The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences

Siu-Kwong Chan and Richard S. Mann

Department of Biochemistry and Molecular Biophysics and Center for Reproductive Sciences, Columbia University College of Physicians and Surgeons, New York, New York 10032 USA

Using an in vivo assay for segment identity, the structural differences that distinguish two Drosophila homeotic selector proteins, Ultrabithorax (Ubx) and Antennapedia (Antp), have been investigated. There are at least two independent parts of Ubx and Antp that contribute to their functional specificities: (1) their homeo domains and (2) residues carboxy-terminal to their homeo domains (C-tails). In the absence of any C-tail, differences in 5 homeo domain amino acids are sufficient to distinguish between the functions of Ubx and Antp. Two of these are at the amino terminus of the homeo domain and could contact DNA directly. A three dimensional model suggests that the other 3 homeo domain residues and the C-tails are unlikely to contact DNA. In addition, we demonstrate that the assay used to measure the segment identity functions of Ubx and Antp is independent of any homeotic selector gene normally active in thoracic and abdominal segments. Therefore, it is likely that this assay measures the coordinate regulation of many downstream target genes. This expectation is confirmed for at least one Ubx target gene, Distal-less.

[Key Words: Homeo domain; segment identity; Ultrabithorax; Antennapedia; Drosophila embryogenesis; homeotic genes]

Received December 11, 1992, revised version accepted March 5, 1993.

The different segments of Drosophila melanogaster obtain their distinct morphological characteristics from the action of a small set of genes known as the homeotic selector genes (Garcia-Bellido 1977; Lewis 1978; Struhl 1982; Kaufman et al. 1990). These genes, which are clustered in either the bithorax or Antennapedia complexes, are collectively referred to as the Homeotic complex (HOM-C). Altering the expression of HOM-C genes results in striking transformations of one body part or segment into another. For example, misexpression of Antennapedia (Antp) or Ultrabithorax (Ubx) in cells that normally generate the adult antennae results instead in the formation of a leg (Schneuwly et al. 1987;2 R; Mann and Hogness 1990). Misexpression of Ubx can also result in the transformation of a wing to a haltere (Lewis 1963,1982; White and Akam 1985). These sorts of homeotic transformations suggest that the HOM-C genes are master regulators that act by controlling groups of subordinate or downstream effector genes which in turn generate segment-specific identities (Garcia-Bellido 1977; Lewis 1978). Accordingly, the different segments of Drosophila have distinct identities because HOM-C proteins select and regulate different, but probably overlapping, sets of target genes. We are investigating how this functional specificity is achieved by HOM-C proteins in vivo.

The molecular characterization of HOM-C genes provides strong support for the idea that they act by regulating the expression of downstream target genes. All HOM-C genes encode proteins that contain closely related 61-amino acid homeo domains. The homeo domain is a protein motif present in many eukaryotic transcription factors and includes a helix–turn–helix reminiscent of those found in prokaryotic regulatory proteins (Laughon and Scott 1984; Gehring 1987; Scott et al. 1989). Consistent with them being transcription factors, HOM-C proteins have been shown to bind specific DNA sequences in vitro by their homeo domains (Beachy et al. 1988; Kuziora and McGinnis 1988; Affolter et al. 1990; Ekker et al. 1991) and can regulate target gene transcription in cell culture, in yeast cells, and in vitro (for review, see Levine and Hoey 1988; Hayashi and Scott 1990).

The homeo domain is a well-studied protein motif. Extensive structural studies on three different homeo domain–DNA complexes show that this 61-amino acid domain has three α-helices and confirm that helices 2 and 3 fold into a structure that is very similar to the prokaryotic helix–turn–helix (Billeter et al. 1990; Kissinger et al. 1990; Otting et al. 1990; Wolberger et al. 1991). As in the prokaryotic helix–turn–helix, the second of these two helices (helix 3 of the homeo domain) makes significant
contacts in the major groove of DNA and is referred to as the recognition helix. In addition, the amino-terminal region of the homeo domain makes contacts with the minor groove of DNA and several residues throughout the homeo domain contact the DNA phosphate backbone.

These structural studies suggest that the third or recognition helix of the homeo domain is an important sequence specificity determinant. Consistent with this, both in vivo and in vitro assays demonstrate that the ninth amino acid of the recognition helix is crucial for distinguishing qualitatively between two classes of DNA-binding sites, the NP class [ATTTAATTTGA] and the Bicoid class [TCTAATCCC] (Hanes and Brent 1989; Treisman et al. 1989). However, all HOM-C homeo domains have very similar recognition helices and all have a glutamate at position 9. Moreover, several different homeo domains specifically bind the same DNA sequence in vitro (Beachy et al. 1988; Desplan et al. 1988; Hoey and Levine 1988; Mihara and Kaiser 1988; Thali et al. 1988; Dessain et al. 1992; Ekker et al. 1992). These results argue that the DNA-binding properties of HOM-C homeo domains are not, by themselves, sufficiently discriminating to distinguish between different sets of target genes in vivo. However, when carefully examined, at least two HOM-C proteins have been shown to bind the same DNA sequence with different affinities (Dessain et al. 1992; Ekker et al. 1992). It is possible that these quantitative differences measured in vitro are exploited to produce functional differences in vivo.

In addition to characterizing the in vitro DNA-binding properties of HOM-C proteins, their in vivo identity functions have also been studied. A very useful way to do this has been to examine the phenotypes resulting from the ubiquitous and, therefore, ectopic expression of HOM-C proteins (Schneuwly et al. 1987a; Gibson and Gehring 1988, Kuziora and McGinnis 1988; Gonzalez-Reyes and Morata 1990; Mann and Hogness 1990). In all cases examined, ubiquitous expression of a particular HOM-C protein causes homeotic transformations toward an identity that is normally determined by that protein. Thus, Ubx, which is required for the parasegment [ps] 6 identity in wild-type embryos, transforms anterior parasegments into additional copies of ps 6 when ubiquitously expressed early in embryogenesis (Gonzalez-Reyes and Morata 1990; Mann and Hogness 1990). Similarly, Antp, which is required for a ps 4 identity in wild-type embryos, transforms anterior parasegments into additional copies of ps 4 after its ubiquitous expression (Gibson and Gehring 1988).

Using these types of in vivo functional assays, the homeo domain and the sequences close to the homeo domain were shown to contain most of the identity functions that distinguish individual HOM-C proteins (Kuziora and McGinnis 1989; Gibson et al. 1990; Mann and Hogness 1990; Lin and McGinnis 1992). In one example, amino-terminal homeo domain residues were shown to be important for distinguishing between the phenotypes generated by Deformed [Dfd] and Dfd–Ubx chimeras (Lin and McGinnis 1992). However, because of the autoregulatory functions present in Dfd, the in vivo assay used in this work actually measures the activation of single target genes [Dfd and Antp, respectively], not the coordinate regulation of a complete set of downstream genes.

In this paper a similar in vivo assay has been used to identify the relevant amino acids that distinguish the identity functions of Ubx and Antp. In contrast to the Dfd–Ubx chimeras, we show that the phenotypes generated by the ubiquitous expression of Ubx, Antp, and Ubx–Antp chimeras are independent of all endogenous homeotic genes that normally act in thoracic and abdominal segments. Therefore, the segment identity functions measured probably represent the coordinate regulation of multiple downstream target genes. Despite the complexity of these phenotypes, a small number of amino acids have been identified that, when changed, is sufficient to switch the identity functions of Ubx into those of Antp. In addition, whereas 5 homeo domain residues are critical, sequences immediately following the homeo domain can also have a profound effect on the segment identity functions of HOM-C proteins. A model based on existing homeo domain–DNA structures suggests that several of the relevant residues are in positions more likely to contact other proteins rather than DNA.

**Results**

The two HOM-C proteins investigated are products of the Antp and Ubx genes [Antp and Ubx proteins, respectively]. Both proteins have large amino-terminal regions with very little sequence similarity [295 residues for Antp, 259 residues for Ubx], very similar homeo domains [only 7/61 amino acid differences], and short carboxy-terminal regions [hereafter referred to as the Ubx or Antp C-tail] with no sequence similarities [23 residues for Antp, 34 residues for Ubx] (Fig. 1). To identify the relevant differences between these two proteins, we analyzed the phenotypic results of ubiquitously expressing various Ubx deletion mutants and Ubx–Antp chimeric proteins. The construction and sequences of these chimeras are detailed in Materials and methods.

The in vivo assay we used to measure the identity functions of HOM-C proteins consists of expressing them under the control of the heat shock-inducible promoter from the hsp70 gene. Thus, each open reading frame (ORF) investigated was cloned into the same CaSpeR-based P-element vector containing the hsp70 promoter and the white gene (Pirrotta 1988; Mann and Hogness 1990). Each of these was introduced into the germ line of flies, and at least two transformants for each construct were isolated and analyzed. Ubiquitous expression of individual HOM-C proteins was induced by applying a single 45-min heat shock to embryos aged from 3 to 6 hr. For most of the proteins studied, Western analysis using an antibody that recognizes an amino-terminal Ubx epitope confirmed the heat shock-induced synthesis of a correctly sized HOM-C protein (Fig. 2, data not shown). Furthermore, for at least one transformant for each construct studied, the chimeric region was
their large amino-terminal regions, both proteins have two short regions of similarity: MXSYF, near their amino-termini, YPWM, amino-terminal and close to their homeo domains. Both the Ubx and Antp homeo domains are located near their carboxyl termini. The sequences of their 61-amino acid homeo domains are shown: the Ubx sequence is written in full (top), and the Antp sequence below; only those amino acids that are different in Antp are shown (7 of 61 identities are indicated by a dot .). The three α-helices of the homeo domain are boxed, and helices 2 and 3, which form the helix-turn-helix motif, are indicated.

**Figure 1.** Ubx vs. Antp. Schematic diagrams of the Ubx [top, light shading] and Antp [bottom, dark shading] proteins. Within their large amino-terminal regions, both proteins have two short regions of similarity: MXSYF, near their amino-termini, YPWM, amino-terminal and close to their homeo domains. Both the Ubx and Antp homeo domains are located near their carboxyl termini. The sequences of their 61-amino acid homeo domains are shown: the Ubx sequence is written in full (top), and the Antp sequence below; only those amino acids that are different in Antp are shown (7 of 61); identities are indicated by a dot (·). The three α-helices of the homeo domain are boxed, and helices 2 and 3, which form the helix-turn-helix motif, are indicated.

reisolated from the genome using polymerase chain reaction (PCR). Each PCR product was sequenced directly to confirm the presence of the correct chimeric ORF in the P element [data not shown]. Once confirmed, segmental transformations were analyzed by observing the cuticle that was secreted -20–25 hr after the heat shock. Unless specified, these transformations were examined in an otherwise wild-type genetic background. In addition, effects on the expression of one Ubx target gene, Distal-less (Dll), were also analyzed [Vachon et al. 1992].

**Effects on ventral cuticle morphology**

The effects of ubiquitous Ubx expression can be distinguished easily from the effects of ubiquitous Antp expression [Fig. 2]. In this assay, expression of wild-type Ubx results in five additional ps 6 denticle belts in place of all three thoracic belts and two head segments [Fig. 2]. In addition, part of a sixth ps 6 denticle belt is observed at the very anterior of transformed embryos, presumably owing to the partial transformation of a third head segment [Gonzalez-Reyes and Morata 1991; Fig. 2]. Ubiquitous Ubx expression also completely suppresses the formation of the thoracic-specific sensory organs [Keilin’s organs and ventral pits]. Finally, ubiquitous Ubx expression has no phenotypic consequences in the cuticle posterior to and including the first abdominal segment [A1].

In contrast, ubiquitous expression of Antp results in a ps 4 denticle belt in place of the denticle pattern normally present in the first thoracic segment [Fig. 2]. Ectopic Antp expression also generates a partial ps 4 belt at the very anterior of transformed embryos. The formation of Keilin’s organs and ventral pits is not suppressed in these embryos, and there are no phenotypic consequences in the cuticle posterior to and including the second thoracic segment [T2]. Because ubiquitous Ubx and Antp expression generate phenotypes that are similar to wild-type ps 6 and ps 4 identities, respectively, they are likely to represent at least part of the wild-type Antp and Ubx regulatory hierarchies and, thus, are a qualitative measure of their in vivo identity functions.

When analyzing the identity functions of a particular mutant protein, we concentrated our attention on the morphology of the denticle belt present in the transformed first thoracic segment [T1]. We chose this segment because its wild-type identity is easily distinguished from all others. Furthermore, both Ubx and Antp efficiently transform this segment to different identities in this assay. In addition, we also analyzed the number of segments transformed and the identity of any partial belts produced in the anterior of transformed embryos. These results are described qualitatively below and are presented numerically in Table 1.

**Changing the homeodomain is sufficient to change identity functions** Previous experiments demonstrated that large portions of the amino-terminal regions of both Ubx and Antp could be deleted without a qualitative effect on their transformation potentials [Gibson et al. 1990; Mann and Hogness 1990]. These experiments also showed that the identity functions of chimeric proteins correlate with the identity of the homeo domain and its neighboring sequences. Here, we repeat and extend those results. Thus, although Ubx produces a transformation to a ps 6 identity, the UAA chimera [Ubx amino-terminal region, Antp homeo domain, and Antp C-tail] generated a ps 4 identity [Fig. 2D, G; Table 1]. UAA differed slightly from Antp in that the formation of Keilin’s organs and ventral pits was suppressed in many of the UAA-transformed embryos, especially when UAA was present in more than two copies [Mann and Hogness 1990] [Fig. 2G]. Except for this effect on the sensory organs, UAA generated a phenotype very similar to that generated by wild-type Antp protein [Table 1]. In contrast, the UUA chimera [Ubx amino-terminal region, Ubx homeo domain, and Antp C-tail] generated a ps 6 identity [Fig. 2F; Table 1]. These experiments demonstrate that the Antp C-tail does not influence the identity functions of a protein that is otherwise Ubx. In addition, these results suggest that the homeo domain is an important identity determinant.

These conclusions were further supported by studying the transformations produced by proteins having carboxy-terminal deletions. Thus, in most of the transformed embryos, UU_{stop} [Ubx amino-terminal region, Ubx homeo domain followed by a stop codon] and UA_{stop} [Ubx amino-terminal region, Antp homeo domain except for residue 61, which was not changed; see Fig. 5A, below] followed by a stop codon] generated ps 6 and ps 4 identities, respectively [Fig. 2I, J; Table 1]. Therefore, in the absence of the C-tail, changing the homeo domain from Ubx to Antp was sufficient to change identity functions. These results also demonstrate that sequences car-
Functional specificity of homeotic proteins

Figure 2. The segment identities generated by Ubx–Antp chimeras. (A) Schematic diagrams of the Ubx–Antp chimeras and a summary of the denticle belt identities that they generate in vivo. Ubx sequences are shown in light shading; Antp sequences are shown in dark shading. The names given to these chimeras are described in the text. (B) Western analysis of extracts from heat-shocked embryos expressing various Ubx–Antp chimeras. The monoclonal antibody used (FP3.38) recognizes an amino-terminal epitope in Ubx. (Lane 1) UAA; (lane 2) UUA; (lane 3) Ubx; (lane 4) UAU; (lane 5) UUstop; (lane 6) UAstop. Approximate molecular weights are indicated. (C–K) Photomicrographs of first-instar larval cuticles resulting from the ubiquitous expression of Ubx, Antp, and Ubx–Antp chimeras. All photographs are ventral or slightly lateral views, with the bottom-most visible segment corresponding to the first abdominal segment (ps 6). For all of the proteins studied here, all segment identities posterior to and including ps 6 were wild type. (C) Wild type; (D) Ubx; (E) Antp; (F) UUA; (G) UAA; (H) UAU; (I) UUstop; (J) UAstop in two copies; (K) UAstop in four copies. Keilin’s organs are indicated with small arrows. Anterior denticles are indicated as either thoracic (t) or abdominal (a) when visible. Ubiquitous UAU expression usually results in the nearly complete transformation of at least one head segment [large arrow in H].

boxy-terminal to the homeo domain are not essential in generating segmental transformations in embryos (see also Gibson et al. 1990).

Although the segment identities generated by UUstop, UUA, and UAstop were typically very similar to their parent proteins [Ubx, Ubx, and UAA, respectively], deleting or changing these C-tails did have phenotypic consequences, especially with regard to the number of segments transformed and the efficiency of transformation. Thus, although Ubx generated five or six ectopic ps 6 belts, UUstop and UAA only generated two or three ectopic ps 6 belts [Fig. 2I,F; Table 1]. Furthermore, UAstop transformed the T1 segment more efficiently than did UAA, as judged by the percentage of transformed embryos having no T1-specific denticle patch or “beard” [Fig. 2J,G; Table 1]. UAstop also suppressed the formation of the Keilin’s organs and ventral pits more efficiently than UAA. In addition, among the UAstop-transformed embryos, a few denticle belts had an intermediate identity, suggesting that this protein may have some weak ps 6 identity function. However, when the heat shock-UAstop gene [HS–UAstop] was present in four instead of the usual two copies, these intermediate identities were observed less frequently and the ps 4 transformation was more obvious [Fig. 2J,K; Table 1].

The Ubx C-tail can influence the identity functions of a heterologous homeo domain. All of the above experiments support the conclusion that sequences carboxy-terminal to the Ubx and Antp homeo domains help in relatively minor ways to specify the segment identity properties of these HOM-C proteins. Therefore, we were surprised to discover that the UAU chimera [Ubx amino-terminal region, Antp homeo domain, Ubx C-tail] gen-
Table 1. Transformations generated by several HOM-C proteins

| ORF | T1 beard gone (%) | Keilin’s organs | Ventral pits | Head segments | T1 identity | Anterior denticles |
|-----|-----------------|-----------------|-------------|--------------|-------------|-----------------|
|     | [n = 25]        | [<30]           | [50]        | [60]         | [ps 4]       | [ps 6]          |
| Ubx |                 | 0               | 0           | 0            | 0           | 0               |
|      | [2]             | 7               | 7           | 11           | 150 (25)    | 150 (25)       |
| UUA |                 | 0               | 0           | 0            | 0           | 0               |
|      | [2]             | 6               | 6           | 13           | 47 (10)     | 47 (10)        |
| UAstop |             | 0               | 0           | 0            | 0           | 0               |
|      | [2]             | 6               | 6           | 13           | 57 (13)     | 57 (13)        |
| UAU |                 | 0               | 0           | 0            | 0           | 0               |
|      | [2]             | 0               | 0           | 0            | 0           | 0               |
| UAA |                 | 0               | 0           | 0            | 0           | 0               |

*For each ORF a population of 3- to 6-hr embryos was given a single 45-min heat shock. After cuticle deposition, 25 embryos showing a blockage of head involution were scored. The number following each ORF represents the number of copies of the HS-ORF gene present in each adult fly in the stock.

*This is an approximate measure of transformation potency. The presence of the T1 beard indicates an inefficient transformation, its absence indicates a strong transformation.

[X] The total number of partial or complete Keilin’s organs (6/embryo or 150 possible), ventral pits (6/embryo or 150 possible), or transformed head segments (2/embryo or 50 possible); [n] the number of embryos that X was distributed over. For example, among the 25 transformed UAU embryos, there were only 5 Keilin’s organs among 4 embryos (21 embryos had no 0 Keilin’s organs).

The morphology of these denticle belts was judged as ps 6-like (ps 6), ps 4-like (ps 4), or intermediate (i).

The relevance of the carboxy-terminal region of Ubx was further supported by a sequence comparison of Ubx homologs from two distantly related invertebrates [Fig. 3]. Significant similarities within the C-tails were observed between the Ubx homologs from the moth Bombyx mori, and the leech, Hirudo medicinalis [Wysocka-Diller et al. 1989; Ueno et al. 1992]. In contrast, except for weak similarities between the Ubx and abdominal-A (abd-A) C-tails, such sequence similarities do not exist between the Drosophila HOM-C proteins, consistent with the suggestion that these residues contribute to their in vivo specificities (Karch et al. 1990).

Two additional constructs were used to address the function of the Ubx C-tail. In the first, the C-tail of UAU
was truncated immediately after the region similar to the Ubx homolog from *B. mori*. This construct, called UAu, deletes the last 17 amino acids from the Ubx C-tail, including a stretch of 9 alanine residues (Figs. 3 and 4A). Heat shock-induced expression of UAu generated a strong ps 6 transformation, demonstrating that the homology region in the Ubx C-tail is sufficient to modify the identity functions of the Antp homeo domain [Fig. 4B].

The modifying influence of the Ubx C-tail could result because it induces subtle conformational changes in the three-dimensional structure of the Antp homeo domain. Alternatively, the Ubx C-tail could function as a domain that is independent from the homodo domain, for example, as a protein interaction domain. To distinguish between these two possibilities, the construct UAalaU, which separates the Antp homeo domain and the Ubx C-tail by inserting 5 alanine residues, was tested [Fig. 4A]. As with UAu, heat shock-induced expression of UAalaU generated a ps 6 transformation, demonstrating that the Ubx C-tail does not need to be immediately adjacent to the Antp homeo domain to influence its identity functions [Fig. 4C].

Changing 5 of the 7 homeo domain residues is sufficient to change identity functions. Given the ps 6 transformation produced by UUstop and the ps 4 transformation produced by UAstop, we constructed further homeo domain chimeras to determine whether fewer amino acid changes could also produce this functional change. The seven-amino acid differences between these two homeo domains are located as follows: two near the amino terminus of the homeo domain (positions 1 and 3), two close to the end of the first α-helix (positions 23 and 25), one in the second helix (position 35), one near the end of the third helix (position 57), and one at the very carboxyl terminus of the homeo domain (position 61) [see Fig. 1]. The UAstop chimera has only the first six of these seven differences changed to the Antp residue, demonstrating that position 61 does not need to be changed to obtain a ps 4 denticle belt. Therefore, the six remaining positions were changed from the Ubx residue to the Antp residue in the context of the UUstop protein to generate additional chimeras with the combinations of changes described below [Fig. 5].

In addition to UAstop, one other chimeric homeo domain generated the same ps 4 transformation phenotype. This was UA1,3,23,25,57stop, with homeo domain positions 1, 3, 23, 25, and 57 all changed to the corresponding Antp residue [Fig. 5B]. Therefore, changing these five homeo domain differences is sufficient to change identity functions, demonstrating that the Met to Ile change at position 35 is not necessary. In contrast, UA1,3,23,25,57stop (with homeo domain positions 1, 3, 25, 35, and 57 all changed to the corresponding Antp residue) generated a transformation phenotype having an apparently mixed identity [Fig. 5D]. Unlike UA1,3,23,25,57stop, this latter chimera does not have the Thr → Phe change at position 23. These results point to this residue as being important for distinguishing between the identity functions of these two homeo domains.

When only the two amino-terminal homeo domain amino acids were changed [UA1,3]stop, no change in identity functions was observed [Fig. 5A,C]. A similar Ubx-like phenotype was observed with UA1,3,23]stop [Fig. 5A; data not shown]. Chimeras that changed only positions 23, 25, and 57 [UA23,25,57stop] or only positions 25, 35, and 57 [UA25,35,57stop] both generated transformation phenotypes having a mixed identity [Fig. 5A,E; data not shown]. Finally, UA1,3,23,25]stop and UA1,3,23,25]stop also generated denticle belts having a mixed identity [Fig. 5A; data not shown]. However, in a few (5/30) of the embryos transformed by UA1,3,23,25]stop, a ps 4 denticle belt was observed in place of the T1 belt. These phenotypes indicate that these subsets of changes are, by themselves, insufficient to switch identity functions consistently. Therefore, these results imply that amino acids at the amino terminus of the homeo domain [residues 1 and/or 3] and residues 23, 25, and 57 are all important for distinguishing between the segment identity functions of Ubx and Antp.

Heat shock-induced transformations are independent of endogenous HOM-C genes. Previous experiments demonstrated that the transformations resulting from either ubiquitous Ubx or UAA expression do not depend on the endogenous Ubx or Antp genes, respectively [Mann and Hogness 1990]. Here, this is tested for the UAstop protein.

![Figure 4](image-url)
Chan and Mann

**Figure 5.** The identity functions generated by proteins with chimeric homeo domains. (A) Schematic diagram of these proteins and a summary of their identity functions. All of these proteins have the identical Ubx-Ia amino-terminal region and all end with a Glu (Q) following the 61-amino acid homeo domain. For UA(t), a few of the transformed belts had a ps 4 identity, the remainder had a mixed identity. (B-E) Photomicrographs of first instar larval cuticles resulting from the ubiquitous expression of several of these chimeras. (B) UA(t,3,23,25,57)stop; (C) UA(1,gl)stop; (D) UA(1,a,25,a5,57)stop; (E) UA(25,35,57)stop. Anterior denticles are indicated as either thoracic (t) or abdominal (a) when visible. In D, denticles with an intermediate characteristics are visible (•) and in E, the denticle belts have a mixture of abdominal (a) and thoracic (t) denticles.

In addition, these results are extended by demonstrating that several of the chimeric HOM-C proteins can generate their transformations in the absence of most endogenous HOM-C genes.

First, because the UAstop chimera produced a segment identity similar to that produced by Antp, we tested whether the endogenous Antp gene plays a role in this transformation. Therefore, a P element containing the HS-UAstop gene was recombined onto a chromosome containing the Antp null mutation, Antp<sup>−/−</sup>. After ubiquitous expression of UAstop in this genetic background, approximately three-quarters of the embryos (n = 50) were transformed to a segment identity indistinguishable from that produced in a wild-type background [data not shown]. Furthermore, after heat shock induction of UAstop, no embryos were observed that had an Antp<sup>−</sup> phenotype. In addition to this genetic experiment, embryos that expressed UAstop ubiquitously were immuno-stained with two antibodies recognizing amino-terminal Antp or Ubx epitopes, respectively. Six hours after heat shock induction of the UAstop protein, all nuclei contained a protein with the Ubx epitope, which is presumably UA<sub>stop</sub> (Fig. 6A,B). In contrast, no ectopic Antp expression was observed. These results argue that the ps 4 transformation produced by the UA<sub>stop</sub> protein is independent of any endogenous Antp gene products.

We then asked whether the HOM-C proteins studied could generate their transformations in the combined absence of five endogenous HOM-C genes. Removing these genes [Sex combs reduced (Scr); Antp; Ubx; abdominal-A (abd-A); and Abdominal-B (Abd-B)] results in a novel segment identity that replaces all wild-type thoracic and abdominal segments (Struhl 1983; Sato et al. 1985) (Fig. 6C). This segment identity, which is unlike any wild-type segment, has elements of head and thoracic segments and may represent the identity that results from the lack of most HOM-C expression. If heat shock-induced expression of HOM-C proteins can generate their respective transformations in this genetic background, it would argue that the expression of individual HOM-C proteins is sufficient to generate their respective segment identities.

The ubiquitous expression of Ubx, Antp, UA<sub>stop</sub>, and UAU in Scr<sup>−</sup> Antp<sup>−</sup> Ubx<sup>−</sup> abd-A<sup>−</sup> Abd-B<sup>−</sup> embryos...
demonstrated that all of these proteins were capable of transforming these embryos on their own (Fig. 6D–F; data not shown). Furthermore, the segment identities produced were nearly identical to those generated in a wild-type background. However, whereas in a wild-type background only thoracic and head segments are transformed by these HOM-C proteins, in Scr−Antp−Ubx−abd-A−Abd-B−embryos all thoracic and abdominal segments and most head segments were transformed. Furthermore, in this genetic background it was especially apparent that Antp supports Keilin’s organ formation but that UAstop does not (Fig. 6D,E).

Effects on the expression of the Ubx target gene Dll

The Dll gene encodes a homeo domain protein required for the development of thoracic-specific Keilin’s organs in the Drosophila embryo (Sunkel and Whittle 1987; Cohen et al. 1989). To prevent Keilin’s organ formation in abdominal segments, Dll expression is repressed by Ubx and abd-A probably in a direct manner (Vachon et al. 1992). Antp, however, is expressed in thoracic segments and does not repress Dll expression. Repression of Dll by Ubx can be observed using a ubiquitous expression system (Vachon et al. 1992). If the segment identities generated by the ubiquitous expression of the chimeras studied here result from the correct regulation of downstream target genes, the expression of Dll should be regulated appropriately. Therefore, the various Ubx–Antp chimeras described above were tested for their ability to repress Dll expression in the thorax.

Following their ubiquitous expression, all proteins that generated a ps 6 denticle belt transformation (Ubx, UUA, UUstop, and UAU) also repressed Dll expression efficiently in the thorax (Fig. 7C,E,F,G). In contrast, ubiquitous expression of UAA or Antp never repressed Dll expression (Fig. 7B; data not shown). UAstop, however, was able to repress Dll expression in the majority of embryos (~80%). In the remaining embryos, weak Dll expression could be observed (Fig. 7D). Because Dll expression is necessary for Keilin’s organ formation, these results are consistent with the cuticle transformation phenotypes showing that UAstop suppresses Keilin’s organ formation more efficiently than UAA (Table 1).

Discussion

The HOM-C proteins of Drosophila control the development of specific and complex cell patterns that distinguish one segment from another. These specific segment identities are probably generated by the precise regulation of downstream target genes by HOM-C proteins. In this investigation, the generation of these complex phenotypes has been used as an assay to measure the in vivo functions of these proteins. By identifying the relevant differences between two HOM-C proteins, Ubx and Antp, we have begun to address how HOM-C proteins select and regulate the correct sets of target genes in vivo.

A similar approach was used to study chimeras between Dfd and Ubx (Kuziora and McGinnis 1989; Lin and McGinnis 1992). However, in this case the potential
for $Dfd$ to positively autoregulate the endogenous $Dfd$ gene provides an explanation for the transformations observed. Whereas wild-type $Dfd$ activates the endogenous $Dfd$ gene, the DUD chimera (for $Dfd$ amino terminus, Ubx homeo domain, and $Dfd$ C-tail) is targeted to and activates the endogenous $Antp$ gene and generates a ps 4 segment identity. Therefore, the transformations generated by ectopic expression of $Dfd$ and DUD can be accounted for by their activation of single target genes ($Dfd$ and $Antp$, respectively), which are themselves HOM-C genes. Although these chimeras can be used to distinguish between the activation of these two target genes, they cannot be used to measure a complete set of segment identity functions that presumably requires the targeting and correct regulation of many downstream genes.

In contrast, Ubx, Antp, and the various Ubx–Antp chimeras studied can generate their specific segment identities in the absence of most endogenous HOM-C genes, including $Ubx$ and $Antp$. This is consistent with the lack of evidence for positive autoregulation by Ubx and Antp. In fact, several experiments demonstrate that Ubx represses its own promoter [Irvine et al. 1993]. This lack of dependency on endogenous HOM-C genes suggests that the transformations generated by Ubx and Ubx–Antp chimeras are the result of the correct regulation of multiple downstream target genes. One such target gene is $Dll$, which we show to be regulated appropriately by the Ubx–Antp chimeras studied here. These experiments also show that despite the complexity of the segment identity assay, only a few differences between Ubx and Antp may be important for discriminating between their two sets of target genes in vivo.

**At least two regions of Ubx have independent ps 6 identity functions**

Conceptually, Ubx and Antp have been divided into three regions: their large amino termini, their homeo domains, and their short C-tails. The initial set of chimeras analyzed were generated by swapping one or two of these regions. The results of these experiments, which are summarized in Figure 8A, demonstrate that both the homeo domain and sequences carboxy-terminal to the homeo domain are important for determining the in vivo functions of these HOM-C proteins. Of the chimeras studied, this conclusion is best supported by the following pair-wise comparisons. First, whereas $UU_{stop}$ generated a ps 6 denticle belt, $UA_{stop}$ generated a ps 4 denticle belt. As the only differences between these two proteins are 6 homeo domain residues, this result demonstrates the relevance of this 61-amino acid domain for determining the in vivo properties of these proteins. Similarly, the ps 4-generating identity of UAA versus the ps 6-generat-

![Figure 7](image-url)
Figure 8. (A) A summary of the various segment identity phenotypes generated by ubiquitous expression of HOM-C and HOM-C chimeric proteins. The drawing at the far left represents the three thoracic segments (T1, T2, and T3) and the first abdominal segment (A1) of a wild-type ventral cuticle. Pairs of Keilin’s organs are shown in each of the three thoracic segments, and the different denticle belts are illustrated. For UAstop, denticle belts with some ps 6 characteristics are indicated by patches of darker shading. For UAstop and UAA, Keilin’s organs are only partially suppressed. UAallstop and UAl5stop generate identities indistinguishable from UAstop. (B) A model of the Ubx homeo domain and C-tail in contact with DNA. The view is down the axis of the DNA helix, for which only the sugar-phosphate backbone is shown. The third, or recognition, helix directly contacts DNA in the major groove, and the amino-terminal homeo domain arm contacts the minor groove. The coordinates for most of this model were taken from the engrailed-DNA crystal structure solved by Kissinger and Pabo (Kissinger et al. 1990). As the crystal and NMR structures only include homeo domain residues 4–58, the positions of the C-tail and amino-terminal most homeo domain residues [indicated with stripes] are highly speculative and are only meant to indicate their possible relation to the rest of the structure. The remainder of the Ubx protein (not shown) continues for 259 amino acids amino-terminal to Leu 1. The residues important for the functional differences between Ubx and Antp are indicated (homeo domain residue number, Ubx/Antp [homeo domain number, Ubx/Antp]; 1, Leu/Glu; 3, Arg/Lys; 23, Thr/Pro; 25 His/Arg; 57, Leu/Trp; and the Ubx/Antp C-tails COOH [34 and 23 amino acids, respectively].

Two comparisons illustrate that the carboxy-terminal 34 amino acids of Ubx (C-tail) is also important for generating a ps 6 identity in vivo. First, the ps 4-generating UAstop chimera was converted into the ps 6-generating UAU chimera by the addition of the Ubx C-tail. Second, although UAA had ps 4 identity functions, UAU generated a ps 6 identity. In this comparison, the identity functions correlate with the C-tail that follows an otherwise identical protein.

Consistent with the Ubx C-tail having important identity functions, its sequence is significantly conserved in the Ubx homologs from two distantly related invertebrates, the annelid H. medicinalis (a leech) and the arthropod B. mori (a moth). Moreover, because UAU18 also generated a ps 6 identity, these conserved amino acids contain the identity functions present in the Ubx C-tail. Conserved sequences also exist between the Antp C-tails from grasshopper and Drosophila (E. Ball and N. Patel, pers. comm.) and between the C-tails of human HOM-C proteins (Acampora et al. 1991). This sequence conservation contrasts with the lack of similarity between the C-tails of the different D. melanogaster HOM-C proteins. The one exception to this rule is a weak similarity (6/9 identities) between the Ubx and abd-A C-tails (Karch et al. 1990; Fig. 3). This similarity is intriguing because the segment identities generated by Ubx and abd-A in wild-type embryos are similar to each other and may indicate that they have overlapping identity functions.

The Ubx amino-terminal region is present in all of the chimeras studied in these experiments; therefore, although these experiments assess the role of the homeo domain and the C-tail, they do not address whether amino-terminal sequences also have identity functions. It is possible, for example, that the ps 6-generating properties of UAU depend on its amino-terminal region in addition to its C-tail. However, large deletions of the amino-terminal regions of either Ubx or Antp do not
effect the segment identities these proteins generate, suggesting that these sequences play a relatively minor role [Gibson et al. 1990; Mann and Hogness 1990]. Moreover, UAA and Antp generate very similar identities with the exception that UAA suppresses Keilin's organ formation more often than does Antp [Mann and Hogness 1990, this paper]. Curiously, however, a chimera predicted to have the amino-terminal region of abd-A fused to the Ubx homeo domain and Ubx C-tail [encoded by the Ubx<sup>C1</sup> mutation] generates a different identity than Ubx, suggesting that amino-terminal sequences can have significant identity functions (Casanova et al. 1988; Rowe and Akam 1988).

The C-tails of other HOM-C proteins are also important for HOM-C function

Although it has been demonstrated most conclusively for the Ubx C-tail, the experiments presented here show that the Antp C-tail also has significant identity functions. This is most apparent when comparing the functions of UA<sub>stop</sub> and UAA. Although UA<sub>stop</sub> generates a ps 4 denticle belt, it differs from Antp in that it often suppresses Keilin's organ formation and weakly represses Dil expression. UAA, however, is much more Antp-like with regard to both of these phenotypes. We conclude that the Antp C-tail is important for Antp function. UA<sub>stop</sub> therefore, has weak identity functions that are intermediate between Ubx and Antp. Addition of the Antp C-tail [to generate UAA] results in an Antp-like protein, whereas addition of the Ubx C-tail [to generate UAU] results in a protein with Ubx-like functions. Similarly, chimeras between Scr and Antp also suggest a role for the Antp C-tail because, in the adult, the AAS chimera [for Antp amino terminus, Antp homeo domain, and Scr C-tail] produced transformations that were different from those produced by wild-type Antp (Gibson et al. 1990). Taken together, these results indicate that sequences carboxy-terminal to the homeo domains of HOM-C proteins may, in general, have significant segment identity functions.

Mapping the homeo domain specificity determinants

As described above, UU<sub>stop</sub> and UA<sub>stop</sub> generated ps 6 and ps 4 denticle belts, respectively. As these two proteins only differ in six homeo domain positions, further chimeras were constructed to define which of these six were necessary to distinguish these functions. Several conclusions can be made from these experiments. First, amino-terminal homeo domain residues 1 or 3, or both, are necessary, but not sufficient, to change identities. This conclusion is interesting in light of the results with the Dfd–Ubx chimeras. These experiments show that changing the amino-terminal homeo domain residues from Dfd to Ubx [in this case, residues 1, 2, 3, 5, 7, and 8] was sufficient to activate the Antp gene instead of the Dfd gene [Lin and McGinnis 1992]. Although both sets of chimeras demonstrate the importance of the amino terminus of the homeo domain, our experiments show that additional homeo domain residues are also important for changing identity functions. These additional homeo domain residues are at positions 23, 25, and 57. The evidence presented suggests that all three of these residues are important in distinguishing between the identity functions of Ubx and Antp. Furthermore, their proximity in the predicted three-dimensional structure [see below] suggests that these 3 amino acids may function together.

Modeling the positions of the relevant amino acids

The existing crystal and nuclear magnetic resonance (NMR) structures of homeo domain–DNA complexes provide the opportunity to model the positions of these amino acids in three-dimensional space [Kissinger et al. 1990; Otting et al. 1990; Wolberger et al. 1991] [Fig. 8B]. Although their structures have not been solved yet, the very amino-terminal portion of the homeo domain and sequences carboxy-terminal to the homeo domain have been included in this model to illustrate where these regions might be located relative to the remaining structure.

Using this model, the residues at the very amino terminus of the homeo domain [residues 1 and 3] may directly interact with DNA because neighboring residues 4 and 6 are known to make minor groove contacts. In contrast, residues 23, 25, and 57 are solvent exposed and are therefore unlikely to contact DNA directly [Fig. 8B].

How might residues that do not directly interact with DNA affect the in vivo specificity of HOM-C proteins? One possibility is that these residues affect the specificity or affinity of the homeo domain–DNA complex by altering its conformation. This seems unlikely for the C-tail because, at least for Antp, the full-length protein binds to the same sequence as its isolated homeo domain [Mihara and Kaiser 1988; Muller et al. 1988; Affolter et al. 1990]. Furthermore, because Ubx and Antp are able to bind the same DNA sequences in vitro, differences in DNA binding attributable to differences in these residues would have to be relatively subtle. Moreover, DNA footprinting experiments using the Ubx–Antp chimeras described here demonstrate that, for at least one set of DNA-binding sites, the specificity of binding is independent of whether the homeo domain or C-tail is derived from Ubx or Antp (S.-K. Chan and R.S. Mann, unpubl.). Additional evidence comes from the three-dimensional structures of three different homeo domain–DNA complexes, which show that the structures of the homeo domains are virtually identical, despite many differences in primary sequence [Kissinger et al. 1990; Otting et al. 1990; Wolberger et al. 1991]. Finally, our in vivo experiments demonstrate that the Ubx C-tail can influence the functions of the Antp homeo domain even when 5 alanine residues separate these two portions of the protein. This result suggests that the Ubx C-tail functions as an independent protein domain and argues against the C-tail subtly modifying the structure of the homeo domain.

An alternative possibility is that residues 23, 25, 57, and the C-tail alter target gene specificity by interacting
with other proteins that may themselves be DNA-binding proteins. In addition to the evidence presented, this suggestion has strong support from other systems where factors that directly interact with homeo domain proteins have already been identified. In the determination of cell type in the yeast *Saccharomyces cerevisiae*, the α2 homeo domain protein either contacts a second homeo domain protein, a1, or a nonhomeo domain protein, MCM1 [Keleher et al. 1988;1989; Smith and Johnson 1992]. Association of these two proteins, which are themselves DNA-binding proteins, alters target gene specificity. Interestingly, the α2 sequences that mediate its association with a1 are located within the flexible tail immediately carboxy-terminal to its homeo domain, in an analogous position to the Ubx C-tail (Johnson 1993; A. Mak and A. Johnson, pers. comm.). Protein–protein interactions have also been documented between homeo domain proteins and serum response factor, a human protein related to MCM1 [Grueneberg et al. 1992], and among the POU subset of homeo domain proteins (K pistic et al. 1989; Stern et al. 1989; Treacy et al. 1991,1992). Single amino acid substitutions in the Oct-1 homeo domain have identified specific residues that are important for its interaction with one such factor, VP16/αTIF [Lai et al. 1992; Pomerantz et al. 1992]. Interestingly, the most critical residue identified in these studies is at the carboxy-terminal end of the first α-helix, in a homologous position to 1 of the 5 homeo domain residues identified here (residue 23). Therefore, by analogy to these other systems, the amino acids that we have identified may also define a protein interaction domain. However, proof that these amino acids are sites of protein–protein interactions will only come from the identification and characterization of such factors in *Drosophila*.

**Materials and methods**

**Fly stocks**

All stocks have the w^1118 mutation to monitor the presence of any CaSpeR P element [containing the mini-white gene] [Pirrotta 1988]. At least two independent P-element insertions were studied for each heat-inducible protein described. The HS–Ubx and the two HS–UAU insertions behaved as homozygous lethal loci without heat shock. All other insertion lines were homozygous viable. The UAstop(four copies) stock contained four copies of the HS–UAstop gene (two homozygous loci on the second and third chromosomes). The chromosome with the Antp–UA stop mutation contained the dominant marker Rinked [Ri]. To construct the UAstop Antp–UA stop/TM6B stock, a recombinant third chromosome was constructed that contained both Ri Antp–UA stop and the HS–UAstop gene balanced by TM6B. The stock containing the five homeotic null mutations Scr^1 Ns^Rc3 Ubx^MX12 abdA^M1 AbdB^M8 stock was constructed by recombining the double point mutation Scr^1 Ns^Rc3 [Struhl 1982] with the triple point mutation Ubx^MX12 abdA^M1 AbdB^M8 [Casanova et al. 1987]. The HS–UAstop Scr^1 Ns^Rc3 Ubx^MX12 abdA^M1 AbdB^M8 and HS–Antp Scr^1 Ns^Rc3 Ubx^MX12 abdA^M1 AbdB^M8 recombinant third chromosomes were constructed by recombining the respective heat shock loci and the quintuple mutant chromosome. To construct the HS–UAU, Scr^1 Ns^Rc3 Ubx^MX12 abdA^M1 AbdB^M8 stock, the HS–UAU gene present on the second chromosome was introduced and balanced (by CyO) in the Scr^1 Ns^Rc3 Ubx^MX12 abdA^M1 AbdB^M8/TM6B stock.

**Plasmids**

The P-element vector used was a CaSpeR vector, which is essentially the same as CaSpeR–la described in Mann and Hoggness [1990], except that the EcoRI and BamHI sites 5' to the hsp70 promoter and the BamHI site 3' to the cDNA-coding sequence were blunted by end-filling with Klenow. The resulting construct was named CaSpeR–Ia[R^1 Bam^2–].

The Ubx cDNA was derived from pl-la, which contained the Ubx–1a ORF and 3' untranslated region (2.2 kb) cloned into pUC-18 [Beachy et al. 1988]. To facilitate subcloning, the proximal SnaBl site within the 3'-untranslated sequence was converted into XbaI by insertion of a linker (dpGCTCTAGAGC). The resulting plasmid was named pl-lax. The NotI–Xbal [715 bp] fragment from pl-lax was then subcloned into pBlueScriptKS+[+] to give KSNX0.7.

The 0.9-kb Xbal fragment containing most of the homeo domain, C-tail, and 3'-untranslated region of Antp was isolated from pARAL–UA [Mann and Hoggness 1990] and subcloned into pBlueScriptKS+[+] to give SK+[+]Antp3' and SK+Antp3' respectively.

The remaining plasmids were constructed using PCR with the oligonucleotides listed in Table 2. Using KSNX0.7 as the template and oligonucleotides T7/SK2, PCR was used to generate a fragment containing the Leu → Glu mutation at homeo domain position 1 and the Arg → Lys mutation at homeo domain position 3. After digestion with Eagl and gel purification, the mutated fragment was used to replace the wild-type Eagl fragment in KSNX0.7. The NotI–Xbal fragment was isolated from this intermediate plasmid and used to replace the wild-type NotI–Xbal fragment in pl-lax to generate pUA123U.

Similarly, using KSNX0.7 as the template and oligonucleotides SK3/SK9, PCR was used to generate a fragment containing a stop codon 1 amino acid after the Ubx homeo domain. In addition, this reaction introduced an XhoI site 17 bp 3' to the stop codon. After digestion with XhoI and Xbal, the gel-purified fragment [125 bp] was used to replace the wild-type fragment in pl-lax to give pUUstop.

To construct pUA123stop, the XhoI–Xbal fragment of pUA123U was replaced by that of pUUAstop. The BamHI–EcoRI fragments from pUUstop and pUA123stop were then cloned into the CaSpeR vector to give CaSpeR–UU stop and CaSpeR–UA123stop respectively.

To change His-25 to Arg, oligonucleotides SK8 and T3 were used. The HIS–Ubx and HIS–Antp loci were isolated from pl-lax and used to replace the wild-type HIS–Ubx and HIS–Antp respectively.

To construct the HIS–UAU allele, HIS–UAU was used to replace the HIS–UAstop allele from pl-lax. To construct the HIS–UAstop allele, HIS–UAstop was used to replace HIS–UAU allele from pl-lax. To construct the HIS–UAU allele, HIS–UAU was used to replace the HIS–UAstop allele from pl-lax.
Table 2. Oligonucleotides used for PCR and cloning

| Oligonucleotide | Sequence (5' → 3') | Mutation | Direction |
|-----------------|---------------------|----------|-----------|
| SK2             | CTGTCGCGCCGCTTTCGCTCACCATTG | Glu-1, Lys-3 | bottom    |
| SK3             | GACAGACATACACCGCTACCC | wild type | top       |
| SK4             |GCCGAAGAAGCTCAGGATCGCGCGAGAGGG | Ile-35 | bottom |
| SK5             | CTCTTCTTCACCTTACCTTCCG | Ile-35 | top       |
| SK6             | GAATGACGCTGACGCGCGCC | Trp-75 | bottom    |
| SK7             | GCCGCGCTGCTACTCTTCCAT | Trp-75 | top       |
| SK8             | ACCCTCGAGCTGAAAGAGGTTCCTCCTGAATGATATCTG | Arg-25 | bottom |
| SK9             | GCTCTAAGCCTCTTCTTGATCTGATCTCCTTG | stop codon | bottom |
| SK17            | ACGCTCGAGCTGGAGAAGGAGTTCCACTTCG | Phe-23 | top       |
| SK22            | GAATGCCAATTGCCTACATTGAG | wild type | bottom |
| SK17b           | GATATCGATTGAAGTGGAACTCC | wild type | bottom |
| SK18            | AGGAGAAGCAGCGAGTACAGAGCTCAAA | Asn-61 | top       |
| SK19            | TTGAGCTCTTCTGTGCGCTGTTCCCTTCCCT | Asn-61 | bottom |
| SK20            | GAATGCAATGCTCAGATCTCCGG | wild type | top       |
| SK21            | ATCTAAAGATCCTAAGGACTAAG | wild type | bottom |
| SK31            | CGCTCTAGACTACTGCGCTTT | stop codon | bottom |
| SK40            | AATCTGCGCTCCTTCCAC | wild type | bottom |
| SK50            | AAGAAGCAGCGAGTACAGAGCTCAAA | [ala]3 insert | top |
| RM14            | GATCTGGTTCCAGAACCGGCGAATCTGAAGAAGGAGATC | Ubx hd/Antp fusion | stranded |

*Top and bottom* From which DNA strand the oligonucleotide sequence is derived. RM14 was double stranded, only the top strand is shown.

2.2-kb Ubx cDNA in CaSpeR-iaR1–, which was a derivative of CaSpeR-ia having the EcoRI site 5' to the hsp70 promoter blunted with Klenow.

To make CaSpeR-UA[1,2,3,25,35]stop, the 251-bp BgII fragment from CaSpeR-UA[1,2,3,25,35]stop was used to replace the equivalent fragment in CaSpeR-UUstop. To make pUA[1,2,3,23]stop oligonucleotides SK17 and SK9 were used for PCR with CaSpeR-UA[1,2,3]stop as the template. The resulting PCR product was digested with XhoI and XbaI and cloned into pUA[1,2,3,25,35,57]stop.

To make the remaining Ubx–Antp chimeras, a two-round PCR strategy was used. In the first round of PCR, two separate reactions (primary-1 and primary-2) were performed (Table 3). After 25 cycles, the products were digested with XhoI and XbaI, gel purified, and subcloned. Products from PCR reactions 1–4 were used to replace the wild-type fragment in pl-1ax to generate pUA[25,35,57]stop, pUA[25,35,37]stop, pUA[25,35,57]stop, and pUA[24,25,35,57]stop, respectively (Table 3). Fragments from reactions 3–7 were used to replace the wild-type fragment in pUA[1,a,23]stop to give pUA[1,a,23,25,35,57]stop, pUA[1,a,23,25,35,57]stop, pUA[1,a,23,25,35,57]stop, pUA[1,a,23,25,35,57]stop, pUA[1,a,23,25,35,57]stop, pUA[1,a,23,25,35,57]stop, and pUA[1,a,23,25,35,57]stop, respectively. Unlike the UAA construct described previously in Mann and Hogness (1990), none of these ORFs depend on the removal of an intron for their expression.

pUAU18 was constructed by (1) a PCR reaction with SK3 and

Table 3. Summary of PCR reactions to generate Ubx/Antp chimeras

| PCR template | oligonucleotides (top/bottom) | Primary 1 | oligonucleotides (top/bottom) | Secondary oligonucleotides (top/bottom) | Product (Xho/Xba) |
|--------------|-------------------------------|-----------|-------------------------------|----------------------------------------|-------------------|
| KS[25]       | T7/SK4                        | KS[25]    | SK5/SK9                       | T7/SK9                                 | [25,35]stop       |
| pUA[25,35,37]stop | SK3/SK6                    | KS[25]    | SK5/SK9                       | T7/SK9                                 | [25,35]stop       |
| pUA[25,35,57]stop | SK3/SK17b                  | pUA[25,35,37]stop | SK22,SK9                      | SK22,SK9                               | [23,25,35,57]stop |
| pUA[23,25,35,57]stop | SK3/SK4                    | pUA[25,35,57]stop | SK5,SK9                       | SK4,T7                                 | [23,25,35,37,57]stop |
| pUA[23,25,35,57]stop | SK3/SK5                    | SK + [Antp3'] | SK18,T3                       | SK3,T3                                 | [23,25,35,37,57]stop |
| pUAU      | SK3/SK5                      | SK3/SK17b | KSNX0.7                       | SK18,T3                                 | [23,25,35,37,57]stop |
| SK3/SK5    | SK3/SK19                      | SK3/SK5    | SK0/SK40                      | SK3/SK40                                | [ala]3 Utail      |

808 GENES & DEVELOPMENT

Cold Spring Harbor Laboratory Press on July 18, 2018 - Published by genesdev.cshlp.org Downloaded from genesdev.cshlp.org on July 18, 2018 - Published by Cold Spring Harbor Laboratory Press
SK31 using pUAU as the template, [2] digesting the PCR product with XhoI and XbaI, and [3] replacing the homologous XhoI–XbaI fragment in pUAU. pUA(ala)sU was constructed using a two-step PCR protocol (Table 3, reaction 8). The PCR product was digested with XhoI and XbaI and used to replace the homologous fragment in pUAU18. The BamHI–EcoRI fragments from pUAU18 and pUA(ala)sU were cloned into CaSpeR–la[Ref]BamHI to generate CaSpeR–la–UAU18 and CaSpeR–la–UA(ala)sU.

All mutagenized fragments were sequenced immediately after their initial subcloning and again after their introduction into the CaSpeR P-element vector to ensure that only the correct mutations were present.

Conditions for PCR

For site-directed mutagenesis, the PCR conditions used were essentially the same for all constructs. The reaction volume was 50 μl with 50 ng of DNA template and 0.2 μl of each oligonucleotide. The denaturing step was at 94°C for 5 min, the renaturation step was at 37°C for 1 min, and the extension step was at 72°C for 1 min. To confirm that the correct ORF sequence was present in each fly transformant, the relevant sequences were amplified from the genome also using PCR. For these reactions, ~200 ng of genomic DNA was used in a 50-μl reaction. The renaturation step was at 45°C for 2 min, and the extension step was at 72°C for 2 min. The oligonucleotides used were SK20/SK21 for constructs containing the Ubx C-tail or SK20/SK40 for all others. As SK20 is derived from the first Ubx exon (~70 kb 5' to the homeo domain in the endogenous Ubx gene), only Ubx sequences in the P element were amplified. The amplified fragments were gel purified and sequenced directly without subcloning using standard dideoxynucleotide methods (Sambrook et al. 1989).

Histochemistry and Western analysis

Monoclonal antibodies FP3.38 (White and Wilcox 1985) for Ubx and 8C.11 for Antp (kindly provided by D. Brower, University of Arizona, Tucson) were used in whole-mount histochemistry experiments. Embryo fixation and staining was done as described in Patel et al. (1989). The DII stains were done as described (Vachon et al. 1992), with antibody kindly provided by S. Cohen (EMBL, Heidelberg, Germany). Three-hour embryo collections were aged for 3 hr, heat-shocked for 45 min, and allowed to recover for 6 hr before fixation. Only embryos still retracting their germ bands were scored; older embryos were ignored because they probably received the heat shock too late to affect DII expression. Antibody staining was visualized with horseradish peroxidase chemistry using the Elite Vectastain ABC kit (Vector Lab).

For Western blots, embryos were collected overnight, heat shocked at 37°C for 1 hr, and dechorionated with 66% bleach. The embryos were then homogenized in an equal volume of 2× sample buffer [Gavis and Hogness 1991]. The soluble fraction (10 μl) was loaded on a 12% SDS-PAGE gel. After transfer to nitrocellulose, the heat shock-induced proteins were detected by the Ubx monoclonal antibody FP3.38 and the signal was visualized by the Immunoblot Assay kit from Biorad.

Larval cuticle preparation

Heat shocks were done for 45 min at 37°C, exactly as described in Mann and Hogness (1990). Embryos were then dechorionated with 66% bleach, devitellinized by hand, and prepared for microscopy as described [Wieschaus and Nusslein-Volhard 1986].

Functional specificity of homeotic proteins

Ubx–DNA model

Figure 8B was generated using the engrafted–DNA coordinates kindly provided by C. Kissinger and C. Pabo [Kissinger et al. 1990], replacing the engrafted residues with the Ubx residues, imaging the structure with the software MOLSCRIPT [Kraulis 1991] on an Iris workstation and later modifying it with CANVAS on an Apple Macintosh.

Acknowledgments

We thank P. Feinstein, G. Morata, M. Scott, and G. Struhl for fly stocks, D. Brower for the Antp antibody, and C. Kissinger and C. Pabo for the engrafted–DNA coordinates. We are very grateful to B. Cohen and S. Cohen for their gift of the Dll antibody and for their help with the antibody stains. We are also indebted to A. Aggarwal, S. Hubbard, J. Hirsch, A. Palmer, and B. Weiss for their help in generating the three-dimensional model of the Ubx–DNA complex and E. Ball, A. Johnson, W. McGinnis, and N. Patel for sharing their unpublished results. Finally, we thank A. Aggarwal, R. Axel, P. Feinstein, L. Jaffe, B. Konforti, and G. Struhl for valuable discussions and their comments on this manuscript. This work was supported by a Searle Scholars, Hirsh, and National Institutes of Health (HD27986) grants to R.S.M.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Acampora, D., A. Simeone, and E. Boncinielli. 1991. Human HOX homeobox genes. Oxf. Surv. Eukaryotic Genes 7: 1–28.
Affolter, M., S.A. Percival, M. Muller, W. Leupin, and W.J. Gehring. 1990. DNA binding properties of the purified Antennapedia homeodomain. Proc. Natl. Acad. Sci. USA 87: 4093–4097.
Beachy, P.A., M.A. Krasnoff, E.R. Gavis, and D.S. Hogness. 1988. An Ultrabithorax protein binds sequences near its own and the Antennapedia P1 promoters. Cell 55: 1069–1081.
Billeter, M., Y. Qian, G. Otting, M. Muller, W.J. Gehring, and K. Wuthrich. 1990. Determination of the three-dimensional structure of the antennapedia homeodomain from Drosophila in solution by 1H nuclear magnetic resonance spectroscopy. J. Mol. Biol. 214: 183–197.
Casanova, J., E. Sanchez-Herrero, A. Busturia, and G. Morata. 1987. Double and triple mutant combinations of the bithorax complex of Drosophila. EMBO J. 6: 3103–3109.
Casanova, J., E. Sanchez-Herrero, and G. Morata. 1988. Developmental analysis of a hybrid gene composed of parts of the Ubx and abd-A genes of Drosophila. EMBO J. 7: 1097–1105. 
Cohen, S.M., G. Bronner, F. Kuttner, G. Jurgens, and H. Jackle. 1989.1081-1090.
Dessain, S., C.T. Gross, M.A. Kuziora, and W. McGinnis. 1992. Antp-type homeo domains have distinct DNA binding specificities that correlate with their different regulatory func-
Ekker, S.C., K.E. Young, D.P. von Kessler, and P.A. Beachy. 1992. Differential DNA sequence recognition is a determinant of specificity in homeotic gene action. *EMBO J.* 11: 4059–4072.

Ekker, S., D. von Kessler, and P. Beachy. 1992. Homoeo boxes in the study of development. *Science* 236: 1245–1252.

Gibson, W.J. 1987. Homeo boxes in Drosophila. *Nature* 332: 858–861.

Garcia-Bellido, A. 1977. Homeotic and atavistic mutations in insects. *Am. Zool.* 17: 613–630.

Gavis, E.R. and D.S. Hogness. 1977. Phosphorylation, expression and function of the Ultrabithorax protein family in *Drosophila melanogaster*. *Development* 112: 1077–1093.

Hanes, S.D. and R. Brent. 1989. DNA specificity of the bicoid homeodomain. *Cell* 59: 563–571.

Hases, S.D. and R. Brent. 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* 57: 1275–1283.

Hayashi, S. and M.P. Scott. 1990. What determines the specificity of the Ubx product interactions with other homeotic products. *Cell* 61: 515–522.

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

Irvine, K.D., J. Botas, S. Jha, R.S. Mann and D.S. Hogness. 1993. Negative autoregulation by Ultrabithorax controls the level and pattern of its expression. *Development* 117: 387–399.

Johnson, A. 1993. A combinatorial regulatory circuit in budding yeast. *In Transcriptional regulation* (ed. K. Yamamoto and S. Kristie, T.M., J.H. LeBowitz, and P.A. Sharp. 1989. The octamer-binding proteins form multiprotein-DNA complexes with the HSV-1 TF regulatory protein. *EMBO J.* 8: 4229–4238.

Kuziora, M.A. and W. McGinnis. 1988. Autoregulation of a Drosophila homeotic selector gene. *Cell* 59: 477–485.

Kuziora, M.A. and W. McGinnis. 1989. A homeo domain substitution changes the regulatory specificity of the Deformed protein in Drosophila embryos. *Cell* 59: 563–571.

Lai, J.-S., M. Cleary, and W. Herr. 1992. A single amino acid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeo domains. *Genes & Dev.* 6: 2058–2065.

Laughon, A. and M.P. Scott. 1984. Sequence of a *Drosophila* segmentation gene: Protein structure homology with DNA-binding proteins. *Nature* 310: 25–31.

Levine, M. and T. Hoey. 1988. Homeobox proteins as sequence-specific transcription factors. *Cell* 55: 537–540.

Lewis, E.B. 1963. Genes and developmental pathways. *Am. Zool.* 3: 33–56.

—. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* 276: 565–570.

—. 1982. Control of body segment differentiation in *Drosophila* by the bithorax gene complex. *Proc. Clin. Biol. Res.* 85: 269–288.

Lin, L. and W. McGinnis. 1992. Mapping functional specificity in the Dfd and Ubx homeodomains. *Genes & Dev.* 6: 1071–1081.

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

Mihara, H. and E.T. Kaiser. 1988. A chemically synthesized Antennapedia homeo domain binds to a specific DNA sequence. *Science* 242: 925–927.

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

Otting, G., Y.Q. Qian, M. Billerter, M. Muller, A. Affolter, W.J. Gehring, and K. Wuthrich. 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *EMBO J.* 9: 3085–3092.

Pomerantz, J., T. Kristie, and P. Sharp. 1992. Expression of engrafted proteins in Arthropods, Annelids, and Chordates. *Cell* 58: 955–968.

Prirnottt, V. 1988. Vectors for P-mediated transformation in *Drosophila*. In *Vectors, a survey of molecular cloning vectors and their uses* [ed. R.L. Rodriguez and D.T. Denhardt], pp. 437–456. Butterworth, Boston, MA.

Pombo, J.T., J. Kristie, and P. Sharp. 1992. Recognition of the surface of a homeo domain protein. *Genes & Dev.* 6: 2047–2057.

Rowe, A. and M. Akam. 1988. The structure and expression of a hybrid homeotic gene. *EMBO J.* 7: 1107–1114.

Sambrook, J., E.F. Fritsh, and T. Maniatis. 1989. *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sato, T., P. Hayes, and R. Denell. 1985. Homeosis in *Drosophila*: Roles and spatial patterns of expression of the Antennapedia and Sex combs reduced loci in embryogenesis. *Dev. Biol.* 111: 171–192.

Schneuwly, S., R. Kleemen, and W.J. Gehring. 1987a. Redesigning the body plan of *Drosophila* by ectopic expression of the...
Functional specificity of homeotic proteins

Schnauwly, S., A. Kuroiwa, and W.J. Gehring. 1987b. Molecular analysis of the dominant homeotic Antennapedia phenotype. EMBO J. 6: 201–206.

Scott, M.P., J.W. Tamkun, and G. Hartzell. 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta. 989: 25–48.

Smith, D. and A.D. Johnson. 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.

Stern, S., M. Tanaka, and W. Herr. 1989. The Oct-1 homeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. Nature 341: 624–630.

Struhl, G. 1982. Genes controlling segmental specification in the Drosophila thorax. Proc. Natl. Acad. Sci. 79: 7380–7384.

—. 1983. Role of the esc+ gene product in ensuring the selective expression of segment-specific homeotic genes in Drosophila. J. Embryol. Exp. Morphol. 76: 297–331.

Sunkel, C.E. and J.R.S. Whittle. 1987. Brista: A gene involved in the specification and differentiation of distal cephalic and thoracic structures in Drosophila melanogaster. Wilhelm Roux’s Arch. Dev. Biol. 196: 124–132.

Thali, M., M.M. Muller, M. DeLorenzi, P. Mathias, and M. Bienz. 1988. Drosophila homeotic genes encode transcriptional activators similar to mammalian OTF-2. Nature 335: 598–600.

Treacy, M.N., X. He, and M.G. Rosenfeld. 1991. I-POU: A POU-domain protein that inhibits neuron-specific gene activation. Nature 350: 577–584.

Treacy, M.N., L.I. Neilson, E.E. Turner, X. He, and M.G. Rosenfeld. 1992. Twin of I-POU: A two amino acid difference in the I-POU homeodomain distinguishes an activator from an inhibitor of transcription. Cell 68: 491–505.

Treisman, J., P. Gonczy, M. Vashishtha, E. Harris, and C. Desplan. 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. Cell 59: 553–562.

Ueno, K., C.C. Hui, M. Fukuta, and Y. Suzuki. 1992. Molecular analysis of the deletion mutants in the E homeotic complex of the silkworm Bombyx mori. Development 114: 555–563.

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

White, R.A.H. and M.E. Akam. 1985. Contrabithorax mutations cause inappropriate expression of Ultrabithorax products in Drosophila. Nature 318: 567–569.

White, R.A.H. and M. Wilcox. 1985. Distribution of Ultrabithorax proteins in Drosophila. EMBO J. 4: 2035–2043.

Wieschaus, E. and C. Nusslein-Volhard. 1986. Looking at embryos. In Drosophila, a practical approach [ed. D.B. Roberts], pp. 199–228. IRL Press Limited, Oxford, England.

Wolberger, C., A.K. Vershon, B. Liu, A.D. Johnson, and C.O. Pabo. 1991. Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. Cell 67: 517–528.

Wysocka-Diller, J.W., G.O. Aisenberg, M. Baumgarten, M. Levine, and E.R. Macagno. 1989. Characterization of a homologue of bithorax-complex genes in the leech Hirudo medicinalis. Nature 341: 760–763.
The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences.

S K Chan and R S Mann

*Genes Dev.* 1993, 7:  
Access the most recent version at doi:10.1101/gad.7.5.796