THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 23, Issue of June 8, pp. 20506–20515, 2001

Printed in U.S.A.

The Recruitment of SOX/OCT Complexes and the Differential Activity of HOXA1 and HOXB1 Modulate the *Hoxb1* Auto-regulatory Enhancer Function*

Received for publication, December 12, 2000, and in revised form, February 23, 2001
Published, JBC Papers in Press, March 1, 2001, DOI 10.1074/jbc.M011175200

Giuliana Di Rocco‡§, Anthony Gavalas¶‖, Heike Pöpperl¶, Robb Krumlauf¶‡‡, Fulvio Mavilio‡§§, and Vincenzo Zappavigna‡¶¶**

From the ‡DIBIT-Istituto Scientifico H. San Raffaele, Via Olgettina 58, 20132 Milano, the §§Department of Biomedical Sciences, University of Modena School of Medicine, Via Campi 287, 41100 Modena, Italy, the ¶Division of Developmental Neurobiology, National Institutes for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom, and the ‡‡Stowers Institute for Medical Research, Kansas City, Missouri 64110

Regionally restricted expression patterns of *Hox* genes in developing embryos rely on auto-, cross-, and para-regulatory transcriptional elements. One example is the *Hoxb1* auto-regulatory element (b1-ARE), which drives expression of *Hoxb1* in the fourth rhombomere of the hindbrain. We previously showed that HOXB1 and PBX1 activate transcription from the b1-ARE by binding to sequences required for the expression of a reporter gene in rhombomere 4 *in vivo*. We now report that in embryonal carcinoma cells, which retain characteristics of primitive neuroectodermal cells, the b1-ARE displays higher basal and HOX/PBX-induced activities than in other cell backgrounds. We have identified a bipartite-binding site for SOX/OCT heterodimers within the b1-ARE that accounts for its cell context-specific activity and is required for maximal transcriptional activity of HOX/PBX complexes in embryonal carcinoma cells. Furthermore, we found that in an embryonal carcinoma cell background, HOXB1 has a significantly higher transcriptional activity than its paralog HOXA1. We map the determinants for this differential activity within the HOXB1 N-terminal transcriptional activation domain. By using analysis in transgenic and HOXA1 mutant mice, we extended these findings on the differential activities of HOXA1 and HOXB1 *in vivo*, and we demonstrated that they are important for regulating aspects of HOXB1 expression in the hindbrain. We found that mutation of the SOX/OCT site and targeted inactivation of *Hoxa1* both impair the response of the b1-ARE to retinoic acid in transgenic mice. Our results show that *Hoxa1* is the primary mediator of the response of b1-ARE to retinoic acid *in vivo* and that this function is dependent on the binding of SOX/OCT heterodimers to the b1-ARE. These results uncover novel functional differences between Hox paralogs and their modulators.

The HOX homeodomain-containing transcription factors control cell fate and developmental patterns in all metazoans, leading to the generation of morphological differences along body axes (reviewed in Ref. 1). In most vertebrates the four *Hox* clusters encode 39 distinct proteins in which the homeodomain (HD)[1] and flanking amino acids dictate the DNA binding specificity by recognizing a restricted set of sites containing the core consensus sequence TNAT(G/T)(G/A) (2–4). Despite the apparent similarity in consensus DNA recognition sites, HOX proteins can modulate their binding properties and specificity through the concomitant activity of an emerging array of cofactors. Interactions with cofactors such as the HD-containing proteins of the EXD/PBX (PBC) (3–6) and MEIS/PREP (MEINOX) (7–11) families can modulate the affinity and stability of DNA binding and regulate transcriptional activity of HOX proteins.

Transgenic analysis in mice has led to the identification of sequences that function as HOX target sites *in vivo* for the auto-, cross-, and para-regulatory interactions among *Hox* genes (6, 12–15). Several of these *in vivo* target sites are composed of bipartite and overlapping HOX- and PBX-binding motifs (HOX/PBC sites) and represent useful models for analyzing the function of HOX-containing transcriptional complexes. Protein-protein contacts involved in functional interactions between the HOX and the EXD/PBX proteins were found to be mediated by both the N-terminal region of the HD (16) and the short conserved hexapeptide or YPWM motif (17–19), located upstream of the HD in a subset of HOX proteins (paralogy groups 1–8). Recent crystal structure analysis of HOX/PBX homeodomain complexes bound to DNA target sites with relevance *in vivo* have provided a useful model for understanding how these interactions can influence specificity of binding (20, 21).

The identification of *bona fide Hox*-responsive enhancers and promoters allowed analysis of the transcriptional properties of these proteins. These studies have shown that HOX proteins share the modular type structure of most eukaryotic transcription factors, featuring separate DNA binding and activator or repressor domains (5, 22–26). One of the HOX-binding elements examined in most detail *in vivo* is contained in a highly conserved auto-regulatory enhancer of the *Hoxb1* gene (b1-ARE). This is a key cis-regulatory element for the normal rhombomere 4 (r4)-restricted expression of *Hoxb1* in the devel-

* This work was supported by grants from the Telethon Foundation (to V. Z. and F. M.), the Italian Association for Cancer Research (to V. Z.), by an EEC Biotechnology Network Grant BIO4 CT-960378, and Medical Research Council (to R. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by an Human Science Frontier Program fellowship. Present address: Dept. of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Ave., Boston, MA 02115.

‖ Supported by an EMBO postdoctoral fellowship.

** Supported by postdoctoral fellowships from EMBO and the Human Science Frontier Program.

¶¶ To whom correspondence should be addressed: DIBIT, Istituto Scientifico H.S. Raffaele, Via Olgettina, 58, 20132 Milano, Italy. Tel.: 39-02-26434805; Fax: 39-02-26434861; E-mail: zappavigna.vincenzo@hsr.it.

[1] The abbreviations used are: HD, homeodomain; r4, rhombomere 4; EC, embryonal carcinoma; RA, retinoic acid; RARE, RA response element; PCR, polymerase chain reaction; aa, amino acids; DTT, dithiothreitol; EMSA, electrophoretic mobility shift assay; dpc, days post-coitum; bp, base pair; FGF, fibroblast growth factor.

oping hindbrain. Genetic analysis has revealed that in an early phase *Hoxa1* and *Hoxb1* function synergistically to establish initial r4 identity by triggering the auto-regulatory loop (6, 27, 28). As a part of this process, the early activation of *Hoxb1* and *Hoxa1* themselves is directly mediated by RA signaling through the presence of RA response elements (RAREs) located at the 3′ ends of both genes (28–31). In later stages *Hoxb1* is required to maintain r4 identity (27, 28). Even though the b1-ARE does not contain a canonical RARE, it indirectly participates in mediating the RA-dependent ectopic activation of *Hoxb1* in anterior regions (r2), through its auto-regulatory ability (6).

Auto- and para-regulation of *Hoxb1* are dependent on the cooperative binding of HOXB1 or HOXA1, and a member of the PBX family on three conserved sequence motifs (R1, R2, and R3) in the b1-ARE related to a consensus bipartite HOX/PBC-binding site (5, 6). Each of these HOX/PBX repeats contributes to the r4-restricted expression of a reporter gene in transgenic mice and to the indirect RA response. The R3 motif, however, makes the largest contribution to the b1-ARE regulatory activity (6). Interestingly, the b1-ARE is not active in other regions of the embryo where *Hoxb1* and *Pbx1* are coexpressed (spinal cord, paraxial mesoderm, endoderm, and limb buds), suggesting that region-specific expression in the hindbrain might be determined by differential activity of additional factors.

We reported previously (5) that the human HOXB1 and PBX1 proteins cooperatively activate transcription from a basal promoter under the control of the b1-ARE in transfected mammalian cells. Comparing the functional selectivity of proteins from a range of *Hox* paralogy groups, we showed (5) that only a restricted subset of HOX proteins (HOXA1, HOXB1, and HOXB2) are able to activate transcription from the b1-ARE in cooperation with PBX1. Selective recognition of the R3 motif by the HOXB1/PBX1 complex is mediated by the N terminus of the HOX homeodomain, whereas the major transcriptional activator domain is provided by the HOXB1 N-terminal region (5, 26).

In this study we show that the basal and HOX/PBX-induced transcriptional activities of b1-ARE, in comparison to other cell lines, are significantly higher in murine or human embryonal carcinoma (EC) cells, which retain phenotypic and molecular characteristics of primitive neuroectodermal cells (32, 33). In this context, HOXB1 displays a stronger transcriptional activation than its paralog HOXA1, in cooperation with PBX1. The differential activity between HOXB1 and HOXA1 is not based on the DNA-binding properties of the HOXB1/PBX1 complex but rather on specific determinants within the HOXB1 transcriptional activation domain that selectively interact with the transcriptional machinery. Furthermore, we show that the full transcriptional activity of HOX/PBX heterodimers on the b1-ARE is dependent on the binding of a SOX/OCT heterodimeric complex to a bipartite site located immediately upstream of the R3 sequence. Mutation of this site, as well as the targeted inactivation of the *Hoxa1* gene, impairs the response of the b1-ARE to RA treatment in transgenic mice and the establishment of a full ectopic auto-regulatory circuit of *Hoxb1 in vivo*. Our results show that the product of the *Hoxa1* gene is required for the response of *Hoxb1* to RA, and this activity requires a functional SOX/OCT-binding site within the b1-ARE. Together, these results have uncovered novel differences among HOX paralogs and their modifiers.

EXPERIMENTAL PROCEDURES

Expression Vectors and Reporter Plasmids—All expression constructs are derivatives of the SV40 promoter-based expression vector pSG5 (34). HOXB1, HOXA1, HOXB2, and PBX1 expression vectors were described previously (5). pSGB1/A1HD was generated by swapping the region spanning the FDWM motif and the homeodomain of HOXB1 (aa 175–257) with the corresponding region of HOXA1 (aa 206–289), whereas pSGA1/B1CT was generated by substituting the C-terminal region of HOXA1 (aa 290–336) with that of HOXB1 (aa 258–296). The mutated cDNAs were cloned into the *Bam*HI site of pSG5. pGAL1–147 (35) contains the DNA binding domain of yeast GAL4 (amino acids 1–147). pGALVP16 was generated by cloning an PCR-amplified region encompassing the C terminus (last 80 aa) of the VP16 protein in frame with the GAL1–147 protein at the *Bam*HI site of the pGAL1–147 vector. PCR was carried out with *Pfu* polymerase (Stratagene). All PCR-generated fragments were sequenced on both strands, and expression of all proteins was preliminarily tested using a T7 polymerase-based transcription and reticulocyte lysate-based translation system (Promega).

The luciferase reporter construct pMLluc is a pXP2-based vector (36) containing the adenovirus major late basal promoter (from −65 to +30). pAdMLARE contains the *Ava*I-*Hae*II fragment of the *Hoxb1* r4-auto-regulatory enhancer (b1-ARE) (6) cloned as a PCR-amplified *Hin*dIII-*Xho*I fragment into pMLluc. pAdMLR3 contains a trimer of repeat 3 of the b1-ARE (sequence, 5′-GATCCGGGGGGTGATGGATGGGCGCT-GGGA-3′) cloned as a *Bam*HI-*Hin*dIII fragment into pML. pAdMLARE Δ1–52, ΔR1, and ΔR1 + R2 have been obtained by cloning PCR-amplified *Hin*dIII-*Xho*I fragments of the b1-ARE (from nucleotides 53–140, 64–140, and 82–140, respectively) into pMLluc. In pAdMLoctm, the Sox/Oct-binding sequence CTTTGTCATGCTAAT in the b1-ARE was changed to GCAAGTCGACTGCCT. The pTUASluc reporter construct was described (37).

For transgenic mouse analysis, the wild type *Hoxb1* 5′-flanking *Stu*I-*Hin*dIII fragment containing the b1-ARE was cloned into the *lacZ* basal reporter vector BGZ40 (38) as described previously (6). Point mutations in the SOX/OCT-binding sequence (TAAT to CCGG and TTTGTC to CGCTGT) were generated by site-directed mutagenesis in M13 or inverse PCR, followed by double-strand sequencing.

Cell Culture and Transfection—COS-7 cells were maintained in Dulbecco's modified Eagle's medium, supplemented with 10% fetal calf serum (Life Technologies, Inc.), 100 IU/ml penicillin, and 100 μg/ml streptomycin. P19 and NT2/D1 cells were maintained in α-minimum Eagle's medium. Transfections were carried out by $CaPO_4$ precipitation (39). In a typical transfection experiment, 8 μg of reporter plasmid, 4–8 μg of expression construct, and 0.2 μg of pCMV-β-gal (CLONTECH) as an internal control were used per 10-cm dish. 48–60 h after transfection, cells were washed and lysed directly on the plate with a solution containing 1% Triton X-100, 25 mM glycyl-glycine, pH 7.8, 15 mM $MgSO_4$, 4 mM EGTA, 1 mM DTT. Extracts were collected, centrifuged to clear the supernatant, and assayed for luciferase and β-galactosidase expression as described (40). To rule out the existence of intrinsic differences between COS-7 and P19 cells in sustaining transcriptional activation, we tested for the capability of a GAL4-VP16 chimera to activate transcription from a GAL4-responding reported in both cell lines. The GAL4-VP16 chimera gave identical results in terms of transactivation in both cell lines (data not shown).

Electrophoretic Mobility Shift Assay (EMSA)—To obtain total cell protein extracts, cells were collected from confluent plates, washed with phosphate-buffered saline, pelleted, frozen with liquid nitrogen, and lysed by resuspension in 5 volumes of Extraction Buffer (10 mM Hepes, pH 7.9, 0.4 M NaCl, 0.1 mM EDTA, 0.5 mM DTT, 5% glycerol, 0.5 mM phenylmethylsulfonyl fluoride, 1% Trasylol). The lysate was then centrifuged for 30 min at 34,000 rpm in a Beckman Ti-50 rotor, and the supernatant was stored in aliquots at −80 °C. To obtain embryonic cell extracts, about 50 mouse embryos were collected at 9.5 dpc, washed with phosphate-buffered saline, and lysed as described above. Gel retardation analysis was performed by preincubating the cell extracts (8 μg) for 30 min on ice in 20 μl of binding buffer (75 mM NaCl, 20% Ficoll, 10 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 10 mM DTT, 3 μg of poly(dG-dC) or poly(dI-dC)), together with 2 μl (0.5 ng, 5×10^4 cpm) of ^{32}P-end-labeled oligonucleotide probe. Competition experiments were carried out by adding 2 μl of α-Oct1 (Santa Cruz Biotechnology), α-Oct3/4 (kindly provided by Hans Shöler), or α-Sox-2 (a gift from Marco E. Bianchi) antisera before adding the probe. The incubation mixture was resolved by electrophoresis on a 5% polyacrylamide gel in 0.25× TBE at 10 V/cm. Gels were dried and exposed to a Kodak X-AR film at −70 °C. The b1-ARE oligonucleotide probe sequence was 5′-AGCTTGTGTCTT-TGTCATGCTAATGATTGGGGGG-3′.

Generation and Analysis of Transgenic Mice—Transgenic mice were produced by microinjection of DNA into fertilized eggs from crosses between F1 (CBA × C57) females and males. Whole mount β-galactosidase reporter activity in founder embryos was performed as described previously (41). RA exposure was achieved by treating pregnant females with embryos at 7.5 dpc by oral gavage with 200 ml of sesame seed oil containing all-*trans*-retinoic acid (Sigma), diluted from a 25

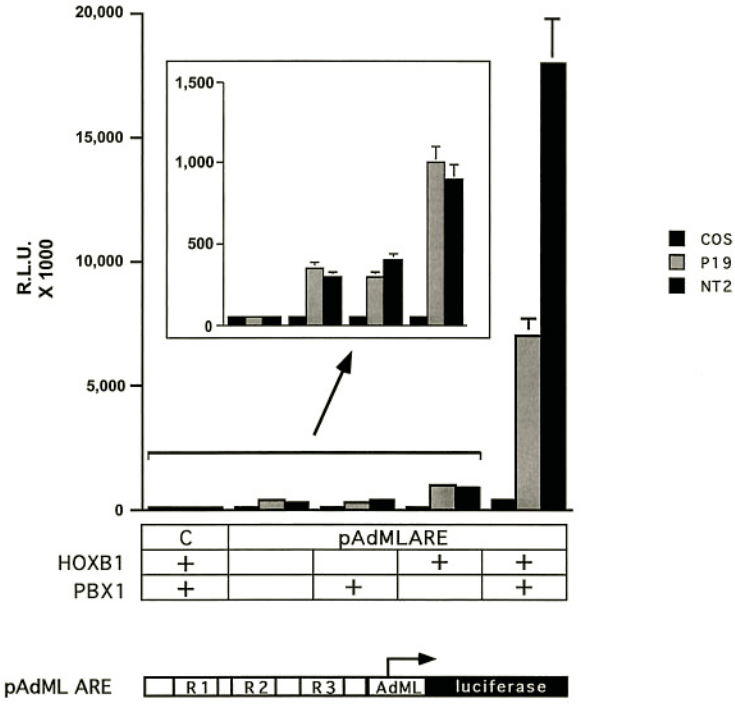

FIG. 1. **The HOXB1/PBX1 complex-mediated transcriptional activation is significantly higher in neuroectodermal *versus* other cell lines.** Luciferase activity, in arbitrary units, was assayed from extracts of the indicated transiently transfected cell lines. The cells were transfected with 4 μg of the SV40-driven HOXB1, and/or PBX1, expression constructs, together with 8 μg of alternatively pAdMLluc (*C*) or pAdMLARE. The *inset* shows the low level luciferase activities in control transfections. 0.2 μg of the pCMVβ-gal plasmid were cotransfected in all experiments as an internal standard. *Bars* represent the mean ± S.E. of at least four independent experiments.

mg/ml stock solution in dimethyl sulfoxide, for a final dose of about 20 mg/kg of maternal body weight (42). The inductive response was assayed in embryos at 9.5 dpc.

RESULTS

The Transcriptional Activity of the HOXB1/PBX1 Complex Is Maximal in a Neuroectodermal Cell Background—We have shown previously that the HOXB1/PBX1 heterodimer is able to activate transcription of a reporter gene under the control of the b1-ARE in mammalian cell lines (5, 26). To investigate the influence of the cell context on this system, we compared the transcriptional activity of the HOXB1/PBX1 complex in different cell lines, including COS-7, HeLa, NIH3T3, P19, and NT2/D1 cells. The murine P19 and the human NT2/D1 EC cells are originally derived from germ cell tumors and retain many characteristics of primitive neuroectodermal cells, including the ability to differentiate into neurons (32, 33) and activate the four *Hox* gene clusters upon induction with retinoic acid (43–45). To measure transcriptional activity, reporter constructs were made by placing a luciferase gene under the control of the adenovirus major late basal promoter either alone (pAdMLluc) or in combination with the 150-bp b1-ARE Hox-responsive region (pAdMLARE). These were cotransfected with expression constructs for HOXB1 and PBX1 into the various cell lines and assayed for reporter activity (Fig. 1). In COS-7, HeLa, or NIH3T3 cells the basal activity of the reporters with and without the b1-ARE enhancer were identical (Fig. 1 and data not shown). In contrast, in both P19 and NT2/D1 EC cells the construct containing the b1-ARE enhancer (pAdMLARE) displayed a 7-fold higher basal level, compared with the control (pAdMLluc). In non-EC cells (*e.g.* COS-7 in Fig. 1), coexpression of HOXB1 and PBX1 led to a 8-fold transactivation of the pAdMLARE reporter. In P19 and NT2 cells, however, coexpression of HOXB1 and PBX1 led to significantly higher levels of transactivation, ~20- and ~60-fold, respectively, over the pAdMLARE basal activity of the reporter, corresponding to a ~140- and ~360-fold increase in activity, respectively, of the enhancer-less pAdMLluc reporter (Fig. 1).

Transfection of the PBX1 expression vector alone had no effect on reporter activity in any cell context, whereas transfection of HOXB1 alone led to a 2–3-fold activation only in EC cells presumably through interaction with the endogenous levels of PBX1 expressed in these cells. The higher activity of the HOXB1/PBX1 complex on the b1-ARE target was found not to be due to a higher transfection efficiency in EC cells, as the percentage of transfected cells was found to be comparable to that of the other cell lines (data not shown). Furthermore, we observed no broad differences in the capability of EC *versus* COS-7 cells to sustain transcriptional activation (see "Experimental Procedures"). Our findings that the basal activity and the ability of the HOXB1/PBX1 complex to act on the b1-ARE are much higher in EC cells, compared with other cell lines, suggest either the presence of EC-specific enhancing factors or COS-7 inhibiting components that modify the b1-ARE regulatory potential.

HOXB1 and HOXA1 Differentially Activate Transcription through the b1-ARE in Neuroectodermal Cells—We previously reported that only HOX proteins belonging to paralogous groups 1 and 2 can bind and activate the full b1-ARE or multimerized versions of R3 in cultured cells in a PBX-dependent manner (5). To compare directly the activity of different HOX/PBX complexes in an EC cell background, expression constructs for HOXA1, HOXB1, and HOXB2 were cotransfected in P19 cells, together with the reporters containing the entire b1-ARE (pAdMLARE) or a trimer of the R3 motif (5). The basal activity of pAdMLARE was higher than that of pAdMLR3 and was only weakly induced by cotransfection with HOX proteins alone (Fig. 2). On the pAdMLR3 reporter, consistent with our previous results using non-EC cells, the activities of the HOXB1/PBX1 and the HOXA1/PBX1 complexes were comparable, resulting in an 80–100-fold transactivation. The HOXB2/PBX1 complex was also active although it remained lower (15-fold transactivation) (Fig. 2). In contrast, on the full b1-ARE reporter (pAdMLARE) while HOXB1 and PBX1 induced a 20-fold transactivation (see also Fig. 1), surprisingly HOXA1 only weakly activated (6-fold) the reporter expression in combination with PBX1 (Fig. 2). HOXB2 displayed an even lower (4-fold) activity with PBX1 (Fig. 2). These results indicate that in the EC cell background HOXB1 activates transcription more efficiently than its paralog HOXA1 in cooperation with PBX1 through the entire b1-ARE element. This differential

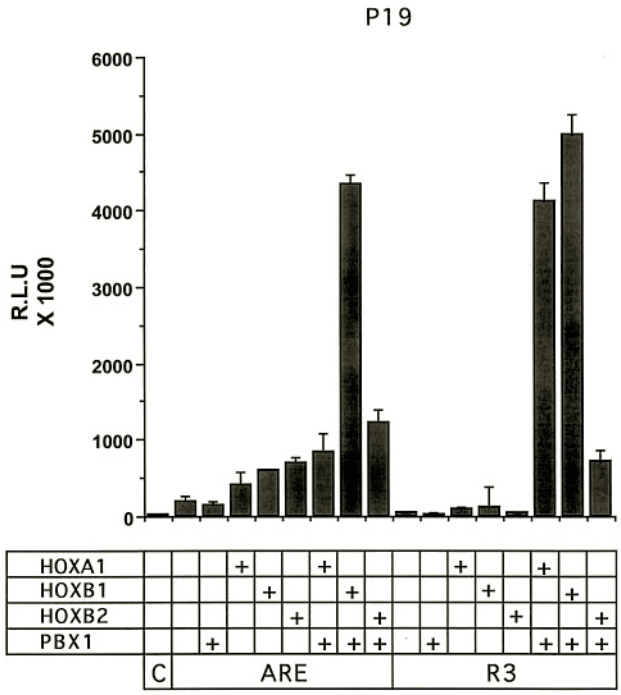

FIG. 2. **Transcriptional activity mediated by b1-ARE is mainly restricted to the HOXB1/PBX1 complex in neuroectodermal cells.** Luciferase activity, in arbitrary units, was assayed from extracts of P19 cells. The cells were transfected with 4 μg of the SV40-driven HOXA1, HOXB1, HOXB2, and/or PBX1 expression constructs together with 8 μg of pAdMLluc (*C*), pAdMLARE (*ARE*), or pAdMLR3 (*R3*). 0.2 μg of the pCMVβ-gal plasmid were cotransfected in all experiments as an internal standard. *Bars* represent the mean ± S.E. of at least five independent experiments.

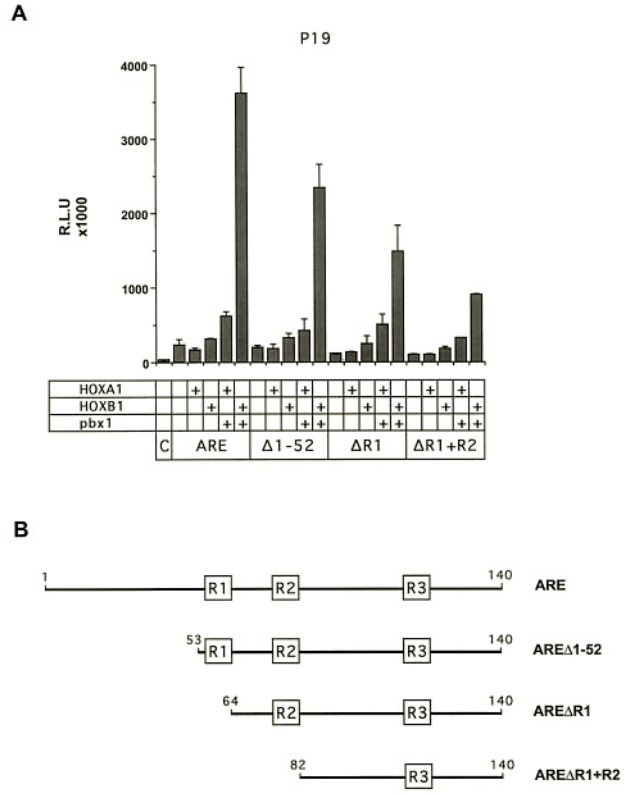

FIG. 3. **The b1-ARE selectivity is unaffected by deletion of the R1 and R2 elements.** *A*, Luciferase activity, in arbitrary units, was assayed from extracts of P19 cells. The cells were transfected with 4 μg of the SV40-driven HOXA1, HOXB1, and/or PBX1 expression constructs, together with 8 μg of pAdMLluc (*C*), pAdMLARE (*ARE*), or its deletion mutants (*Δ1–52, ΔR1, and ΔR1 + R2*). 0.2 μg of the pCMVβ-gal plasmid were cotransfected in all experiments as an internal standard. *B,* schematic representation of the *Hoxb1*-ARE and of its deletion mutant derivatives. *Numbers* indicate the nucleotide positions.

activity was never observed in non-neuroectodermal cell lines.

As the use of a multimerized version of the HOX/PBC repeat 3 target sequence apparently relaxes the paralog selectivity, allowing HOXA1 and HOXB2 to function as effectively as HOXB1, we considered the possibility that other enhancer sequences within the b1-ARE might influence the activity of the different HOX/PBX heterodimers. In order to analyze the contribution of different regions of the enhancer in restricting transcriptional activation, we cotransfected reporter constructs containing sequential 5′ deletions of the b1-ARE (Fig. 3*B*) with HOXB1, HOXA1, and PBX1 in P19 cells. In deletions with increasing size there was a progressive reduction in the overall level of transactivation that progressively reduced expression to 70, 50, and 30% of the entire b1-ARE (Fig. 3*A*). However, the differential ability of HOXB1/PBX1 *versus* HOXA1/PBX1 to activate these sites was maintained even on a reporter containing only repeat 3 and its surrounding sequences (*ARE*Δ*R1 + R2*, Fig. 3*A*). Considering that a reporter containing a multimer of the R3 motif alone responds equally well to HOXB1 or HOXA1 (Fig. 2), this result suggests that sequences flanking R3 are important in mediating the selective preference of the b1-ARE for HOXB1.

A SOX/POU-OCT-binding Site Is Necessary for Full Transcriptional Activity of the b1-ARE in EC Cells—In the b-1ARE, located between the R2 and R3 HOX/PBC sites, there is a bipartite sequence motif that is highly conserved in the human, mouse, chicken and pufferfish *Hoxb1* locus (6). In this motif (designated *SOct* in Fig. 4), the 5′-most part (TCTTTGTC) closely resembles the target sequence for the HMG box protein SOX-2 (46–48), whereas the 3′ part (ATGCTAAT) shows a high degree of similarity with the consensus recognition sequence for POU/Octamer-binding proteins (49). To test the role of this putative SOX/OCT heterodimer-binding site in the function of the b1-ARE in EC cells, we generated a 11-bp mutation (ARESOctm) encompassing both the SOX and the POU/OCT sites in the context of the whole b1-ARE sequence (Fig. 4*B*). In cotransfection experiments this SOX/OCT site mutation reduced transactivation of the HOXB1/PBX1 complex by 50% and that of HOXA1-PBX1 by 77% (Fig. 4*A*). Interestingly, this SOX/OCT mutation completely abolished the EC cell-specific basal activity of the b1-ARE, as well as the weak but detectable transactivation by HOXB1 in the absence of PBX1 (Fig. 4*A*). To help distinguish between the contribution of the SOX-specific *versus* the POU/OCT-specific sequence on the b-1ARE activity, an additional mutation (*ARE*Δ*R1+R2+Sox,* Fig. 4*B*) was made in the context of the smallest fragment (*ARE*Δ*R1+R2*) showing differential response to transactivation by HOXB1/PBX1 and HOXA1/PBX1 (see Fig. 3). This mutation, like the ARE SOctm mutation, completely abolished the basal activity of the reporter in P19 cells, reduced transactivation by the HOXB1/PBX1 complex by 30%, and almost completely suppressed HOXA1/PBX activity (Fig. 4*A*).

These results indicate that the increased basal activity of the b1-ARE in EC cells compared with COS cells is due to EC-specific or EC-enriched factors interacting with the SOX/OCT site and that the SOX-specific hemisite is necessary for this activity. The same factors appear to affect the overall transcriptional activity of HOX/PBX heterodimers on the b1-ARE through this site. Furthermore, our data show that transactivation by HOXA1/PBX1 is more dependent upon the SOct motif than that by HOXB1/PBX1.

The SOX/OCT Site Is Cooperatively Bound by SOX-2/OCT1

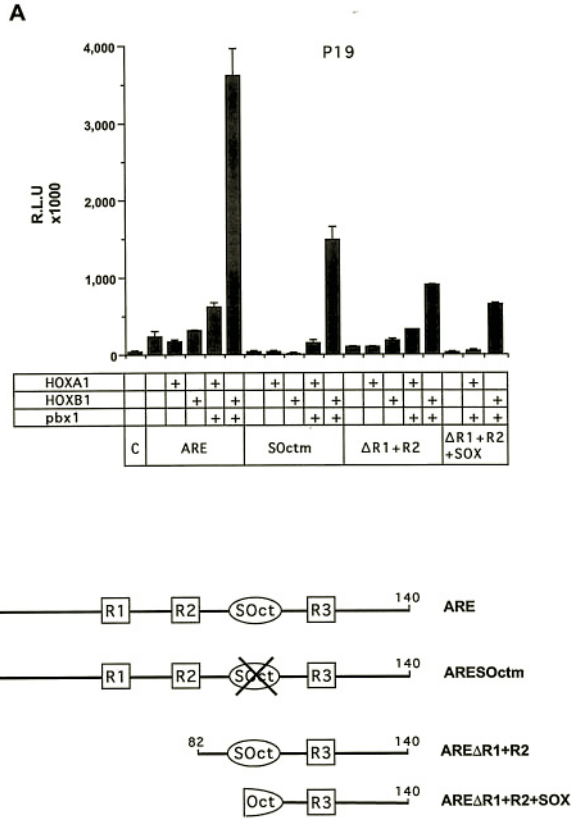

FIG. 4. **The bipartite SOX/OCT-binding site is necessary for basal and HOXB1/PBX1-induced activity of b1-ARE in P19 cells.** *A,* luciferase activity, in arbitrary units, was assayed from extracts of P19 cells. The cells were transfected with 4 μg of the SV40-driven expression constructs for HOXA1, HOXB1, and/or for PBX1, together with 8 μg of pAdMLluc (*C*), pAdMLARE (*ARE*), pAdMLR3 (*R3*), pAdMLSOctm (*SOctm*), pAdMLAREΔR1+R2 (*R1 + R2*), or pAdMLAREΔR1 + R2 + SOX (Δ*R1 + R2 + SOX*). 0.2 μg of the pCMVβ-gal plasmid were cotransfected in all experiments as an internal standard. *B,* schematic representation of the *Hoxb1*-ARE and of its mutant derivatives. *Numbers* indicate the nucleotide positions.

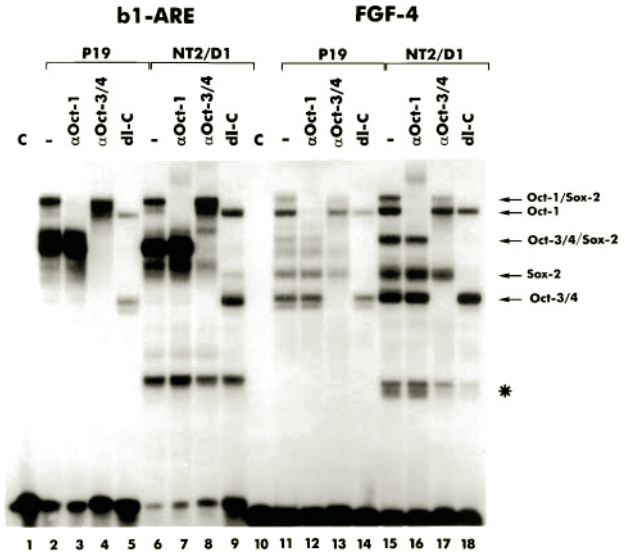

FIG. 5. **SOX/OCT1 and SOX/OCT3/4 complexes bind to the b1-ARE in P19 and NT2/D1 cells.** EMSA of double-stranded oligonucleotides representing a region of 34 bp encompassing the SOct site within b1-ARE (b1-ARE, *left side*) or the SOX/OCT bipartite-binding site found within the FGF-4 promoter (FGF-4, *right side*). Nuclear extracts from P19 or NT2/D1 cells were challenged with the labeled oligonucleotides as described under "Experimental Procedures." Anti-Oct1 and anti-Oct3/4-specific antisera were used to characterize the two retarded complexes. An *asterisk* indicates the position of a nonspecific retarded complex found in NT2/D1 cells.

and SOX-2/OCT3/4 Heterodimers—Since the conserved SOX/OCT site has a functional role in the b1-ARE enhancer activity in EC cells, we looked for EC cell-specific binding activities on this site by EMSA. The octamer-binding proteins OCT1 and OCT3/4 and the HMG box protein SOX-2 are expressed at high levels in EC cells and were previously shown to bind and regulate cooperatively the activity of developmentally regulated enhancers such as that of FGF-4 (50, 51). Hence, we used a double-stranded oligonucleotide spanning the SOX/OCT site of the FGF-4 enhancer as a control. The b-1ARE EMSAs were carried out using as probe a double-stranded oligonucleotide containing the SOX/OCT site (SO-ARE) in combination with total cell extracts obtained from either P19 or NT2 EC cells (Fig. 5). Except where indicated, poly(dG-dC) was used instead of poly(dI-dC) as nonspecific DNA competitor to allow binding by HMG box proteins (50). As shown in Fig. 5 (*lanes 2* and *6*), in both P19 and NT2/D1 cell extracts a major SOX-2/OCT3/4 complex and a minor SOX-2/OCT1 complex bind the b1-ARE probe. These were specifically competed and/or supershifted by the addition of ant-OCT1 and anti-OCT3/4 antibodies to the binding reaction (Fig. 5, *lanes 3* and *4* and *7* and *8*). Addition of poly(dI-dC) competed out the binding of the heterodimeric complexes and allowed the formation of monomeric OCT1 and OCT3/4 complexes (Fig. 5, *lanes 5* and *9*).

In comparison, at least five complexes with similar intensity were detected on the control FGF-4 probe, and they correspond to the individual monomers and the SOX-2/OCT1 and SOX-2/OCT3/4 heterodimers (Fig. 5, *lanes 11* and *15*). The identity of all complexes was confirmed by anti-OCT1 and anti-OCT3/4 antibodies (Fig. 5, *lanes 12* and *13* and *16* and *17*), and again the addition of poly(dI-dC) allowed only binding of the monomeric OCT complexes (Fig. 5, *lanes 14* and *18*). These results indicate that OCT1 and OCT3/4 bind exclusively as heterodimers with SOX-2 to the SOX/OCT sequence in the b1-ARE.

To check whether binding of the SOX-2/OCT complexes on the b1-ARE SOX/OCT site also occurs in embryonic tissues, we used total cell extracts obtained from the central nervous system of 9.5-dpc mouse embryos. As shown in Fig. 6 (*left panel*), strong binding of both SOX-2/OCT3/4 and SOX-2/OCT1 complexes to the b1-ARE sequence was observed in embryo extracts, with a pattern similar to that observed with EC cell extracts. The identity of the complexes was also confirmed with anti-OCT and anti-SOX-2 antibodies and by the differential use of poly(dG-dC) or poly(dI-dC) as nonspecific competitors (Fig. 6). A much weaker binding was observed on the FGF-4 probe, particularly for the SOX-2/OCT3/4 complex (Fig. 6, *right panel*). These results show that the b1-ARE sequence binds protein complexes containing SOX and OCT proteins present in embryonic tissues at a stage corresponding to the maximal activity of the b1-ARE in the developing hindbrain (6).

SOX-2 and OCT1 Cooperatively Enhance the Activity of the b1-ARE in COS Cells—Taken together, the experiments above suggest that the different activity of the b1-ARE in EC cells compared with COS cells could be mediated by differences in the availability of SOX and OCT proteins. To assess this issue directly, we cotransfected expression plasmids for the SOX-2, OCT1, and OCT3/4 into COS-7 cells together with the b1-ARE reporter in the presence or absence of HOXB1 and PBX1 (Fig. 7). No activation of reporter expression was observed by transfection of any of the individual plasmids in the absence of HOXB1 and PBX. Cotransfection of SOX-2 and OCT1 caused a 4-fold induction of the reporter basal activity bringing it closer to that observed in EC cells. More important, the activity of

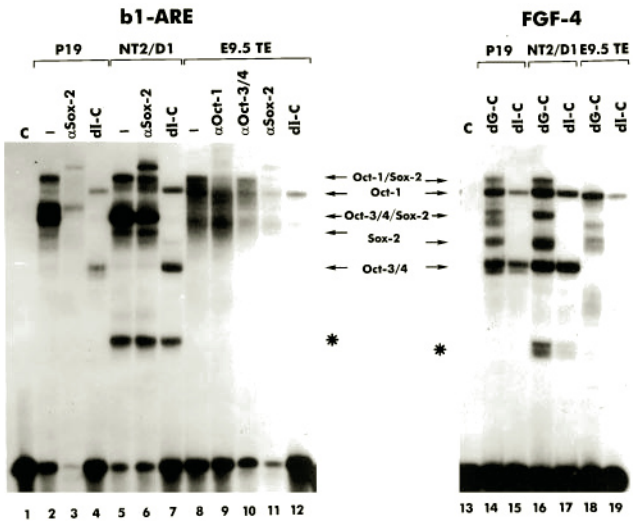

FIG. 6. **SOX/OCT1 and SOX/OCT3/4 complexes bind to the b1-ARE in mouse E9.5 embryonic extracts.** EMSA of double-stranded oligonucleotides representing a region of 34 bp encompassing the SOct site within b1-ARE (b1-ARE, *left side*) or the SOX/OCT bipartite-binding site found within the FGF-4 promoter (FGF-4, *right side*). Nuclear extracts from P19 or NT2/D1 cells and whole cell extracts from embryonic day 9.5 (*E.9.5*) mouse central nervous system were challenged with the labeled oligonucleotides as described under "Experimental Procedures." Anti-Oct1 and anti OCT3/4-specific antisera were used to characterize the two retarded complexes. An *asterisk* indicates a nonspecific binding complex in NT2/D1 cell extracts.

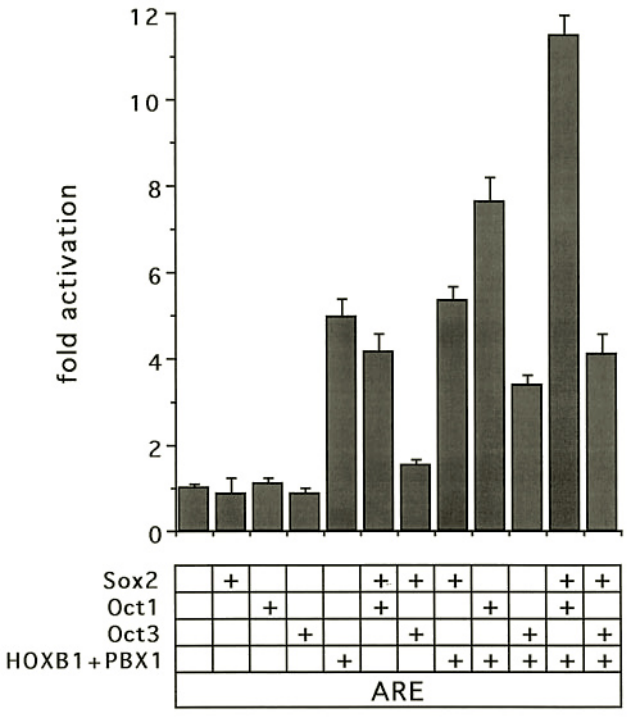

FIG. 7. **SOX-2 activates transcription from the b1-ARE in cooperation with OCT1 but not with OCT3/4 in COS-7 cells.** Results represent the fold activation over the b1-ARE basal luciferase activity assayed from extracts of transiently transfected COS-7 cells. The cells were transfected with 8 μg of pAdMLARE together with 4 μg of the SV40-driven SOX-2, OCT1, OCT3/4, or HOXB1 and PBX1 expression constructs as indicated. 0.2 μg of the pCMVβ-gal plasmid were cotransfected in all experiments as an internal standard. *Bars* represent the mean ± S.E. of at least four independent experiments. An *asterisk* indicates a nonspecific binding complex in NT2/D1 cell extracts.

SOX-2/OCT1 had an additive effect on that of the HOXB1/PBX1 complex on the b1-ARE and caused a further 2–3-fold increase in its transcriptional activity (Fig. 7). OCT1 alone, but not SOX-2, had some effect on the activity on the HOXB1/PBX1 complex, whereas OCT3/4 had no effect, neither alone nor in combination with SOX-2. These results confirm that the SOX/OCT site within the b1-ARE significantly contributes to its enhancer activity by recruiting SOX/OCT heterodimers in the appropriate cell context that are able to increase the transcriptional activity of HOX/PBX complexes.

The SOX/OCT-binding Site Is Necessary for Ectopic Activation of b1-ARE by Retinoic Acid in Vivo—In view of our data indicating that the SOX/OCT-binding sequence flanking R3 is important both for the basal activity of the b1-ARE and for the activation by HOX proteins in cell culture, we investigated the role of these sequence using transgenic analysis. Deletions 3′ of R3 had no effect on the *in vivo* regulatory activity of the enhancer (data not shown). Two additional mutants were generated in the SOct sites in the context of a highly conserved 331-bp *Stu*I/*Hin*dIII fragment containing the b1-ARE (Fig. 8). In the first case, where we introduced a mutation (TAAT to CCGG) into the more 3′ OCT-specific hemisite, *lacZ* reporter expression appeared normal (Fig. 8). At 9.5–10.0 dpc reporter staining was correctly restricted to r4 and second arch neural crest in embryos not exposed to RA and ectopically induced in r2 in all embryos (n = 10) that had been exposed to RA at an earlier time (Fig. 8, *A–D*). This is identical to the properties of the wild type b1-ARE and shows that mutation of the POU homeodomain-binding sequence is still compatible with full enhancer activity. In contrast, a second variant that alters both the OCT (TAAT to CCGG) and the SOX (TTTGTC to CGCTGT)-binding hemisites completely abolished (4/5 embryos), or dramatically reduced (1/5 embryo), the RA-induced ectopic stripe of reporter expression in r2 of embryos exposed to RA (Fig. 8, *E-G*). The same mutation had no obvious effect (7/7 embryos) on normal r4 expression (Fig. 8, *E—G,* and data not shown). However, there may be a quantitatively reduced level of activity in this mutated b1-ARE element that is hard to detect, since β-galactosidase staining allows only a qualitative estimate of transgene expression. These data show that a complete SOX/OCT site, and presumably efficient binding of SOX/OCT heterodimer(s), in combination with the HOX/PBC repeats are necessary *in vivo* for mediating the RA response controlled by the conserved auto-regulatory elements in the b1-ARE.

Hoxa1 Is Required for RA-induced Ectopic Activation of b1-ARE in Vivo—Our results obtained by EC cell transfection indicate that the transactivation of the b1-ARE by a HOXA1/PBX1 complex is weaker and more dependent on the presence of a SOct site as that of the HOXB1/PBX1 complex. This differential activity, given that the response of the b1 ARE to retinoids in embryos requires the SOct site, might reveal HOXA1 as being the primary component involved in mediating the response of Hoxb1 to RA *in vivo*. To test this idea, we crossed a transgenic reporter line including the b1-ARE (31) into a *Hoxa1* mutant genetic background (52) and assayed its response to RA in mutant and wild type embryos. This *Hoxb1/lacZ* transgene does not contain the 3′-RAREs responsible for *Hoxb1* activation in the ectoderm (28, 31) and in the endoderm (53), and therefore any response to RA would be indirect. Untreated wild type and heterozygous $Hoxa1^{+/-}$ embryos show a robust r4 *lacZ* expression (Fig. 9*A*), whereas homozygous $Hoxa1^{-/-}$ untreated embryos show a significantly reduced expression in r4 corresponding to the reduced r4 territory in these mutants (Fig. 9*C*). Upon treatment with RA, 6/6 *Hoxa1* wild type and 9/10 $Hoxa1^{+/-}$ heterozygous embryos responded by ectopically expressing the transgene in the anterior hindbrain (Fig. 9*B*), whereas 5/5 homozygous *Hoxa1* mutant embryos failed to induce ectopically transgene expression (Fig. 9*D*).

Our results *in vivo* indicate that the ectopic activation of this

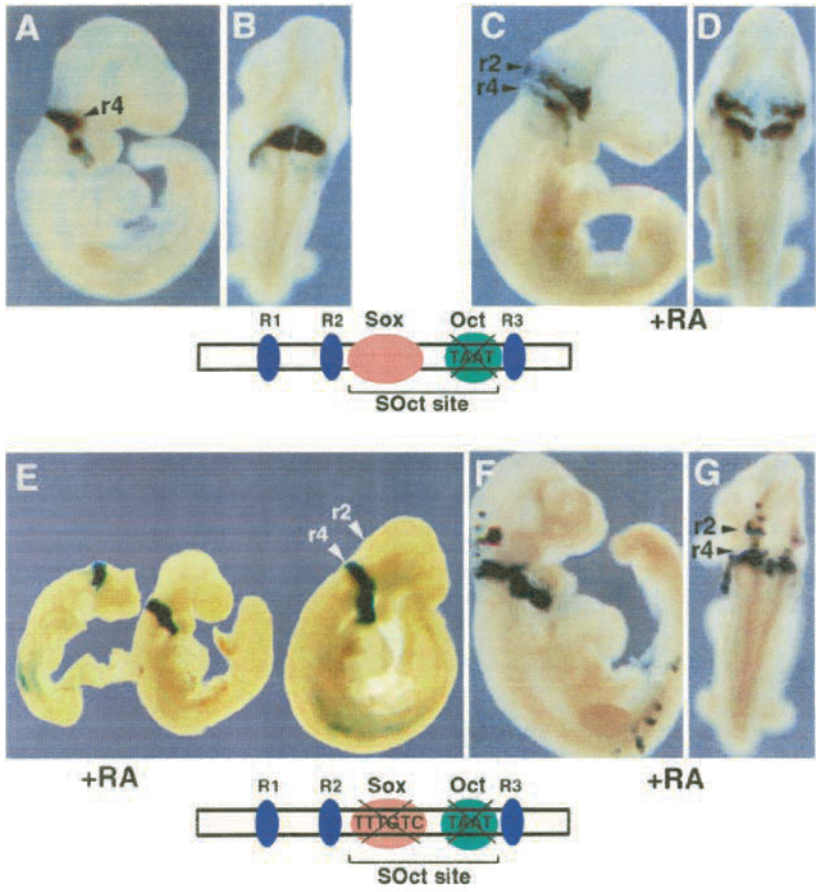

FIG. 8. **Mutations in the SOX/OCT-binding site within the *Hoxb1* r4 enhancer (b1-ARE) influence the response to RA but not r4-restricted expression.** *A–D,* lateral (*A* and *C*) and dorsal (*B* and *D*) views of transgene expression in 9.5–10.0 dpc. embryos carrying a *lacZ* reporter construct with a mutated form of the OCT-specific hemisite (TAAT to CCGG) in the context of the highly conserved 331-bp *Stu*I-*Hin*dIII *Hoxb1* r4-autoregulatory enhancer (6). The embryo in *A* and *B* was not exposed to ectopic RA, whereas that in *C* and *D* was isolated from a female given RA at 7.75 dpc by oral gavage. This mutation does not impair the ability of the r4 enhancer to direct expression to r4 (*A* and *B*) or to mediate a response to ectopic doses of RA as indicated by the induction of a second stripe of expression in r2 (*C* and *D*). *E–G,* lateral (*E* and *F*) and dorsal (*G*) views of reporter expression in 9.0–10.5-dpc embryos treated with RA and carrying the combined mutations in the SOX- (TTTGTC to CGCTGT) and OCT (TAAT to CCGG)-specific hemisites. In both untreated and RA-treated cases r4 expression is not affected (*E–G* and data not shown). However in the majority of cases (4/5) the double mutant construct fails to mediate a response to ectopic RA (*E*), and in the single example of an RA response the induction in r2 is very patchy and incomplete (*F* and *G*), compared with wild type or the single TAAT mutation (*C* and *D*). A schematic diagram of the transgenic constructs and mutations is indicated *below* the panels. The three bipartite HOX/PBX repeats are indicated as R1–R3.

element by RA requires both the presence of a functional SOX/OCT site and the presence of a functional *Hoxa1* gene. This implies that for HOXA1 to function efficiently on the b1-ARE, it is important that SOX and OCT proteins are recruited to the SOct site. It is interesting that even though HOXA1 and HOXB1 both participate in regulating normal r4 expression (27, 28), we find that, unlike the RA response, transgene expression is maintained in r4 despite mutation of the SOct site. This suggests that *in vivo* as in the EC cells, HOXB1 is not as dependent as HOXA1 upon SOX and OCT proteins and is able to activate the b1-ARE in r4.

Determinants for the Differential Activity of HOXB1 and HOXA1 Reside within the HOXB1 N-terminal Domain—To examine whether the differential activation of HOXA1 and HOXB1 on the b1-ARE is related to intrinsic structural differences between these proteins, we used a domain swap approach. The HOXA1 and HOXB1 proteins have almost identical homeodomains (54), and they share a conserved FDWM motif just N-terminal to the homeodomain that serves as the HOX/PBX interaction surface, necessary for cooperative DNA binding and transcriptional activation (reviewed in Ref. 3). HOXA1 and HOXB1 bind the b1-ARE R3 site with the same affinity *in vitro* and have similar transcriptional activity on the full b1-ARE element in non-EC cells and in transgenic mice (5, 6, 55). However, the two proteins differ completely in their C-terminal and N-terminal regions (56), and for HOXB1 the N-terminal region contains the major transcriptional activation domain used in the context of heterodimers with PBX1 (5, 26). Therefore, we designed two chimeric mutants by replacing either the N terminus (aa 1–205) and the C terminus, or only the C terminus (aa 290–336) of HOXA1, with the corresponding regions of HOXB1 (aa 1–174 and 258–296) (Fig. 10*B*), leaving the HOXA1 homeodomain and FDWM regions intact (Fig. 10*B*).

Expression vectors encoding the HOXA1/B1CT and the HOXB1/A1HD chimeric proteins were cotransfected with PBX1 in P19 cells together with reporters under the control of either the entire b1-ARE (pAdMLARE) or R3 (pAdMLR3). In controls, both chimeras bind the b1-ARE R3 element *in vitro* with comparable affinity, and both activate transcription from the pAdMLR3 reporter in EC cells and COS-7 cells (Fig. 10*A* and data not shown). Testing the response using the entire b1-ARE control region, the transcriptional activity of the chimera carrying only the C-terminal replacement remained low and comparable to that of unmodified HOXA1. In contrast, the activity of the chimera containing the combined C- and N-terminal replacement was enhanced and comparable to that of wild type HOXB1 (Fig. 10*A*).

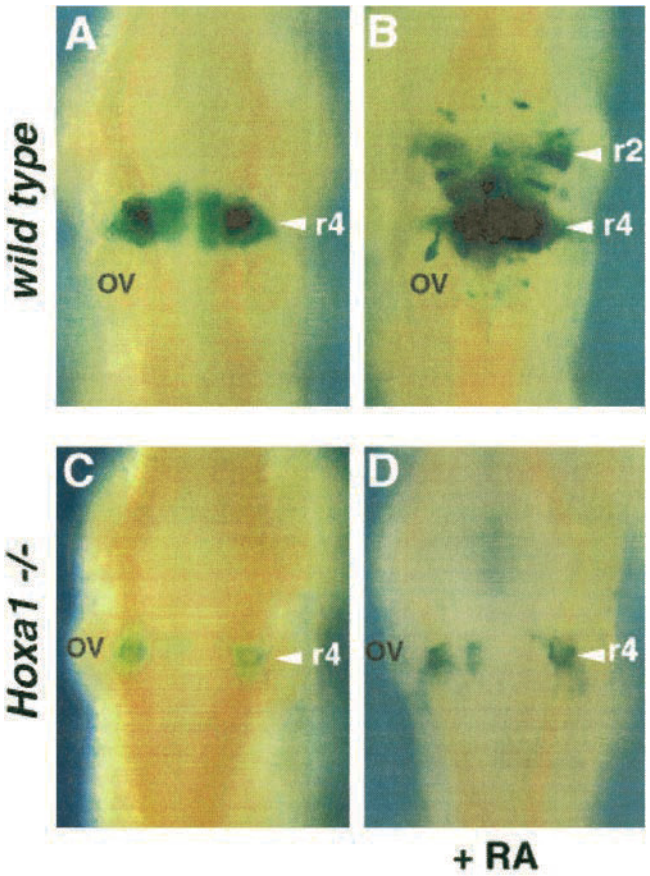

FIG. 9. **Targeted inactivation of *Hoxa1* impairs the RA-induced ectopic activation of b1-ARE.** Dorsal (*A–D*) views of transgene expression in 9.5–10.0-dpc. embryos carrying a *Hoxb1* r4-autoregulatory enhancer-driven *lacZ* reporter construct (6). *A* and *B, Hoxa1* wild type background. The embryo in *A* was not exposed to ectopic RA, whereas that in *B* was isolated from a female given RA at 7.75 dpc by oral gavage. The Hoxb1-ARE *lacZ* reporter is ectopically activated by RA treatment and displays a stripe of expression in r2 in 6/6 embryos (*B*). *C* and *D, Hoxa1−/−* background. The *Hoxb1/lacZ* transgene fails to be expressed ectopically in r2 in 5/5 embryos isolated from females treated with RA at 7.75 dpc by oral gavage (*D*). In both untreated (*C*) and RA treated (*D*) cases r4 expression is not affected. *ov*, otic vesicle.

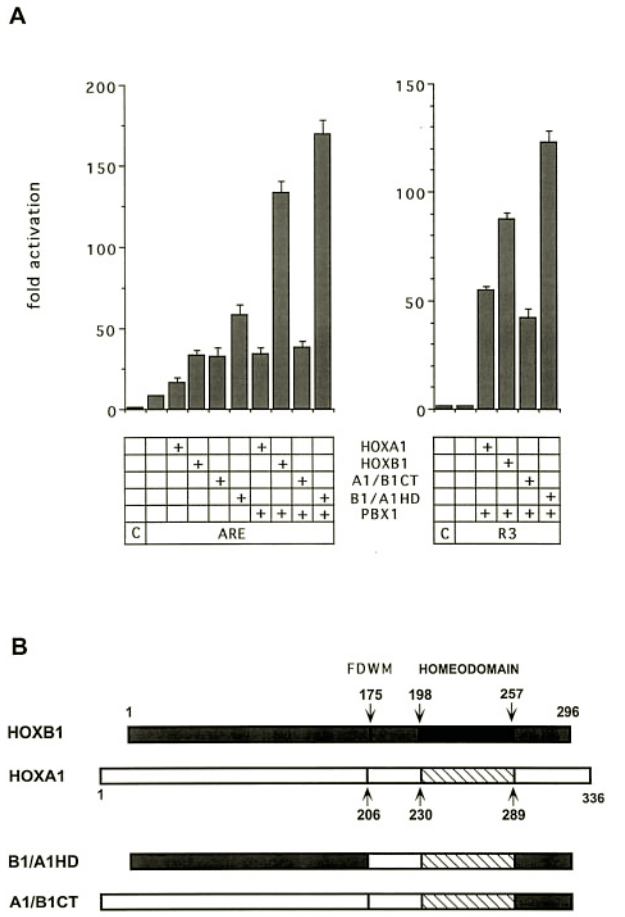

FIG. 10. **The determinants for the selective transactivation of b1-ARE reside within the N-terminal transcriptional activation domain of HOXB1.** *A,* results represent the fold activation over the b1-ARE basal luciferase activity assayed from extracts of transiently transfected P19 cells. The cells were transfected with 4 μg of the SV40-driven expression constructs for HOXA1, HOXB1, or their chimeric derivatives, and/or for PBX1, together with 8 μg of pAdMLluc (*C*), pAdMLARE (*ARE*), or pAdMLR3 (R3). 0.2 μg of the pCMVβ-gal plasmid were cotransfected in all experiments as an internal standard. *B,* schematic representation of the HOXB1 and the HOXA1 proteins and of their chimeric derivatives. *Numbers* indicate amino acid positions.

These results show that specific DNA recognition and cooperative binding with PBX1, provided by the region encompassing the homeo- and FDWM domains, are not responsible for the differential activity of HOXA1 and HOXB1 on the b1-ARE. Conversely, the HOXB1 N-terminal region, comprising the transcriptional activation domain, appears to contain specific determinants that are able to respond to cues from the cell background leading to a higher activity with respect to HOXA1.

DISCUSSION

Mammalian HOX proteins are a large family of transcription factors that control cell identity, differentiation, and patterning in animal embryonic development. Due to the similarities in their structure, *in vitro* binding abilities, and *in vivo* function, it is important to understand the factors and components that serve to regulate the selectivity and specificity of HOX transcription complexes during activation of downstream target genes. In this study we have addressed some of these issues by examining the regulatory properties and functional requirements of an *in vivo Hox*-responsive target sequence represented by an auto-regulatory enhancer (b1-ARE) from the *Hoxb1* gene. We have used human or murine EC cells, which retain most of the characteristics of primitive neuroectodermal cells (32, 33), as a model system to test the overall transcriptional activity of the b1-ARE. We found that the b1-ARE has both higher basal and higher HOX/PBX-induced activities in neuro-ectodermal EC cells as compared with other cell backgrounds. The enhanced EC activity correlates with the presence of a bipartite motif in the b1-ARE that binds SOX and OCT proteins and is necessary for the optimal response of b1-ARE to transcriptional activation by HOX/PBX heterodimers both *in vitro* and *in vivo*. Surprisingly, despite recent evidence suggesting that Hox paralogs are functionally equivalent (57), our analysis revealed that there are differences in the way that the b1-ARE responds to HOXB1 and HOXA1, which correlate with differences in their N-terminal domains. Together, our results have uncovered some novel aspects of how HOX proteins interact with their *in vivo* target sequences, identified cis-elements that can modulate HOX/PBX complex activities, and raised a number of important issues on the functional differences between paralogous HOX proteins.

Cis-determinants in the b1-ARE—Our analysis has revealed that a highly conserved sequence motif located immediately upstream of a HOX/PBX sequence (R3) is an important determinant of the cell type-specific restriction of the b1-ARE activity. This bipartite element mediates high affinity binding of a SOX/OCT heterodimer that contributes to enhancer activity in addition to, and in combination with, HOXA1/PBX1 and

HOXB1/PBX1 heterodimers. This motif (SOct) allows cooperative binding of SOX-2/OCT1 and SOX-2/OCT3/4 heterodimers present in extracts of murine or human EC cells and mouse embryos. The binding of the SOX-2/OCT1 complex to the b1-ARE site actually appears to be stronger than that observed with the SOX/OCT-binding sequence located in the developmentally regulated FGF-4 enhancer (50, 51). This difference could be explained by the different spacing between the SOX- and OCT-specific hemisites on the two bipartite sequences. In the b1-ARE sequence, the SOX and Octamer sites actually overlap by 1 base pair, whereas in the FGF-4 enhancer the two sites are spaced apart by 2 base pairs. Since it was previously shown that the distance between the two sites is critical for the assembly of the SOX/OCT heterodimer in the context of the FGF-4 enhancer (46), it seems likely that the configuration of the b1-ARE site allows a different interaction between SOX-2 and OCT1 or OCT3/4, leading to a stronger binding of the SOX-2/OCT complexes. The slightly different mobility of the SOX-2/OCT3/4 complex on the b1-ARE *versus* the FGF-4 sequence (see Fig. 5) might be caused by a differential bending of the DNA induced by SOX-2 that might also affect binding strength. High affinity binding of a SOX-2/OCT3/4 complex was recently reported on a closely spaced bipartite sequence found in the upstream regulatory element of the murine UTF1 transcription factor (58).

The *Hoxb1* SOX/OCT (SOct) site contains at its 5′ end a potential binding site (TGACAA) for the TALE homeodomain proteins MEIS and PREP. This is adjacent to R2 and separated from R3 by 17 nucleotides. A similar combination of a TALE site and a HOX/PBX site is also found in an r4-restricted cross-regulatory enhancer from the *Hoxb2* gene, except that the sites are separated by 8 base pairs (11, 15). In the case of *Hoxb2* there is no overlapping SOX/OCT site, but the TALE and HOX/PBX sites synergize *in vitro* and *in vivo* to allow the formation of a trimeric and transcriptionally active HOXB1-PBX1-MEIS/PREP complex required for enhancer activity in transgenic assays (10, 11). In the context of the b1-ARE, however, MEIS1 or PREP1 is not necessary for DNA binding of HOX/PBX heterodimers and, accordingly, is not required to bind DNA to enhance the transcriptional activity of the HOXB1/PBX1 complex in transfected cells (8). Indeed, in EMSAs none of the retarded complexes formed on the complete SOct sequence by either EC cell or mouse embryo extracts appeared to contain MEIS or PREP proteins (data not shown).

The Role of OCT and SOX Proteins in b1-ARE Enhancer Activity in EC Cells—Transient transfection analysis showed that the SOct site is responsible for the high basal transcriptional activity of the b1-ARE sequence in EC cells and is also required for fully activated transcription in the presence of HOXB1 and PBX1. Interestingly, transfection of SOX-2 and OCT1 in non-EC cells (COS-7), which do not express endogenous SOX proteins, significantly increases both the basal activity of the b1-ARE reporter and the HOXB1/PBX1-mediated transactivation. The nature of the OCT partner seems to be important, as transfection of SOX-2 and OCT3/4 had no effect on the b1-ARE. Conversely, a SOX-2/OCT3/4 complex but not a SOX-2/OCT1 complex was reported to have a positive effect on the transcriptional activity of the FGF-4 (51), osteopontin (47), and UTF1 (58) enhancers in HeLa cells. These results indicate that the b1-ARE SOX/OCT site might have unique binding properties *in vivo* compared with previously reported bipartite sites and facilitates functional interaction specifically for the SOX-2/OCT1 complex. Therefore, both sequence and spacing between the two hemisites appear to be important general factors in dictating specific binding of heterodimeric complexes between SOX-2 and alternative OCT proteins to this class of target elements.

In Vivo Role for the SOX/OCT Site in b1-ARE Regulation—These experiments suggest a model whereby the bipartite SOX/OCT site contributes to the activity of the b1-ARE enhancer *in vivo* by recruiting SOX-2/OCT1 heterodimers. This in turn could either increase the stability or the affinity of binding by HOXB1-PBX1 and HOXA1/PBX1 complexes and/or activate transcription in synergy with such complexes. In support of this, combined mutations in both parts of the SOX/OCT site show that it is necessary for full *in vivo* activity of the b1-ARE in transgenic mouse embryos.

Genetic and regulatory analyses have previously shown that *Hoxa1* and *Hoxb1* function synergistically to regulate the early r4 expression of *Hoxb1* (7.5–8.5 dpc) and that this segmental expression is maintained from 8.5 dpc and onwards by *Hoxb1* itself (6, 27, 28, 59). Mutation of both the SOX- and OCT-specific hemisites in the b1-ARE does not abolish reporter expression in r4 at 9.5 dpc (Fig. 8). This expression in later stages implies that endogenous HOXB1/PBX complexes are still able to activate the mutant b1-ARE reporter. This is consistent with our findings in EC cells that showed HOXB1/PBX complexes could stimulate transcription in the absence of the SOct site, although not at maximal levels, whereas the HOXA1/PBX complexes displayed an absolute requirement for this site in transactivation (Fig. 4). In contrast, mutation of the SOct site impairs the ability of the b1-ARE reporter to generate a stripe of ectopic expression in r2 in response to *in utero* RA treatment (Fig. 8). Furthermore, this ectopic RA response is also abolished in *Hoxa1* mutant embryos (Fig. 9). These results suggest that, as in EC cells, the interaction of HOXA1/PBX complexes with factors binding to the SOct site is essential for generating complexes with sufficient activity to trigger the b1-ARE auto-regulatory loop in r2. This underscores the importance of the SOct site for *in vivo* activity and differential dependence of HOXA1/PBX *versus* HOXB1/PBX complexes upon this site for modulating transcriptional activity. It is interesting that the ectopic response of the transgene to RA depends upon HOXA1 and that HOXB1 does not compensate. We have found that the reason for this is that while both the *Hoxa1* and *Hoxb1* genes have 3′-RAREs necessary for early neural expression, the Hoxb1 3′-RARE does not respond to *in utero* RA treatment (data not shown).

A mutation in the OCT-specific hemisite, which is recognized by the POU homeodomain of the Octamer factor, has little or no effect in transgenic mice. This suggests that the SOX/OCT complex can assemble *in vivo* on a suboptimal target sequence, through interaction of the SOX factor with the POU-specific domain of the Octamer factor. Similar data on the relative importance of the POU-specific sequence for binding of OCT1 and OCT3/4 to a bipartite SOX/OCT site *in vitro* was also observed in the case of the UTF1 upstream regulatory element (58).

Selectivity and HOX Protein Determinants—The homeodomains of HOXB1 and HOXA1 are almost identical, and in combination with PBX proteins display a virtually indistinguishable binding specificity. Therefore, the differential activity of HOXB1 or HOXA1 complexes with PBX most likely resides in specific interactions with other DNA-binding factors and/or with the transcriptional machinery. Indeed, our analysis demonstrates that a critical determinant of the difference between the HOXB1/PBX1 complex and that of HOXA1/PBX1 resides in the HOXB1 N-terminal transcriptional activation domain. A chimeric protein, where the N-terminal domain of HOXA1 is replaced with that of HOXB1, makes it indistinguishable from wild type HOXB1 in activating the b1-ARE in cooperation with PBX1. This suggests that DNA recognition is not the key variable involved in modulating the activity of

group-1 HOX proteins on the b1-ARE target and that factors interacting with the N terminus of these HOX proteins are responsible for mediating their differential activities in a tissue- or cell-specific manner.

Although we have focused on how distinct HOX proteins in the heterodimers differentially interact with cofactors, diversity in PBX partners may also contribute to the selectivity. The existence of both the PBX and other novel HOX cofactors might explain the observed tissue-specific restriction of some *Hox*-responsive enhancers *in vivo*. A recent report (57) has challenged the view that paralogous HOX proteins have gained intrinsic functional diversity in the course of evolution. They propose that paralogous proteins are functionally equivalent, and it is only the relative levels or domains of their expression that govern their unique activities. Our results, conversely, show the existence of intrinsic differences in function between the HOXA1 and HOXB1 proteins. In this regard, the difference between HOXB1 and HOXA1 in their ability to interact functionally with the same *Hoxb1* auto-regulatory element in a specific cell context is a clear example of non-redundancy between products of *Hox* genes belonging to the same paralogous group. We feel that this observation may also extend to other paralogous groups, and it will be important to characterize the nature of the proteins interacting with the N-terminal domains of the HOX proteins.

Acknowledgment—We thank Hans Schoeler for the kind gift of anti-OCT3/4 antiserum.

REFERENCES

1. Krumlauf, R. (1994) *Cell* **78,** 191–201
2. Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G., and Wuthrich, K. (1994) *Cell* **78,** 211–223
3. Mann, R. S., and Chan, S.-K. (1996) *Trends Genet.* **12,** 258–262
4. Mann, R. (1996) *BioEssays* **17,** 855–863
5. Di Rocco, G., Mavilio, F., and Zappavigna, V. (1997) *EMBO J.* **16,** 3644–3654
6. Pöpperl, H., Bienz, M., Studer, M., Chang, S.-K., Aparicio, S., Brenner, S., Mann, R., and Krumlauf, R. (1995) *Cell* **81,** 1031–1042
7. Mann, R. S., and Affolter, M. (1998) *Curr. Opin. Genet. & Dev.* **8,** 423–429
8. Berthelsen, J., Zappavigna, V., Ferretti, E., Mavilio, F., and Blasi, F. (1998) *EMBO J.* **17,** 1434–1445
9. Berthelsen, J., Zappavigna, V., Mavilio, F., and Blasi, F. (1998) *EMBO J.* **17,** 1423–1433
10. Jacobs, Y., Schnabel, C. A., and Cleary, M. L. (1999) *Mol. Cell. Biol.* **19,** 5134–5142
11. Ferretti, E., Marshall, H., Popperl, H., Maconochie, M., Krumlauf, R., and Blasi, F. (2000) *Development* **127,** 155–166
12. Zappavigna, V., Renucci, A., Izpisua-Belmonte, J. C., Urier, G., Peschle, C., and Duboule, D. (1991) *EMBO J.* **10,** 4177–4187
13. Pöpperl, H., and Featherstone, M. S. (1992) *EMBO J.* **11,** 3673–3680
14. Gould, A., Morrison, A., Sproat, G., White, R. A., and Krumlauf, R. (1997) *Genes Dev.* **11,** 900–913
15. Maconochie, M. K., Nonchev, S., Studer, M., Chan, S. K., Popperl, H., Sham, M. H., Mann, R. S., and Krumlauf, R. (1997) *Genes Dev.* **11,** 1885–1895
16. Lu, Q., and Kamps, M. P. (1997) *Oncogene* **14,** 75–83
17. Chang, C.-P., Shen, W.-F., Rozenfeld, S., Lawrence, H. J., Largman, C., and Cleary, M. (1995) *Genes Dev.* **9,** 663–674
18. Knoepfler, P., and Kamps, M. (1995) *Mol. Cell. Biol.* **15,** 5811–5819
19. Phelan, M. L., Rambaldi, I., and Featherstone, M. S. (1995) *Mol. Cell. Biol.* **15,** 3989–3997
20. Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999) *Nature* **397,** 714–719
21. Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., and Wolberger, C. (1999) *Cell* **96,** 587–597
22. Rambaldi, I., Kovacs, E. N., and Featherstone, M. S. (1994) *Nucleic Acids Res.* **22,** 376–382
23. Zappavigna, V., Sartori, D., and Mavilio, F. (1994) *Genes Dev.* **8,** 732–744
24. Zhao, J. J., Lazzarini, R. A., and Pick, L. (1996) *EMBO J.* **15,** 1313–1322
25. Schnabel, C. A., and Abate-Shen, C. (1996) *Mol. Cell. Biol.* **16,** 2678–2688
26. Viganò, A., Di Rocco, G., Zappavigna, V., and Mavilio, F. (1999) *Mol. Cell. Biol.* **18,** 6201–6212
27. Studer, M., Lumsden, A., Ariza-McNaughton, L., Bradley, A., and Krumlauf, R. (1996) *Nature* **384,** 630–634
28. Studer, M., Gavalas, A., Marshall, H., Ariza-McNaughton, L., Rijli, F. M., Chambon, P., and Krumlauf, R. (1998) *Development* **125,** 1025–1036
29. Frasch, M., Chen, X., and Lufkin, T. (1995) *Development* **121,** 957–974
30. Dupe, V., Davenne, M., Brocard, J., Dolle, P., Mark, M., Dierich, A., Chambon, P., and Rijli, F. M. (1997) *Development* **124,** 399–410
31. Marshall, H., Studer, M., Pöpperl, H., Aparicio, S., Kuroiwa, A., Brenner, S., and Krumlauf, R. (1994) *Nature* **370,** 567–571
32. Andrews, P. W. (1988) *Biochim. Biophys. Acta* **948,** 17–36
33. Pleasure, S. J., and Lee, V. M. Y. (1993) *J. Neurosci. Res.* **35,** 585–602
34. Green, S., Issemann, I., and Sheer, E. (1988) *Nucleic Acids Res.* **16,** 369–373
35. Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988) *Nature* **335,** 563–564
36. Nordeen, S. K. (1988) *BioTechniques* **6,** 454–457
37. Zappavigna, V., Falciola, L., Helmer Citterich, M., Mavilio, F., and Bianchi, E. M. (1996) *EMBO J.* **15,** 4981–4991
38. Yee, S.-P., and Rigby, P. W. J. (1993) *Genes Dev.* **7,** 1277–1289
39. Di Nocera, P. P., and Dawid, I. B. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 7095–7098
40. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
41. Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P., Stott, D., and Allemann, R. (1991) *Genes Dev.* **5,** 2048–2059
42. Marshall, H., Nonchev, S., Sham, M. H., Muchamore, I., Lumsden, A., and Krumlauf, R. (1992) *Nature* **360,** 737–741
43. Deschamps, J., De Laaf, R., Joosen, L., Meijlink, F., and Destrée, O. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 1304–1308
44. Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E., and Mavilio, F. (1990) *Nature* **346,** 763–766
45. Simeone, A., Acampora, D., Nigro, V., Faiella, A., D'Esposito, M., Stornaiuolo, A., Mavilio, F., and Boncinelli, E. (1991) *Mech. Dev.* **33,** 215–228
46. Ambrosetti, D. C., Basilico, C., and Dailey, L. (1997) *Mol. Cell. Biol.* **17,** 6321–6329
47. Botquin, V., Hess, H., Fuhrmann, G., Anastassidis, C., Gross, M. K., Vriend, G., and Schöler, H. R. (1998) *Genes Dev.* **12,** 2073–2090
48. Kamachi, Y., Sockanathan, S., Liu, Q., Breitman, M., Lovell-Badge, R., and Kondoh, H. (1995) *EMBO J.* **14,** 3510–3519
49. Ryan, A. K., and Rosenfeld, M. G. (1997) *Genes Dev.* **11,** 1207–1225
50. Dailey, L., Yuan, H., and Basilico, C. (1994) *Mol. Cell. Biol.* **14,** 7758–7769
51. Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995) *Genes Dev.* **9,** 2635–2645
52. Lufkin, T., Dierich, A., LeMeur, M., Mark, M., and Chambon, P. (1991) *Cell* **66,** 1105–1119
53. Huang, D., Chen, S. W., Langston, A. W., and Gudas, L. J. (1998) *Development* **125,** 3235–3246
54. Duboule, D. (ed) (1994) *Guidebook to the Homeobox Genes*, Oxford University Press, Oxford
55. Zhang, M., Kim, H.-J., Marshall, H., Gendron-Maguire, M., Lucas, A. D., Baron, A., Gudas, L. J., Gridley, T., Krumlauf, R., and Grippo, J. F. (1994) *Development* **120,** 2431–2442
56. Acampora, D., D'Esposito, M., Faiella, A., Pannese, M., Migliaccio, E., Morelli, F., Stornaiuolo, A., Nigro, V., Simeone, A., and Boncinelli, E. (1989) *Nucleic Acids Res.* **17,** 10385–10402
57. Greer, J. M., Puetz, J., Thomas, K. R., and Capecchi, M. R. (2000) *Nature* **403,** 661–665
58. Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999) *Mol. Cell. Biol.* **19,** 5453–5465
59. Gavalas, A., Studer, M., Lumsden, A., Rijli, F. M., Krumlauf, R., and Chambon, P. (1998) *Development* **125,** 1123–1136