The EMBO Journal vol.13 no.14 pp.3339–3347, 1994

The thyroid transcription factor-1 gene is a candidate target for regulation by Hox proteins

Stefania Guazzi, Renata Lonigro[1], Luisa Pintonello, Edoardo Boncinelli, Roberto Di Lauro[1,2] and Fulvio Mavilio

Department of Biology and Biotechnology, Istituto Scientifico H.S.Raffaele, Milano, [1]Department of Science and Biomedical Technologies, University of Udine, Udine and [2]Stazione Zoologica 'A.Dohrn', Napoli, Italy

Communicated by E.Boncinelli

Vertebrate Hox homeobox genes are transcription factors which regulate antero–posterior axial identity in embryogenesis, presumably through activation and/or repression of downstream target genes. Some of these targets were reported to code for molecules involved in cell–cell interactions, whereas no relationship has yet been demonstrated between Hox genes and other transcription factors involved in determining and/or maintaining tissue specificity. The thyroid transcription factor-1 (TTF-1) is a homeodomain-containing protein required for expression of thyroid-specific genes. A 862 bp 5′ genomic fragment of the rat TTF-1 gene, conferring thyroid-specific expression to a reporter gene, was sufficient to mediate transactivation by the human HOXB3 gene in co-transfection assay in both NIH3T3 or HeLa cells. HOXB3 is expressed in early mammalian embryogenesis in the anterior neuroectoderm, branchial arches and their derivatives, including the area of the thyroid primordia and thyroid gland. Transcription of the TTF-1 promoter is induced only by HOXB3, while its paralogous gene HOXD3 or other Hox genes expressed more posteriorly (HOXA4, HOXD4, HOXC5, HOXC6, HOXC8 and *Hoxd-8*) have no effect. Transactivation by HOXB3 is mediated by two binding sites containing an ATTA core located at −100 and +30 from the transcription start site. DNase I footprinting experiments show that the two sites bind HOXB3 protein synthesized in both *Escherichia coli* and eukaryotic cells, as well as nuclear factor(s) present in protein extracts obtained from mouse embryonic tissues which express group 3 Hox genes and TTF-1. Some of the DNA–protein complexes formed by the embryonic extracts are indistinguishable from those generated by HOXB3. These data suggest that HOXB3 might be a transcriptional regulator of the TTF-1 gene in early embryogenesis, and could therefore participate in the specification and development of the thyroid gland.

Key words: homeobox/Hox genes/thyroid development/transcriptional regulation/TTF-1

Introduction

The homeobox is a 180 bp DNA sequence motif originally found in the homeotic (HOM) genes that control morphogenesis and segment identity in the fruit fly *Drosophila melanogaster*. The homeobox encodes a 60 amino acid (aa) residue domain, the homeodomain, which is capable of sequence-specific DNA binding (reviewed in Gehring *et al.*, 1990). Hox genes have been cloned in a broad spectrum of animals, from simple invertebrates to mammals (reviewed in Gehring, 1987; McGinnis and Krumlauf, 1992). In higher vertebrates, Hox genes are grouped into four, unlinked gene clusters derived through duplication from a putative, ancestral Hox complex (see Schubert *et al.*, 1993). Both HOM and Hox genes are expressed along the antero–posterior axis of developing embryos according to spatially-restricted patterns which correlate with their physical arrangement in the clusters, following a 3′ anterior-proximal, 5′ posterior-distal colinearity rule (reviewed in Gehring, 1987; Gruss and Kessel, 1991; Duboule, 1992; McGinnis and Krumlauf, 1992; Krumlauf, 1993). These patterns strongly suggest a causal role for Hox genes in specifying positional identity in vertebrate embryonic development, where progressively more posterior structures might be defined by activation of progressively more 5′ genes (see also Hunt and Krumlauf, 1991; Kessel and Gruss, 1991). However, the molecular and cellular mechanisms by which this putative Hox code is translated into positional specification remain unknown to date.

Homeodomain proteins are able to act as transcriptional activators or repressors (Jaynes and O'Farrell, 1988; Thali *et al.*, 1988; Krasnow *et al.*, 1989; Hayashi and Scott, 1990), and presumably play their developmental role through transcriptional regulation of distinct sets of downstream target genes. In *Drosophila*, HOM genes are known to auto- and cross-regulate their own transcription to establish correct spatial and temporal expression patterns (Appel and Sakonju, 1993; reviewed in Hayashi and Scott, 1990; McGinnis and Krumlauf, 1992). Auto- and cross-regulatory interactions between vertebrate Hox proteins and Hox gene regulatory elements have also been demonstrated, at least in cell culture (Zappavigna *et al.*, 1991; Arcioni *et al.*, 1992; Pöpperl and Featherstone, 1992). On the other hand, only a few downstream target genes have been so far identified for both HOM and Hox proteins, the best characterized being the *Drosophila decapentaplegic* gene, which encodes a member of the TGFβ family of peptide growth factors (Capovilla *et al.*, 1994). Recently, several genes encoding for morphoregulatory molecules have been suggested as candidate targets for transcriptional regulation by HOM/Hox genes, i.e. the *Drosophila connectin* (Gould *et al.*, 1990; Gould and White, 1992) and β-3 tubulin (Hinz *et al.*, 1993), and the mouse neural cell

adhesion molecule (N-CAM) (Jones *et al.*, 1992, 1993) and *mgl-1*, also encoding a putative cell adhesion molecule (Tomotsune *et al.*, 1993). Conversely, no relationship has yet been demonstrated between Hox genes and other transcriptional regulators directly involved in determining and/or maintaining tissue specificity, although they might represent a natural link between positional information cues and development of a specific cell type. In *Drosophila*, the gene for the transcription factor *salm* is apparently a target of *Antennapedia* (Wagner-Bernholz *et al.*, 1991).

We have cloned previously a cDNA encoding a tissue-specific transcriptional regulator, the thyroid transcription factor-1 (TTF-1) (Guazzi *et al.*, 1990). TTF-1 is required for the expression of at least two thyroid-specific genes, i.e. thyroglobulin and thyroperoxidase, and contains a divergent homeodomain which is necessary and sufficient for its DNA binding activity (Civitareale *et al.*, 1989; Guazzi *et al.*, 1990; Damante and Di Lauro, 1991; Francis-Lang *et al.*, 1992). TTF-1 is expressed also in developing lungs and in restricted areas of the brain. In rat embryogenesis, TTF-1 expression becomes detectable at day 10.5 p.c., and is associated with the onset of both thyroid and lung organogenesis (Guazzi *et al.*, 1990; Lazzaro *et al.*, 1991). The rat TTF-1 gene promoter and 5′ flanking sequences have been identified recently, and shown to confer tissue specificity to a reporter gene (R.Lonigro and R.Di Lauro, manuscript in preparation).

Hox genes belonging to the third paralogous group, i.e. *Hoxa-3* (= *Hox-1.5*), *Hoxb-3* (= *Hox-2.7*) and *Hoxd-3* (= *Hox-4.1*), are expressed in mouse embryos in the anterior neuroectoderm, branchial arches and their derivatives, including the thyroid and lung rudiments, starting from day 9.5 p.c. (Gaunt, 1987; Hunt *et al.*, 1991a,b; Sham *et al.*, 1992). In particular, a *Hoxb-3–lacZ* transgene, expressed in a pattern similar to that of the endogenous *Hoxb-3* gene, was shown to direct β-galactosidase activity specifically in the thyroid and the lung (Sham *et al.*, 1992). Mouse embryos homozygous for a *Hoxa-3* knockout mutation showed complex morphological defects localized to the head and thorax, including abortive thyroid development (Chisaka and Capecchi, 1991), whereas genetic ablation of the paralogous *Hoxd-3* gene caused a dramatic remodeling of the craniocervical joint but did not apparently affect thyroid development (Condie and Capecchi, 1993). Taken together, these data suggest that expression of group 3 Hox genes is necessary for the specification and/or development of specific anterior structures, which may include the thyroid gland. Given that TTF-1 and *Hoxb-3* genes are expressed at the same time and in the same structures in mouse and rat embryogenesis, we decided to investigate whether the human HOXB3 gene might be a transcriptional regulator of the TTF-1 gene. Co-transfection of a HOXB3 expression vector caused a 4- to 6-fold increase in the transcriptional activity of the rat TTF-1 promoter driving a reporter gene, whereas HOXD3 and HOX genes belonging to several other paralogy groups had no effect. Two different binding sites for the HOXB3 protein were identified in close proximity to the TTF-1 transcription start site, which are necessary for transactivation in transfected cells. These sites are specifically recognized by protein extracts obtained from 12.5 day-old mouse embryos. Our data suggest that group 3 Hox genes might be directly involved in the regulation of TTF-1 gene expression in early embryogenesis, thereby participating in thyroid gland development.

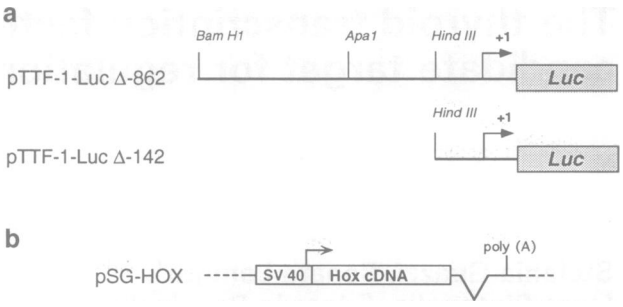


Fig. 1. Schematic representation of reporter and expression constructs used in cell co-transfection assays. (**a**) The reporter constructs pTTF-1–Luc Δ-862 and pTTF-1–Luc Δ-142 were generated by cloning two different upstream genomic sequences of the rat TTF-1 gene, extending from position –945 or –142, respectively, to position +103 from the transcription start site, upstream of the luciferase (Luc) coding region in the pBS–Lucα vector. (**b**) Full-length, human or mouse Hox cDNA sequences (Hox cDNA) were cloned in the mammalian expression vectors pSG-5, under the control of the SV40 early promoter (pSG–HOX). Arrows indicate transcription start sites in all constructions.

Results

HOXB3 transactivates the rat TTF-1 promoter in HeLa and NIH3T3 cells

Two overlapping, 965 and 243 bp 5′-flanking DNA fragments encompassing the major transcription start site of the rat TTF-1 gene were cloned in front of the luciferase reporter gene in the pBS–Lucα expression vector (pTTF-1–Luc Δ-862 and pTTF-1–Luc Δ-142, respectively; see Figure 1a). In transfection assays in a rat differentiated thyroid cell line (FRTL-5), pTTF-1–Luc Δ-862 and pTTF-1–Luc Δ-142 showed a 6- and 3-fold higher activity, respectively, when compared with that detected in a rat fibroblast cell line (Rat-1), suggesting that at least part of the information necessary to confer tissue specificity to TTF-1 gene expression resides in this upstream DNA sequence (R.Lonigro and R.Di Lauro, manuscript in preparation).

A full-length cDNA coding for the human HOXB3 protein (Acampora *et al.*, 1989) was cloned in the SV40 early promoter-based mammalian expression vector pSG5 to generate the pSG–HOXB3 construct (Figure 1b). Co-transfection of pSG–HOXB3 in HeLa and NIH3T3 cells caused a 4- to 6-fold increase in the transcriptional activity of both TTF-1 reporter constructs (Figure 2 and data not shown). pTTF-1–Luc Δ-142 was therefore used in all subsequent experiments. To assess the specificity of the HOXB3 transactivation on the TTF-1 promoter, expression vectors for several other Hox cDNA (i.e. HOXD3, HOXD4, HOXC5, HOXC6, HOXC8 and the murine *Hoxa-4* and *Hoxd-8*) were co-transfected in the same cell lines, together with pTTF-1–Luc Δ-142. HOXD4 caused a very modest (~2-fold) increase in the activity of the reporter construct in HeLa cells, whereas all other proteins caused either no variation or even a decrease of the TTF-1 promoter basal activity (Figure 2a and b). These results indicate that the ability to transactivate the TTF-1 promoter

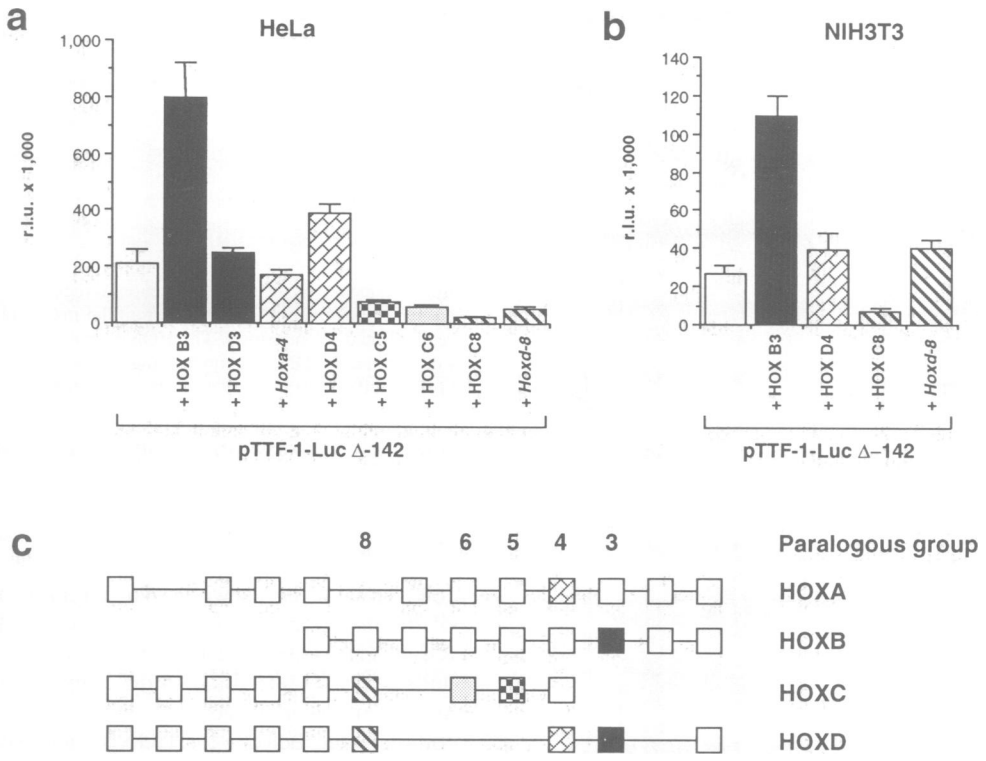


Fig. 2. Transactivation of the rat TTF-1 −142 to +103 upstream region by the human HOXB3 protein. 10 μg of the pTTF-1–Luc Δ-142 reporter construct were co-transfected in HeLa (**a**) or NIH3T3 (**b**) cells together with 5 μg of expression vectors encoding a number of homeoproteins belonging to different paralogous groups. The open bars indicate the basal activity of the TTF-1 upstream region in HeLa or NIH3T3 cells. Data from co-transfection assays with Hox expression constructs are represented by filled bars. The position of each co-transfected gene in the four mammalian Hox clusters is shown in (**c**), where each square represents an individual gene, and co-transfected genes are filled following the same code used for bars in (a) and (b). The numbers on top of the clusters indicate the paralogous group of the transfected genes. Luciferase activity is expressed in arbitrary light units.

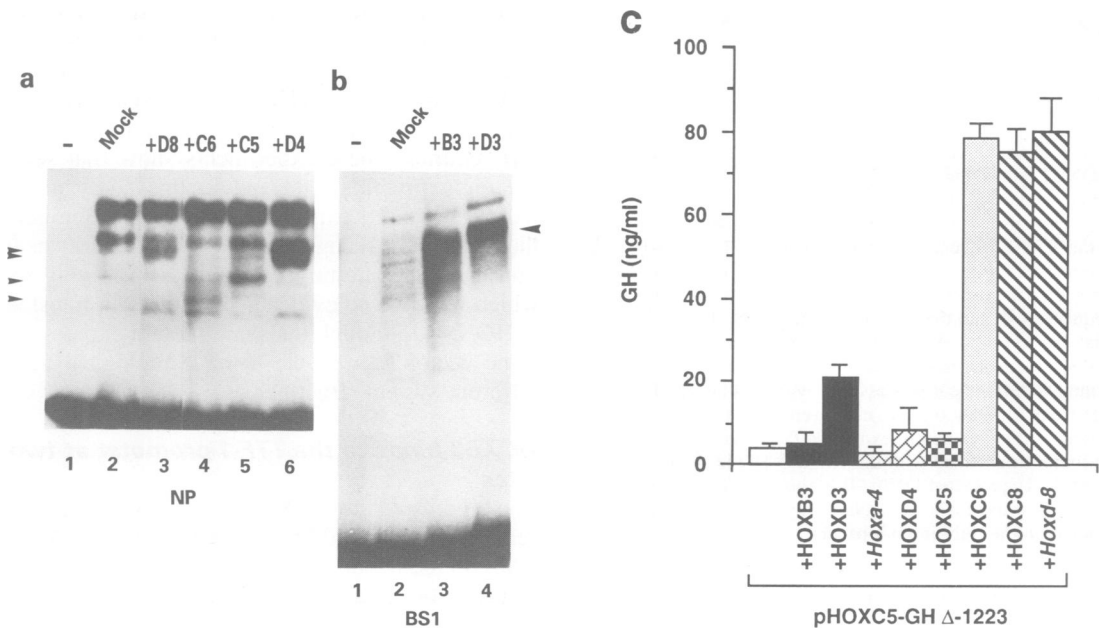


Fig. 3. Production and activity of Hox proteins after transfection of expression vectors in HeLa cells. (**a**) Mobility shift analysis of nuclear extracts from HeLa cells transfected with pSG-5 (lane 2) or pSG-based expression vectors for *Hoxd-8*, HOXC6, HOXC5 and HOXD4 (lanes 3–6), using a double-stranded oligonucleotide containing the NP Hox consensus binding sequence (see Materials and methods). (**b**) Same as in (a), using the TTF-1 promoter BS1 binding site as probe (see Figure 4). (**c**) Transactivation of the HOXC5 promoter and upstream regulatory sequences linked to the human growth hormone (hGH) reporter gene (pHOXG5–GH Δ-1223; see Arcioni *et al.*, 1992) by co-transfection of Hox expression vectors. The bar filling code is the same as in Figure 2. Activity is expressed in ng/ml of hGH secreted in the medium.

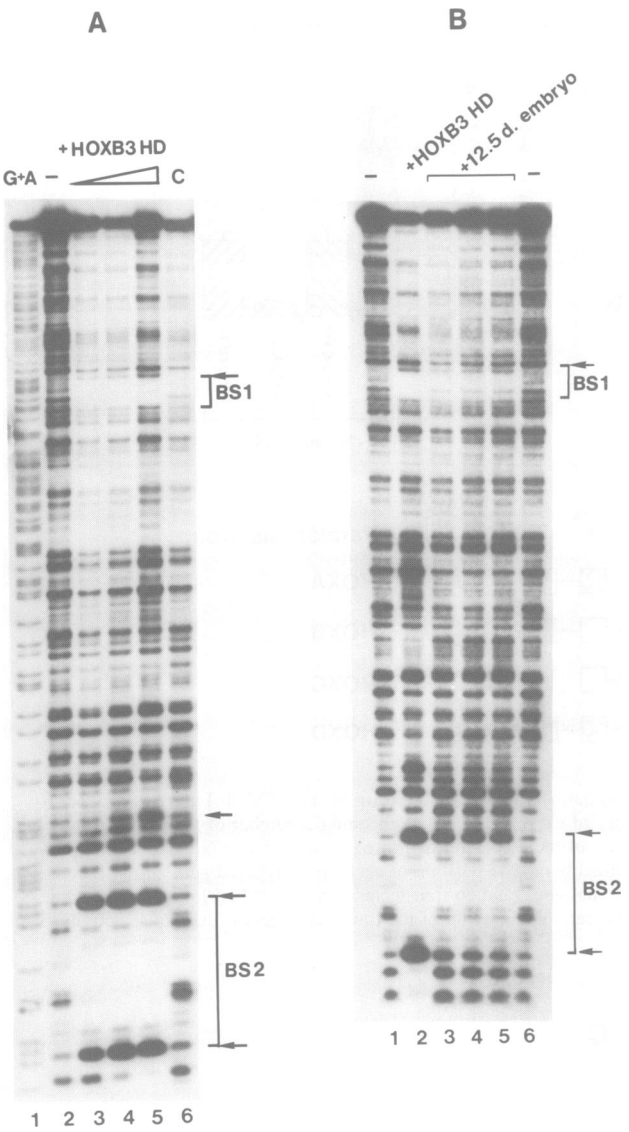


Fig. 4. DNase I footprinting analysis of the binding of a bacterially synthesized HOXB3 homeodomain (HD) and nuclear extracts from 12.5 day-old mouse embryos to the −142 to +103 TTF-1 upstream region. (**A**) A 5′ end-labeled 323 bp *Hin*dIII–*Xba*I fragment, containing all the TTF-1 sequence present in the pTTF-1–Luc Δ-142 construct, was incubated with 0.25, 0.50 and 1.00 μl of crude, HOXB3 HD-containing bacterial extracts (+HOXB3 HD, lanes 3–5). The unprotected sequence (–), and the sequence protected by 1 μl of control bacterial extracts (C) are indicated in lanes 2 and 6. G+A, Maxam and Gilbert sequence reaction is shown as a size marker (lane 1). (**B**) The same DNA fragment was assayed with 2, 4 and 6 μl of nuclear extracts from 12.5 day-old p.c. mouse embryos (lanes 3–5), enriched in anterior structures before extraction (see Materials and methods). The footprint obtained with the HOXB3 HD is shown for comparison (lane 2). Unprotected DNA (–) is shown in lanes 1 and 6. Protected regions (BS1 and BS2) and DNase I hypersensitive sites (arrows) are shown on the right in both panels.

is not a feature shared by all members of the Hox gene family, but is apparently specific for HOXB3.

Efficient production in HeLa cells of the protein that showed no effect on the TTF-1 promoter was checked by mobility shift assay after binding of nuclear extracts from transfected HeLa cells to an oligonucleotide containing the NP consensus homeodomain binding site (Desplan *et al.*, 1988; Hoey and Levine, 1988). Specific DNA–protein complexes were formed with extracts obtained from cells transfected with four out of the eight Hox proteins, i.e. HOXD4, HOXC5, HOXC6 and *Hoxd-8* (Figure 3a). The difference in intensity observed between the retarded complexes may reflect different rates of protein synthesis, as well as differential affinity for the NP binding site. We favor the latter explanation, since HOXB3, HOXD3 and HOXC8, which showed no retarded bands with this site, bind to different targets in the same experimental conditions, i.e. the BS1 site in the TTF-1 promoter (see below) for HOXB3 and HOXD3 (Figure 3b), and the HCR sequence in the HOXD9 promoter for HOXC8 (Zappavigna *et al.*, 1994). Production and activity of the Hox proteins which had no effect on the TTF-1 promoter were also checked by co-transfection with an alternative reporter construct, i.e. the human growth hormone (GH) gene driven by the promoter and upstream regulatory sequences of the HOXC5 gene. We showed previously that the HOXC5 promoter contains binding sites for, and is transactivated by, a number of Hox proteins in co-transfection assay (Arcioni *et al.*, 1992). As shown in Figure 3c, the Δ-1223 HOXC5–GH construct is activated ~5-fold by HOXD3, and ~20-fold by HOXC6, HOXC8 and *Hoxd-8*. HOXB3, as well as *Hoxa-4*, HOXD4 and HOXC5, had apparently no activity on the HOXC5–GH reporter. These experiments show that seven out of the eight proteins analyzed in this study are produced from pSG-based expression vector in HeLa cells, where they are able to recognize, bind to and/or activate transcription from specific target sequences. The only protein for which we have no evidence of activity in transfected cells is HOXA4, although the same protein expressed from the same vector was shown to have biological activity in a different system (Buettner *et al.*, 1991).

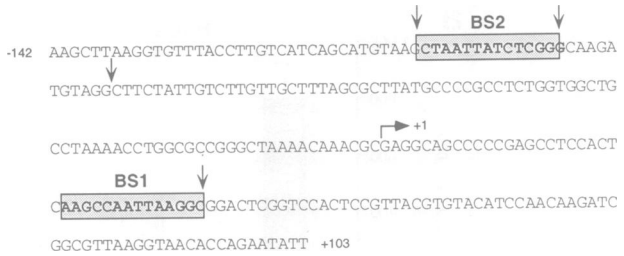


Fig. 5. Sequence of the upstream genomic region of the rat TTF-1 gene, from position −142 to +103 from the major transcription start site (indicated by an arrow). The HOXB3 binding sites BS1 and BS2, as determined by DNase I footprinting (Figure 3) are boxed. DNase I hypersensitive sites are indicated by vertical arrows.

HOXB3 binds to the TTF-1 promoter at two ATTA sites

The binding of the HOXB3 protein to the TTF-1 promoter sequence was tested by a DNase I footprinting assay. The HOXB3 homeodomain (HD) and immediately flanking residues (4 aa at the N-terminus and 10 aa at the C-terminus) produced in *Escherichia coli* by a T7 promoter-based expression system (see Materials and methods), showed two strong protection signals around positions −100 and +30 from the major transcription start site in the TTF-1 243 bp promoter fragment (Figure 4A). Binding of the HOXB3 homeodomain to these two sites (BS1 and BS2) generated four DNase I hypersensitive sites, located

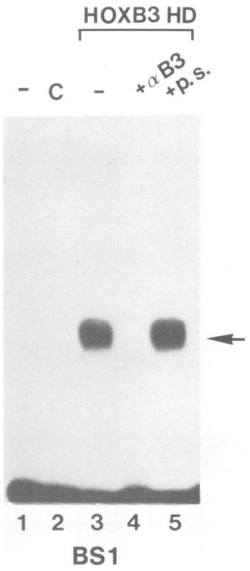
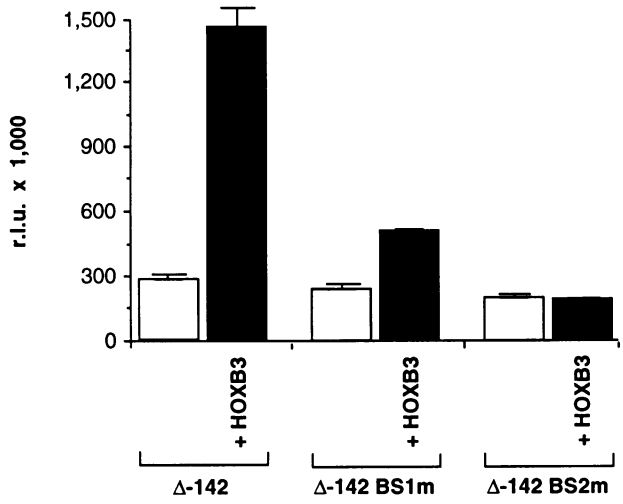


Fig. 6. Mobility shift assay of a 25mer, double-stranded oligonucleotide containing the BS1 sequence (see Materials and methods), after *in vitro* binding to a bacterial extract containing the HOXB3 homeodomain (HOXB3 HD, lane 3). A polyclonal, anti-HOXB3 HD rabbit antiserum (αB3), or a control pre-immune serum (p.s.) were added to the binding reactions run in lanes 4 and 5, respectively. Naked DNA (–) and DNA bound to a control bacterial extract (c) are in lanes 1 and 2. The specific, retarded complex is indicated by an arrow.




```
BS1      +25 AAGCCAATTAAGGC +38
BS1m     +25 AAGCCCGGGTAGGC +38

BS2     -109 AAGCTAATTATCTC -96
BS2m    -109 AAGCTATCAGTCTC -96
```

Fig. 7. Transactivation by HOXB3 of the pTTF-1–Luc Δ-142 reporter construct, containing the wild type –142 to +103 TTF-1 genomic sequence (Figures 1 and 4), and two mutant constructs (Δ-142 BS1m and Δ-142 BS2m) in which the BS1 or the BS2 sequences were mutated in BS1m and BS2m, respectively, as indicated at the bottom. Luciferase activity is expressed in arbitrary light units.

at the 3′ end of BS1, at the 5′ and 3′ ends of BS2, and 12 bp downstream of BS2 (Figures 4A and 5). Both sites contain an ATTA core sequence, a characteristic feature of most HOM/Hox protein binding sites (see Hayashi and Scott, 1990), whereas the immediately adjacent sequences are different in the two sites. In particular, BS2 contains a palindromic TAATTA sequence. On the basis of footprinting information, we designed two 25mer double-stranded oligonucleotides (see Materials and methods) containing the BS1 and BS2 protected sequences. Both were specifically retarded by the HOXB3 HD in a gel-shift assay (Figure 6 and data not shown). Formation of the retarded complex was competed by the addition of a 500× molar excess of unlabeled specific oligonucleotide, but not by an unrelated one (data not shown). BS1 and BS2 oligonucleotides were able to compete each other's binding to the HOXB3 HD to approximately the same extent, indicating no significant difference in the affinity of HOXB3 for the two sites *in vitro*. A polyclonal antiserum raised against the HOXB3 HD specifically interfered with the HOXB3–BS1 complex formation, confirming the presence of HOXB3 in the retarded band (Figure 6, lane 4). The observed interference did not depend on whether the antibody was added before or after the DNA–protein complex formation in the binding assay, suggesting that binding of the HOXB3 HD to the DNA or to the antibody is mutually exclusive. At higher concentrations of HOXB3 HD-containing crude extracts, the DNA–protein complex could not be completely abolished, and a faint, slower migrating DNA–protein–antibody complex was detected (data not shown).

BS1 and BS2 mediate transactivation by HOXB3 in transfected cells

To test the role of the BS1 and BS2 sites in the transactivation of the TTF-1 promoter by HOXB3, the ATTA cores of the two sequences were individually mutagenized to GGGT and TCAG, respectively, in the pTTF-1–Luc Δ-142 vector to generate the pTTF-1–Luc Δ-142 BS1m and pTTF-1–Luc Δ-142 BS2m plasmids, which were then tested in co-transfection assay together with pSG–HOXB3 in NIH3T3 cells. The mutation at the BS1 site caused a >50% decrease in the extent of the HOXB3 transactivation of the TTF-1 promoter, whereas the mutation at the BS2 site completely abolished the transactivation (Figure 7). These data show that the HOXB3 activity of the TTF-1 promoter is directly mediated by the BS1 and BS2 sites. The loss of even one of the two sites, particularly BS2 which contains the palindromic TAATTA sequence, is sufficient to abolish the HOXB3-induced activation, suggesting the existence of a cooperative effect, at least in cell culture.

HOXB3 and nuclear factors from mouse embryos bind to the same sites in the TTF-1 promoter

To assess whether HOXB3 or other HOX proteins expressed in embryonic tissues are able to recognize the same sequences bound by the bacterially synthesized HOXB3 homeodomain, nuclear extracts from 12.5 day p.c. mouse embryos were prepared and used in a footprinting assay. Both TTF-1 and HOXB3 are expressed in the mouse at this developmental stage (Lazzaro *et al.*, 1991; Sham *et al.*, 1992). The embryonic preparations were 'enriched' in regions expressing group 3 Hox proteins by surgical removal of fore- and mid-head, anterior limbs and all the posterior (i.e. post thoracic) body before nuclear protein extraction. The extracts were allowed to react with

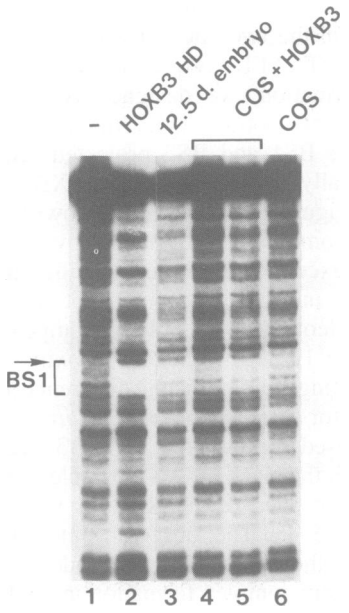


Fig. 8. DNase I footprinting analysis of the TTF-1 upstream genomic sequence around the BS1 site after *in vitro* binding to 4 and 6 µl of nuclear extracts from COS-7 cells transfected with the HOXB3 expression vector (lanes 4 and 5). Binding to a control extract from untransfected COS cells is shown in lane 6. Footprinting of the same sequence by a bacterial extract containing the HOXB3 HD (lane 2) and a nuclear extract from 12.5 day-old mouse embryos (lane 3) is shown for comparison. Unprotected DNA (–) is run in lane 1. Location of the BS1 sequence is indicated on the left. The arrow indicates the DNase I hypersensitive site generated by binding to the HOXB3-containing extracts.

the 243 bp TTF-1 5′ genomic fragment where they protected from DNase I digestion virtually the same regions recognized by the HOXB3 HD. In particular, three of the four hypersensitive sites flanking the BS1 and BS2 sites were as well generated by the embryonic extracts (Figure 4B). Nuclear extracts were prepared also from COS cells transfected with the pSG–HOXB3 vector. These extracts protected the BS1 site from DNase I digestion generating a footprinting pattern virtually identical to that obtained with the embryonic extracts (Figure 8), indicating that a native, full-length HOXB3 protein is able to bind to the same site recognized by the bacterially expressed homeodomain and the embryonal extracts (Figure 8). Binding to the BS2 site could not be analyzed in COS cells due to an extensive footprint encompassing the BS2 sequence observed also with untransfected cell extracts (data not shown).

To further investigate the binding properties of the HOXB3 protein, transfected COS cell nuclear extracts were tested in band-shift experiments, using the BS1 and BS2 oligonucleotides as probes. With both binding sites, nuclear extracts from HOXB3-transfected COS cells formed multiple retarded DNA–protein complexes, all specifically competed by the addition of cold BS1 and BS2 DNA (Figure 9). Three of the complexes formed by both BS1 and BS2 showed the same relative mobility, and were not present in control, untransfected COS cell nuclear extracts (indicated by arrows in Figure 9). Formation of all three complexes was specifically competed by addition of the anti-HOXB3 antiserum, suggesting that they all contain the transfected HOXB3 protein (Figure

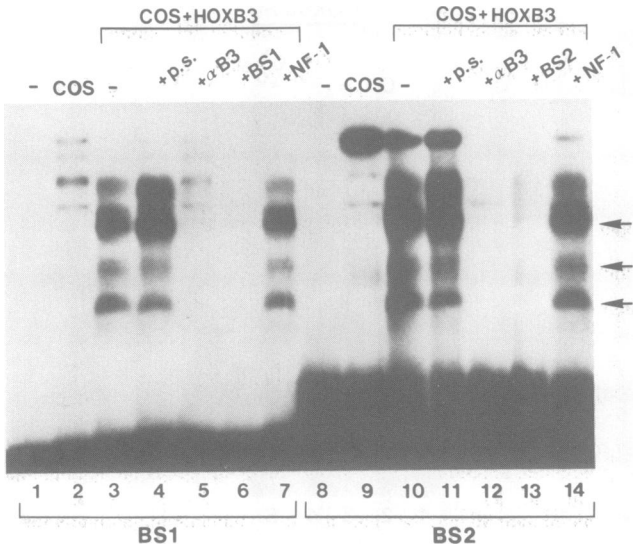


Fig. 9. Mobility shift assay of double-stranded oligonucleotides containing the BS1 and BS2 sequences, after binding to nuclear extracts from COS-7 cells transfected with the HOXB3 expression vector (lanes 3 and 10). A polyclonal, anti-HOXB3 antiserum (αB3) or a pre-immune serum (p.s.) were added to the binding reactions (lanes 4, 5, 11 and 12). Competition by a 500× molar excess of cold, specific BS1 and BS2 oligonucleotides (lanes 6 and 13), or by the same excess of an unrelated oligo containing the NF-1 binding site (lanes 7 and 14), is also shown. Control lanes contain unbound DNA (lanes 1 and 8) or DNA bound to nuclear extracts from untransfected COS cells (lanes 2 and 9). DNA–protein complexes specifically formed in the presence of the HOXB3 protein are indicated by arrows.

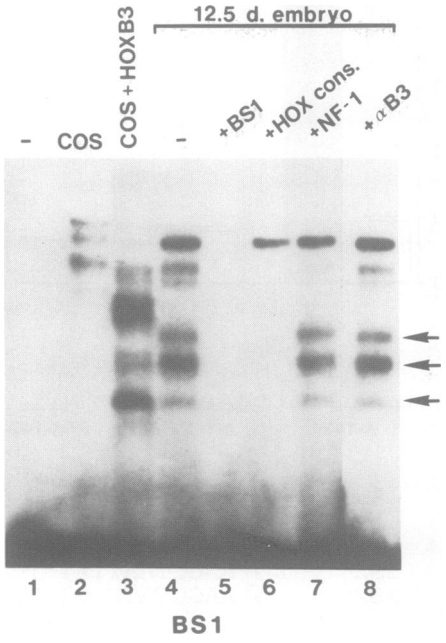


Fig. 10. Mobility shift assay of the BS1 oligonucleotide after binding to 12.5 day-old mouse embryo extracts (lane 4), together with a 500× molar excess of cold competitor oligonucleotides containing the BS1 sequence (lane 5), a HOM/Hox consensus binding site sequence (see Materials and methods, lane 6), or an unrelated NF-1 binding site (lane 7). The reaction was carried out also in the presence of an anti-HOXB3 antiserum (+αB3, lane 8). Binding reactions with nuclear extracts from COS-7 cells, untransfected (lane 2) or transfected with the HOXB3 expression vector (lane 3), are shown for comparison. The unbound oligonucleotide is run in lane 1. DNA–protein complexes specifically formed in the presence of embryonic extracts are indicated by arrows.

9, lanes 4, 5, 11 and 12). A very slow-migrating complex, present in both transfected and untransfected COS cells, was detected only with the BS2 site. This binding activity might account for the extensive protection observed around the BS2 sequence in the DNase I footprinting assay (see above).

Nuclear extracts from 12.5 day-old mouse embryos generated at least six different DNA–protein complexes in a gel-shift assay using the BS1 oligonucleotide as probe (Figure 10). Although all these complexes were competed out by an excess of cold BS1 oligo, only four of them were not competed by an unrelated DNA sequence (Figure 10, lanes 5 and 7). Three of these four bands were also competed by an oligonucleotide containing a consensus binding site for HOM/Hox proteins (arrows in Figure 10), and two of them shared the same relative mobility as two of the specific complexes generated by transfection of HOXB3 in COS cells (Figure 10, compare lane 3 with lane 4). These data indicate that the BS1 sequence is able to bind to protein complexes contained in embryonic nuclear extracts, at least some of which are likely to contain *Hoxb-3*. The presence of *Hoxb-3* in these extracts was confirmed by Western blotting using the anti-HOXB3 antiserum as probe (not shown). However, addition of anti-HOXB3 antiserum did not interfere with any of the binding activities to the BS1 site in the embryonic extracts (Figure 10, lane 8). We interpret these data by suggesting that either the antibody does not recognize the embryonic *Hoxb-3* protein, or it is titrated out by cross-reacting Hox protein, or else the *Hoxb-3*–BS1 complex comigrates with other, more abundant, Hox–BS1 complexes.

Discussion

Vertebrate Hox genes encode a large family of proteins evolutionarily related to the *Drosophila* HOM gene products involved in the specification of the segmental body plan in the embryonic development of insects. The temporal and regional restrictions in the developmentally regulated expression of Hox genes suggest that they might have a conserved role as a transducer of positional specification also in vertebrate development. The analysis of mouse knockout mutants for a number of Hox genes, which show complex developmental abnormalities sometimes resembling homeotic mutations, confirms this hypothesis (reviewed in Duboule, 1992; McGinnis and Krumlauf, 1992; Krumlauf, 1993). The function of Hox proteins as transcriptional regulators, established in cell culture studies or by analysis of their functional homology to HOM gene products in *Drosophila* embryos, allows us to postulate that Hox genes might play their developmental role by regulating a number of downstream target genes. However, apart from a well-documented number of auto- and cross-regulatory interactions within the Hox gene family (see Mavilio, 1993), the nature of the Hox targets in development remains largely unknown, and even the analysis of the knockout mutants failed to provide indications of putative candidates. Indeed, the few non-Hox, putative targets emerged so far were all identified by *in vitro* studies based on binding of Hox proteins to regulatory elements and/or by transfection studies showing positive or negative Hox regulation of those elements (Jones *et al.*, 1992, 1993; Tomotsune *et al.*, 1993).

In this report, we provide evidence that the gene for a transcription factor required for cell type-specific expression of thyroid genes, i.e. TTF-1, is a candidate target for regulation by the HOXB3 gene. We show that the rat TTF-1 5′ genomic region, containing the promoter and regulatory elements required for thyroid-specific gene expression, is transactivated by co-transfection of an expression vector for HOXB3 in murine or human cell lines. Transactivation is mediated by two HOXB3 binding sites flanking at short distance the transcription start site. There is no apparent difference in the binding affinity of HOXB3 for the two sites *in vitro*, and transfection experiments with mutants of either site indicate that they are both necessary for transactivation of the TTF-1 promoter. Moreover, the two sites are able to bind nuclear factors from mouse embryonic tissues expressing HOXB3 and TTF-1, generating footprinting and gel-shift patterns indistinguishable from those obtained with the HOXB3 protein expressed in eukaryotic cells.

Nuclear extracts from both mouse embryos and HOXB3-transfected cells give rise to multiple DNA–protein complexes in gel-shift experiments with both HOXB3 binding sites. This indicates that HOXB3 may be part of multiprotein complexes *in vivo*, or that it may be post-translationally modified in more than one form. The possibility that some of these complexes represent HOXB3 homodimeric or heterodimeric forms in association with other Hox proteins is considered unlikely, since Hox proteins have been generally reported to bind DNA as monomers (Affolter *et al.*, 1990), and the two oligonucleotides contain only one Hox binding site.

HOXB3 was able to upregulate the TTF-1 promoter by 4- to 6-fold, depending on the cell line. Although this may be considered an overall weak effect, it is in the order of magnitude of all the observed Hox-mediated transactivations, including both auto- and cross-regulations (Arcioni *et al.*, 1992; Pöpperl and Featherstone, 1992; Zappavigna *et al.*, 1994) and activation of other putative target genes (Jones *et al.*, 1992, 1993). Moreover, we cannot rule out that additional sequences located elsewhere in the TTF-1 gene, and/or transcription factors not present in the examined cell background, are necessary for full activity of the promoter, and consequently also for a stronger induction by HOXB3. Interestingly, a number of different Hox proteins, which were reported previously to transactivate other promoters containing Hox binding sites (Zappavigna *et al.*, 1991, 1994; Arcioni *et al.*, 1992; Pöpperl and Featherstone, 1992), failed to activate the TTF-1 promoter, indicating that transactivation of TTF-1 is not a general feature of homeodomain proteins, at least in our cell system. In particular, activation of the TTF-1 promoter was specifically exerted by HOXB3, and not by the paralogous HOXD3 or by Hox proteins belonging to more 5′ paralogy groups, expressed more posteriorly in embryonic development and in regions which do not overlap those where TTF-1 is expressed. Expression of *Hoxb-3*, as well as the other group 3 proteins *Hoxa-3* and *Hoxd-3*, precedes that of TTF-1 and is observed in the same structures in which TTF-1 is first activated and then constitutively expressed, including the thyroid primordia, the thyroid gland and the lungs (Gaunt, 1987; Guazzi *et al.*, 1990; Hunt *et al.*, 1991a,b; Lazzaro *et al.*, 1991; Sham *et al.*, 1992). Our data therefore indicate that at

least one of the Hox genes that are co-expressed with TTF-1 has a potential regulatory role on its tissue-specific promoter. Although largely circumstantial, these data allow us to postulate that HOXB3 might participate in the regulatory events leading to activation and/or maintenance of TTF-1 expression in thyroid and lung tissues.

It has been suggested previously that TTF-1, the expression of which is first activated in the lung/thyroid primordia before overt thyroid differentiation and expression of thyroid-specific genes, might have a developmental role in the actual determination of the thyroid gland (Lazzaro *et al.*, 1991). Therefore, HOXB3 might play a direct role in the determination of the lung/thyroid primordia by participating in the activation of the TTF-1 gene. The selective activity of HOXB3 on the TTF-1 promoter, as suggested by our data, might be the molecular basis for providing anterior identity to developing structures determined by the activation of the TTF-1 gene. It is suggestive that knockout mutants of the paralogous gene *Hoxa-3* are characterized by reduced, even if not completely abolished, thyroid development (Chisaka and Capecchi, 1991), whereas genetic ablation of *Hoxd-3*, which has no effect on the TTF-1 promoter in transfected cells, has no consequence on the development of the thyroid gland (Condie and Capecchi, 1993). Although defective thyroid development might be secondary and not directly determined by the absence of *Hoxa-3* function, it seems reasonable to suggest that the absence of *Hoxa-3* might have a profound effect on the development of the thyroid gland by causing a reduced or abnormal expression of TTF-1. More direct evidence is obviously required to prove this assumption. Nevertheless, although we were so far unable to clone a full-length cDNA for HOXA3 to test the activity of this protein on the TTF-1 promoter, and a *Hoxb-3* knockout mutant is not yet available, it can be proposed that at least two of the group 3 Hox genes might be involved in the developmental regulation of the TTF-1 gene.

TTF-1 is the first transcription factor to be identified as a candidate target of Hox gene function. However, it has been reported recently that *Hoxa-5* (formerly *Hox-1.3*), homologous to the *Drosophila* Sex comb reduced (*Scr*) HOM gene, is able to functionally substitute *Scr* by regulating the *Scr* target gene *fork head* (*fkh*) (Weigel *et al.*, 1990; Zhao *et al.*, 1993). Interestingly, *fkh* has been reported to be the *Drosophila* counterpart of the mammalian transcriptional regulator HNF-3α, originally purified from rat liver nuclear extracts (Lai *et al.*, 1991). *Hoxa-5* and HNF-3α are expressed in the same structures in embryonic development, suggesting that *Hoxa-5* may participate in liver organogenesis by activating HNF-3α (Zhao *et al.*, 1993). Taken together, these findings indicate that tissue-specific transcription factors are likely targets of Hox gene regulation in development, thereby representing a hierarchical link between positional specification, provided by the regionally-restricted Hox gene products, and development of specific tissues or organs.

Materials and methods

Reporter plasmids and cDNA expression vectors

To obtain the pTTF-1–Luc Δ-862 construct, a 965 bp *Bam*HI–*Ssp*I fragment of the rat TTF-1 genomic upstream sequence, extending to position +103 in its 3′ end, was cloned into the *Bam*HI and *Hin*dIII (filled-in) sites of the pBS–Luc α luciferase reporter plasmid. The pTTF-1–Luc Δ-142 construct was obtained by internal deletion of a 722 bp *Bam*HI–*Hin*dIII fragment. The BS1 and BS2 sites on the TTF-1 promoter were mutagenized by PCR with the mismatched primers BS1m (5′-TCAAGCCCGGGTAGGCGGACTCGGT-3′), and BS2m (5′-CAGCATGTAAGCTATCAGTCTCGGGCAAGATG-3′), to obtain the pTTF-1–Luc Δ-142 BS1m and pTTF-1–Luc Δ-142 BS2m constructs. The pHOXC5–GH Δ-1223 has been described previously as pHOX3D Δ-1223–GH (Arcioni *et al.*, 1992).

The 1935 bp *Eco*RI–*Eco*RI HOXB3 cDNA (Acampora *et al.*, 1989) was cloned at the *Eco*RI site of the mammalian expression vector pSG-5 (Green *et al.*, 1988). To increase the level of expression of the HOXB3 transfected protein, most of the 3′-UT cDNA sequence was deleted by digesting and filling-in the *Kpn*I site (nucleotide 1764 of the cDNA sequence) and *Bgl*II site (in the vector) and then recircularizing the plasmid. A mismatched primer around the AUG codon, containing an *Eco*RI cloning site, and a primer around the unique *Pst*I site at position 527 of the cDNA sequence, were used to amplify a 180 bp fragment, which was then used to replace the cDNA *Eco*RI–*Pst*I fragment containing a long 5′-UT sequence. The mismatched 5′ primer (5′-GGAATTCCACCATGCAGAAAGCCACC-3′) contained also a consensus sequence for translation (Kozak, 1987). The pSG–HOXD3 construct was generated by cloning into the *Bam*HI site of pSG-5 the full-length cDNA of HOXD3, a gift from L.Cianetti (Istituto Superiore di Sanità, Roma). An *Apa*I cDNA fragment containing only the coding region of the HOXD4 gene (Mavilio *et al.*, 1986) was cloned into a Klenow-filled *Eco*RI site of the pSG-5 vector to obtain the pSG–HOXD4 plasmid. The pSG–HOXC8 construct was generated by cloning into the *Bam*HI and *Bgl*II sites of pSG-5 the complete coding sequence of HOXC8, reconstructed from a partial cDNA clone and genomic sequences (kindly provided by M.Pannese). The pCT–HOXC6, pCT–HOXC5 and pSG–*Hoxd-8* constructs have been described previously as pCT–H3C, pCT–H3D and pSG–H4.3, respectively (Arcioni *et al.*, 1992). The pSG–*Hoxa-4* construct has been described previously (Buettner *et al.*, 1991).

Cell culture and transfection assays

Cell lines were cultured in high-glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum and 20 mM HEPES buffer (all from Gibco) in 5% CO_2 in air humidified atmosphere. HeLa, NIH3T3 and COS-7 cells were transfected at 50% confluency by calcium phosphate co-precipitation with 5–8 μg of expression vector, 10 μg of reporter plasmid, and 1–2 μg of pRSV–β-gal (Bonnerot *et al.*, 1987) as internal control, per 10 cm dish. Cells were harvested 48–60 h after transfection, lysed by three to four freeze–thaw cycles and centrifuged at 4°C in a microfuge for 5 min. Luciferase assays were carried out as described (de Wet *et al.*, 1987). β-Galactosidase assays were performed according to Sambrook *et al.* (1989). Expression of hGH was evaluated as amount of protein secreted in the medium 48 h after transfection by a commercial RIA (Nichols Institute). All transfections were carried out in duplicate batches, using at least two different DNA preparations of each expression construct.

Preparation of nuclear and bacterial extracts

Crude nuclear extracts were prepared from COS-7 cells and embryonic tissues as described by Dignam *et al.* (1983). Nuclear extracts from transfected cells were prepared 48 h after transfection by the same procedure. Embryonic nuclear extracts were obtained from a pool of 12.5 day-old mouse embryos from which forehead and the posterior (abdominal) part of the body had been surgically removed. Frozen embryos (10–15) were ground to a powder with a pestle, resuspended in 10 mM HEPES pH 7.9, 1.5 M $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF and 0.2 mM EGTA, passed through an 18 G syringe needle and then processed as for cell lines.

A 220 bp *Sma*I–*Apa*I fragment of the HOXB3 cDNA sequence (encompassing the entire homeodomain and extending 12 bp on the 5′ side and 30 bp on the 3′ side) was cloned into the *Bam*HI site of the pT7.7 prokaryotic expression vector (Studier *et al.*, 1990). This construct, named pT7.7–HOXB3 HD, and a control pT7.7 vector, were used to transform the BL 21+ bacterial strain, containing the pLysE plasmid (Studier *et al.*, 1990). Bacterial cells were grown and the recombinant proteins induced and extracted as described by Zappavigna *et al.* (1991), except for the lysozyme reaction which is not required for pLysE-carrying bacterial cells.

Gel retardation and DNase I footprinting assays

The pTTF-1–Luc Δ-862 plasmid was digested with *Hin*dIII, end-labeled with a standard T4 polynucleotide kinase reaction (Sambrook *et al.*, 1989), digested at the *Xba*I site in the luciferase gene and purified by PAGE. This 323 bp fragment, containing the entire TTF-1 promoter sequence plus 80 bp from the luciferase gene, was used in DNase I footprinting experiments which were carried out as described previously (Arcioni *et al.*, 1992) in the presence of 3–6 μl of cell nuclear extracts or 0.1–0.5 μl of bacterial extracts. ^{32}P-labeled double-stranded oligonucleotides BS-1 (5′-CATTCAAGCCAATTAAGGCGGACTC-3′) and BS-2 (5′-TGTAAGCTAATTATCTCGGGCAAGA-3′) were used as probes in gel retardation assays as described previously (Arcioni *et al.*, 1992). Double-stranded oligonucleotides containing either a HOX consensus binding site (5′-AAATATCAATTAAATCTTAATTATAA-3′) or an unrelated NF-1 binding site were used in competition experiments.

HOXB3 polyclonal antisera

HOXB3-containing bacterial extract (200–400 μg) was run on a preparative, 12.5% SDS–polyacrylamide gel. The HOXB3-containing band was excised from the gel, crushed and used to inject rabbits and raise polyclonal antisera by standard methods (Harlow and Lane, 1988). 1 μl of immune or pre-immune rabbit serum was added to the DNA binding reaction in the gel retardation assays.

Acknowledgements

We wish to thank Luciano Cianetti for the gift of the HOXD3 cDNA, Michael A.Tainsky for the HOXA4 expression vector, Giovanna Viale for the anti-HOXB3 antiserum and Valeria Ganazzoli for the expert assistance in the manuscript preparation. This work was supported by grants from the Italian Association of Cancer Research (AIRC), the CNR Projects 'Ingegneria Genetica' and 'ACRO', and the European Community Project 'BIOTECH'.

References

Acampora,D. *et al.* (1989) *Nucleic Acids Res.*, **17**, 10385–10402.
Affolter,M., Schier,A. and Gehring,W.J. (1990) *Curr. Opin. Cell Biol.*, **2**, 485–495.
Appel,B. and Sakonju,S. (1993) *EMBO J.*, **3**, 1099–1109.
Arcioni,L., Simeone,A., Guazzi,S., Zappavigna,V., Boncinelli,E. and Mavilio,F. (1992) *EMBO J.*, **11**, 265–277.
Bonnerot,C., Rocancourt,D., Briand,P., Grimber,G., Nicolas,J.F. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 6795–6799.
Buettner,R., Yim,S.O., Hong,Y.S., Boncinelli,E. and Tainsky,M.A. (1991) *Mol. Cell. Biol.*, **11**, 3573–3583.
Capovilla,M., Brandt,M. and Botas,J. (1994) *Cell*, **76**, 461–475.
Chisaka,O. and Capecchi,M.R. (1991) *Nature*, **350**, 473–479.
Civitareale,D., Lonigro,R., Sinclair,A. and Di Lauro,R. (1989) *EMBO J.*, **8**, 2537–2542.
Condie,B.G. and Capecchi,M.R. (1993) *Development*, **119**, 579–595.
Damante,G. and Di Lauro,R. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 5388–5392.
de Wet,J.R., Wood,K.V., DeLuca,M., Helsinki,D.R. and Subramani,S. (1987) *Mol. Cell. Biol.*, **7**, 725–737.
Desplan,C., Theis,J. and O'Farrell,P.H. (1988) *Cell*, **54**, 1081–1090.
Dignam,J.R., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
Duboule,D. (1992) *BioEssays*, **14**, 375–384.
Francis-Lang,H., Price,M., Polycarpou-Schwharz,M. and Di Lauro,R. (1992) *Mol. Cell. Biol.*, **12**, 576–588.
Gaunt,S.J. (1987) *Development*, **101**, 51–60.
Gehring,W.J. (1987) *Science*, **236**, 1245–1252.
Gehring,W.J., Müller,M., Affolter,M., Percival-Smith,A., Billeter,M., Qian,Y.Q., Otting,G. and Wüthrich,K. (1990) *Trends Genet.*, **6**, 323–329.
Gould,A.P. and White,R.A.H. (1992) *Development*, **116**, 1163–1174.
Gould,A.P., Brookman,J.J., Strutt,D.I. and White,R.A.H. (1990) *Nature*, **348**, 308–312.
Green,S., Issemann,I. and Sheer,E. (1988) *Nucleic Acids Res.*, **16**, 369–373.
Gruss,P. and Kessel,M. (1991) *Curr. Opin. Genet. Dev.*, **1**, 204–210.
Guazzi,S., Price,M., De Felice,M., Damante,G., Mattei,M.-G. and Di Lauro,R. (1990) *EMBO J.*, **9**, 3631–3639.
Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Hayashi,S. and Scott,M. (1990) *Cell*, **63**, 883–894.
Hinz,U., Wolk,A. and Renkavitzpohl,R. (1993) *Development*, **116**, 543.
Hoey,T. and Levine,M. (1988) *Nature*, **332**, 858–861.
Hunt,P. and Krumlauf,R. (1991) *Cell*, **66**, 1075–1078.
Hunt,P., Gulisano,M., Cook,M., Sham,M.H., Faiella,A., Wilkinson,D., Boncinelli,E. and Krumlauf,R. (1991a) *Nature*, **353**, 861–864.
Hunt,P., Wilkinson,D. and Krumlauf,R. (1991b) *Development*, **112**, 43–50.
Jaynes,J.B. and O'Farrell,P.H. (1988) *Nature*, **336**, 744–749.
Jones,F.S., Prediger,E.A., Bittner,D.A., De Robertis,E.M. and Edelman,G.M. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 2086–2090.
Jones,F.S., Holst,B.D., Minowa,O., De Robertis,E.M. and Edelman,G.M. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 6557–6561.
Kessel,M. and Gruss,P. (1991) *Cell*, **67**, 89–104.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8132.
Krasnow,M.A., Saffman,E.E., Kornfeld,K. and Hogness,D.S. (1989) *Cell*, **57**, 1031–1043.
Krumlauf,R. (1993) *Trends Genet.*, **9**, 106–112.
Lai,E., Prezioso,V.R., Tao,W., Chen,W.S. and Darnell,J.E.,Jr (1991) *Genes Dev.*, **5**, 416–427.
Lazzaro,D., Price,M., De Felice,M. and Di Lauro,R. (1991) *Development*, **113**, 1093–1104.
Mavilio,F. (1993) *Eur. J. Biochem.*, **212**, 273–288.
Mavilio,F. *et al.* (1986) *Nature*, **324**, 664–666.
McGinnis,W. and Krumlauf,R. (1992) *Cell*, **68**, 283–302.
Pöpperl,H. and Featherstone,M.S. (1992) *EMBO J.*, **11**, 3673–3680.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schubert,F.R., Nieselt-Struwe,K. and Gruss,P. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 143–147.
Sham,M.H., Hunt,P., Nonchev,S., Papalopulu,N., Graham,A., Boncinelli,E. and Krumlauf,R. (1992) *EMBO J.*, **11**, 1825–1836.
Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) *Methods Enzymol.*, **185**, 60–89.
Thali,M., Muller,M.M., DeLorenzi,P., Matthias,P. and Bienz,M. (1988) *Nature*, **336**, 598–600.
Tomotsune,D., Shoji,H., Wakamatsu,Y., Kondoh,H. and Takahashi,N. (1993) *Nature*, **365**, 69–72.
Wagner-Bernholz,J.T., Wilson,C., Gibson,G., Schuh,R. and Gehring,W.J. (1991) *Genes Dev.*, **5**, 2467–2480.
Weigel,D., Seifert,E., Reuter,D. and Jäckle,H. (1990) *EMBO J.*, **9**, 1199–1207.
Zappavigna,V., Renucci,A., Izpisua-Belmonte,J.C., Urier,G., Peschle,C. and Duboule,D. (1991) *EMBO J.*, **10**, 4177–4187.
Zappavigna,V., Sartori,D. and Mavilio,F. (1994) *Genes Dev.*, **8**, 732–744.
Zhao,J.J., Lazzarini,R.A. and Pick,L. (1993) *Genes Dev.*, **7**, 343–354.

Received on November 16, 1993; revised on April 29, 1994