Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 387–391, January 1995
Developmental Biology

A retinoic acid-triggered cascade of *HOXB1* gene activation

TOSHIHIKO OGURA* AND RONALD M. EVANS*†‡

†Howard Hughes Medical Institute, *The Salk Institute for Biological Studies, La Jolla, CA 92037

Communicated by Wylie Vale, The Salk Institute, La Jolla, CA, September 23, 1994

ABSTRACT **Retinoic acid (RA) has been proposed to be a direct regulator of *HOX* gene complexes. However, the molecular mechanism of the RA signaling pathway during normal development is unclear. We have identified an RA-responsive element in the promoter of *HOXB1* gene composed of two functionally separable sites: (*i*) a DR-2 sequence, which is the direct target of the RA receptor·retinoid X receptor heterodimer; and (*ii*) a motif for an RA-inducible and tissue-specific coactivator termed retinoid-inducible protein. Though neither enhancer alone is functional, this combined element strongly activates the *HOXB1* promoter in a cell-specific and retinoid-dependent manner. Finally, this activation is potentiated by a proximal autoregulatory site for *HOXB1* gene itself. These data define a tripartite cascade leading to the establishment of *HOXB1* gene activation.**

The mammalian homeobox genes (*HOX*) encode a family of >30 related proteins that share the common "homeo box" motif originally identified in *Drosophila Antennapedia/Bithorax* homeotic complex (1). Human homeobox gene clusters designated *HOX A, B, C,* and *D* were mapped on chromosomes 7, 17, 12, and 2, respectively, and retain a linear gene arrangement similar to their *Drosophila* counterparts (2). Expression of mammalian homeobox genes is strictly regulated temporally and spatially during embryonic development (3–5). Strikingly, it has been shown that the position of mammalian and *Drosophila* homeobox genes in each cluster follows the order of the anterior boundaries of their expression along the anterior–posterior axis of the developing embryo (6, 7).

Retinoic acid (RA), a natural metabolite of vitamin A, has been proposed to be a vertebrate morphogen and a regulator of the *HOX* gene clusters (8, 9). Systemic treatment of vertebrate embryos with RA results in severe developmental deformities, whereas local application of RA to the chicken limb bud produces digit duplication that is accompanied by a change of homeobox gene expression (10–12).

In the human embryonal carcinoma cell line NT2/D1 (13), homeobox genes are sequentially activated by RA in a graded fashion from the 3′ → 5′ direction (14). Activation of the 3′ *HOXB1* gene is not dependent on protein synthesis and thus is a candidate for direct regulation by RA (15). However, the precise molecular link between *HOXB1* and RA signaling and the mechanism establishing graded chromosomal expression remain obscure. The actions of RA are mediated by the RA receptors (RARs) and retinoid X receptors (RXRs), both members of the nuclear receptor family (16–22). These receptors have been shown to function through a heterodimer to create a sequence-specific DNA binding complex that activates target genes through an RA-responsive element (RARE) in a hormone-dependent pathway. Mouse *Hoxa-1* and *Hoxb-1* genes are located at the 3′ end of each cluster (*labial* paralogues) and have a similar pattern of expression in the posterior half of the 7.5- to 8.0-day mouse embryo (23). It is important to note that both genes have been reported to have RAREs in the 3′ regions (24, 25). Our preliminary data suggest that this RARE is required for the establishment of early expression in the posterior half of the embryo. However, by day 8.5 of gestation, expression of the *Hoxa-1* and *Hoxb-1* genes has retreated caudally with the exception of the persistence of *Hoxb-1* in rhombomere 4. As reported, this rhombomere 4-specific expression is sensitive to RA, a second RA signaling pathway for midgestational activation. We provide evidence for a 5′ promoter proximal set of regulatory elements that controls *HOXB1* gene expression in a cell-specific and retinoid-dependent manner.

MATERIALS AND METHODS

Plasmids. A DNA fragment containing the thymidine kinase (*TK*) luciferase gene was ligated into the *Sal* I site in pBluescript SK (+) vector (Stratagene) using the *Sal* I linker (pBS.TK.Luc). *TK* promoter was replaced by various genomic DNA fragments derived from the 5′ region of *HOXB1* gene of human cosmid clone cwd3. DR-2A (5′-TCACTCCCTGAACTCTTGCCCTCCTGGACT-3′), DR-2B (5′-CGGGCTGACCTTTTTACCTCGAAGCG-3′), upstream response element (URE) (5′-CAGGCAGACACCCTGACAGGTTACAAATGA GCGTGG-3′), and UREmut (5′-CAGGCAGACACACTAGTAGGTTACAAATGAGCGTGG-3′) oligonucleotides were used for the construction.

Cell Cultures and Transfection. Embryonal carcinoma cell lines P19 and NT2/D1 (13) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (Irvine Scientific). Twenty-four hours (1 hr for P19 cells) before transfection, cells were split in fresh medium described above. Transfection was performed by using the calcium phosphate precipitation method. Five micrograms of reporter luciferase plasmid and 7 μg of pCMX-βGAL (as an internal control) were transfected. NT2/D1 and P19 cells were incubated for 12 hr in the presence of DNA precipitates. After DNA precipitates were washed, cells were cultured with fresh medium containing 1 μM RA for another 24 hr. Cells were harvested and the luciferase assays were carried out. Transfection efficiency was normalized by using β-galactosidase activity derived from pCMX-βGAL.

Gel Retardation Assay. pCMX-hRARα and pCMX-hRXRα were linearized, and capped mRNA was synthesized *in vitro* using T7 RNA polymerase (Stratagene) according to the manufacturer's instruction. Aliquots of mRNA were incubated with rabbit reticulocyte lysate (Promega) for *in vitro* translation. For gel retardation assay, 5 μl of *in vitro* translated proteins was preincubated in binding buffer [10 mM Tris, pH 8.0/40 mM KCl/0.05% Nonidet P-40/6% glycerol/1 mM dithiothreitol/5 μg of poly(dI·dC) per ml] on ice for 15 min. For competition assays, a 20-fold molar excess of competitor oligonucleotides was mixed at this step. Oligonucleotides containing DR-5 (RARβ2 RARE), *HOXB* DR-2A, and DR-2B were used as competitors and probes.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RA, retinoic acid; RAR, RA receptor; RXR, retinoid X receptor; RARE, RA-responsive element; URE, upstream response element; RIP, retinoid-inducible protein; TK, thymidine kinase.

‡To whom reprint requests should be addressed.

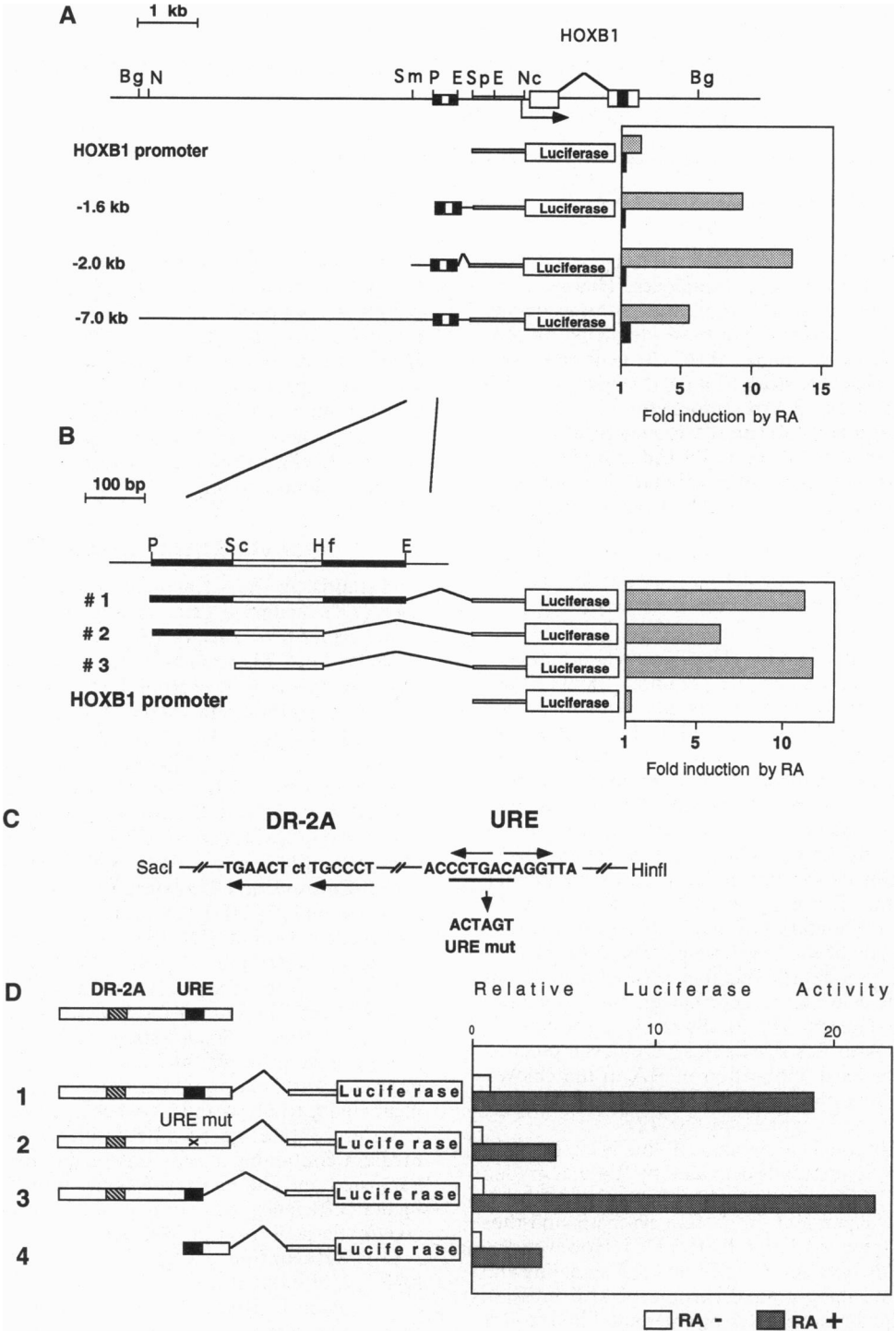

FIG. 1. (*A*) RA-dependent activation of *HOXB1* promoter. Various 5′ deletion constructs were transfected to P19 cells (shaded bars) and NT2/D1 cells (solid bars). The box indicates an RA-responsive site in the *Pst* I–*Eco*RI region. The transcription initiation site is indicated by an arrow. Fold induction by 1 μM RA is shown at the right of each construct. Note that a 1-kb *Spe* I–*Nco* I fragment (double lined) was used as the basal *HOXB1* promoter. Enzyme designations are as follows: Bg, *Bgl* II; E, *Eco*RI; Hf, *Hin*fI; N, *Nde* I; Nc, *Nco* I; P, *Pst* I; Sc, *Sac* I; Sm, *Sma* I; Sp, *Spe* I. (*B*) More detailed transfection analysis localizes the 160-bp *Sac* I–*Hin*fI region (indicated by open box) as an RA-responsive site (P19 cells). (*C*) Sequences of DR-2A and the URE. The palindromic sequence of the URE is indicated by arrows. A mutation (UREmut) was introduced as indicated and used for the construction of expression plasmids and the gel retardation assay. The binding site for RAR·RXR heterodimer is indicated by arrows (DR-2A). (*D*) Activation profiles of deletion and UREmut constructs are shown (constructs 1–4). The basal *HOXB1* promoter was used as in *A* and *B*.

Subsequently, ^{32}P-labeled oligonucleotide probes were added to the reaction mixtures and incubated on ice for 15 min. The same oligonucleotides as those used for the construction of luciferase plasmids were used. Reaction mixtures were resolved by 5% polyacrylamide gel electrophoresis in 0.5× TBE (1× TBE = 90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3). The dried gels were autoradiographed at −70°C. Nuclear extracts of P19 and NT2/D1

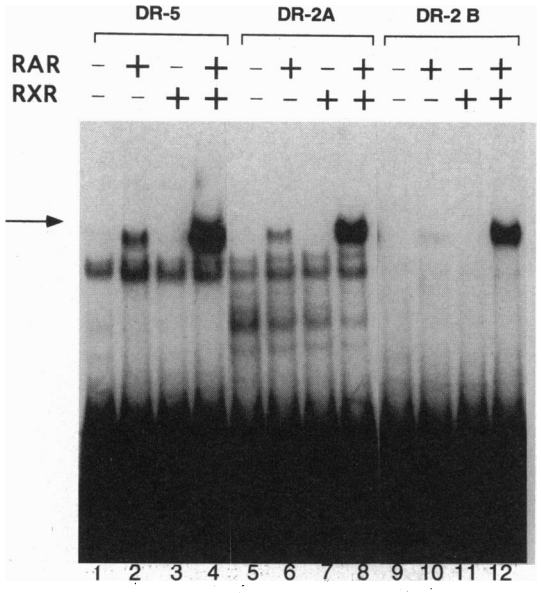

FIG. 2. RAR·RXR heterodimer formation on DR-2A and DR-2B sequences. The gel retardation assay was conducted using *in vitro* translated RARα and/or RXRα proteins. ^{32}P-labeled DR-5 probes derived from RARβ2 promoter (lanes 1–4), *HOXB1* DR-2A (lanes 5–8), and DR-2B (lanes 9–12) were used. Specific heterodimer complexes are indicated by the arrow. Unprogramed reticulocyte lysate was used as a control (lanes 1, 5, and 9).

cells cultured in the absence of or in the presence of 1 μM RA for 2 days were prepared according to the method of ref. 26 and stored at −80°C. Three to 5 μg of protein was used for each reaction.

RESULTS

To analyze the regulatory elements of the *HOXB1* promoter, an overlapping set of large chromosomal fragments was isolated and ligated to a luciferase reporter gene. As shown in Fig. 1*A*, a 7-kb fragment spanning the *HOXB1* promoter was transfected into two embryonal carcinoma cell lines: P19 (shaded bars) and NT2/D1 (solid bars). In P19 cells, but not in NT2/D1 cells, this plasmid was strongly induced by 1 μM RA. Reporter constructs containing either −2.0 kb or −1.6 kb of upstream sequence retained full enhancer activity, whereas deletions containing −1.2 kb or less produced only marginal induction. These data indicate the presence of a specific enhancer in an ≈400-bp region between *Pst* I and *Eco*RI restriction endonuclease sites (Fig. 1*B*). This region was examined in more detail. RA activation was localized to a 160-nt *Sac* I–*Hin*fI subfragment that confers a 13-fold induction to RA (Fig. 1*B*). Transfection experiments using a series of mutations within this small region and gel retardation assay revealed two candidate enhancer motifs in this region (Fig. 1*C*): (*i*) a direct repeat of the sequence AGGTCA, which forms the core binding site for the RAR and the RXR and is observed in the antisense position (TGAACTctTGCCCT; designated DR-2A) and (*ii*) the URE, which serves as the binding site for RA-inducible cofactor RIP (*r*etinoid-*i*nducible *p*rotein; see below).

The ability of RAR and RXR to form either homo- or heterodimers on the *HOXB* DR-2A and DR-2B (RARE found in the 3′ side of *HOXB1* gene) was examined by using a gel retardation assay. Human RARα and RXRα proteins were synthesized by *in vitro* translation using rabbit reticulocyte lysate and mixed with ^{32}P-labeled response elements from the RARβ2 RARE (a DR-5) and the *HOXB* DR-2A and DR-2B followed by gel electrophoresis. Fig. 2, lanes 1–4, show high affinity binding of RAR and RXR heterodimers to the DR-5

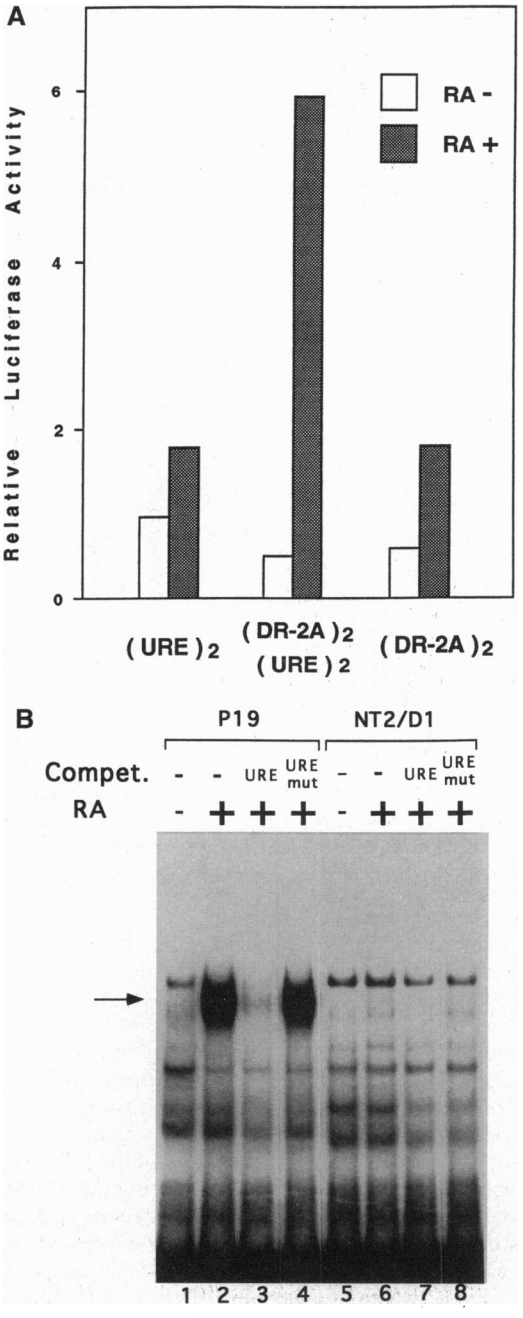

FIG. 3. (*A*) Synergistic activation by the URE and DR-2A motifs. Three luciferase constructs in which two copies of the URE and/or DR-2A sequences were inserted at the 5′ end of the basal *HOXB1* promoter (*Spe* I–*Nco* I) were transfected to P19 cells. Note that RA (1 μM, 24 hr) strongly activates the *HOXB1* promoter in the presence of both elements. (*B*) Gel retardation assay identifies an RA-inducible URE binding protein (RIP) in P19 cells. Nuclear extracts were prepared from P19 and NT2/D1 cells cultured for 2 days with (+) or without (−) 1 μM RA (lanes 2–4 and 6–8 and lanes 1 and 5, respectively). Three to 5 μg of nuclear extracts was used for each reaction. Unlabeled URE and UREmut (shown in Fig. 1*C*) oligonucleotides were used as competitors (lanes 3 and 7 and lanes 4 and 8, respectively). Specific DNA–protein complex is indicated by an arrow. Note that the URE binding protein is not detected in NT2/D1 cells even after RA treatment (lanes 5 and 6).

motif. Similar heterodimer binding was observed on the DR-2A probe (lanes 5–8) and the downstream 3′ DR-2B probe (lanes 9–12). Although the specific activities of the probes were the same, the DR-2 complex was less intense than the DR-5 band. This is in agreement with previous results

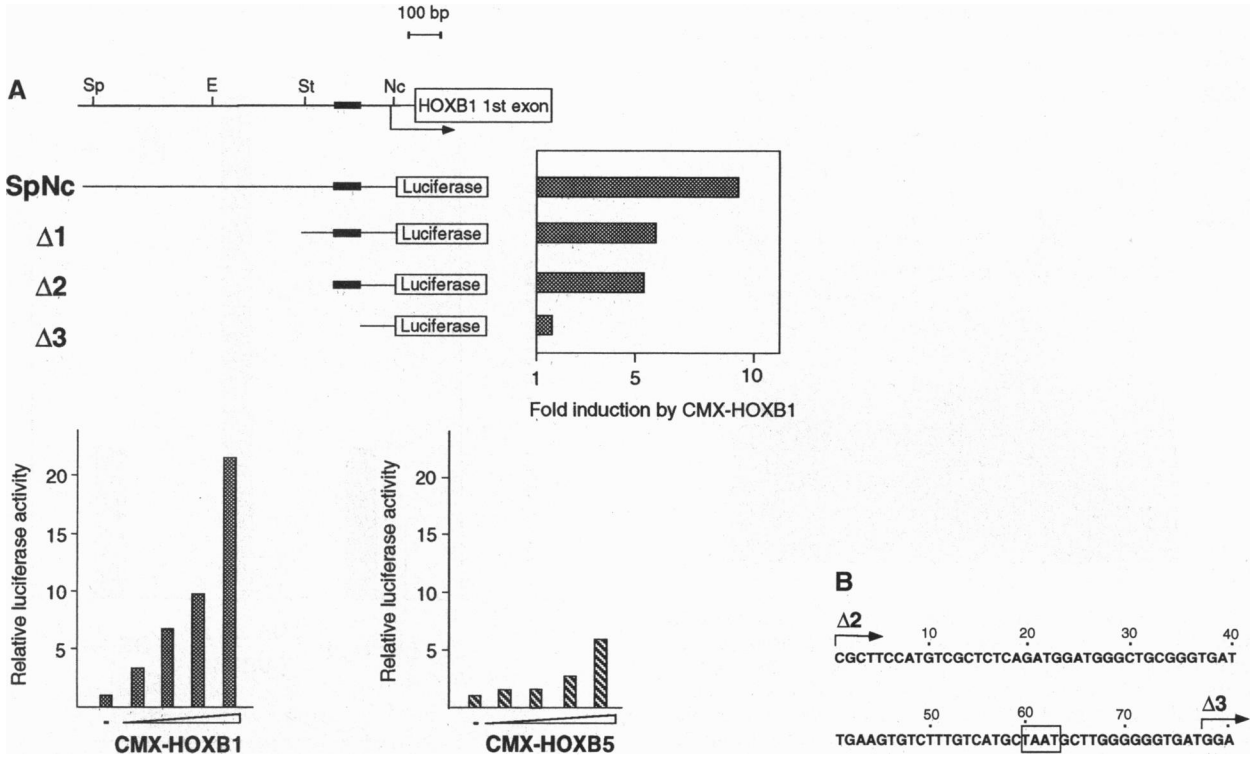

FIG. 4. (*A*) Cotransfection of the series of *HOXB1* promoter deletion plasmids along with a *HOXB1* expression vector identifies a *HOXB1*-responsive site (indicated as solid boxes in the map shown here). NT2/D1 cells were used in this experiment. The transcription initiation site is indicated by an arrow. *HOXB1* promoter plasmids were cotransfected along with increasing amounts of CMX-HOXB1 or CMX-HOXB5 expression vectors. E, *Eco*RI; Nc, *Nco* I; Sp, *Spe* I; St, *Stu* I. (*B*) Sequence of *HOXB1*-responsive region indicated by the solid boxes in *A*. The putative *HOXB1* binding site is boxed.

suggesting that the RAR·RXR heterodimers may form lower affinity complexes on the DR-2 motif. These results indicate that RAR and RXR bind cooperatively to the DR-2A and DR-2B motifs as heterodimer.

Though the URE contains no homology to known RAREs, it is apparently necessary for RA response of the *HOX* reporter. The 160-bp *Sac* I–*Hin*fI fragment confers robust RA induction on *HOXB1* promoter, whereas activation is severely repressed (to ≈25%) following mutation of the URE (Fig. 1*D*, constructs 1 and 2). Removal of DR-2A sequence also severely repressed this activation (Fig. 1*D*, constructs 1, 3, and 4). Interestingly, the URE and the RARE when tested independently as single copies show virtually no activity (data not shown). Even two tandem copies of the URE show only a low-level basal activity and are unresponsive to RA treatment (Fig. 3*A*). Similarly, two copies of the DR-2A site show only modest RA responsiveness when tested in the absence of the URE. However, when combined, full retinoid inducibility is regained. This suggests that RA activation requires the co-stimulatory activity of the URE, the RARE, and their associated binding proteins.

Gel retardation analysis was used to identify the URE binding activity in nuclear extracts from P19 cells. In the absence of RA, only background level activity for binding to the URE-containing fragment is observed (Fig. 3*B*, lane 1). However, following RA treatment there is a nearly 100-fold induction of this activity referred to as RIP (lane 2). RIP binding is specific for the URE, as shown by competition (lane 3), and is not competitively inhibited by the URE mutant sequence shown in Fig. 1*C* (lane 4). In contrast, no URE binding is seen in nuclear extracts from NT2/D1 cells either before or following RA treatment (lanes 5 and 6). This would suggest that in these cells, the isolated *HOX* promoter would be unresponsive to RA induction. As shown in Fig. 1*A*, the *HOX* reporter is indeed unresponsive to RA, even in the presence of the URE and DR-2A (Fig. 1*A*, solid bars). Interestingly, these cells are RA responsive as indicated by the control TK reporter containing the RARβ2 RARE (discussed below). Furthermore, the NT2/D1 cells are known to undergo differentiation in response to RA (13). These results suggest that the deficiency of RIP precludes the *HOX* promoter from responding to RA. This is consistent with the results of Fig. 3*A*, indicating that the DR-2A site alone is incapable of producing a sustained RA response, and supports the suggestion that RA responsiveness of the *HOXB1* promoter is based on a cooperative interaction with the URE and its associated binding protein (RIP). Thus, in this case, the RIP is functioning as an RA-dependent cofactor.

As is frequently observed, many genes encoding transcription factors like *HOXB1* also contain binding sites for those transcription factors that would generate an autoregulatory loop for sustained activation. As shown in Fig. 4*A*, cotransfection of the series of reporter vectors along with a *HOXB1* expression vector identified significant inducibility, which mapped to a discrete promoter proximal region (Fig. 4*A*, solid boxes). Cotransfection of increasing amounts of CMX-HOXB1 expression vector produced up to a 20-fold activation (Fig. 4*A*). This activation is specific to the B1 product as cotransfection of equal amounts of the *HOXB5* expression vector is capable of only low-level induction even at the highest levels of plasmid. Inspection of the 80-nt fragment identifies a candidate *HOXB1* binding site at nt 60 (Fig. 4*B*) whose properties have been confirmed by using mobility shift assays (data not shown). Accordingly, these data suggest that activation of the *HOXB1* locus by RA leads to the activation of an autoregulatory loop that might further potentiate the RA effects.

DISCUSSION

HOXB1 gene is expressed in ectoderm and mesoderm in the posterior half of the embryo by the primitive streak stage. By

the early somite stage, *HOXB1* expression has become divided into the prospective rhombomere 4 domain and the posterior half of embryo (23, 27). Teratogenic doses of RA induce the expansion of both domains of *HOXB1*. However, after the early somite stage, the rhombomere 4 expression becomes refractory to RA, whereas expression in the lateral foregut maintains its sensitivity (23, 27). These lines of evidence strongly suggest that RA is one of important determinants controlling the expression of *HOXB1* gene and that the sensitivity to RA may be regulated in a tissue-specific and time-specific manner. Preliminary data from transgenic mice harboring *LacZ* reporter gene controlled by the entire *HOXB1* region strongly suggest that the RARE found in the promoter region is essential in the establishment of rhombomere 4 expression and regulation of its sensitivity to RA. Our *in vitro* data suggest a theoretical cascade to explain the establishment of rhombomere 4 expression. First, the RAR·RXR heterodimer is activated by endogenous retinoids, which, in turn, activate the coactivator RIP. Together, these proteins induce *HOXB1* gene through the DR-2A element and the URE found in the promoter region. We have identified an autoregulatory site in the proximal region of the promoter that is strongly activated by the *HOXB1*-encoded protein. These data suggest that after the initial activation, expression is maintained by this autoregulatory loop, at which point the dependence and sensitivity of the *HOXB1* gene to RA will be reduced or eliminated.

This work provides the evidence for a complex but nonetheless central role of retinoid receptors in the activation and regulation of the *HOXB1* gene. A key feature is the dependence on the highly inducible RIP gene, which is an actual component in the cascade. Accordingly, the purification and identification of RIP is anticipated to be essential in understanding the earliest events in *HOX* gene control.

We thank Drs. Henry Sucov, Bruce Blumberg, and Kazuhiko Umesono for critical reading of the manuscript. We also thank Dr. Kenneth K. Kidd for providing the opportunity to clone the entire *HOXB* locus. R.M.E. is an Investigator of the Howard Hughes Medical Institute at The Salk Institute for Biological Studies. T.O. was a Research Associate of the Howard Hughes Medical Institute. This work was supported by the Howard Hughes Medical Institute and National Institutes of Health Grant 5RO1HD27183.

1. Akam, M. (1989) *Cell* **57,** 347–349.
2. Acampora, D., D'Espositi, M., Faiella, A., Pannese, M., Migliaccio, E., Morelli, F., Strornaiuolo, A., Nigro, V., Simeone, A. & Boncinelli, E. (1989) *Nucleic Acids Res.* **17,** 10385–10402.
3. Wilkinson, D. G., Bhatt, S., Cook, M., Boncinelli, E. & Krumlauf, R. (1989) *Nature (London)* **341,** 405–409.
4. Hunt, P., Gulisano, M., Cook, M., Sham, M.-H., Faiella, A., Wilkinson, D., Boncinelli, E. & Krumlauf, R. (1991) *Nature (London)* **353,** 861–864.
5. Graham, A., Maden, M. & Krumlauf, R. (1991) *Development (Cambridge, U.K.)* **112,** 255–264.
6. Dressler, G. R. & Gruss, P. (1989) *Differentiation* **41,** 193–201.
7. Graham, A., Papalopulu, N. & Krumlauf, R. (1989) *Cell* **57,** 367–378.
8. Eichele, G. (1989) *Trends Genet.* **5,** 246–251.
9. Morriss-Kay, G. M., Murphy, P., Hill, R. E. & Davidson, D. R. (1991) *EMBO J.* **10,** 2985–2995.
10. Izpisua-Belmonte, J.-C., Tickle, C., Dolle, P., Wolpert, L. & Duboule, D. (1991) *Nature (London)* **350,** 585–589.
11. Yokouchi, Y., Sasaki, H. & Kuroiwa, A. (1991) *Nature (London)* **353,** 443–445.
12. Nohno, T., Noji, S., Koyama, E., Ohyama, K., Myokai, F., Kuroiwa, A., Saito, T. & Taniguchi, S. (1991) *Cell* **6,** 1197–1205.
13. Anderws, P. W., Damjanov, I., Simon, D., Banting, G., Carlin, C., Dracopoli, N. & Fogh, J. (1984) *Lab. Invest.* **50,** 147–162.
14. Simeone, A., Acampora, D., Arcioni, L., Andrews, P. W., Boncinelli, E. & Mavilio, F. (1990) *Nature (London)* **346,** 763–767.
15. Simeone, A., Acampora, D., Nigro, V., Faiella, A., D'Esposito, M., Stornauiolo, A., Mavilio, F. & Boncinelli, E. (1991) *Mech. Dev.* **33,** 215–228.
16. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. (1987) *Nature (London)* **330,** 444–450.
17. Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. (1987) *Nature (London)* **330,** 624–629.
18. Brand, N., Petkovich, M., Krust, A., Chambon, P., de The, H., Marchio, A., Tiollais, P. & Dejean, A. (1988) *Nature (London)* **332,** 850–853.
19. Benbrook, D., Lernhardt, E. & Pfahl, M. (1988) *Nature (London)* **333,** 669–672.
20. Mangelsdorf, D. J., Ong, E. S., Dyck, J. A. & Evans, R. M. (1990) *Nature (London)* **345,** 224–229.
21. Krust, A., Kastner, Ph., Petkovich, M., Zelent, A. & Chambon, P. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5310–5314.
22. Mangelsdorf, D. J., Borgmeyer, U., Heyman, R. A., Yang Zhou, J., Ong, E. S., Oro, A. E., Kakizuka, A. & Evans, R. M. (1992) *Genes Dev.* **6,** 329–344.
23. Murphy, P. & Hill, R. E. (1991) *Development (Cambridge, U.K.)* **111,** 61–74.
24. Langston, A. W. & Gudas, L. J. (1992) *Mech. Dev.* **38,** 217–228.
25. Ogura, T. & Evans, R. M. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 392–396.
26. Digman, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
27. Conlon, R. A. & Rossant, J. (1992) *Development (Cambridge, U.K.)* **116,** 357–368.