim), 20 μ l 5 M NaOAc, pH 5.2, 400 μ l ethanol. Samples were re-precipitated twice from 200 μ l 0.5 M NaOAc, pH 5.2. For treatment with KMnO₄, DNA samples were diluted to 60 μ l with 30 mM Tris-HCl, pH 8.0 and denatured for 2 min at 95°C before quickly cooling in an ice water bath and adding 60 μ l 0.25 mM KMnO₄. After 12.5 min at 20°C, reactions were stopped and precipitated on dry ice with 20 μ l allyl alcohol (Aldrich), 20 μ l 2.5 M NaOAc, pH 5.2, 1 μ l glycogen (20 μ g) and 400 μ l ethanol. The reactions were then reprecipitated as for the ENU treated DNA.

Two aliquots each containing ~500,000 c.p.m. of the modified DNAs were incubated with protein under the high stringency binding conditions and run on gels as described above. After an overnight exposure at 4°C, gel fragments containing bound and unbound DNA samples were shaken with 3 ml QO1 buffer (100 mM NaCl, 50 mM MOPS, 15% ethanol, pH 7.0; Qiagen) for 6 h at 37°C, filtered through a 45 μ m syringe filter and purified over a QIAGEN-tip 5 column (Qiagen) following standard protocols, eluted with 1.5 ml QF (1.25 M NaCl, 50 mM MOPS, 15% ethanol, pH 8.2) and precipitated with glycogen and isopropanol. The KMnO₄ treated DNAs were resuspended in 10% piperidine, the ENU treated DNA samples were resuspended in 0.1 M NaOH and both were incubated for 30 min at 90°C before ethanol precipitation (no additional glycogen). The NaOH treated samples were reprecipitated to reduce the pellet size. Samples were suspended in urea loading dye (8 M urea, 10 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue) and run on 15% denaturing polyacrylamide gels.

Heat shock and staining of embryos

HS:UBX-1a22 is a *Drosophila* strain which is heterozygous for a heat shock inducible *Ubx* gene (isoform 1a; O'Connor *et al.*, 1988; Kornfeld *et al.*, 1989). HS:UBX-FS40 is heterozygous for a heat shock inducible *Ubx* gene which contains a frameshift mutation in the homeobox and has no function in embryos. Both strains were constructed and provided by Richard Mann (Mann and Hogness, 1990). Embryos were collected for 3 h, aged for 2 h, heated to 37°C for 1 h and collected and stained 5 h later. Heat shocks were performed by placing the Parafilm-sealed collection plate directly into a 37°C water bath. *Dfd* expression is strongly reduced or abolished in ~70% of the embryos of the strain heterozygous for the functional *Ubx* gene and normal in the control embryos. Embryos were stained with antibodies to Dfd protein using horseradish peroxidase as has been described (Jack *et al.*, 1988) and were photographed at 300 $\times$ using brightfield optics.

Acknowledgements

We are grateful to Mike Regulski for providing the HOX4B (Hox 4.2) expression vector and for much advice and help. We thank Steve Ekker and Phil Beachy for providing Ubx binding site oligonucleotides, Ubx protein and an *E. coli* strain expressing Ubx protein, as well as plasmid DNA containing the *Antp* 472 bp *Hinf*I fragment, John Bermingham and Matt Scott for providing *Antp* P1 upstream DNA, Richard Mann for the *hsp70-Ubx Drosophila* strains, Phil DeJager, Mark Biggin, Steve Schultz and Michael Carey for help with protein induction and purification protocols, and Rachel Raulerson for lab support. We also are grateful to many of our colleagues at Yale for instructive comments, especially Alan Garen, Don Crothers and Joe Coleman. This research was supported by grants to W.M. from the NSF (Presidential Young Investigator Award) and NIH (PO1 GM39813), to C.T.G. from the NSF (Graduate Fellowship Program) and to S.D. from the NIH Medical Scientist Training Program.

References

- Affolter, M., Percival-Smith, A., Müller, M., Leupin, W. and Gehring, W.J. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4093–4097.
- Beachy, P.A., Krasnow, M.A., Gavis, E.R. and Hogness, D.S. (1988) *Cell*, **55**, 1069–1081.
- Bender, A. and Sprague, G.F., Jr (1987) *Cell*, **50**, 681–691.
- Bergson, C. and McGinnis, W. (1990) *EMBO J.*, **9**, 4287–4297.
- Bermingham, J.R., Jr, Martinez-Arias, A. and Scott, M.P. (1990) *Development*, **109**, 553–566.
- Biggin, M.D. and Tjian, R. (1989) *Cell*, **58**, 433–440.
- Bradford, M. (1976) *Anal. Biochem.*, **72**, 248–253.
- Campos-Ortega, J.A. and Hartenstein, V. (1985) *The Embryonic Development of Drosophila melanogaster*. Springer-Verlag, Berlin.
- Damante, G. and DiLauro, R. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 5388–5392.
- Desplan, C., Theis, J. and O'Farrell, P.H. (1985) *Nature*, **318**, 630–635.
- Desplan, C., Theis, J. and O'Farrell, P.H. (1988) *Cell*, **54**, 1081–1090.
- Driever, W. and Nüsslein-Volhard, C. (1989) *Nature*, **337**, 138–143.
- Ekker, S.C., Young, K.E., von Kessler, D.P. and Beachy, P.A. (1991) *EMBO J.*, **10**, 1179–1186.
- Florence, B., Handrow, R. and Laughon, A. (1991) *Mol. Cell. Biol.*, **7**, 3613–3623.
- Fried, M.G. and Crothers, D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
- Garcia-Bellido, A. (1977) *Am. Zool.*, **17**, 613–629.
- Gibson, G., Schier, A., LeMotte, P. and Gehring, W.J. (1990) *Cell*, **62**, 1087–1103.
- Glass, C.K., Devary, O.V. and Rosenfeld, M.G. (1990) *Cell*, **63**, 729–738.
- González-Reyes, A. and Morata, G. (1990) *Cell*, **61**, 515–522.
- Goutte, C. and Johnson, A.D. (1988) *Cell*, **52**, 875–882.
- Han, K., Levine, M.S. and Manley, J.L. (1989) *Cell*, **56**, 573–583.
- Hanes, S.D. and Brent, R. (1989) *Cell*, **57**, 1275–1283.
- Hanes, S.D. and Brent, R. (1991) *Science*, **251**, 426–430.
- Halazonetis, T.D., Georgopoulos, K., Greenberg, M.E. and Leder, P. (1988) *Cell*, **55**, 907–915.
- Himmelfarb, H.J., Pearlberg, J., Last, D.H. and Ptashne, M. (1990) *Cell*, **63**, 1299–1309.
- Hoey, T., Warrior, R., Manak, J. and Levine, M. (1988) *Mol. Cell. Biol.*, **8**, 4598–4607.
- Hoey, T. and Levine, M. (1988) *Nature*, **332**, 858–861.
- Jack, T., Regulski, M. and McGinnis, W. (1988) *Genes Dev.*, **2**, 635–651.
- Jaynes, J.B. and O'Farrell, P.H. (1988) *Nature*, **336**, 744–749.
- Johnson, F.B. and Krasnow, M.A. (1990) *Genes Dev.*, **4**, 1044–1052.
- Johnson, P.F. and McKnight, S.L. (1989) *Annu. Rev. Biochem.*, **58**, 799–839.
- Kadonaga, J.T. and Tjian, R. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 5889–5893.
- Kao-Huang, Y., Revzin, A., Butler, A.P., O'Conner, P., Noble, D.W. and von Hippel, P.H. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4228–4232.
- Kissinger, C.R., Liu, B., Martin-Blanco, E., Kornberg, T.B. and Pabo, C.O. (1990) *Cell*, **65**, 579–590.
- Kornfeld, K., Saint, R.B., Beachy, P.A., Harte, P.J., Peattie, D.A. and Hogness, D.S. (1989) *Genes Dev.*, **3**, 243–258.
- Krasnow, M.A., Saffman, E.E., Kornfeld, K. and Hogness, D.S. (1989) *Cell*, **57**, 1031–1043.
- Krause, H.M., Klemenz, R. and Gehring, W.J. (1988) *Genes Dev.*, **2**, 1021–1036.
- Kristie, T.M. and Sharp, P.A. (1990) *Genes Dev.*, **4**, 2383–2396.
- Kuziora, M.A. and McGinnis, W. (1988) *Cell*, **55**, 477–485.
- Kuziora, M.A. and McGinnis, W. (1989) *Cell*, **59**, 563–571.
- Kuziora, M.A. and McGinnis, W. (1990) *Mech. Dev.*, **33**, 83–94.

- Laughon,A. and Scott,M.P. (1984) *Nature*, **310**, 25–31.
- Laughon,A., Boulet,A.M., Bermingham,J.R.,Jr, Laymon,R.A. and Scott,M.P. (1986) *Mol. Cell Biol.*, **6**, 4676–4689.
- Lewis, E.B. (1978) *Nature*, **276**: 565–570.
- Malicki,J., Schugart,K. and McGinnis,W. (1990) *Cell*, **63**, 961–967.
- Mann,R. and Hogness,D.S. (1990) *Cell*, **60**, 597–610.
- Mavilio,F., Simeone,A., Giampaolo,A., Faiella,A., Zappavigna,V., Acampora,D., Poiana,G., Russo,G., Peschle,C. and Boncinelli,E. (1986) *Nature*, **324**: 664–667.
- Maxam, A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**: 560–564.
- McGinnis,W., Levine,M., Hafen,E., Kuroiwa,A. and Gehring, W.J. (1984a) *Nature*, **308**, 428–433.
- McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring W.J. (1984b) *Cell*, **37**, 403–408.
- McGinnis,N., Kuziora,M.A. and McGinnis,W. (1990) *Cell*, **63**, 969–976.
- McKay,R.D.G. (1981) *J. Mol. Biol.*, **145**, 471–488.
- Mott,J.E., Grant,R.A., Ho,Y.-S. and Platt,T. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 88–92.
- Müller,M., Affolter,M., Leupin,W., Otting,G., Wüthrich,K. and Gehring,W.J. (1988) *EMBO J.*, **7**, 4299–4304.
- Nakabeppu,Y., Ryder,K. and Nathans,D. (1988) *Cell*, **55**, 907–915.
- Nowock,J. and Sippel, A.E. (1982) *Cell*, **30**, 607–615.
- O'Connor,M.B., Binari,R., Perkins,L.A. and Bender,W. (1988) *EMBO J.*, **7**, 435–445.
- Ohkuma,Y., Horikoshi,M., Roeder,R.G. and Desplan,C. (1990) *Cell*, **61**, 475–484.
- Otting,G., Qian,Y., Müller,M., Affolter,M., Gehring,W. and Wüthrich,K. (1988) *EMBO J.*, **7**, 4305–4309.
- Otting,G., Qian,Y.Q., Billeter,M., Müller,M., Affolter,M. Gehring,W.J. and Wüthrich, K. (1990) *EMBO J.*, **9**, 3085–3092.
- Percival-Smith,A., Müller,M., Affolter,M. and Gehring,W.J. (1990) *EMBO J.*, **9**, 3967–3974.
- Plashne,M. (1986) *A Genetic Switch*. Cell Press and Blackwell Scientific Publications, Cambridge, MA and Palo Alto, CA.
- Qian,Y.Q., Billeter,M., Otting,G., Müller,M., Gehring,W.J. and Wüthrich,K. (1989) *Cell*, **59**, 573–580.
- Rauscher,F.J.,III, Voulalas,P.J., Franza,B.R.,Jr and Curran,T. (1988) *Genes Dev.*, **2**, 1687–1699.
- Regulski,M., Harding,K., Kostriken,R., Karch,F., Levine,M. and McGinnis,W. (1985) *Cell*, **43**, 71–80.
- Regulski,M., Dessain,S., McGinnis,N. and McGinnis,W. (1991) *Genes Dev.*, **5**, 278–286.
- Rubin,C.M. and Schmid, C.W. (1980) *Nucleic Acids Res.*, **8**, 4613–4619.
- Riggs,A.D., Suzuki,H. and Bourgeois,S. (1970) *J. Mol. Biol.*, **48**, 67–83.
- Samson,M., Jackson-Grusby,L. and Brent,R. (1989) *Cell*, **57**, 1045–1052.
- Scott,M.P., Tamkun,J.W. and Hartzell,G.W.,III. (1989) *Biochim. Biophys. Acta*, **989**, 25–48.
- Scott,M.P. and Weiner,A. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4115–4119.
- Siebenlist,U. and Gilbert,W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 122–126.
- Senear,D.F., Brenowitz,M., Shea,M.A. and Ackers,G.K. (1986) *Biochemistry*, **25**, 7344–7354.
- Shuey,D.J. and Parker, C.S. (1986) *J. Biol. Chem.*, **261**, 7934–7940.
- Stern,S., Tanaka,M. and Herr,W. (1989) *Nature*, **341**, 624–630.
- Studier,F.W. and Moffatt, B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.
- Struhl,G. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7380–7384.
- Struhl,G., Struhl,K. and Macdonald,P.M. (1989) *Cell*, **57**, 1259–1273.
- Sturm,R.A., Das,G. and Herr,W. (1988) *Genes Dev.*, **2**, 1582–1599.
- Treisman,J., Gönczy,P., Vashishtha,M., Harris,E. and Desplan,C. (1989) *Cell*, **59**, 553–562.
- Truss,M., Chalepkis,G. and Beato,M. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 7180–7184.
- Vincent,J.-P.,Kassis,J.A. and O'Farrell,P.H. (1990) *EMBO J.*, **9**, 2573–2578.
- von Hippel,P.H., Revzin,A., Gross,C.A. and Wang,A.C. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 4808–4812.
- Wakimoto,B.T. and Kaufman, T.C. (1981) *Dev. Biol.*, **81**, 51–64.
- Winslow,G.M., Hayashi,S., Krasnow,M., Hogness,D.S. and Scott,M.P. (1989) *Cell*, **57**, 1017–1030.
- Zink,B., Engstrom,Y., Gehring,W.J. and Paro,R. (1991) *EMBO J.*, **10**, 153–162.

Received on November 15, 1991; revised on December 12, 1991