MOLECULAR AND CELLULAR BIOLOGY, Dec. 1994, p. 8143–8154
0270-7306/94/$04.00+0

Vol. 14, No. 12

Multiple Positive and Negative Regulatory Elements in the Promoter of the Mouse Homeobox Gene *Hoxb-4*

ALEJANDRO GUTMAN, JONATHAN GILTHORPE, AND PETER W. J. RIGBY*

Laboratory of Eukaryotic Molecular Genetics, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, England

Received 2 June 1994/Returned for modification 6 July 1994/Accepted 25 August 1994

Mouse *Hoxb-4* (*Hox-2.6*) is a homeobox gene that belongs to a family which also includes *Hoxa-4*, *Hoxc-4*, and *Hoxd-4* and that is related to the *Deformed* gene in *Drosophila melanogaster*. We have determined the sequence of 1.2 kb of 5′ flanking DNA of mouse *Hoxb-4* and by nuclease S1 and primer extension experiments identified two transcription start sites, P1 and P2, 285 and 207 nucleotides upstream of the ATG initiator codon, respectively. We have shown that this region harbors two independent promoters which drive CAT expression in several different cell lines with various efficiencies, suggesting that they are subject to cell-type-specific regulation. Through detailed mutational analysis, we have identified several *cis*-regulatory elements, located upstream and downstream of the transcription start sites. They include two cell-type-specific negative regulatory elements, which are more active in F9 embryonal carcinoma cells than in neuroblastoma cells (regions a and d at −226 to −186 and +169 to +205, respectively). An additional negative regulatory element has been delimited (region b between +22 and +113). Positive regulation is achieved by binding of HoxTF, a previously unknown factor, to the sequence GCCATTGG (+148 to +155) that is essential for efficient *Hoxb-4* expression. We have also defined the minimal promoter sequences and found that they include two 12-bp initiator elements centered around each transcription start site. The complex architecture of the *Hoxb-4* promoter provides the framework for fine-tuned transcriptional regulation during embryonic development.

Homeobox genes encode a large family of transcription factors and are found in a wide range of species, both invertebrate and vertebrate (reviewed in references 12, 17, and 23). In *Drosophila melanogaster*, the genes of the *Antennapedia/Ultrabithorax* complexes, which are clustered together, specify the phenotypes of individual segments along the anteroposterior axis of the body. In mammals, there are four clusters of *Antennapedia*-class homeobox genes, called *Hox* genes, each on a different chromosome, which are presumed to have arisen by duplication and expansion of an ancestral complex. The genes in each of the four clusters that derive from the same ancestral gene are called paralogs. One of the most remarkable features of both the invertebrate and the vertebrate *Hox* clusters is that there is a direct relationship between the position of a gene in the cluster and its expression domain. The genes at the 5′ end, in the transcriptional sense, have the most posterior domains, and the genes at the 3′ end have the most anterior.

The role of *Hox* genes in mammalian development has been demonstrated by both ectopic expression (3, 16, 22, 39) and by the introduction of loss-of-function mutations (5, 6, 18, 21, 27). These data make it clear that the proteins encoded by these genes are involved, as in *D. melanogaster*, in the specification of structures along the anteroposterior axis. Thus, if a gene is ectopically expressed in regions more anterior than its normal expression domain, transformations to more posterior structures are observed, while if a gene is inactivated, it can lead to anterior transformations. It is therefore clear that the exact setting of the boundaries of expression of the *Hox* genes along the anteroposterior axis is crucial to orderly development.

We, and others, have therefore sought to identify the DNA sequences that control transcription of *Hox* genes during development. Experiments using transgenic animals have shown that it is possible to recapitulate the normal expression pattern. In two cases, *Hoxa-7* and *Hoxb-4*, multiple regulatory elements located downstream of the transcription start site are required (26, 38). However, these elements operate in opposite fashions. In the case of *Hoxa-7*, the promoter, the 5′ flanking DNA, is active in all regions of the embryo, and this activity is restrained by the regulatory elements (26). In the case of *Hoxb-4*, the 5′ flanking DNA is essentially inactive, and proper expression is seen only in the presence of the regulatory elements. We have, moreover, shown that the regulatory elements of *Hoxb-4* act as spatially specific enhancers in that they are capable of imposing correct boundaries of expression along the anteroposterior axis on heterologous promoters (38).

Although the 5′ flanking DNA of *Hoxb-4* is inactive in the mid-gestation embryo in the absence of the enhancers, it is clear that the promoter must contain elements involved in responding to them and in achieving proper levels of transcription. In order to characterize such elements, we have investigated the transcriptional regulation of the gene in cultured cells in which the 5′ flanking DNA is active. These experiments have enabled us to define the architecture of the promoter and to identify, in addition to several negative regulatory elements, a binding site for a previously unknown activator which is essential for efficient transcription of the gene.

MATERIALS AND METHODS

DNA sequencing. The nucleotide sequence of both strands of the DNA was determined, at least twice, by the dideoxynucleotide method (29), using ^{35}S-dATP.

Recombinants. The −926/+285 reporter recombinant was constructed by inserting a *Pst*I-*Ase*I fragment (the *Ase*I site was blunted with Klenow) of the mouse *Hoxb-4* promoter into the *Pst*I-*Bgl*II restriction sites (the latter made blunt by filling in with Klenow) of the pBLCAT3 vector (20). The +285/−926 construct was made by cloning this same fragment into the

* Corresponding author. Phone: 44 81 913 8533. Fax: 44 81 913 8527.

blunt (*Hin*dIII)-*Pst*I sites of pBLCAT3. Recombinants −312/+285, −226/+285, −186/+285, and +21/+285 were constructed from −926/+285 by deleting a fragment delimited by a *Hin*dIII site in the vector polylinker and *Esp*I, *Dra*III, *Apa*I, or *Sac*II restriction sites in the *Hoxb-4* promoter, respectively (see Fig. 1). −280/+285 was generated by digesting −926/+285 with *Pst*I and partially with *Nco*I and subsequent treatment with exonuclease III, S1 nuclease, T4 polynucleotide kinase, and T4 DNA ligase (the precise boundaries of the clone were determined by sequencing). An oligonucleotide was used to replace the *Hin*dIII-*Sac*II fragment of −926/+285 by the −32 to +21 sequences, giving −32/+285. +21/+285 mut was constructed by inserting a *Nru*I linker into the single *Pml*I site of +21/+285. Another oligonucleotide was inserted into the *Msc*I (partial digestion)-*Bgl*II sites of −926/+285, generating −926/+205. −926/+169, −926/+149, and −926/+113 were created by cutting the −926/+285 reporter with *Bgl*II and *Pml*I, *Msc*I or *Bsp*MI, respectively, followed by filling in and religation. −186/+205 and +21/+205 were derived from −926/+205 by deletion of a *Hin*dIII-*Apa*I or a *Hin*dIII-*Sac*II fragment, respectively. A similar strategy was used for creating −186/+169 and +21/+169 from −926/+169, −186/+149 and +21/+149 from −926/+149, and −186/+113 and +21/+113 from −926/+113. Oligonucleotides with *Hin*dIII-*Bgl*II ends were used to construct −32/+21, +46/+86, −5/+7, and +75/+86. Mutations in the Sp1-like binding sites of −32/+21 and +46/+86 were introduced by cloning the following oligonucleotides in pBLCAT3 (mutated bases are underlined; compare with the wild-type sequence in Fig. 1):

5′-AGCTTGTTCATCATTTTTCTCTGTTGTGGAGAAAGCCAGGCGAGGAAAAGTCTCCCCA-3′
5′-AGCTTGTTGTTTTTGTGGTTATTGGTAGAGAAGGGGAAATAAACCTA-3′

HSPCAT was constructed by inserting the 0.3-kb *Sal*I-*Bam*HI fragment of the heat shock gene *hsp68* promoter into the same sites of pBLCAT3. pBLCAT2 and pBLCAT3 vectors have been previously described (20). A double-stranded oligonucleotide containing the +135 to +169 sequences of *Hoxb-4* (Fig. 1), with added *Hin*dIII ends, was cloned in both orientations into the *Hin*dIII sites of the following recombinants: B-CAT (19), pBLCAT2 (20), −5/+7 (see Fig. 3), and −32/+21 (see Fig. 3), producing ACTB-CATd and ACTB-CATr, ACTPBLCAT2d and ACTPBLCAT2r, ACT-INd and ACT-INr, and LACT-MPd and LACT-MPr, respectively. Another double-stranded oligonucleotide containing the +143 to +169 sequences of *Hoxb-4* (Fig. 1) was cloned into the *Hin*dIII site of −32/+21, producing in the direct and reverse orientations ACT-WT and ACT-WTr, respectively. HoxTFmut, YY1mut, and Double mut were obtained by cloning the M8, M5, and M3 double-stranded oligonucleotides (see Fig. 7) in the *Hin*dIII site of −32/+21 (previously filled in with Klenow).

Transfections. Subconfluent 9-cm-diameter plates were transfected by the calcium phosphate method (14) with 10 to 15 μg of a reporter recombinant, 5 μg of pBluescript KS+, and 4 μg of the internal control plasmid RSV-βgal (13) or 0.25 to 1 μg of pBCTBX2 (15) in S1 nuclease experiments. Cells were incubated with the calcium phosphate precipitate in culture medium containing 10% (vol/vol) fetal calf serum (FCS) for 20 h and then washed and grown for an additional 28 h in complete medium. Each experimental point was determined in duplicate or triplicate in at least three independent transfections. We have calculated the standard deviations on all of our experiments. These are generally between 10 and 20% of the stated value and never exceed 25%. β-Galactosidase and chloramphenicol acetyltransferase (CAT) activities were measured as described previously (28, 34). Quantitation of CAT activity was performed with a phosphoimager or by scintillation counting. F9 embryonal carcinoma (EC) stem cells were differentiated by incubation for 3 days in medium supplemented with 20% (vol/vol) fetal calf serum, 1 mM cyclic AMP (cAMP), 0.1 mM isobutyl methylxanthine and 5×10^{-8} M retinoic acid. PC12 cells were first grown in suspension and later attached to collagen-coated plates for 2 weeks, when they were differentiated with 50 ng of mouse 2.5S nerve growth factor (Sigma) per ml for 10 days.

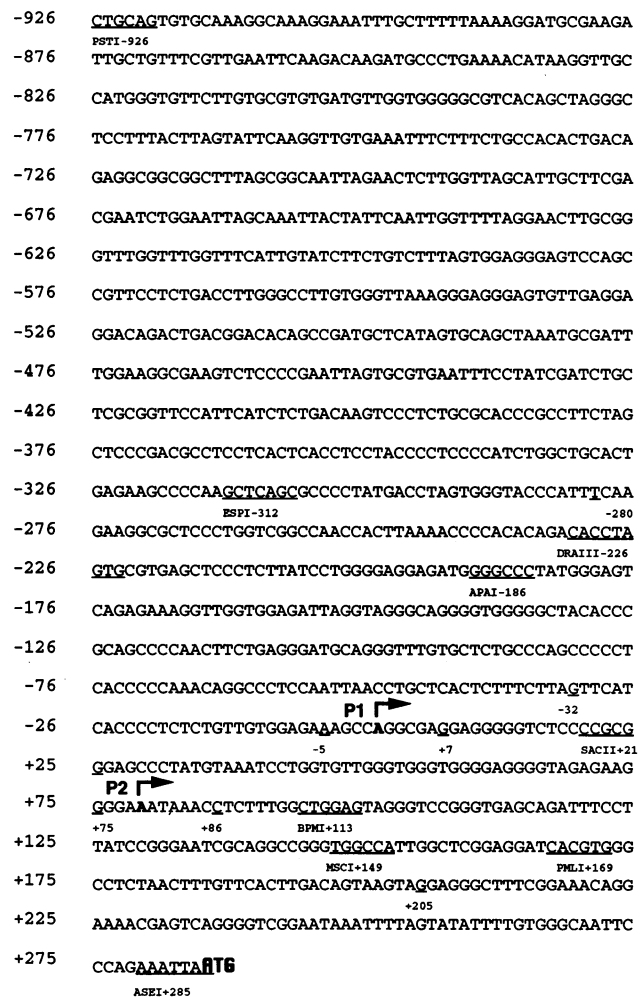

FIG. 1. Nucleotide sequence of the mouse *Hoxb-4* gene promoter. Transcription start sites P1 and P2 are indicated by arrows. Numbering is from P1 (+1). The boundaries of different deletion recombinants and several restriction enzyme sites employed in their construction are underlined. The ATG initiator codon is shown in large, boldface type.

S1 nuclease analysis. A *Sac*I-*Eae*I DNA restriction fragment encompassing nucleotides −211 to +151 (Fig. 1) of mouse *Hoxb-4* was dephosphorylated with 0.1 U of calf intestinal phosphatase and then heated at 85°C for 30 min to inactivate the enzyme and labelled with [γ-^{32}P]ATP. The labelled fragment was purified by a minicolumn (Magic Clean-up System, Promega Inc.) and denatured in 50% (vol/vol) dimethyl sulfoxide–0.5× TBE (0.045 M Tris-borate, 0.001 M EDTA [pH 7.5]), and both strands were separated in a 6% (wt/vol) polyacrylamide gel (60:1, acrylamide-bisacrylamide), in 0.5× TBE and run at 8 V/cm. An excess of probe (10^4 cpm; 10 fmol) was hybridized overnight at 52°C with 150 μg of total RNA in 30 μl of a solution of 0.04 M piperazine-*N*,*N*′-bis-2-ethanesul-

fonic acid (PIPES; pH 7), 0.4 M NaCl, and 1 mM EDTA. Digestion with S1 nuclease was performed at 37°C, as described by Sambrook et al. (28). The reaction products were resolved in a 7 M urea–6% (wt/vol) polyacrylamide gel. To quantitate CAT mRNA in transfected cells (see Fig. 6B), an *Eco*RI-*Hin*dIII restriction fragment from pBLCAT3 was employed under the same conditions described above. In addition, an 80-mer oligonucleotide labelled with [γ-^{32}P]ATP was used as a probe to detect mRNA from the internal control plasmid pBCTBX2 (15).

Primer extension. An oligonucleotide primer (+115/+179 [Fig. 1]) was labelled with [γ-^{32}P]ATP (20,000 cpm/μl; 3×10^6 cpm/pmol) and incubated with 100 μg of total RNA in 30 μl of a solution of 0.04 M PIPES (pH 7), 0.4 M NaCl, 1 mM EDTA, and 80% (vol/vol) formamide for 10 min at 85°C. This mixture was then incubated overnight at 30°C and treated with reverse transcriptase, as previously described (28). The reaction products were resolved in a 7 M urea–6% (wt/vol) polyacrylamide gel.

DNA mobility shift assays. Oligonucleotides were labelled with [γ-^{32}P]ATP, annealed, purified by native polyacrylamide gel electrophoresis, and used as probes in gel retardation experiments (1,000 cpm per fmol, and 10 fmol per reaction mixture). Each reaction mixture contained, in a 20-μl volume, 10 to 20 μg of whole cell extract, 1 μg of calf thymus DNA, specific competitor as indicated, 50 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 0.2 mM EDTA, 1 mM dithiothreitol, and 15% (vol/vol) glycerol and was incubated at 0°C for 15 min. The probe was then added, and incubation proceeded for 10 additional min at 30°C. Afterwards, the samples were loaded onto a 5% (wt/vol) polyacrylamide gel, in 0.25× TBE, and electrophoresed at 150 V for 2 h. The gel was then dried and autoradiographed.

Nucleotide sequence accession number. The DNA sequence presented here has been deposited in the EMBL database under accession number X71912.

RESULTS

Sequence of 1.2 kb of *Hoxb-4* 5′ flanking DNA and identification of the transcription start sites. Our transgenic experiments showed that the recapitulation of the expression pattern of *Hoxb-4* in the embryo requires only 1.2 kb of DNA 5′ to the translation initiation codon (38). We therefore began by determining the sequence of this DNA (Fig. 1). We then identified the transcription start sites by performing nuclease S1 protection experiments using RNA from F9 cells differentiated to parietal endoderm-like cells (F9-PE) with cAMP and retinoic acid. We observed two major bands, of 151 and 73 nucleotides, which correspond to initiation 285 and 207 nucleotides upstream of the ATG, respectively (Fig. 2A). Similar results were obtained with RNA from F9 EC stem cells (data not shown). Analysis by primer extension experiments confirmed the nuclease protection data. We obtained two bands (Fig. 2B), the sizes of which, 179 and 101 nucleotides, again correspond to initiation 285 and 207 bases, respectively, upstream of the ATG. We therefore confirmed that there are two major transcription start sites in the *Hoxb-4* promoter. The more distal of these is designated P1, and it is taken to define +1 in the sequence numbering (and as a consequence, the first nucleotide of the region under study is −926); the other promoter, P2, is thus located at +79 (Fig. 1). We also observed, in both the S1 and primer extension experiments, two weaker bands which map downstream of P2 and probably correspond to minor start sites. Additional primer extension experiments, using primers located close to the ATG, provided

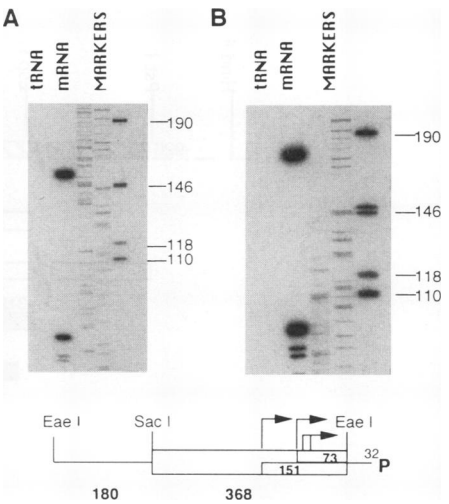

FIG. 2. Identification of the transcription start site. (A) Nuclease S1 experiments. An *Eae*I restriction fragment containing the mouse *Hoxb-4* promoter (−217 to +151) and plasmid sequences was end labelled with [γ-^{32}P]ATP, and the antisense strand was isolated (see Materials and Methods) and used as a probe in S1 nuclease mapping experiments. A total of 150 μg of tRNA or total RNA from F9 cells differentiated with cAMP and retinoic acid was hybridized to this probe, and the products were treated with nuclease S1, fractionated in a 7 M urea–6% (wt/vol) polyacrylamide gel, and autoradiographed, as described in Materials and Methods. pBluescript KS+/*Hpa*II radiolabelled fragments and products of sequencing reactions were used as markers. A diagram of the probe is shown: box, *Hoxb-4* sequences; plain line, plasmid sequences. Major transcription start sites P1 and P2 are indicated by arrows, and minor transcription start sites are shown by shorter arrows. Sizes of probe and protected fragments as well as restriction enzyme sites employed are shown. (B) Primer extension experiments. A 65-mer oligonucleotide primer (+114/+179) was labelled with [γ-^{32}P]ATP, hybridized to 100 μg of total RNA from F9 cells differentiated with cAMP and retinoic acid, and extended with reverse transcriptase, as described in Materials and Methods. The reaction products were fractionated in a 7 M urea–6% (wt/vol) polyacrylamide gel and autoradiographed. pBluescript KS+/*Hpa*II radiolabelled fragments and products of sequencing reactions were used as markers.

no evidence for other start sites. The sequences surrounding P1 and P2 contain no obvious TATA box or any GC-rich sequences. Database searches revealed that the 5′ untranslated region had extensive homology with the chicken *Hoxb-4* gene and with the mouse paralog *Hoxd-4* (see Fig. 10A). In contrast, homology with the paralog *Hoxc-4* was much more limited, and there was none to the paralog *Hoxa-4*.

***Hoxb-4* 5′ flanking DNA contains two independent promoters.** In order to identify the sequences required for *Hoxb-4* transcription, we cloned the DNA from −926 to +285 upstream of a CAT reporter gene (Fig. 3) and transfected this construct into various cultured cell lines. CAT activity was determined and normalized with respect to that of pBLCAT2, a minimal thymidine kinase (TK) promoter construct, which was taken as 100% in each cell type. As an additional reference point the activity of a heat shock promoter construct (HSP-CAT) was measured. The −926/+285 construct was expressed in all cell lines tested (Fig. 4A), showing that this region has promoter activity. When cloned in the reverse orientation (+285/−926), this DNA was inactive (data not shown). The level of expression of the −926/+285 construct varied markedly in the different cell lines tested. Figure 4A shows that expression was low, 5% of TK, in HeLa cells; intermediate, 12

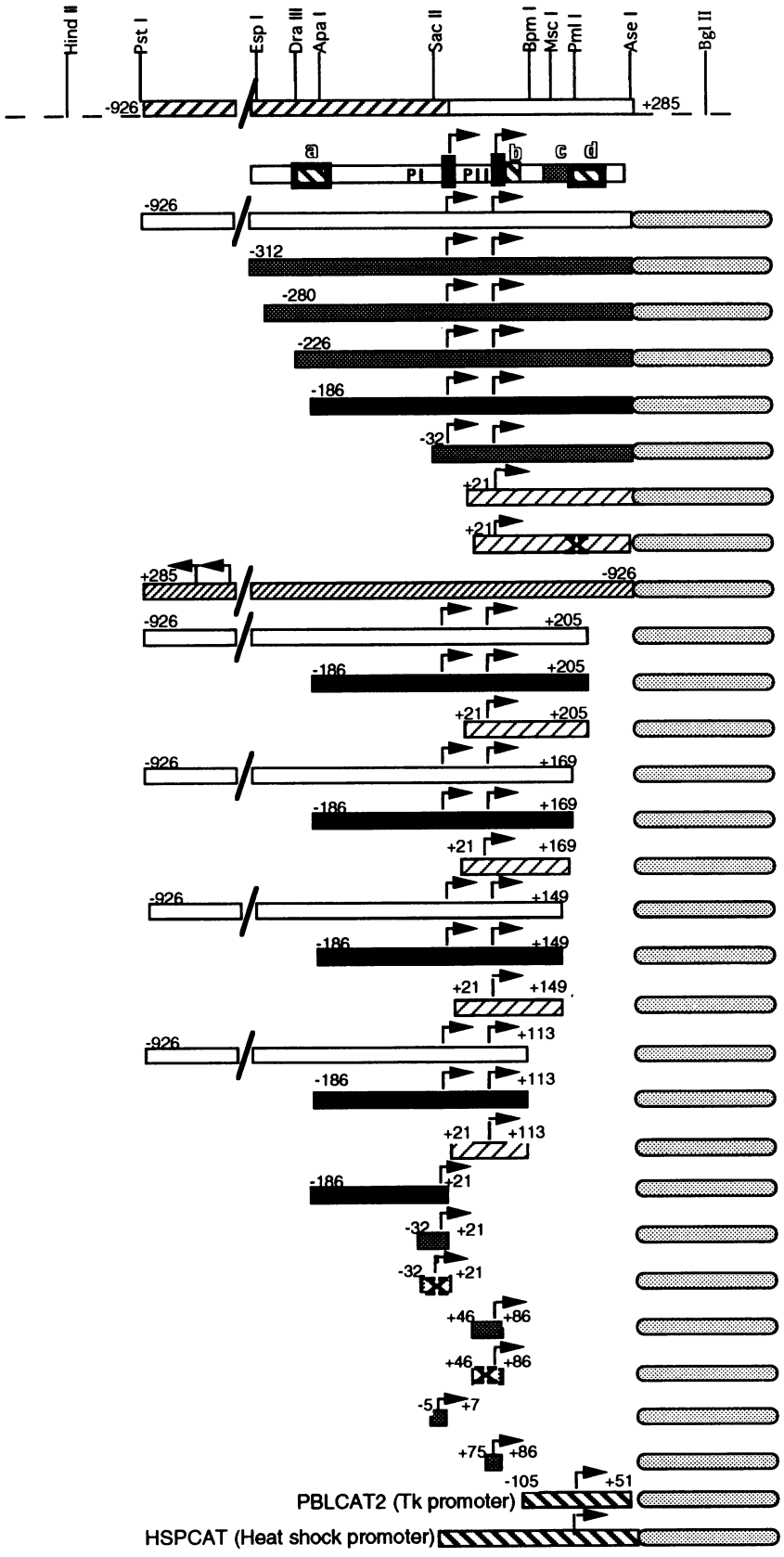
Hind III
Pst I
Esp I
Dra III
Apa I
Sac II
Bpm I
Msc I
Pml I
Ase I
Bgl II
-926
+285
a
b
c
d
PI
PII
-926
-312
-280
-226
-186
-32
+21
+21
+285
-926
-926
+205
-186
+205
+21
+205
-926
+169
-186
+169
+21
+169
-926
+149
-186
+149
+21
+149
-926
+113
-186
+113
+21
+113
-186
+21
-32
+21
-32
+21
+46
+86
+46
+86
-5
+7
+75
+86
-105
+51
PBLCAT2 (Tk promoter)
HSPCAT (Heat shock promoter)

FIG. 3. Reporter recombinants. In the first line, a schematic diagram of the mouse *Hoxb-4* gene promoter and 5′ noncoding region is shown (dashed lines, plasmid sequences; striped box, promoter sequences; white box, 5′ noncoding region). Numbering is from the first transcription start site (P1). Several restriction enzyme sites employed in the construction of reporter recombinants are also shown. In the second line, a model of the *Hoxb-4* promoter and its regulatory elements a, b, c, and d is included. Striped boxes, negative regulatory elements; grey box, positive regulatory element; thick boxes, cell-type-specific elements; arrows, P1 and P2 transcription start sites; PI and PII, promoter I and promoter II, respectively; thin black boxes, transcriptional initiators. Below the second line, the schematic structures of the reporter recombinants are depicted. Arrows, transcription start sites; white boxes, −926 series of recombinants; black boxes, −186 series of recombinants; boxes with thin stripes, +21 series of recombinants; dark grey boxes, other *Hoxb-4* recombinants; boxes with broad stripes, non-*Hoxb-4* recombinants; light grey boxes, CAT gene; X, point mutations.

to 15% of TK, in fibroblasts (NIH 3T3 and 10T1/2), F9 EC cells, and differentiated PC12 (PC12D) cells; and high, 30 to 45% of TK, in F9-PE cells and the neuroblastoma cell lines SK-N-SH (SK) and Neuro2A (NEU). It is thus apparent that the *Hoxb-4* promoter is subject to cell-type-specific regulation.

We next asked whether the two start sites, P1 and P2, reflect the activities of two independent promoters. To do this, we generated constructs containing (i) both P1 and P2 but lacking a negative regulatory element (−186/+285, see below), (ii) only P1 (−186/+21), and (iii) only P2 (+21/+285). As expected, the constructs −926/+285 and −186/+285, which carry P1 and P2, were expressed in F9 EC and Neuro2A cells (Fig. 4B). Interestingly, the constructs carrying only P1 (−186/+21) or only P2 (+21/+285) were also efficiently expressed in both cell lines, showing that the promoter region of *Hoxb-4* contains two separable promoters that can act independently.

A cell-type-specific, negative regulatory element is present between −226 and −186. Having defined the basic architecture of the *Hoxb-4* promoter, we searched for regulatory elements, initially by constructing a series of 5′ deletions based on the −926/+285 construct (Fig. 3). Figure 5A shows that sequences between −926 and −226 had little, if any, effect on expression. However, deletion of the sequences between −226 and −186 led to a sixfold increase in expression in F9 EC cells, indicating the presence of a negative regulatory element (named element a). The effect of deleting this element is less marked in F9-PE (3-fold), SK (2.6-fold), and Neuro2A (1.8-fold) cells, showing that it is regulated in a cell-type-specific fashion. Further deletions, up to +21, had little effect.

The 5′ untranslated region contains an element essential for promoter activity and a second, cell-type-specific, negative regulatory element. In order to explore the role of the 5′ untranslated region in the regulation of the *Hoxb-4* promoter, we constructed three series of recombinants carrying 3′ deletions in this region. The −926 series contains all of the 5′ flanking DNA required for correct expression in the embryo; the −186 series lacks the negative regulatory element a defined above but includes both promoters, PI and PII; and the +21 series lacks both element a and PI but contains PII (Fig. 3). Analysis of the −926 series in F9 EC cells showed that sequences between +205 and +285 did not significantly affect expression (Fig. 5B). However, deletion of sequences between +169 and +205 led to a fivefold increase in CAT activity, thus defining a second regulatory element, element d. Deletion of the 20-bp fragment between +149 and +169 had a marked effect, reducing expression 10-fold, to very low levels, and thus defining an element, element c, that is essential for efficient promoter activity.

The −186 series of constructs exhibited essentially the same pattern of expression as the −926 series, although in each case the level of expression was higher because of the absence of the negative regulatory element a. These data show that elements c and d do not depend on element a for their function. The overall pattern of expression of the +21 series was again

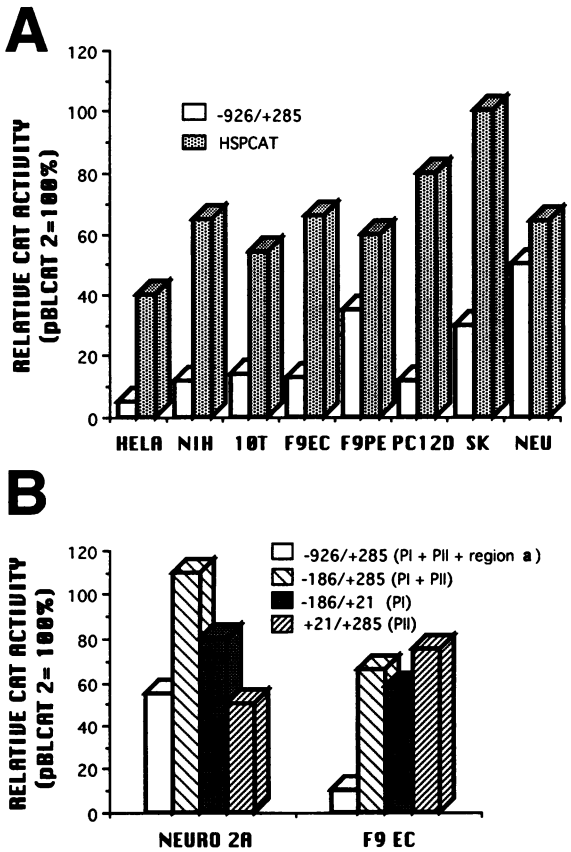

FIG. 4. (A) Activity of the *Hoxb-4* promoter in several cultured cell lines. A total of 12 μg of each of the reporter recombinants −926/+285, +285/−926, pBLCAT2 (TK promoter), HSPCAT (heat shock promoter), and CAT vector pBLCAT3 was cotransfected with 4 μg of the internal control RSV-βgal by the calcium-phosphate method, into the following cell lines: HeLa, NIH 3T3 (NIH), 10T1/2 (10T), F9 EC, F9 cells differentiated with retinoic acid and cAMP (F9-PE), PC12 cells differentiated with nerve growth factor (PC12D), SK-N-SH (SK), and Neuro2A (NEU), as described in Materials and Methods. Cell extracts were made, and CAT activity was measured. CAT activity of −926/+285, HSPCAT, and pBLCAT2 was corrected by subtraction of background levels of either the antisense promoter construct +285/−926 or the promoterless pBLCAT3. Relative CAT activity was normalized with respect to that of pBLCAT2, taken as 100%. (B) Two independent promoters drive CAT expression in F9 EC and Neuro2A neuroblastoma cell lines. A total of 12 μg of each reporter recombinant was cotransfected with 4 μg of the internal control RSV-βgal, by the calcium-phosphate method, into F9 EC and Neuro2A neuroblastoma cell lines. Cell extracts were made, and CAT activity was measured and corrected by subtraction of background levels of the promoterless vector pBLCAT3. Relative CAT activity was normalized with respect to that of pBLCAT2 (a minimal TK promoter construct), taken as 100%.

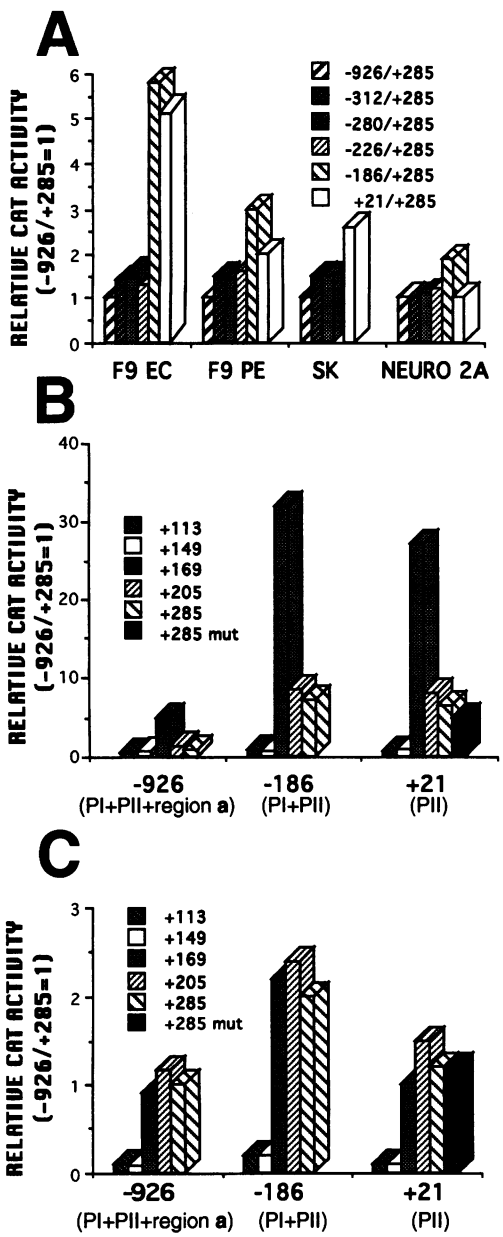

FIG. 5. (A) A cell-type-specific, negative *cis*-regulatory element is present between −226 and −186. A total of 12 μg of each reporter recombinant was cotransfected with 4 μg of the internal control plasmid RSV-βgal, by the calcium-phosphate method, into F9 EC cells, F9 cells differentiated with retinoic acid and cAMP (F9-PE), and the neuroblastoma cell lines SK-N-SH (SK) and Neuro2A. Cell extracts were prepared, and CAT activity was measured and corrected by subtraction of background levels of the promoterless vector pBLCAT3. Relative CAT activity was normalized with respect to that of −926/+285, taken as 1. (B and C) Identification of negative and positive *cis*-regulatory elements in the 5′ noncoding region of *Hoxb-4*. Three series of recombinants, −926, −186, and +21 (depicted in Fig. 3 and described in the text), were transfected in F9 EC cells (B) and the neuroblastoma cell line Neuro2A (C). A total of 12 μg of each reporter recombinant was cotransfected with 4 μg of the internal control plasmid RSV-βgal, by the calcium-phosphate method. Cell extracts were prepared, and CAT activity was measured and corrected by subtraction of background levels of the promoterless vector pBLCAT3. Relative CAT activity was normalized with respect to that of −926/+285, taken as 1.

similar, showing that elements c and d can regulate the PII promoter in the absence of the PI promoter. We noticed the presence of an E box (CACGTG), a consensus binding site for the proteins of the basic helix-loop-helix family of transcription factors, at nucleotides +166 to +172. However, mutation of this binding site had no effect on expression (compare +21/+285 mut with +21/+285 in Fig. 5B).

When the same series of recombinants were analyzed by using Neuro2A cells (Fig. 5C), we found that deletion of region d had no effect, indicating that its activity, like that of the other negatively acting element, element a, is regulated in a cell-type-specific fashion. However, the positively acting element c was, as in F9 EC cells, essential for efficient expression.

Both PI and PII contain transcriptional initiators, the activity of which is modified by neighboring sequences. In order to identify the minimal sequences required for transcriptional activity, additional mutants were derived from the virtually inactive reporters −186/+149 and +21/+149, and their activities were assayed by using Neuro2A cells (Fig. 6A). Removal of sequences between +113 and +149 had no effect, but deletion of the region between +21 and +113 restored expression, and consequently, a third negative element, element b, is delimited. As deletion of sequences from +21 to +46 and those from +86 to +113 rendered PII active again (see construct +46/+86 [Fig. 6A]), element b must be located in one of these two regions. Smaller constructs encompassing only the PI promoter showed lower levels of expression (see −32/+21 and −5/+7), but even a 12-bp element (−5/+7) was active, suggesting that it acts as a transcriptional initiator (35, 36). Similarly, the +75/+86 construct encompassing 12 bp of PII gave a lower level of expression, suggesting that PII also includes an initiator. The region surrounding these two initiators contains potential binding sites for the transcription factor Sp1. However, mutation of these sites had no effect on expression (see −32/+21 mut and +46/+86 mut).

Elements b, c, and d in the *Hoxb-4* 5′ untranslated region regulate mRNA levels. The location of elements b, c, and d in the 5′ untranslated region raises the possibility that they regulate expression at the level of translation rather than transcription. We therefore performed quantitative nuclease S1 experiments using RNA extracted from Neuro2A cells (for elements b and c) or from F9 EC cells (for element d) that had been transfected with the appropriate reporter (see Fig. 8B). The promoterless vector pBLCAT3 gave rise to no detectable mRNA, while the data with the other constructs paralleled those obtained from the CAT assays. In F9 EC cells, deletion of region d increased the amount of transcript sixfold. In Neuro2A cells, deletion of region c reduced the level of transcript 15-fold, while deletion of region b increased it 6-fold. These data show that the three elements in the 5′ untranslated region act by regulating mRNA levels.

Regulatory element c activates transcription of heterologous promoters. To further characterize positive regulatory element c, we cloned it into CAT-based reporters, upstream of four different promoters, in both orientations. Two of these promoters are heterologous. One is essentially a TATA box (19) which drives a very low level of transcription (the TATA box series [Fig. 7A]); the other is the TK minimal promoter (20) with a much higher level of basal transcription (the TK promoter series). The other two promoters belong to the *Hoxb-4* gene and are characterized in this report. One is initiator 1, which contains the −5 to +7 sequences of the gene (Fig. 1 and 6A) and which drives a very low level of transcription, comparable to that of the TATA box-based recombinants (the *Hoxb-4* initiator 1 series). The other is the minimal promoter that comprises the −32 to +21 sequences (Fig. 1 and

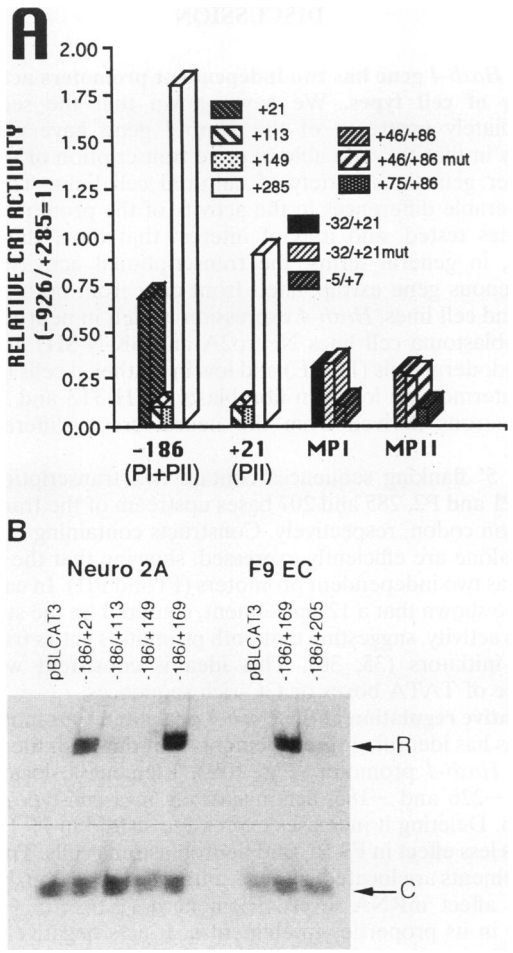

FIG. 6. (A) Delimitation of minimal promoter sequences. Four different kinds of recombinants were used: the −186, the +21, and the minimal promoter I and II series (MPI and MPII, respectively), which were transfected into Neuro2A cells. Transfection and CAT assays were as described in the legend to Fig. 5B. (B) Elements b, c, and d in the 5′ noncoding region of *Hoxb-4* regulate mRNA levels. Neuro2A and F9 cells were transfected with 20 μg of the indicated recombinants and 0.25 μg (Neuro2A) or 1 μg (F9 EC) of the internal control plasmid pBCTBX2 (15). A total of 150 μg (Neuro2A) or 200 μg (F9 EC) of total RNA was hybridized with a probe specific for CAT mRNA (*Eco*RI-*Hin*dIII fragment from pBLCAT3) and an 80-mer oligonucleotide probe specific for the mRNA of the internal control, treated with S1 nuclease and fractionated in a 7 M urea–6% (wt/vol) polyacrylamide gel. Arrows: fragments protected by mRNA from the CAT reporter recombinants (R) or by mRNA from the internal control (C).

6) of the gene and has an intermediate level of activity (the *Hoxb-4* minimal promoter series). CAT assays showed that in all four series the recombinants containing element c exhibited a moderate increase (two- to threefold) in activity compared with those without it, irrespective of the orientation of the insert (compare, in Fig. 7A, ACTB-CATd and ACTB-CATr with B-CAT, ACT-PBLCAT2d and ACT-PBLCAT2r with pBLCAT2, ACT-INd and ACT-INr with −5/+7, and LACT-MPd and LACT-MPr with −32/+21). A smaller version of element c, containing only 27 bases (+143 to +169), was also cloned upstream of the *Hoxb-4* minimal promoter and tested in the same cell line. Its activity was about seven times higher, in both orientations (ACTB-CATd and ACTB-CATr), compared with that of the minimal promoter alone (−32/+21) and 2.5 times higher than the longer version of element c, suggesting that the +135 to +143 sequences have a negative effect on transcription. These experiments further delimit element c (+143 to +169) and show that it activates transcription in an orientation- and promoter context-independent manner.

Two different factors bind independently to overlapping binding sites in element c. To characterize the factor(s) that activates through element c, we performed DNA mobility shift experiments with oligonucleotides comprising the +135 to +169 or the +143 to +169 sequences (see Fig. 7B). With the smaller probe (ACT-WT) we detected two major bands, A and B (lanes 2 and 5), that are specifically competed for by an excess of cold oligonucleotide (lanes 3 and 4 and 6 and 7). With the larger probe (LACT), band A is partly replaced by band X, which migrates slightly faster (lane 1) and which was not further characterized. Factors A and B have a wide distribution. Both are present in Neuro2A (lanes 2 to 4) and F9 EC (lanes 5 to 7) cells. Interestingly, they are also present in 10.5-day mouse embryos in a variety of different structures: forebrain (lane 8), hindbrain (lane 9), branchial arches (lane 10), liver (lane 11), and heart (lane 12), as well as in 8.5- and 14-day embryos (data not shown).

To determine the exact binding sites of factors A and B, we systematically mutated the +143 to +169 sequences and performed mobility shifts with the mutant oligonucleotides (Fig. 8). These experiments revealed two overlapping binding sites with similar but not identical specificities. While M3 is unable to bind any of the factors, M5 specifically binds factor A but not factor B. In contrast, M8 binds factor A with greatly diminished affinity, but the binding of factor B remains unchanged (or is even enhanced). These results, and the existence of similar binding sites in the myogenin and myoD1 promoters, allowed us to establish a consensus for each factor (Fig. 8): G C C A T/G T G/C G/C for factor A and C/A C A T t T T for factor B.

We also competed for binding to the wild-type probe with sites that showed some sequence similarity to those defined above. While NF-1 and NFY sites did not compete (data not shown), a YY1 binding site did, but only for the binding of factor B (Fig. 9B, compare lanes 1 and 4) strongly suggesting that factor B is indeed YY1 and is unrelated to factor A. In contrast, the myogenin binding site competed for the binding of only factor A (compare lanes 1 and 3), showing again that the two factors bind independently to the *Hoxb-4* sequences. In conclusion, element c binds two factors: one is likely to be the YY1 transcription factor; the other does not seem to be related to any known DNA-binding protein, and we will call it HoxTF (for homeobox gene transcription factor).

The HoxTF binding site mediates activation of transcription of the *Hoxb-4* minimal promoter. To distinguish which of the two factors, YY1 or HoxTF, is the activator, we introduced mutations specific for each of them and for both in combination. The double mutation completely abolished activation by element c (Fig. 9A, compare Double mut with ACT-WT). Mutation of the HoxTF site had a similar although less drastic effect (compare HoxTF mut with ACT-WT). The residual activity of HoxTF mut is most likely explained by the residual binding of the factor by the mutant sequence (10 to 20% of wild type; see M8 in Fig. 8), as can be noticed in Fig. 9B, which shows that HoxTF mut, but not Double mut, partially inhibits the HoxTF band. In contrast to HoxTF mut, mutation of the YY1 site had only a slight effect. These experiments show that the HoxTF binding site activates transcription of the *Hoxb-4* minimal promoter.

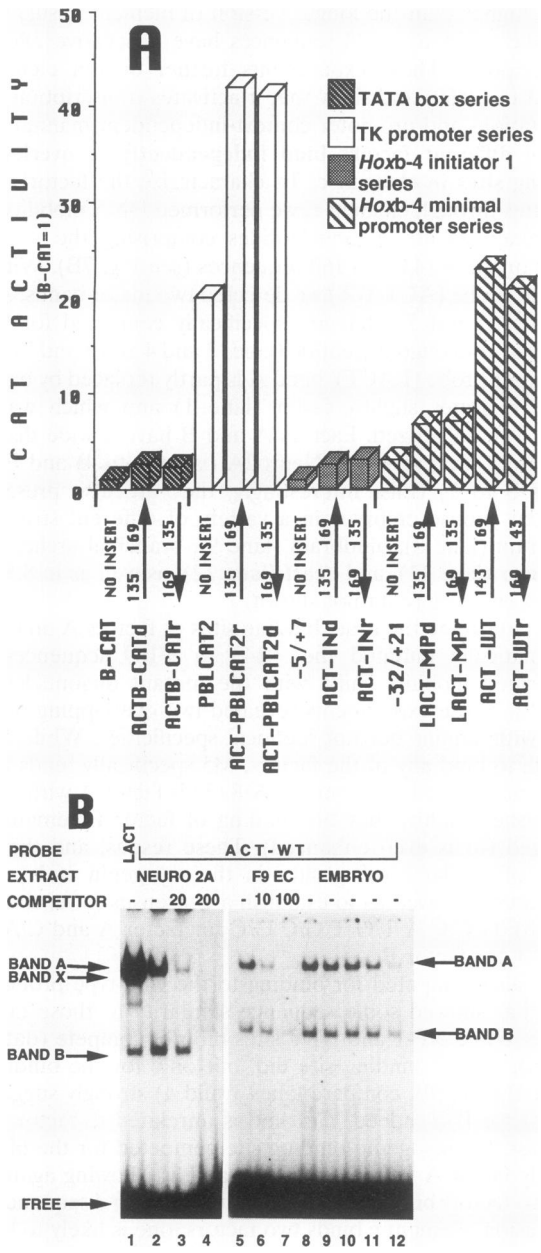

FIG. 7. (A) Regulatory element c activates transcription of heterologous promoters. Four kinds of reporter recombinants were transfected into the neuroblastoma cell line Neuro2A: the TATA box, TK minimal promoter, and *Hoxb-4* initiator 1 and *Hoxb-4* minimal promoter series. The founder reporters were, respectively, (i) B-CAT that contains the adenovirus E1B TATA promoter (19), (ii) pBLCAT2 that contains the minimal TK promoter (20), (iii) −5/+7 that contains the corresponding sequences of *Hoxb-4*, and (iv) −32/+21 that encompasses the corresponding sequences of *Hoxb-4*. Sequences +135 to +169 of the *Hoxb-4* gene (Fig. 1) were cloned in a single copy, in both orientations, upstream of these reporters producing ACTB-CATd and ACTB-CATr, ACT-PBLCAT2d and ACT-PBLCAT2r, ACT-INd and ACT-INr, and LACT-MPd and LACT-MPr, respectively. ACT-WT and ACT-WTr contain the +143 to +169 sequences of *Hoxb-4* upstream of −32/+21. A total of 16 μg of each reporter recombinant was cotransfected with 4 μg of the internal control plasmid RSV-βgal, by the calcium-phosphate method. Cell extracts were prepared, and CAT activity was measured and corrected by subtraction of background levels of the promoterless vector pBLCAT3. Relative CAT activity was normalized with respect to that of B-CAT, taken as 1. (B) Protein binding to positive regulatory element c. The LACT or ACT-WT end-labelled, double-stranded oligonucleotides (containing the +135 to +169 or the +143 to +169 sequences of *Hoxb-4*, respectively) were used as probes in gel shift experiments as described in Materials and Methods. Whole-cell extracts prepared from tissue culture cell lines (neuroblastoma Neuro2A or F9 EC) or from different structures from 10.5-day mouse embryos (hindbrain, lane 8; forebrain, lane 9; liver, lane 10; branchial arches, lane 11; and heart, lane 12) were used as a source of proteins. The molar excess of specific competitor, whenever used, is indicated. The three major retarded bands are shown. FREE, unbound probe.

DISCUSSION

The *Hoxb-4* gene has two independent promoters active in a variety of cell types. We have shown that the sequences immediately upstream of the *Hoxb-4* gene have promoter activity in that they are able to drive transcription of the CAT reporter gene in a variety of cultured cell lines. There are considerable differences in the activity of the promoter in the cell lines tested, and it is of interest that these differences reflect, in general terms, the transcriptional activity of the endogenous gene extrapolated from data from both the embryo and cell lines. *Hoxb-4* expression is high in neuronal cells (neuroblastoma cell lines Neuro2A and SK-N-SH) and parietal endoderm cells (F9-PE) and low in epithelial cells (HeLa), with intermediate levels in fibroblasts (NIH 3T3 and 10T1/2) and neurons derived from the neural crest (differentiated PC12).

The 5′ flanking sequences contain two transcription start sites, P1 and P2, 285 and 207 bases upstream of the translation initiation codon, respectively. Constructs containing either P1 or P2 alone are efficiently expressed, showing that the *Hoxb-4* gene has two independent promoters (PI and PII). In each case we have shown that a 12-bp element, centered on the start site, retains activity, suggesting that both promoters act as transcriptional initiators (35, 36). This idea is consistent with the absence of TATA boxes or GC-rich sequences.

Negative regulation of the *Hoxb-4* promoter. Our mutational analysis has identified three elements that diminish the activity of the *Hoxb-4* promoter (Fig. 10B). Element a, located between −226 and −186, acts negatively in a cell-type-specific fashion. Deleting it increases expression sixfold in F9 EC cells but has less effect in F9-PE and neuroblastoma cells. The other two elements are located in the 5′ untranslated region, but they clearly affect mRNA levels. Element d (+169 to +205) is similar in its properties to element a. It acts negatively in F9 EC cells but is without effect in neuroblastoma cells. In contrast, element b does negatively regulate expression in neuroblastoma cells. The negative elements a and d act additively and both can regulate each of the promoters. They are presumably responsible for the inefficient expression of the *Hoxb-4* promoter in F9 EC cells compared with F9-PE and neuroblastoma cells. Arcioni et al. have shown that the *HOXC-5* promoter contains a negatively acting element, of similar cell specificity, located upstream of the transcription start site (1). Moreover, Puschel et al. have demonstrated that the *Hoxa-7* gene contains elements within the 5′ untranslated region that restrict the activity of the promoter in the embryo (26). The presence of negative regulatory elements could explain the activation of *Hox* gene expression in EC cells by inhibitors of protein synthesis (32).

Positive regulation of the *Hoxb-4* promoter. Our deletion mutants revealed an element, named c (+143 to +169), that is required for efficient expression in all cell types tested. Removing it reduces expression markedly, 10- to 30-fold depending

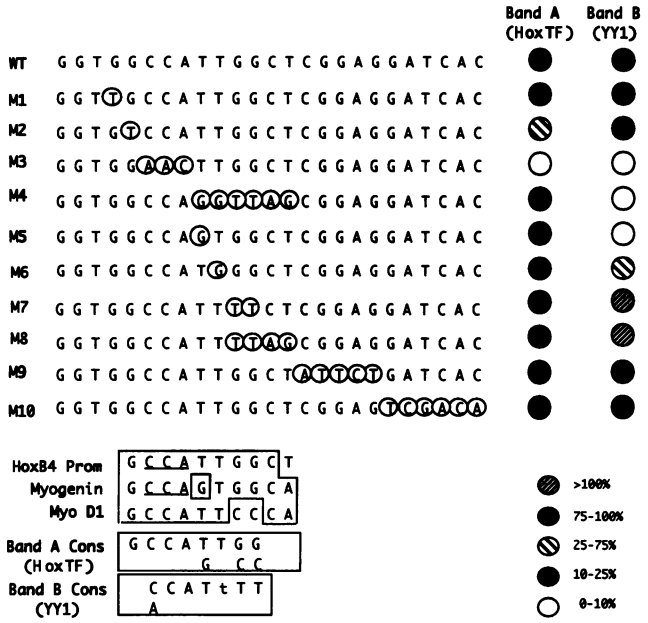

FIG. 8. Binding activity of element c. End-labelled, double-stranded oligonucleotides containing wild-type (WT) or mutated sequences (M1 to M10) of *Hoxb-4* (+143 to +169) were used as probes in gel shift experiments as described in Materials and Methods. Mutated bases are circled. Whole-cell extracts prepared from the neuroblastoma cell line Neuro2A were used as a source of proteins. Binding activities of the mutant sequences, compared with that of the wild type, are indicated for both complexes A and B. Sequences of similar binding sites in the mouse *myogenin* and *myoD1* genes and consensus binding sites for bands A and B are also presented.

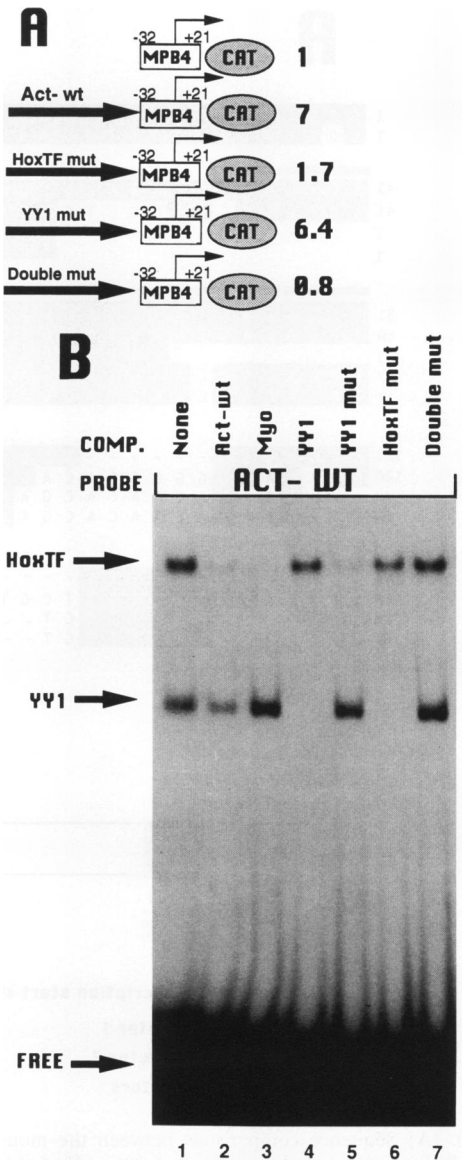

FIG. 9. Mutation of the HoxTF binding site abolishes activation of transcription by element c. (A) Transcription. Several recombinants were transfected into the Neuro2A cell line: the −32/+21 and ACT-WT reporters were described in the legend to Fig. 7. In the HoxTF mut, YY1 mut, and Double mut reporters, the mutated sequences M8, M5, and M3 (Fig. 8) replaced the wild-type sequences of ACT-WT. A total of 16 μg of each reporter recombinant was cotransfected with 4 μg of the internal control plasmid RSV-βgal, by the calcium-phosphate method. Cell extracts were prepared, and CAT activity was measured and corrected by subtraction of background levels of the promoterless vector pBLCAT3. Relative CAT activity was normalized with respect to that of −32/+21, taken as 1. MPB4, *Hoxb-4* minimal promoter. Broken arrows indicate the start site of transcription. (B) Binding activity. An end-labelled, double-stranded oligonucleotide containing wild-type sequences of *Hoxb-4* (+143 to +169) was used as a probe in gel shift experiments as described in Materials and Methods. Whole-cell extracts prepared from the neuroblastoma cell line Neuro2A were used as a source of proteins. Specific competitor, as indicated, was used at 100-fold molar excess. Act-wt, YY1 mut, HoxTF mut, and Double mut are wild-type and mutated sequences (+143 to +169) from the *Hoxb-4* promoter (WT, M5, M8, and M3, respectively, in Fig. 7). Myo comprises sequences from +9 to +35 of the mouse myogenin promoter (40); the YY1 binding site is described by Ellis et al. (10). HoxTF and YY1 indicate the two main retarded complexes (the same as bands A and B in Fig. 7 and 8). FREE, unbound probe.

on cell type, although in its absence the promoter retains detectable activity (4% of TK). This element can activate transcription of its own promoter as well as that of heterologous promoters, in an orientation-independent manner and in different contexts. It is able to interact with single elements like a TATA box or an initiator, but it is also functional when placed upstream of more complex promoters that exhibit higher levels of basal transcription and, presumably, contain several regulatory elements. Its versatility is also exemplified by the fact that its normal location is in the 5′ untranslated region, downstream of the transcription start site.

A long version of element c (+135 to +169) is less effective than a shorter one (+143 to +169) in activating transcription, indicating that the +135 to +143 sequences have a negative effect. Two factors bind to overlapping sites within the +143 to +169 region. One seems to be related to the YY1 transcription factor, as shown by our mutational analysis and by specific competition with a well-characterized YY1 binding site. The other factor is unrelated to YY1 and binds independently to its own site, as shown by competition experiments. By mutational studies and by comparison with similar binding sites in the myogenin and myoD1 promoters, we established a consensus binding site for this factor: G C C A T/G T G/C G/C. Binding sites that had some sequence similarity with this failed to compete, suggesting that we are likely to have identified a new factor that we name HoxTF (for homeobox gene transcription factor). Indeed, as revealed by the effect of mutations specific for each factor, the HoxTF binding site mediates activation of transcription of the *Hoxb-4* minimal promoter. In contrast, the YY1 site has no significant

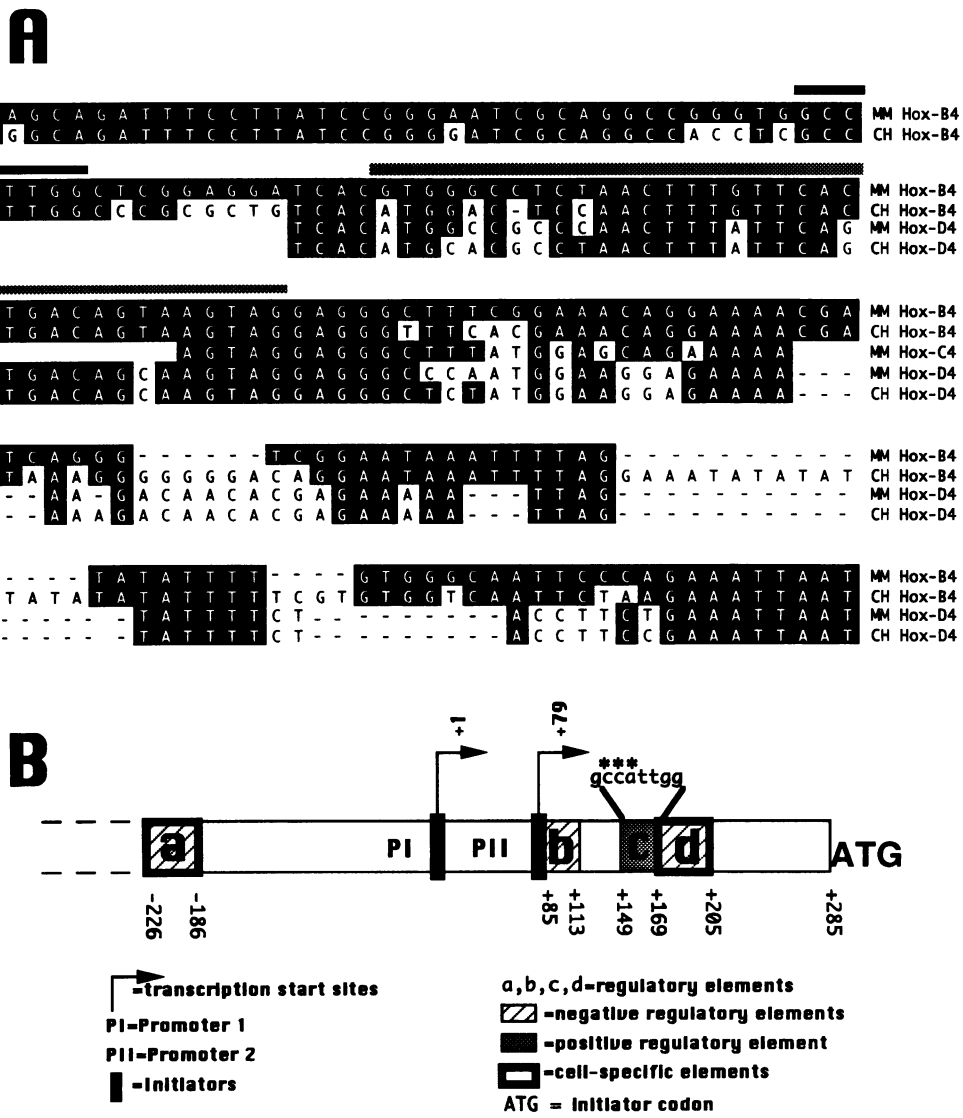

FIG. 10. (A) Sequence comparisons between the mouse *Hoxb-4* 5′ noncoding region and homologous and paralogous genes. Sequence alignment between 5′ noncoding regions of mouse *Hoxb-4* (MM Hox-B4), chicken *Hoxb-4* (CH Hox-B4 [30]), mouse *Hoxc-4* (MM Hox-C4 [13]), chicken *Hoxd-4* (CH Hox-D4 [31]), and mouse *Hoxd-4* (MM Hox-D4 [25]). The conserved regulatory elements (see Results) are indicated: black bar, HoxTF binding site; grey bar, element d. (B) Model of the mouse *Hoxb-4* gene promoter. Schematic diagram of the mouse *Hoxb-4* gene promoter and its regulatory elements (a, b, c, and d). Location of element b is uncertain; it could be between +21 and +47 or (as shown) between +85 and +113. Numbering is from the P1 transcription start site. The HoxTF binding site is shown in its approximate location. Asterisks indicate essential for binding.

effect. In conclusion, HoxTF is the active factor for the positive regulatory element c and is essential for efficient expression of the *Hoxb-4* gene. The element does not show any cellular specificity, which is in agreement with the wide distribution of the factor, in tissue culture cell lines as well as in the mouse embryo.

Our transgenic experiments showed that the 1.2-kb region of *Hoxb-4* under study here has no activity in the mid-gestation embryo, with the exception of ectopic expression in the superior colliculi (38). However, it is activated by spatially specific enhancers located both within the gene and downstream of it. In contrast, as reported here, this region is active in a variety of cell lines. Perhaps these cell lines reflect, in part, the situation in the later embryo or in the adult. A factor(s) that binds within the 1.2-kb region must be able to activate the gene. Among the regulatory elements we have characterized, the HoxTF binding site, the only one that is essential for efficient transcription of the *Hoxb-4* gene, is the most likely candidate. Interestingly, analogous switches seem to exist in some *D. melanogaster* homeobox genes (see next section).

Do homeobox gene promoters have common features? The main structural and functional features of the *Hoxb-4* promoter are diagrammed in Fig. 10B and can be summarized as follows. (i) The minimal promoter contains two initiators which act independently to dictate specific, albeit inefficient, initiation of transcription. (ii) The upstream sequences contain a cell-type-specific, negative regulatory element (element a) and sequences around each initiator which increase efficiency. (iii) The downstream sequences contain two negative regulatory elements (elements b and d), the latter of which is cell type

specific, and a powerful positive regulatory element (element c) which acts in all cell types tested.

The promoters of three *D. melanogaster* homeobox genes, *Ultrabithorax*, *Engrailed*, and *Antennapedia* (P2), lack TATA boxes and Sp1-like binding sites and appear to depend on initiators. All three have positive regulatory elements in the 5′ nontranslated region that are essential for transcription in vitro in embryo extracts (4, 24, 37). These general similarities to the *Hoxb-4* promoter are not reflected in any obvious sequence homologies.

The promoters of mammalian homeobox genes are not well characterized. Nonetheless, it appears that there are several different types of minimal promoter. Some genes, e.g., *HOXC-5* (1) and *HOXD-4* (8), have putative TATA boxes, while others, e.g., *Hoxa-4* (11) and *Hoxc-4* (13), that lack TATA boxes contain, between −60 and −30, potential Sp1 binding sites that are known to be able to mediate the initiation of transcription at multiple, closely spaced sites. A third class of genes, e.g., *Hoxc-6* (9) and *Hoxc-8* (2), lack such sequence motifs and thus may, like *Hoxb-4*, depend on initiators. It is striking that promoter architecture appears to be quite different even among the genes of a paralogous set. However, it is noteworthy that the negative element d is conserved in the paralogs *Hoxc-4* and *Hoxd-4*, although not in *Hoxa-4*, suggesting that there are regulatory mechanisms common to some genes in the paralogous group.

Homeobox gene promoters do not appear to share a common organization. Nevertheless, it is quite clear that a large subset of them share some unusual features. They are often multiple, sometimes close together, sometimes spaced up to several kilobases apart (7, 8, 33). A relatively high proportion contain minimal promoters that do not have TATA boxes but have initiators instead. They can be regulated by elements situated upstream or, interestingly, downstream of the transcription start site. The *Hoxb-4* promoter has all three of these characteristics. Its complex and unusual organization is appropriate for a developmental control gene, exhibiting precise spatial and temporal expression patterns, and is thus subject to sophisticated transcriptional regulation.

Conclusion. The experiments reported here have enabled us to define the structural and functional architecture of the *Hoxb-4* promoter and give insights into how other developmentally regulated genes may be organized. We can now investigate, with transgenic animals, how the numerous regulatory elements and the factors that bind to them interact with the spatially specific enhancers of the gene, which must set the precise boundaries of expression along the anteroposterior axis in both the mesoderm and the hindbrain, and can ensure that the levels of gene product made are appropriate to its role in the patterning process.

ACKNOWLEDGMENTS

We thank Robb Krumlauf for the gift of the *Hoxb-4* clone and for helpful comments on the manuscript. We also thank Rudi Allemann for subclones of the *Hoxb-4* promoter and for partial sequence data, James Ellis for inspiration regarding YY1, and the members of the Rigby laboratory for useful discussions.

J.G. holds a Research Studentship from the Medical Research Council, which also paid for this work.

REFERENCES

1. **Arcioni, L., A. Simeone, S. Guazzi, V. Zappavigna, E. Boncinelli, and F. Mavilio.** 1992. The upstream region of the human homeobox gene HOX3D is a target for regulation by retinoic acid and HOX homeoproteins. EMBO J. **11**:265–277.
2. **Awgulewitsch, A., C. Bieberich, L. Bogarad, C. Shakishant, and F. Ruddle.** 1990. Structural analysis of the Hox-3.1 transcription unit and the Hox-3.2 Hox-3.1 intergenic region. Proc. Natl. Acad. Sci. USA **87**:6428–6432.
3. **Balling, R., G. Mutter, P. Gruss, and M. Kessel.** 1989. Craniofacial abnormalities induced by ectopic expression of the homeobox gene Hox-1.1 in transgenic mice. Cell **58**:337–347.
4. **Biggin, M. D., and R. Tjian.** 1988. Transcription factors that activate the ultrabithorax promoter in developmentally staged extracts. Cell **53**:699–711.
5. **Chisaka, O., and M. R. Capecchi.** 1991. Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene Hox-1.5. Nature (London) **350**:473–479.
6. **Chisaka, O., T. S. Musci, and M. R. Capecchi.** 1992. Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene Hox-1.6. Nature (London) **355**:516–520.
7. **Cho, K., J. Goetz, C. Wright, A. Fritz, J. Hardwicke, and E. De Robertis.** 1988. Differential utilization of the same reading frame in a *Xenopus* homeobox gene encodes two related proteins sharing the same DNA-binding specificity. EMBO J. **7**:2139–2149.
8. **Cianetti, L., A. Di Cristofaro, V. Zappavigna, L. Bottero, G. Boccoli, J. Testa, G. Russo, E. Boncinelli, and C. Peschle.** 1990. Molecular mechanisms underlying the expression of the human Hox-5.1 gene. Nucleic Acids Res. **18**:4361–4368.
9. **Coletta, P. L., S. M. Shimeld, C. Chaudhuri, U. Muller, J. P. Clarke, and P. T. Sharpe.** 1991. Characterisation of the murine Hox-3.3 gene and its promoter. Mech. Dev. **35**:129–142.
10. **Ellis, J., D. Talbot, N. Dillon, and F. Grosveld.** 1993. Synthetic human β-globin 5′HS2 constructs function as locus control regions only in multicopy transgene concatamers. EMBO J. **12**:127–134.
11. **Galliot, B., M. Dolle, M. Vigneron, M. Featherstone, A. Baron, and D. Duboule.** 1989. The mouse Hox-1.4 gene: primary structure, evidence for promoter activity and expression during development. Development **107**:343–359.
12. **Gaunt, S. J.** 1991. Expression patterns of mouse Hox genes: clues to an understanding of developmental and evolutionary strategies. Bioessays **13**:505–512.
13. **Geada, A. M. C., S. J. Gaunt, M. Azzawi, S. M. Shimeld, J. Pearce, and P. T. Sharpe.** 1992. Sequence and embryonic expression of the murine Hox-3.5 gene. Development **116**:497–506.
14. **Gorman, C., R. Padmanabhan, and B. H. Howard.** 1983. High efficiency DNA-mediated transformation of primate cells. Science **221**:551–553.
15. **Gutman, A., C. Wasylyk, and B. Wasylyk.** 1991. Cell-specific regulation of oncogene-responsive sequences of the c-*fos* promoter. Mol. Cell. Biol. **11**:5381–5387.
16. **Kessel, M., R. Balling, and P. Gruss.** 1990. Variations of cervical vertebrae after expression of a Hox-1.1 transgene in mice. Cell **61**: 301–308.
17. **Krumlauf, R.** 1992. Evolution of the vertebrate Hox homeobox genes. Bioessays **14**:245–252.
18. **Le Mouellic, H., Y. Lallemand, and P. Brulet.** 1992. Homeosis in the mouse induced by a null mutation in the Hox-3.1 gene. Cell **69**:251–264.
19. **Lillie, J., and M. R. Green.** 1989. Transcription activation by the adenovirus E1a protein. Nature (London) **338**:39–44.
20. **Luckow, B., and G. Schutz.** 1989. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. **15**:417–420.
21. **Lufkin, T., A. Dierich, M. LeMeur, M. Mark, and P. Chambon.** 1991. Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression. Cell **66**:1105–1119.
22. **Lufkin, T., M. Mark, C. Hart, P. Dolle, M. LeMeur, and P. Chambon.** 1992. Homeotic transformation of the occipital bones of the skull by ectopic expression of a homeobox gene. Nature (London) **359**:825–841.
23. **McGinnis, W., and R. Krumlauf.** 1992. Homeobox genes and axial patterning. Cell **68**:283–302.
24. **Perkins, K. K., G. M. Dailey, and R. Tjian.** 1988. *In vitro* analysis of the antennapedia P2 promoter: identification of a new *Drosophila* transcription factor. Genes Dev. **2**:1615–1626.
25. **Popperl, H., and M. S. Featherstone.** 1992. An autoregulatory

element of the murine Hox-4.2 gene. EMBO J. **11**:3673–3680.
26. **Puschel, A. W., R. Balling, and P. Gruss.** 1991. Separate elements cause lineage restriction and specify boundaries of Hox-1.1 expression. Development **112**:279–287.
27. **Ramirez-Solis, R., H. Zheng, J. Whiting, R. Krumlauf, and A. Bradley.** 1993. Hoxb-4 (Hox-2.6) mutant mice show homeotic transformation of a cervical vertebra and defects in the closure of the sternal rudiments. Cell **73**:279–294.
28. **Sambrook, J., E. F. Fritsch, and T. A. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
29. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
30. **Sasaki, H., and A. Kuroiwa.** 1989. The nucleotide sequence of the cDNA encoding a chicken Deformed family homeobox gene, Chox-Z. Nucleic Acids Res. **18**:184.
31. **Sasaki, H., E. Yokoyama, and A. Kuroiwa.** 1990. Specific DNA binding of the two chicken Deformed family homeodomain proteins, Chox-1.4 and Chox-a. Nucleic Acids Res. **18**:1739–1747.
32. **Simeone, A., D. Acampora, V. Nigro, A. Faiella, M. D'Esposito, A. Stornaiuolo, F. Mavilio, and E. Boncinelli.** 1991. Differential regulation by retinoic acid of the homeobox genes of the four HOX loci in human embryonal carcinoma cells. Mech. Dev. **33**: 215–228.
33. **Simeone, A., F. Mavilio, D. Acampora, A. Giampaolo, A. Faiella, V. Zappavigna, M. D'Esposito, M. Pannese, G. Russo, E. Boncinelli, and C. Peschle.** 1987. Two human homeobox genes, C1 and C8: structure analysis and expression in embryonic development. Proc. Natl. Acad. Sci. USA **84**:4914–4918.
34. **Sleigh, M. J.** 1986. A nonchromatographic assay for expression of the chloramphenicol acetyltransferase gene in eucaryotic cells. Anal. Biochem. **156**:251–256.
35. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57**:103–113.
36. **Smale, S. T., M. C. Schmidt, A. J. Berk, and D. Baltimore.** 1990. Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proc. Natl. Acad. Sci. USA **87**:4509–4513.
37. **Soeller, W. C., S. J. Poole, and T. Kornberg.** 1988. *In vitro* transcription of the *Drosophila* engrailed gene. Genes Dev. **2**:68–81.
38. **Whiting, J., H. Marshall, M. Cook, R. Krumlauf, P. W. J. Rigby, D. Stott, and R. K. Allemann.** 1991. Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. Genes Dev. **5**:2048–2059.
39. **Wolgemuth, D. G., R. R. Behringer, M. P. Mostoller, R. L. Brinster, and R. D. Palmiter.** 1989. Transgenic mice overexpressing the mouse homeobox-containing gene Hox-1.4 exhibit abnormal gut development. Nature (London) **337**:464–467.
40. **Yee, S.-P., and P. W. J. Rigby.** 1993. The regulation of myogenin gene expression during the embryonic development of the mouse. Genes Dev. **7**:1277–1289.