The homeotic selector (HOM) genes of Drosophila melanogaster and the related mammalian Hox genes encode DNA binding transcription factors that specify developmental fates of different body segments by differentially regulating the activity of downstream target genes. A central question is how the HOM proteins achieve their developmental specificity despite the very similar DNA binding specificities of isolated HOM proteins in vitro. Specificity could be achieved by differential interactions with protein cofactors. The extradenticle gene might encode such a cofactor because it interacts genetically in parallel with Ultrabithorax, abdominal-A, and perhaps other HOM genes. By using a yeast two-hybrid system, we demonstrate selective interaction of the extradenticle homeodomain protein with certain Ultrabithorax and abdominal-A proteins but not with an Antennapedia protein or a more distant homeodomain protein. Strong interaction with Ultrabithorax proteins requires only the Ultrabithorax homeodomain and a 15-residue N-terminal extension that includes Tyr-Pro-Trp-Met (YPWM), a tetrapeptide motif found near the homeodomain in most HOM proteins and their mammalian Hox counterparts. The size and sequence of the region between the YPWM element and the homeodomain differ among Ultrabithorax isoforms, and this variable region appears to affect the interaction detected in the assay.

The homeotic selector (HOM) genes of Drosophila melanogaster and the related mammalian Hox genes encode DNA binding transcription factors that are expressed in and specify the identities of different body regions along the anterior–posterior axis of the developing animal (1). For example, the Antennapedia (Antp) gene of Drosophila controls development of the winged mesothorax while the development of the neighboring metathorax with its small wing-like halteres is specified by the Ultrabithorax (Ubx) gene and development of anterior abdominal segments is under abdominal-A (abd-A) gene control. Mutations in these genes lead to homeotic transformations in which a particular segment or body region develops as the likeness of another.

HOM and Hox genes are clustered in the genome, and each contains a conserved homeobox sequence encoding a 61-residue motif, called the homeodomain, which is the DNA binding domain of the proteins (2). Homeodomains are also found in hundreds of other known or suspected transcription factors, but HOM and Hox proteins compose a distinct functional and structural group with a greater degree of sequence similarity in the homeodomain (3). The HOM/Hox family can be further subdivided by the sequence of the homeodomain, with Antp, Ubx, Abd-A, and their mammalian Hox6, -7, and -8 cognates showing ≈90% sequence identity in this region. Although some HOM and Hox proteins show additional similarities outside the homeodomain with their cognates in other species, the only other common feature of the HOM/Hox group (except for Abdominal-B and its cognates) is the tetrapeptide Tyr-Pro-Trp-Met (YPWM) or a closely related sequence that in general is located in the region just N-terminal to the homeodomain (3). The function of the YPWM motif is unknown. It does not appear to influence DNA binding specificity in vitro (4); it has been speculated to play a role in interactions with other proteins (5, 6). Interestingly, for several HOM genes including Ubx and Antp, the region between the YPWM motif and the homeodomain differs in different protein isoforms as a result of alternative RNA splicing (see ref. 7).

While it is becoming clear that the distinct developmental identities specified by the various HOM genes result from their differential regulation of target gene expression, the mechanisms by which they achieve this differential regulation are not known (8, 9). It is significant in this regard that several isolated HOM proteins bind to very similar or identical DNA sequences in vitro (10, 11). While some of their functional specificity appears to result from different transcriptional effects of these proteins acting through the same sites (11, 12), it has also been widely suggested that differential interactions with protein cofactors might enhance the target gene specificity of the HOM proteins. Several distantly related homeodomain proteins use protein cofactors to enhance target specificity. One of the best understood is the yeast homeodomain protein MATa2, a critical regulator of cell type that recognizes different target sequences depending on whether it is expressed with the homeodomain protein a1 or the MADS box protein MCM1 (6, 13).

The extradenticle (exd) gene product (EXD) is a candidate for such a HOM protein cofactor because reduction in exd function causes homeotic transformations similar to those caused by mutations in the HOM genes (14). And, unlike all other known genes that can cause homeotic phenotypes (aside from the HOM genes themselves), exd is not a regulator of HOM gene expression nor does its function depend on its regulation by HOM genes. Thus, exd and the HOM genes act in parallel to regulate segmental identity. The finding that exd encodes a distantly related homeodomain protein, very similar to DNA binding proteins of the pbx-1 protooncogene family, strongly suggested that EXD itself is a DNA binding transcription factor and raised the possibility that EXD might interact directly with HOM proteins in the regulation of target gene expression (15). Consistent with this idea, exd and certain HOM genes are required for the regulation of some of the same target genes and in one case were shown to function through the same enhancer element (16). Here we test for physical association of EXD with Ubx proteins (UBX) and
other HOM proteins by using a yeast two-hybrid system (17, 18). We demonstrate highly selective interaction of EXD with certain UBX isoforms and with the abd-A protein (ABD-A), and we map UBX sequences critical for the interaction to UBX-specific sequences in the homeodomain and to a 15-residue N-terminal extension of the homeodomain that includes the YPWM sequence. The region between YPWM and the homeodomain, which varies among different UBX isoforms, appears to influence the interaction.

MATERIALS AND METHODS

Plasmids. Vectors used to express hybrid proteins were pAS2 (18) [for GAL4 DNA binding domain (DBD) hybrids] and pGAD-GH [for GAL4 activation domain (ACT) hybrids; constructed by G. Hannon, Cold Spring Harbor Laboratory]. Full-length EXD, ABD-A, eve protein (EVE), and Antp protein (ANTP) hybrids and the ANTP254–378 hybrid were constructed from exd cdNA td48 (15), pPacABD-A [provided by S. Sakonju (12)], pAR-EVE (19), and pET3a-ANTPII (a derivative of Antp cdNA G1100) (20), respectively. The various UBX hybrids were constructed by polymerase chain reaction (PCR) using custom oligonucleotide primers and plPUBX plasmids (21) corresponding to the different UBX isoforms as templates. The ANTP, EVE, and EXD homeodomains and UBX NH2A were constructed similarly using pPacANTP (22), pAR-EVE, exd cdNA td48, and pUA2A (23) as templates, respectively. UBX NH2A was made using a two-round PCR strategy (23) with Ubx and Scr cdNA sequences. The EXD ΔHD hybrid was generated by a similar strategy using exd cdNA sequences. The IaAAAM mutant was generated using a custom 5′ PCR primer in which the codons for the YPWM amino acid sequence were all changed to alanine codons. The sequences of the coding regions of all clones generated by PCR and all new junctions created in subcloning were confirmed by chain-termination sequencing. Details of the constructions are available upon request from the authors. The c-myc hybrid construct pMTADZ was kindly provided by S. Hays and P. Berg (Stanford University).

Yeast Transformations, β-Galactosidase Assays, and Immunoblot Analysis. DBD-hybrid and ACT-hybrid plasmids were cotransformed into the Saccharomyces cerevisiae Y190 (18) reporter strain by using a standard lithium acetate procedure (24) with minor modifications. β-Galactosidase activities were determined in crude extracts (24) of logarithmic-phase cultures grown in selective liquid medium; two extracts were prepared from cultures of each of three individual yeast transformants for each recombinant construct (n = 6), except where noted. One unit of β-galactosidase activity is 1 nmol of o-nitrophenyl β-D-galactoside hydrolyzed per min per mg of extract protein (24).

The DBD-hybrid proteins contain a hemagglutinin epitope tag. Their expression in transformant yeast strains was analyzed on immunoblots by using the anti-hemagglutinin antibody 12CA5 (Boehringer Mannheim) and enhanced chemiluminescence detection. DBD-hybrid proteins of the expected size were expressed at similar low levels for all constructs except the UBX ΔHΔ and UBX ΔHIC, which were more abundant. Also, UBX236–389, UBX NΔA, ABD-A, ANTP, and MYC were below the limit of detection; however, each of these showed an interaction with EXD or stimulated reporter gene activity on its own, indicating that functional protein was made.

RESULTS

In the yeast two-hybrid system a transcriptional activation domain is fused to one test protein and a DNA binding domain is fused to the other; if the test proteins bind one another, the activation domain is effectively tethered to the DNA binding domain and thus can activate a reporter gene containing sites for the DNA binding domain. A UBX protein (UBX236–389) lacking the N-terminal transcriptional activation domain but containing the homeodomain and flanking sequences was expressed as a fusion with the DBD and tested for interaction with full-length EXD expressed as a fusion with an ACT (Fig. 1a). (The corresponding full-length UBX constructs were toxic to yeast and could not be tested.) High-level activity of a reporter gene containing a DNA binding site for GAL4 was observed only in the presence of both hybrid proteins (Fig. 1b). A similar effect was observed when the reciprocal hybrids were tested (Fig. 1b) and when the hybrid proteins were expressed in another yeast strain (SFYS27) (26) containing a different GAL4 reporter gene (data not shown). The results indicate that the two hybrid proteins interact to form an active GAL4 regulator, and because the UBX and EXD portions of the hybrids are specifically required for the effect (see below), we conclude that the interaction is mediated by the physical association of UBX and EXD. Furthermore, the interaction does not appear to require any specific sequences in the target gene except the GAL4 DNA binding site.

EXD shows striking selectivity in its interactions with Drosophila homeodomain proteins. Like UBX, ABD-A interacted with EXD in the same assay (Fig. 2). In contrast, a full-length ANTP did not interact significantly, nor did a deletion derivative lacking the N-terminal activation domain (ANTP254–378), nor did the ANTP homeodomain alone. The Drosophila homeodomain protein EVE is the product of the even-skipped (eve) segmentation gene that helps establish HOM gene

---

**FIG. 1.** Interaction between UBX and EXD in a yeast two-hybrid system. (a) Structures of a hybrid protein with the DBD fused to the C-terminal half of UBX (DBD-UBX) and of a second hybrid with the ACT fused to full-length EXD (ACT-EXD). The DBD-UBX hybrid contains residues 236–389 of the UBX Ia isoform (see Fig. 4), which includes the homeodomain (H) and N-flanking (N) and C-flanking (C) regions; the major transcriptional activation domains have been removed. Residue numbering for UBX is as described (25). The ACT-EXD hybrid contains the entire 376-residue EXD; the position of its 64-residue homeodomain (H) is shown. (b) β-Galactosidase (β-gal) activity of extracts of S. cerevisiae Y190 transformants (n = 6; mean and SEM are shown) carrying a UASG-lacZ reporter gene and plasmids expressing the hybrid proteins shown in a or hybrid proteins containing the reciprocal fusions with EXD fused to DBD and UBX fused to ACT. (+), Expressor plasmid for ACT alone without an adjoining fusion; (−), no DBD expressor plasmid. The ACT–UBX fusion used here contains residues 236–354 of UBX (like UBX NH2Δ in Fig. 3).
expression and appears not to require exd function in vivo. EVE did not interact with EXD in the assay, nor did the EVE homeodomain or an unrelated c-myc protein fragment.

We asked which parts of UBX are required for interaction with EXD (Fig. 3a). No interaction was detected with the UBX homeodomain alone (UBX AHΔ), and only a low-level interaction was observed with a derivative containing the homeodomain and C-terminal sequences (UBX AHC). A derivative containing the homeodomain and 49 N-flanking residues (UBX NHΔ), however, interacted as well as the original UBX construct (called UBX NHC in Fig. 3a). Both the N-flanking region and the homeodomain are critical for the interaction with EXD, as deletion of the homeodomain leaving the N-flanking region alone (UBX NDA) or substitution of the homeodomain sequence with that of ANTP (UBX NHΔA) abolished the interaction. Since the ANTP homeodomain differs from that of UBX at only 7 of 61 positions (Fig. 3b), the precise sequence of the homeodomain is critical for the interaction with EXD. A similar result was obtained with substitution of the homeodomain of another HOM protein, the Sex combs reduced gene product, which differs from the UBX homeodomain at 11 positions (UBX NHΔ). The homeodomain of EXD also appears to be required but insufficient for binding to UBX as the EXD homeodomain alone (EXD ΔHD) and an EXD derivative lacking just the baking the domain (EXD ΔHDΔ) both failed to show significant interaction with UBX in the assay (Fig. 3c). This last conclusion should be considered tentative, however, because we have not been able to confirm expression of EXD–ACT hybrids by immunoblot analysis.

The critical 49-residue region flanking the UBX homeodomain includes the region that distinguishes the five UBX protein isoforms that arise from differential RNA splicing (25). It also includes the YPWM motif. The isoforms differ in the combination of three short (9–17 residues) optional elements, called the b element, ml, and mlI, located between the YPWM element and the homeodomain as shown in Fig. 4a. The UBX isoforms have different temporal and tissue distri-
Fig. 3. Requirement for the homeodomains and the flanking regions in the interaction between UBX and EXD. (a) Experiments were as in Fig. 2 except the DBD-UBX hybrid was modified by precise removal or replacement of the homeodomain (H) or the N-flanking (N) or C-flanking (C) regions shown in Fig. 1a. NH, original unmodified DBD-UBX hybrids; Δ, the corresponding region in UBX was removed; Hβ and Hδ, replacement of the UBX homeodomain with the 61-residue ANTP and SCR homeodomains, respectively. (b) Comparison of homeodomain amino acid sequences of the HOM proteins UBX, ABD-A, ANTP, and SCR. Residues different from UBX are indicated. Residues that distinguish UBX from ANTP are marked with symbols: solid squares, residues on the surface of the DNA-bound homeodomain; open circle, a residue in the hydrophobic core of the homeodomain; open squares, residues that are structurally disordered or for which high-resolution structural information is not available (27). (c) Experiments were as in a except the ACT-EXD hybrid was modified to contain just the 64-residue EXD homeodomain (EXD HD) (see ref. 15) or the entire EXD protein with the homeodomain precisely removed (EXD ΔHD). The data from a with the unmodified full-length EXD are included for comparison.

The 15-residue N-terminal extension of the UBX homeodomain contains a YPWM sequence that is critical for the interaction with EXD (Fig. 4). The YPWM motif has long been noted in HOM and Hox proteins. It is generally located in a region just upstream of the homeodomain and within a region that is highly conserved among closely related HOM genes in different species. Our results suggest that the YPWM element defines an additional site of contact with the EXD cofactor. There is a close analogy between the UBX–EXD interaction and that of the yeast MATα2 homeodomain protein with its cofactor MCM1. The latter interaction also requires (in MATα2) a short N-terminal extension of the homeodomain (6). Proteolysis studies indicate that this N-terminal extension is flexible (31). Although the analogous region of UBX has not been analyzed, NMR studies of an ANTP fragment containing the homeodomain and a 13-residue N-terminal extension (including YPWM) show that it too is unstructured (5). Thus, N-flanking regions of the homeodomain may in general be disordered but disposed to interact with cofactors. However, unlike the MATα2 N-terminal extension, which can confer upon a heterologous homeodomain the ability to interact with MCM1 (6), the UBX N-terminal extension is unable to confer interaction with EXD on even the highly similar ANTP homeodomain (Fig. 3).

The YPWM element is located 7 residues N-terminal to the homeodomain in the UBX IVA isoform, but in the other isoforms the region separating the YPWM element and the homeodomain is increased in length and variable in sequence. Isoforms containing one or both of the 17-residue I and II elements interact at least as well with EXD, but the two isoforms containing the 9-residue b element do not interact with EXD in our assay. A simple interpretation is that the b element impairs the interaction between UBX and EXD, for example by perturbing the neighboring YPWM element. But because of the activation function of the b element, more complicated interpretations are possible. For example, the
which is detected in the assays used so far. If ANTP does not interact with EXD, what is the role of its YPWM element? Perhaps YPWM is not an element dedicated solely to interactions with EXD but is also involved in interactions with other as yet unidentified cofactors. The demonstration that HOM proteins can interact selectively with EXD in the yeast two-hybrid system indicates that it may be possible to use this assay to screen for other cofactors.

We are grateful to C. Rauskolb and E. Wieschaus for providing the exd cDNA and for discussions. We thank P. Bartel, S. Elledge, D. Fiorentino, S. Hays, J. Lopez, A. McCormick, and W. Zeng for plasmids and strains. We also thank J. Manak and the members of this laboratory for stimulating and helpful discussions. F.B.J. is supported by the Medical Scientist, A. (1994) EMBO J. 13, 3551-3556. supported by a National Science Foundation President Young Investigator Award. This work was supported by a grant from the National Institutes of Health to M.A.K.