Hox proteins control developmental patterns and cell differentiation in vertebrates by acting as positive or negative regulators of still unidentified downstream target genes. The homeodomain and other small accessory sequences encode the DNA-protein and protein-protein interaction functions which ultimately dictate target recognition and functional specificity in vivo. The effector domains responsible for either positive or negative interactions with the cell transcriptional machinery are unknown for most Hox proteins, largely due to a lack of physiological targets on which to carry out functional analysis. We report the identification of the transcriptional activation domains of three human Hox proteins, HOXB1, HOXB3, and HOXD9, which interact in vivo with the autoregulatory and cross-regulatory enhancers of the murine Hoxb-1 and human HOXD9 genes. Activation domains have been defined both in a homologous context, i.e., within a Hox protein binding as a monomer or as a Hox-Pbx heterodimer to the specific target, and in a heterologous context, after translocation to the yeast Gal4 DNA-binding domain. Transfection analysis indicates that activation domains can be identified in different regions of the three HOX proteins depending on the context in which they interact with the DNA target. These results suggest that Hox proteins may be multifunctional transcriptional regulators, interacting with different cofactors and/or components of the transcriptional machinery depending on the structure of their target regulatory elements.

For the mammalian system, and more generally for vertebrates, genetic analysis has so far failed to identify the downstream targets of Hox gene function in development or cell differentiation. Analysis of mutant Drosophila embryos showed that in some circumstances, the products of mammalian Hox genes can substitute for the function of the corresponding Drosophila proteins, but it provided no clues about the target genes which are regulated in such a heterologous context (22, 24, 25, 30, 48). Transgenic-mouse analysis, on the other hand, has led to the identification of only a few autoregulatory and cross-regulatory elements within the Hox clusters (14, 23, 32, 33, 43). The lack of “true” target genes and therefore of bona fide Hox-responsive sequences is the single most important factor which has so far limited the analysis of the transcriptional functions of vertebrate Hox proteins. Despite this limitation, a few studies showed the existence of domains with positive or negative transcriptional activity in some Hox proteins, as defined by their ability to regulate the activity of reporter sequences in vitro or in vivo (9, 17, 34, 38, 44, 47). In this respect, Hox proteins thus seem to share the modular structure of most eukaryotic transcription factors, featuring separate DNA-binding domain (DBD) and effector domain.

In this paper, we report the identification of functional domains in the human HOXD9, HOXB1, and HOXB3 proteins, which interact with and activate transcription from the autoregulatory and cross-regulatory elements of the human HOXD9 and murine Hoxb-1 genes (32, 43). Transcriptional activation domains have been identified by deletion analysis in all three proteins and defined by their ability to activate a specific target in a “homologous” context, i.e., within a Hox protein binding either as a monomer to an ATTA-containing sequence or as a Hox-Pbx dimer on a TGAT(T/G)NAT-containing sequence, or in a “heterologous” context, after translocation to the yeast Gal4 DBD. Activation domains have been
identified in different regions of the three Hox proteins, depending on the context in which they are brought onto the DNA target. We speculate that Hox proteins may be multifunctional transcriptional regulators, which interact with different cofactors and/or components of the transcriptional machinery depending on the structure of their target regulatory elements.

MATERIALS AND METHODS

**Protein expression and reporter plasmids.** All the expression constructs used are derivatives of the simian virus 40 (SV40) promoter-based expression vector pSG5. The HOXD9 expression plasmid and the pTHCR luciferase reporter plasmid. Patterned boxes indicate the HD.

**Bacterial translation assay in HeLa cells.** Cells were subjected to bacterial translation in E.coli-based translation system (Promega, Madison, Wis.). A 10-

**Cell culture and transfection.** HeLa and COS7 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (Gibco), 100 IU per ml penicillin and 100 μg of streptomycin per ml, and 0.1% (v/v) nonessential amino acids. The cells were maintained under the same conditions with alpha minimal essential medium. Transfections were carried out by CaPO4 precipitation. Typically, for a 6-mm-diameter dish, 10 μg of total DNA was added to cells that had reached one-half confluency. A 48 h after transfection, the cells were harvested and lysed for Western blotting. The proteins were separated by electrophoresis, blotted to nitrocellulose filters, and probed with a monoclonal antibody recognizing the HOXD9 homeodomain. M, molecular mass markers (in kilodaltons).

**Protein production, Western blotting, and DNA-binding assays.** HOXD9 and PBX proteins were produced in vitro from the corresponding pSG5-derived expression vectors by using a T7 polymerase-based transcription and a reticulocyte lysate-based translation system (Promega, Madison, Wis.). A 10-μg portion of transfected nuclear extracts was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis, blotted to nitrocellulose membranes, and probed with a monoclonal antibody against the HOXD9 homeodomain. The bound immunocomplexes were then revealed with the ECL peroxidase detection kit (Amersham). A synthetic oligonucleotide containing the 17-bp recognition sequence of the yeast Gal4 protein was end labeled and incubated for 15 min at room temperature with about 4 μg of nuclear extracts from transfected cells in a 20-μl reaction mixture containing 10 mM Tris-HCl (pH 7.4), 50 mM KCl, 1 mM EDTA, 5 mM diithiothreitol (DTT), 50 μg of bovine serum albumin, and 4 μg of poly(dI-dC), and 4% Ficoll. The reaction mixtures were then loaded onto a 5% polyacrylamide-0.5× TBE sequencing gel and electrophoresed for about 2 h. The gels were dried and exposed to a XAR Kodak film with an intensifying screen at -70°C.

To detect Hox protein binding in an electrophoretic mobility shift assay (EMSA), a previously described consensus oligonucleotide (43) was end labeled and incubated in the same binding buffer as described above with 6 μl of reticulocyte lysate and 2 μg of poly(dI-dC). The binding reactions were incubated at 4°C for 30 min, and the reaction mixtures were electrophoresed as de-
scribed above. The HOX-PBX complex was detected in EMSA as described previously (9), with 4 μl of reticulocyte lysate for each protein.

RESULTS

Identification of activation domains in the N terminus of the HOXD9 protein. The human HOXD9 gene codes for a 337-aa protein, with 270 aa at the N terminus and 8 aa at the C terminus of the HD (43). We had previously shown in cotransfection experiments that the HOXD9 protein is able to transactivate a luciferase reporter construct (pTHCR) in which the minimal (–81) thymidine kinase (TK) promoter from herpes simplex virus is placed downstream of the Hox control region (HCR), a 90-bp, ATTA-rich autoregulatory element identified in the HOXD9 locus (43, 44). cDNAs encoding the full-length HOXD9 protein and four proteins with N-terminal deletions were cloned under the control of the SV40 promoter-enhancer and cotransfected in 0.5- to 5.0-μg amounts in HeLa cells together with pTHCR. As shown in Fig. 1B, transfection of 4.0 μg of the HOXD9 plasmid led to a 20-fold activation of the pTHCR basal activity. Deletion of the N-terminal 75 aa had no effect on the transcriptional activity of HOXD9 (D9(1–75)-Gal4), while further deletions up to aa 142, 222, and 264 (D9(1–142)-Gal4, D9(1–222)-Gal4, and D9(1–264), respectively) progressively abolished the activity. All proteins were expressed at comparable levels, as assayed by Western blotting of nuclear extracts obtained from transfected cells with a monoclonal antibody specifically recognizing the HOXD9 HD (Fig. 1C). These experiments indicate that the functional domain necessary for transcriptional activation of the HCR element is spread over a large region of the HOXD9 protein, extending from aa 75 to 222.

To test the activity of the HOXD9 N-terminal region in a heterologous context, we fused the regions from 1 to 265, 75 to 265, 142 to 265, and 222 to 265 at the N terminus of the DNA-binding domain (aa 1 to 147) of the yeast transcription factor Gal4 (Gal4-DBD). The fusion proteins were cotransfected in HeLa cells with a luciferase reporter gene in which the TK promoter was placed under the control of a 5-mer Gal-4-responsive element (pTUAS). As shown in Fig. 2B, the chimera containing the full-length N-terminal domain of HOXD9 [D9(1–265)-Gal4] was able to activate the pTUAS reporter 40- to 50-fold over the basal level in a dose-dependent fashion. Removal of the first 75 aa of HOXD9 [D9(75–265)-Gal4] caused a 80% reduction in the transcriptional activity of the chimera, while further deletions [D9(142–265)-Gal4 and D9(222–265)-Gal4] showed the same activity of the Gal4 DBD alone on the pTUAS reporter (less than threefold activation, see Fig. 2B). Biosynthesis of the chimeric proteins was tested by EMSA of nuclear extracts obtained from transfected HeLa cells with a 17-bp double-stranded oligonucleotide containing one copy of the Gal4 recognition sequence as the probe, which showed that the D9(1–265)-Gal4, D9(142–265)-Gal4, and
D9(222–265)-Gal4 proteins are synthesized at comparable levels, whereas the D9(75–265)-Gal4 protein and the Gal4 DBD are synthesized at 10-fold-higher levels or bind to the probe with a higher affinity (results not shown). After normalization for the DNA-bound protein levels, these experiments indicate that in the context of a protein binding to the Gal4-responsive element, the N-terminal 75 aa of HOXD9 contain a potential transcriptional activator domain, whereas the region containing most of the activity in the context of the native HOXD9 protein (aa 75 to 222) is virtually inactive.

To maintain a HOX-like geometry in the HOXD9-Gal4 chimeras, the Gal4 DBD was placed in the same position of the HD, i.e., at the C terminus of the fusion protein. In the Gal4 protein, however, the DBD is at the N terminus of the protein and the activation domain is at its C terminus. To test the activity of the HOXD9 N-terminal region in a Gal4-like conformation, we constructed a series of chimeric constructs in which the regions from 1 to 298, 75 to 298, 142 to 298, and 222 to 298 were fused at the C terminus of the Gal4 DBD, which were tested by cotransfection with the pTUAS reporter. Although the Gal4-HOXD9 proteins were 50% less active in activating the reporter construct than were their HOXD9-Gal4 counterparts, the region from 1 to 75 contained most of the activity when tested at the C terminus of the DBD (results not shown).

A comparative analysis of all known group 9 human, murine, and Xenopus Hox proteins indicates a modest conservation in the N-terminal regions, with most conserved residues concentrated in the first 130 positions (Fig. 3).

Identification of the activation domain of the HOXB1-PBX complex. The human HOXB1 gene codes for a 296-aa protein with 197 aa at the N terminus and 39 aa at the C terminus of the HD (1). We had previously reported (9) that the HOXB1 protein can cooperatively activate transcription, together with PBX1, from an autoregulatory element directing spatially restricted expression of the murine Hoxb-1 gene (b1-ARE) in the developing hindbrain (32). Selective recognition of the b1-ARE and transcriptional activation are mediated by HOXB1, while DNA-binding and protein-protein interaction functions of both HOXB1 and PBX1 are required for the assembly of a transcriptionally active complex (9). To localize the transcriptional activation domain of the complex, constructs coding for...
in the absence of PBX1 (column 4). An N-terminal deletion up to aa 155, which does not affect the FDWM domain necessary for cooperative interaction with PBX1 (9), further reduced the activity of the complex, down to three times the reporter basal activity (column 9). All HOX B1 mutants bound cooperatively to the R3 core element (TGATGGATGAG) of the b1-ARE together with PBX1, as checked by EMSA with in vitro-translated proteins (reference 9 and data not shown). These data indicate that the transcriptional activation domain of the HOXB1 protein in the context of the HOXB1-PBX1a complex resides between aa 38 and 90, a Ser-Pro-rich (20%) region only slightly conserved in the N termini of other vertebrate group 1 Hox genes (Fig. 5). The N-terminal region from aa 38 to 90 also contains most of the activating functions when tested in the presence of PRepl, a recently identified PBX1 cofactor forming a HOXB1-PBX1-Prep1 ternary complex on the b1-ARE (5).

The transcriptional activity of the HOXB1 N-terminal region was also tested as an N-terminal fusion to the Gal4-DBD, by cotransfection in COS7 cells together with the pTUAS reporter. As shown in Fig. 6B, the chimera containing the full-length N-terminal domain of HOXB1 [B1(1–164)-Gal4] was able to activate the pTUAS reporter 300- to 600-fold over the basal level. Removal of the first 38 aa [B1(38–164)-Gal4] or 90 aa [B1(90–164)-Gal4] caused a 40 and 80% reduction, respectively, in transcriptional activity (Fig. 6A). Biosynthesis of the HOXB1-Gal4 chimeric protein in COS7 cells was tested by EMSA as described for the HOXD9-Gal4 chimeras. The full-length N-terminal chimera and the B1(38–164)-Gal4 protein were synthesized at a level comparable to that of the Gal4 DBD, while the B1(90–164)-Gal4 protein accumulated at significantly lower levels in transfected cell nuclei (results not shown). These data indicate that the N terminal of the HOXB1 protein also contains a strong activator domain in the context of a Gal4 DNA-binding protein. After normalization for the DNA-bound protein levels, however, the activator domain appears to be spread over the entire N-terminal region, extending also to the region from aa 90 to 164, which is virtually inactive in the context of the HOXB1-PBX1 complex (Fig. 4).

Identification of two activation domains in the N terminus and C terminus of the HOXB3 protein. The human HOX3 gene codes for a 431-aa protein with 187 aa at the N terminus and 184 aa at the C terminus of the HD (1), which is able to transactivate a reporter gene driven by a promoter containing one or more ATTA core sequences (15). Cotransfection of 1 to 6 μg of an expression construct for the full-length HOXB3 protein in COS7 cells led to a 20- to 30-fold, dose-dependent trans-activation of a luciferase reporter construct in which the −109 TK promoter was placed under the control of an ATTA-rich Hox consensus binding site (pTCBS) (44) (Fig. 7A). Partial or total deletion of the N-terminal (B3Δ2–182, B3Δ1–182) or C-terminal (B3Δ273–360, B3Δ373–431) region of HOXB3 had little or no effect on the activity of the protein, while deletion of both regions (B3Δ1–182;Δ273–431) reduced the activity by almost 75% (Fig. 7B). All HOXB3 proteins were synthesized at comparable levels and were able to bind in vitro to the CBS sequence, as shown by EMSA of in vitro-translated proteins (Fig. 7C). These results indicate that both the N terminus and the C terminus of HOXB3 can promote transcriptional activation of an ATTA-containing target element.

The HOXB3 N terminus and C terminus were also tested as fusions to the Gal4 DBD by cotransfection in COS7 cells together with the pTUAS reporter. Each domain was tested both in HOX-like (i.e., N-terminal to the DBD) and in Gal4-like (i.e., C-terminal) configurations (Fig. 8A). As shown in Fig. 8B, the HOXB3 N terminus was unable to activate tran-
The HOXB3 C terminus is the only activation domain in the context of a heterodimeric complex with PBX1. We had previously shown that HOXB3 can cooperatively bind and activate the b1-ARE element together with PBX1 if the N terminus of the HD is replaced with that of HOXB1 (9). To test the activity of the HOXB3 protein in the context of a HOX-PBX heterodimer, expression constructs encoding a full-length HOXB3-HOXB1 chimera (HOXB3/B1, in which the HOXB1 portion encompasses the FDWM motif and the HD N terminus [9]) and five mutants containing a partial N-terminal deletion including the FDWM motif (B3/B1 Δ72–150), a complete N-terminal deletion (B3/B1 Δ1–123), a partial or a complete C-terminal deletion (B3/B1 Δ238–325 and B3/ B1 Δ328–396), and a combined N-terminal and C-terminal deletion (B3/B1 Δ1–123/Δ238–396) were cotransfected in P19 cells together with the PBX1a expression plasmid and the pAdMLARE reporter (Fig. 9A). The full-length HOXB3/B1 chimera induced a PBX-dependent, 20-fold transactivation of the pAdMLARE reporter (Fig. 9B, column 7), as previously reported (9). Deletion of the HOXB3 N-terminal domain from positions 1 to 23 significantly increased the activity of the chimera, resulting in transactivation levels that were 40-fold higher than the basal reporter activity (column 9), while removal of the FDWM HOX-PBX interaction domain (Δ72–150) completely abolished its activity (column 8). Internal deletion of the C-terminal domain from positions 238 to 325 also increased the activity of the chimera. Interestingly, this mutant was able to activate the reporter at significant (>10-fold) levels even in the absence of PBX (column 4). Deletion of the entire C terminus (positions 238 to 396) or of both the N terminus and the C terminus (positions 1 to 123 and positions 238 to 396), completely abolished the activity of the complex (columns 11 and 12). All mutants were translated in vitro and tested for their ability to bind a double-stranded oligonucleo-
DISCUSSION

Hox proteins are presumed to function as transcriptional regulators of the early steps of vertebrate embryonic development. Although the DNA-binding properties of this family of proteins are characteristically overlapping and nonselective in vitro, the homeodomains are known to mediate functional specificity in vivo (6, 13, 20, 29, 46). Specific target recognition by Hox, or closely related HD proteins such as the Drosophila

*Ft*, may in fact require the activity of cofactors, regulating both high-affinity DNA-binding and transcriptional activity (9, 16, 26–28, 41). The complexity of the system is further increased by the facts that at least some of the functional specificity of Hox proteins is mediated by protein-protein rather than DNA-protein interactions (2, 8, 10, 36, 37, 39, 44) and that some Hox proteins may work as both activators and repressors of transcription, depending on the context in which their function is tested (36, 38, 44).

While specific functional properties of Hox proteins, such as DNA binding, nuclear localization, and target recognition, have all been assigned to the HD, little is known about the function of the N- and C-terminal regions. Conservation of these regions among vertebrate Hox genes and between these genes and the *Drosophila* orthologs is minimal and essentially restricted to the YPWM motif and a few N-terminal amino acids. Nevertheless, the regions outside the HDs of two vertebrate Hox proteins, the mouse *Hoxa-5* and the chicken *Hoxb-1*, turned out to be absolutely essential in rather stringent functional tests, such as induction of homeotic transformations (47) or rescue of homeotic mutations (22) in *Drosophila* embryos. Thus, paradoxically, protein regions that are poorly conserved or not conserved at all during evolution are apparently capable of exerting specific functions across species in vivo, presumably at the level of transcription. Very few studies have addressed the biochemical properties of these “effector” domains and the nature of their interaction with the transcriptional machinery, mainly due to a lack of well-defined target genes and regulatory elements. A significant exception is represented by a few “autoregulatory” enhancers identified by genetic analysis upstream of some Hox promoters, which allowed functional analysis of the transcriptional properties of at least some Hox proteins on bona fide natural target elements (4, 9, 34, 43, 47).

In this study, we used two of these elements, the murine *Hoxb-1* and the human HOXD9 autoregulatory enhancers, to carry out a functional dissection of the human HOXD9, HOXB1, and HOXB3 proteins.

**Hox proteins contain potentially alternative activation domains.** A conventional deletion analysis on the 270-aa N terminus of HOXD9 showed that the first 75 residues contain a potential transcriptional activator when tested in the context of a Gal4 chimeric protein. In contrast, this region is dispensable when the activity of the protein is tested on the HCR, a context in which most of the activating function appears to be located within residues 76 to 264. The regions identified by the two alternative assays share no obvious characteristics with canonical eukaryotic activator domains and are only loosely conserved among different vertebrate species (Fig. 3). In a previous study, we showed that the activation domain of another posteriorly expressed Hox protein, HOXD8, can be localized to a similar sub-N-terminal region (44). HOXD8 and HOXD9 bind the multiple ATTA-containing sites within the HCR as monomers in a noncooperative fashion (42, 43), while Gal4-DBD chimeras bind the Gal4-responsive element (UAS) as a homodimer, a context which could force the HOXD9 N-terminal region to assume a different structural conformation and unmask a potential activating function in the N-terminal 75 residues. For the HOXB1-PBX heterodimer, the analysis carried out on the natural ARE identified a transcriptional activation domain in a Ser-Pro-rich, 52-residue sub-N-terminal region. This region also contained most of the HOXB1 transcriptional activity when tested as a Gal4-DBD chimera, a possible indication that the 52-residue region assumes a similar conformation or activates transcription by a similar mechanism, either in the context of a homodimer or in that of a HOXB-PBX heterodimer. The activity of HOXB3 was tested in

**FIG. 6.** (A) Schematic representation of the Gal4 fusion proteins containing the HOXB1 N-terminal domain (positions 1 to 164) or its deletion mutants (positions 38 to 164 and 90 to 164), and of the pTUAS luciferase reporter plasmid. Solid boxes indicate the Gal4 1-147 DBD. (B) Transcriptional activity of the mutants were able to bind cooperatively with PBX1 to the ARE R3 element by EMSA. As shown in Fig. 9C, the *In vivo* interaction between HOXB1 and PBX1 is restricted to the subset of AREs containing one copy of the HOX-PBX-binding site from Group 1 genes.
three different contexts, i.e., upon binding DNA as a monomer to an ATTA-containing element, as a HOX-PBX heterodimer to a bipartite HOX-PBX core element, and as a Gal4-DBD chimeric homodimer to the Gal4-responsive element. Although in the context of a monomer the transcriptional activity was spread over the entire protein sequence, only the C terminus contained a potent activator domain in the context of a Gal4 homodimer or of a PBX heterodimer. The 71-residue C terminus is relatively highly conserved in the mammalian group 3 Hox proteins (Fig. 10).

Our data indicate that the identification of transcriptionally active regions in Hox proteins is highly dependent on the context in which the activity of the protein is analyzed and possibly depends on the conformation that the different regions of the proteins assume when they are brought onto the DNA targets. It is therefore crucial that functional analysis of these proteins be carried out in the appropriate context, that is, in a native conformation on a natural, HD-binding target, and upon interaction with a natural DNA-binding partner. Previous attempts to identify active domains in the mammalian HOXD4 or Hoxa-7 proteins either gave inconsistent results or identified potential activators and repressors in somewhat unexpected regions of those proteins, such as the HD, when these regions were brought onto DNA via heterologous DBD (34, 38). Interestingly, a Hox protein binding DNA through the HD, for instance HOXB3, has the potential to activate a re-
FIG. 8. (A) Schematic representation of the fusion proteins between the HOXB3 N terminus or C terminus and the Gal4 1–147 DBD and of the pTUAS reporter plasmid. (B) Transcriptional activity of the HOXB3-Gal4 chimeras in COS7 cells transfected with 2 μg of reporter plasmid (UAS) and cotransfected with 2 to 6 μg of B3(1–182)-Gal4 and B3(273–431)-Gal4, 4 to 6 μg of Gal4-B3(1–182) and Gal4-B3(273–431), and 6 μg of Gal4-DBD expression plasmids. Luciferase activity is expressed as fold activation over the basal activity of the promoter-only reporter construct (see the legend to Fig. 1 for details).
porter gene through the activity of alternative regions (N terminus plus C terminus or C terminus only), depending on whether it binds DNA as a monomer or as a Hox-Pbx heterodimer. This would suggest that Hox proteins may be multifunctional transcriptional regulators, interacting with different cofactors and/or components of the transcriptional machinery depending on the context in which they bind DNA and therefore on the nature of the target elements on which they exert their regulatory function.

How do Hox proteins regulate transcription? The discrepancy between the results obtained by using Hox-binding and Gal4-binding elements as targets may be an indication that Hox proteins exert their function in a way which is intrinsically different from that of classical enhancer-binding transcription factors, such as Gal-4, VP16, or (p65)NF-κB. These factors contain acidic and/or serine/threonine-rich activator domains and may activate transcription by establishing direct contacts with general components of the transcriptional machinery (40, 45). It is conceivable that protein regions acting as activators when tested as chimeras with DBD of this type of factor are only those fitting particular requirements, such as net charge or presence of specific side chains, or those able to assume a restricted set of structural conformations. The effector domains of Hox proteins, on the other hand, might play a different role in gene regulation, such as providing a “positionally” restricted function in the context of regulatory elements, like the Hoxb-1 enhancer used in this study, on which this information is integrated by the interaction with tissue-specific, inducible, or structural factors. These complexes might in turn recruit coactivators or adapter molecules to signal the general transcriptional machinery, as in the case of the homeodomain-containing Oct2 protein (3). In this framework, the function of Hox-containing complexes could be to “open” genes, or sets of
FIG. 10. Best-fit alignment of the N-terminal (A) and C-terminal (B) regions of group 3 human (all capitals) and mouse (m) Hox proteins. Numbers indicate amino acid positions within the HOXB3 protein. Amino acids at the borders of the deletions generated in the HOXB3 N-terminus (Fig. 7A) are indicated in boldface type.


The specificity of HOX protein function depends on DNA-protein and protein-protein interactions, both mediated by the homeo domain. Genes Dev. 8:753–262.


The mouse Hox-4.2 gene is functionally equivalent to the Drosophila sex combs reduced gene. Genes Dev. 7:343–354.