Journal of Cellular Biochemistry 59:486–497 (1995)

rHox: A Homeobox Gene Expressed in Osteoblastic Cells

Yunshan Hu, Judith Flanagan, Denise P. Brennan, Hong Zhou, Kong Wah Ng, John A. Eisman, and Nigel A. Morrison

Garvan Institute of Medical Research, St. Vincent's Hospital, Sydney 2010, NSW, Australia (Y.H., J.F., D.P.B., J.A.E., N.A.M.); Department of Biotechnology, University of New South Wales, Kensington, Sydney 2052, Australia (Y.H., N.A.M.); and University of Melbourne, Department of Medicine, St. Vincent's Hospital, Victoria 3065, Australia (H.Z., K.W.N.)

Abstract Homeodomain proteins are characterized by a conserved domain with a helix-turn-helix motif. These proteins act as regulatory factors in tissue differentiation and proliferation. However, their role in the regulation of osteoblast differentiation is unknown. In this study we have identified and characterized a homeobox gene in osteoblast-like cells. This gene, termed rHox, was isolated from a cDNA library derived from rat osteoblast-like cells. The nucleotide sequence of the 1,375 base pair (bp) cDNA contains a noncoding leader sequence of 329 bp, a 735 bp open reading frame, and 312 bp of 3′ noncoding sequence. Sequence comparison demonstrates that rHox is identical to the mouse Pmx gene (also called MHox) at the amino acid level and 90% homologous at the nucleotide level. Both Southwestern blotting and gel shift analyses indicate that rHox has potential to bind both the collagen I α1 and the osteocalcin promoters. Transfection experiments using an rHox expression vector showed a strong repression of target promoter activity, regardless of whether the target promoters contained homeodomain binding response elements. These data suggest that rHox is a potent negative regulator of gene expression, although the specific role of rHox in bone gene regulation remains to be determined.

Key words: homeobox gene, rHox, rat, osteocalcin, collagen I α1

Homeobox genes were first identified in *Drosophila* by analysis of homeotic gene mutations [Ingham, 1988; Gehring, 1987]. These genes encode proteins containing a conserved 60 amino acid homeodomain with a helix-turn-helix motif [Scott et al., 1989; Laughon, 1991]. Homeodomain proteins bind with high affinity to A-T-rich elements containing an ATTA core sequence [Kessel et al., 1990; Catron et al., 1993]. Homeobox genes encode a large family of transcription regulators in both invertebrates and vertebrates [Gaunt, 1991; Krumlauf, 1992]. In *Drosophila*, the genes of Antennapedia/Ultrabithorax complexes play a major role in segment identity along the anteroposterior axis of the body [De Robertis et al., 1991; McGinnis et al., 1992]. In mammals, the *Hox* genes form four clusters, each on a different chromosome [Duboule et al., 1989], and are also involved in anteroposterior specification. In addition to the *Hox* genes, other homeobox genes have been identified in vertebrates which are also important in pattern formation and regulation of cell differentiation and proliferation during developmental processes. For instance, the chicken homeobox genes Msx-1 [Yokouchi et al., 1991] and Prx-1 [Nohno et al., 1993] are expressed in the developing limb buds and contribute to limb bud formation. Since a mutation of the human Msx-2 gene causes autosomal dominant craniosynostosis, a developmental abnormality of the skull bones [Jabs et al., 1993], homeodomain proteins are good candidates for developmental and cellular regulators of bone. In addition, homeodomain proteins have been described as tissue-specific regulators in the pituitary [Castrillo et al., 1991], liver [Frain et al., 1989], and thyroid [Guazzi et al., 1994].

Osteoblasts play a crucial role in bone formation, bone turnover, and calcium homeostasis. Mature osteoblasts synthesize large amounts of collagen and several noncollagenous proteins, including osteonectin and osteocalcin, which

Received June 20, 1995; accepted July 21, 1995.

Address reprint requests to Nigel A. Morrison, Garvan Institute of Medical Research, St. Vincent's Hospital, Sydney, NSW 2010, Australia.

form the organic matrix of bone. Collagen type 1 represents more than 85% of the total collagen expressed by osteoblastic cells [Martin et al., 1993], and is necessary for normal bone formation. Osteocalcin, the major noncollagenous protein of bone, is expressed in a developmental tissue-specific manner [reviewed by Stein and Lian, 1993, 1995] and is associated with bone mineralization [Price et al., 1982; Pastoureau et al., 1993]. Osteocalcin may play a role in recruitment of bone cell precursors [Mundy et al., 1983; Glowacki et al., 1991; Chenu et al., 1994]. Thus, identification of genes controlling collagen type I and osteocalcin transcription should increase understanding of bone metabolism and bone formation. It was hypothesized that the engrailed response element represents a promiscuous homeodomain binding site capable of detecting homeodomain proteins, although no such target genes of bone regulated by homeodomain proteins are known. Using this approach a new homeobox gene, termed rHox, from rat osteoblast-like cells has been isolated. The protein encoded by this gene binds to both the collagen I α1 and osteocalcin promoters.

MATERIALS AND METHODS

Library Screening

A lambda gt 11 cDNA library derived from rat osteoblast-like cells (ROS 17/2.8) was used for expression screening as described by Vinson et al. [1988]. The library was plated at a density of 2×10^4 recombinants per dish, and 4×10^5 recombinants were screened. The *E. coli* Y1090 strain was transformed and incubated at 42°C until plaques just appeared. The lawn was then overlaid with an isopropyl thiogalactoside (IPTG)-soaked nitrocellulose membrane and incubated for a further 6 h at 37°C. Duplicate membranes were applied. The proteins on the filters were denatured in 6 M guanidine hydrochloride for 5 min at 4°C and renatured with decreasing concentrations of guanidine hydrochloride in binding buffer (25 mM HEPES, pH 7.9, 3 mM $MgCl_2$, 4 mM KCl, 1 mM DTT). These membranes were then incubated with 2×10^6 cpm/ml DNA probe in probing buffer [binding buffer, 0.25% (w/v) nonfat dry milk, 100 μg/ml denatured salmon sperm DNA] overnight at 4°C. The probe used for screening was a 12 bp oligonucleotide of the engrailed promoter homeodomain protein response element (TCAATTAATTGA) which was concatenated to a pentamer. Positive clones from the initial screening were radiolabeled for use as probes to rescreen the same rat osteosarcoma cDNA library, by standard DNA-DNA hybridization. After second and third round screening, a purified full-length positive clone was obtained.

DNA Sequencing and Sequence Analysis

Inserts from all positive lambda gt 11 plaques were subcloned into the plasmid pUC18 at the EcoR1 site. The sequences of the subcloned DNA inserts were determined by the dideoxynucleotide chain termination method [Sanger et al., 1977] using the T_7 polymerase sequencing system (Promega, Madison, WI). Both strands were sequenced with significant overlap.

rHox–Maltose Binding Protein (MBP) Fusion Protein Preparations

The open reading frame (ORF) of rHox was obtained by PCR amplification and cloned into the EcoR1 site of pMAL-c (New England Biolabs, MA) in both the sense and antisense orientations. TB-1 *E. coli* were transformed with these constructs, and the corresponding fusion proteins were induced with isopropyl thiogalactoside (IPTG) according to the manufacturer's instructions. The fusion protein was partially purified on an ion exchange column (DEAE-Sephadex, Whatman, NJ) by adsorption in lysis buffer [25 mM Tris.HCl, pH 8.0, 10 mM EDTA, 50 mM NaCl, 10% (w/v) glycerol, 2 mM DTT, 1 mM PMSF, 500 μg/ml lysozyme] and elution with increasing concentrations of NaCl from 100 mM to 300 mM in the same buffer. Fractions were examined for the fusion protein by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Southwestern Analysis of DNA–rHox Interactions

Southwestern blotting was carried out using the following method [Brennan et al., in preparation]. MBP–rHox crude fusion protein extracts were subjected to SDS-PAGE in 10% acrylamide and electroblotted onto nitrocellulose membranes overnight at 180 mA. The blots were incubated in blocking buffer (10 mM Tris.HCl, pH 7.4, 1 mM EDTA, 50 mM NaCl, 2 mM 2-mercaptoethanol, 10% nonfat dry milk, 0.1% Nonidet P-40) for 1–3 h, and probed with double-stranded oligonucleotide probes in blocking buffer overnight at room temperature. The oligonucleotide probes were end labeled, annealed, and concatamerized. For competition experiments, unlabeled double-stranded, concatamer-

ized oligonucleotides were used at 200-fold molar excess over the labeled probes. The oligonucleotides used were collagen I α1 box (Collbox), rat OC box (rOC box), mutant collagen I α1 box (Collbox-mut), mutant rOC box (rOCbox-mut), nonspecific oligonucleotide (NS), and human OC box (hOC box) (Table I). The DNA–protein complexes were visualized by autoradiography for 1–16 h at room temperature or at −70°C with intensifying screens.

Gel Mobility Shift Assay

The double-stranded oligonucleotides Collbox and rOC box (Table I) were end labeled with ^{32}P using T_4 DNA kinase, and purified from a 15% acrylamide gel. DNA–protein binding was carried out in a total volume of 40 μl, with 20 μg of partially purified rHox–MBP fusion protein from *E. coli* or ROS 17/2.8 nuclear extracts, 10 μg of double-stranded poly dI-dC, and 0.5–2 ng of labeled double-stranded oligonucleotide (10–50 × 10^3 dpm), and incubated for 10 min on ice. Protein–DNA complexes were separated in a 5% polyacrylamide gel with 0.5 × TBE buffer (25 mM Tris.HCl, pH 8.0, 20 mM boric acid, 0.5 mM EDTA) as described previously [Morrison et al., 1992], and shifted bands were visualized by autoradiography with Kodak X-ray film for 14–16 h at −70°C with intensifying screens.

Cell Culture

Rat osteosarcoma cells (ROS 17/2.8) were grown in Ham-F12 medium with 7.5% fetal bovine serum (FCS) (Commonwealth Serum Laboratory, Australia) and UMR 201 cells in α-modified Eagle's medium (α-MEM) with 10% FCS. NIH 3T3 fibroblast cells were cultured in Dulbecco's modified Eagle's medium (DMEM) containing 10% FCS. All cell lines were cultured as monolayers in flasks and maintained at 37°C either in a warm room or in a humidified incubator at 37°C with 5% CO_2. Medium was replaced every 2–3 days during expansion of cell lines.

TABLE I. Oligonucleotide Sequences Used as Response Elements in This Study*

Oligonucleotide name	Sequence
En	TCAATTAATTGA
Collbox	GGACTTCATCAATTATCCCAACCAT
rOC box	CATGACCCCCAATTAGTCCTGGCAGCA
Collbox-mut	GGACTTCATCctagATCCCAACCAT
rOCbox-mut	CATGACCCCCctagAGTCCTGGCAGCA
hOC box	GATGACCCCCAAaTAGCCCTGGCAGAT
NS	GACAGCCTGTATGCCTTGTACCTG

*Only the top strand in 5′–3′ orientation is shown. The engrailed homeodomain protein response element, collagen I α1 promoter, rat OC box, and human OC box are designated En, Collbox, rOC box, and hOC box, respectively. The oligonucleotides containing a mutated ATTA core sequence were termed Collbox-mut and rOCbox-mut corresponding to mutations of the Collbox and rOC box oligonucleotides. The mutated bases are indicated in lowercase letters. NS is a nonspecific oligonucleotide.

RNA Isolation and Northern Analysis

Expression of rHox in different rat osteoblastic cell lines was analyzed by Northern blot analysis. Some UMR 201 cells were treated with 10^{-6} M retinoic acid (RA) or 1 ng/ml transforming growth factor (TGF)-β1 for 24 h before RNA extraction. Total cellular RNA was prepared by the guanidinium isothiocyanate–cesium chloride method. RNA was separated in a 1.5% agarose–formaldehyde gel and transferred to nylon membranes (Hybond N+) [Sambrook et al., 1989]. The 1.3 kb rHox cDNA was random primer labeled with ^{32}P. The membranes were hybridized overnight in hybridization buffer as previously described [Zhou et al., 1991]. After washing in 2× SSPE (18 mM NaH_3PO_4, pH 7.4, 30 mM NaCl, 2 mM EDTA), 0.1% SDS at 42°C for 15 min, followed by 30 min in 1× SSPE, 0.1% SDS at 65°C and 15 min in 0.1× SSPE, 0.1% SDS at room temperature, the membranes were exposed to Kodak X-ray film XAR-5 without intensifying screens.

Transient Transfection in NIH 3T3 Cells

The open reading frame of rHox, obtained by PCR amplification, was cloned into the pRcCMV (Invitrogen Co.) eukaryotic expression vector at the XbaI site, in which expression is driven by the cytomegalovirus promoter. The engrailed homeodomain protein response element (En), rat OC box (rOC box), and collagen I α1 promoter (Collbox) oligonucleotides were inserted upstream of the thymidine kinase promoter in the plasmid pTKCAT. pOSCAT2 was constructed as described previously [Morrison et al., 1989]. pOSCAT17 is a minimal promoter consisting of 41 bp upstream of the osteocalcin transcription start site (−41–+34 bp). The pSV40CAT plasmid contains the simian virus 40 early promoter (−135–+61 bp) driving CAT gene expression. NIH 3T3 cells plated at a density of 2 × 10^6 per T150 flask were transfected by a calcium phosphate precipitation method

[Sambrook et al., 1989]. The amounts of pRcCMV–rHox expression vectors (sense, antisense) and empty vector controls were maintained constant with an equimolar amount of promoter–CAT constructs (15 μg) per flask plus 15 μg pRSV β-GAL [Gorman et al., 1982] as an internal control for transfection efficiency. Precipitates were left on the cells for 14–16 h or 24 h post-transfection, and then cells were subcultured into six-well plates with 2×10^5 cells/well. CAT assays were performed as previously described [Morrison et al., 1989]. Samples from each well were incubated with 1.6 μCi of [^{14}C] acetyl CoA (Dupont, Wilmington, DE). The CAT activity was quantitated by scintillation counting, and β-galactosidase activity was determined as described by Sambrook et al. [1989]. CAT activity was standardized to β-galactosidase levels to account for differences in transfection efficiency, and all experiments were performed in triplicate. In dose–response experiments 0.15–16.5 μg pRcCMV–rHox sense and antisense were cotransfected with 15 μg prOCbox–TKCAT or pOSCAT2. The exogenous DNA was kept constant by adding empty vector. CAT assays were performed as described above.

RESULTS

Isolation and Structure of rHox

Using the homeodomain protein binding site of the engrailed gene promoter (En) from *Drosophila* to screen the rat osteosarcoma (ROS 17/2.8) cDNA expression library, several clones were isolated and one was characterized in detail. This partial cDNA clone was sequenced and used to probe the same ROS 17/2.8 cDNA library. Five positive cDNA clones were picked. One of these clones, which was purified and sequenced, encoded a novel homeodomain protein, termed rHox. The complete nucleotide sequence (1,375 bp) is shown in Figure 1. The open reading frame consists of 735 bp (nucleotide position 329–1,063 bp). A consensus polyadenylation signal (AATAAA) was not present at the 3′ end of the cDNA. There are two ATG start codons in frame. According to Kozak [1984], the consensus sequence CCA/GCCATGG surrounding the translation start is essential for efficient translation. The first ATG meets this criterion and therefore is likely to be the initiation codon, which would yield a putative rHox protein of 245 amino acids. Sequence analysis of rHox shows 90% homology at the nucleotide level to mouse Pmx [Kern et al., 1994] or MHox [Cserjesi et al., 1992], identified recently as a transcription cofactor regulating muscle creatine kinase. Moreover, rHox and Pmx have identical amino acid sequences (Fig. 2). rHox is also highly related to chicken Prx-1, which is expressed in the developing limb [Nohono et al., 1993]. Most of the differences between rHox and Prx-1 are in the amino-terminal region (Fig. 2).

Northern Blot Analysis

Northern blot analysis of mRNA from the rat preosteoblast cell line UMR 201 and the more mature osteoblastic cell line ROS 17/2.8 probed with ^{32}P-labeled rHox cDNA showed three bands in the ROS 17/2.8 cell line (A, B, and C in Fig. 3) and two bands in the UMR 201 line (B and C in Fig. 3). These transcripts correspond to approximately 4.7 kb (A), 3.8 kb (B), and 3.2 kb (C) in ROS 17/2.8 cells. The transcript A, which corresponds to the size of the mouse Pmx transcript, is the most abundant in ROS 17/2.8 cell lines. The transcripts B and C are of approximately equal abundance in UMR 201 cells. Treatment with both retinoic acid and TGF β1, which increase the levels of procollagen I α1 and osteopontin [Zhou et al., 1993, 1994] in UMR 201 cells, alone or in combination did not increase the level of any of these transcripts in the UMR 201 cells.

DNA Binding Activity of rHox

Collagen I α1 and osteocalcin are the major proteins expressed in osteoblast cells. In the osteocalcin promoter, the osteocalcin box is thought to direct osteoblast cell type–specific transcriptional activity [Towler et al., 1994]. The bone-specific expression of the collagen I α1 promoter is reported to be regulated by a sequence element that has similarity to a homeodomain binding site (GGACTTCATCA*ATTA*TCCCAACCAT). This sequence was synthesized and termed Collbox. Sequence comparison of the collagen I α1 promoter element (Collbox) and the rOC box in the osteocalcin promoter (rOC box) with the engrailed homeodomain protein response element (En) shows that they all contain a motif, ATTA, which is the core sequence of homeodomain protein binding sites.

In order to test whether the rHox protein can bind these bone-specific promoter motifs, the rHox ORF (amino acids 1–245) was expressed as a 77 kD fusion protein with maltose binding protein (MBP) using the pMAL-c system. Southwestern analysis was used to examine the bind-

```
CTTTTTTTTTTTTTCTTTCTGATTTAAGCTTAGCGTGTTTGTTTGTTTGGTTTGGTTTTT          60

AAACTTTTTTTGGTGTGGATTATCTCTCTGGACCACGCCGGACTTGGCTTCAGCGAACGA          120

ACCGAAGCTGGGAGAAAGTTACTTGAGCTGAAAGTTCTGCACTTCAGAAGGGGGTTTCTC          180

TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCTCCACTCCCCCCTCTTTCTTCCC            240

CACTCGGCTCCTCTCCCCCCTCACGCCCACAGCGTTTGGTGTTGATTCGAGCGTTAAGAG          300

GGGGGTGGGTGGGATTGGAGGGAAGACCATGACCTCCAGCTACGGGCACGTTCTGGAGCG          360
                              M  T  S  S  Y  G  H  V  L  E  R

GCAACCGGCTCTGGGCGGCCGCTTGGATAGCCCTGGCAACCTCGACACCCTGCAGGCGAA          420
  Q  P  A  L  G  G  R  L  D  S  P  G  N  L  D  T  L  Q  A  K

AAAGAACTTCTCCGTCAGTCACCTGCTAGACCTGGAGGAGGCCGGGGACATGGTGGCGGC          480
  K  N  F  S  V  S  H  L  L  D  L  E  E  A  G  D  M  V  A  A

ACAGGCGGACGAAAGCGTGGGCGAGGCGGGCCGGAGCCTGCTGGAGTCACCGGGACTGAC          540
  Q  A  D  E  S  V  G  E  A  G  R  S  L  L  E  S  P  G  L  T

CAGTGGCAGCGACACCCCTCAGCAGGACAATGATCAGCTGAACTCTGAGGAGAAGAAGAA          600
  S  G  S  D  T  P  Q  Q  D  N  D  Q  L  N  S  E  E  K  K  K

GAGAAAGCAACGGAGAAACAGGACAACCTTCAATAGCAGCCAGCTGCAGGCCTTGGAGCG          660
  R  K  Q  R  R  N  R  T  T  F  N  S  S  Q  L  Q  A  L  E  R

TGTCTTCGAGAGGACCCATTACCCGGATGCTTTTGTACGGGAAGATCTTGCACGTCGGGT          720
  V  F  E  R  T  H  Y  P  D  A  F  V  R  E  D  L  A  R  R  V

GAACCTCACGGAGGCCAGAGTGCAGGTGTGGTTTCAGAACCGAAGAGCCAAGTTCCGCAG          780
  N  L  T  E  A  R  V  Q  V  W  F  Q  N  R  R  A  K  F  R  R

GAATGAGCGAGCCATGCTGGCCAATAAAAACGCTTCTCTCCTCAAGTCCTACTCAGGAGA          840
  N  E  R  A  M  L  A  N  K  N  A  S  L  L  K  S  Y  S  G  D

CGTGACTGCCGTGGAACAACCCATTGTACCTCGTCCTGCTCCCAGACCCACCGATTATCT          900
  V  T  A  V  E  Q  P  I  V  P  R  P  A  P  R  P  T  D  Y  L

CTCCTGGGGGACAGCCTCTCCGTACAGCGCCATGGCTACTTATTCTGCCACATGTGCCAA          960
  S  W  G  T  A  S  P  Y  S  A  M  A  T  Y  S  A  T  C  A  N

CAATAGCCCTGCGCAGGGTATCAACATGGCCAACAGCATTGCCAACCTGAGACTGAAGGC          1020
  N  S  P  A  Q  G  I  N  M  A  N  S  I  A  N  L  R  L  K  A

CAAGGAATATAGTTTACAGAGGAACCAGGTGCCAACAGTCAACTGAGGAAAAAAATAATT          1080
  K  E  Y  S  L  Q  R  N  Q  V  P  T  V  N  *

AAACAGGCCTAAGAAGAAATCAAAAACCATAAGACACCTATCCTGTTCTGTCATTTCTTC          1140

ATCCGCTGGAAAAAAAAGATAAAAGCAAACAAACAAAGCCAGAACTAAAATATTGGGACC          1200

ATGGCGGAGAAAAGCAGGAGAGGAACAAAATGAAAATTAGTTAACAAATGTTTCTTCTTC          1260

CTCTTGGATACCATCACCATTTGTTGTGTGTGTATTTTTTTTCCTTCTACCCATGCTTTG          1320

CTAAGTATACTTCAGGCTTCTTCTATTAGGCTCAACCCACCCATTCCTATAATTG                1375
```

Fig. 1. Nucleotide and putative amino acid sequence of rHox. The in frame stop codon is shown by an asterisk.

ing potential of this protein for different DNA sequences. The fusion protein bound not only to the engrailed promoter homeodomain protein response element (En) but also to the rat collagen I α1 promoter (−1,763/−1,742 bp) (Collbox) and the OC box of the rat osteocalcin promoter (−100/−83 bp) (rOC box) (Fig. 4a, lanes 6–8). Empty vector and antisense controls showed no binding (Fig. 4a, lanes 1–5). Binding to the Collbox appeared more intense than to the rOC box. In competition experiments, a 200-fold molar excess of the unlabeled Collbox and rOC box oligonucleotides efficiently inhibited binding to the rHox–MBP fusion protein by their respective labeled probes (Fig. 4b, c, lane 2), indicating that the binding was specific. In addition, competition experiments were carried out with oligonucleotides lacking the ATTA motif (nonspecific oligonucleotide, Collbox-mut, rOCbox-mut, and hOC box oligonucleotides; see Table I). The nonspecific oligonucleotide did not inhibit rHox–MBP binding to the Collbox (Fig. 4b, lane 4). Similarly, the Collbox-mut and rOCbox-mut oligonucleotides, which were mutated at the first three nucleotides of the ATTA core sequence, did not compete for rHox–MBP binding to the collagen I α1 promoter (Collbox) or rOC box, respectively (Fig. 4b, c, lane 3). The

```
  1  MTSSYGHVLERQPALGGRLDSPGNLDTLQAKKNFSVSHLL        rHox
     ----------------------------------------        Pmx
     -A---A-AM---AL-PA---G-AG--N-------------        Prx

 41  DLEEAGDMVAAQADESVGEAGRSLLESPGLTSGSDTPQQD        rHox
     ----------------------------------------        Pmx
     ------------G--GG--P-------------------        Prx

 81  NDQLNSEEKKKRKQRRNRTTFNSSQLQALERVFERTHYPD        rHox
     ----------------------------------------        Pmx
     ----------------------------------------        Prx

121  AFVREDLARRVNLTEARVQVWFQNRRAKFRRNERAMLANK        rHox
     ----------------------------------------        Pmx
     ---------------------------------*---S-        Prx

161  NASLLKSYSGDVTAVEQPIVPRPAPRPTDYLSWGTASPYS        rHox
     ----------------------------------------        Pmx
     ----------------------------------------        Prx

201  AMATYSATCANNSPAQGINMANSIANLRLKAKEYSLQRNQ        rHox
     ----------------------------------------        Pmx
     ------T--T-A-----M-----------------------        Prx

241  VPTVN                                           rHox
     -----                                           Pmx
     -----                                           Prx
```

Fig. 2. Comparison of the entire amino acid sequence of rHox with those of mouse Pmx and chicken Prx-1, which are closely related homeodomain proteins. Conserved amino acids are dashed. The homeodomain is underlined. A gap is shown by * to improve sequence alignment.

hOC box, which contained an A instead of a T at the second nucleotide of the ATTA core sequence, did not compete for rHox–MBP binding (Fig. 4c, lane 4). Thus the binding of rHox to promoters is both sequence and species specific.

To further examine the DNA binding properties of the rHox protein, gel mobility shift assays were performed using the partially purified rHox–MBP fusion protein. Specific DNA–protein complexes were formed with both the ^{32}P-labeled Collbox and the rOC box (Fig. 5a). DNA–protein complexes were not observed with the homeodomain protein response element of the engrailed promoter (En), perhaps due to the small size of the fragment (12 bp or its affinity). The rHox fusion protein bound more strongly to the Collbox than the rOC box, supporting the result from the Southwestern analysis. These data suggest that rHox has a greater affinity for the collagen promoter element than the osteocalcin promoter element. The collagen I α1 promoter response element (Collbox) was used in gel shift analysis with ROS 17/2.8 nuclear extracts to assess whether any ROS 17/2.8 nuclear proteins were capable of binding to this sequence. At least seven DNA–protein complexes were formed, of which the three larger complexes were competed completely by a 200-fold molar excess of the unlabeled Collbox (Fig. 5b).

Regulatory Function of rHox in NIH 3T3 Cells

Transient transfection assays were used to assess whether rHox can regulate promoter activity via the rOC box, En, and Collbox elements. We initially examined pCollbox–TKCAT promoter activity in NIH 3T3 fibroblast cells. pRcCMV–rHox sense, antisense orientation expression vectors and empty vector controls were cotransfected with pCollbox–TKCAT into NIH 3T3 cells. The transfection of the pRcCMV–rHox sense construct resulted in approximately 60% repression of CAT activity from the pCollbox–TKCAT compared to the empty vector (Fig. 6a). Similarly, repression was observed in cotransfection experiments with pEn–TKCAT and prOCbox–TKCAT (sense constructs) (Fig. 6a). Cotransfection of the pRcCMV–rHox sense construct with pTKCAT (which does not contain a known homeodomain protein binding site) resulted in similar significant repression of CAT activity (Fig. 6a). These results suggest that rHox can repress promoter activity in the absence of homeodomain DNA binding sites. To

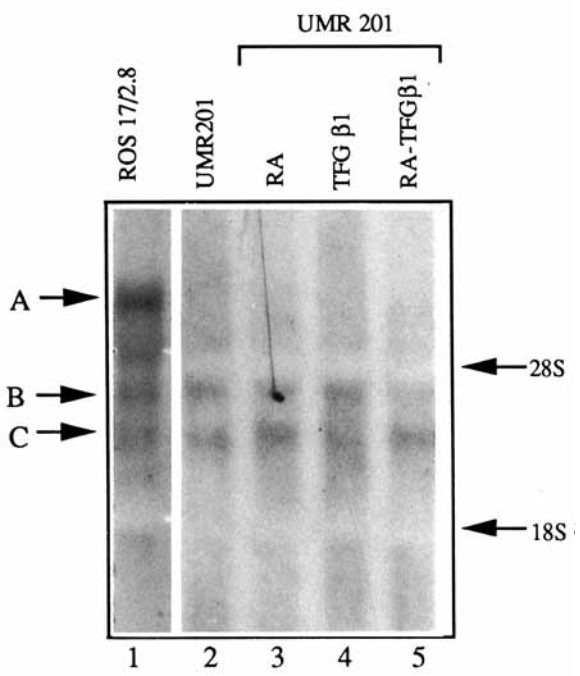

Fig. 3. Northern analysis of rHox expression in bone cells. Twenty micrograms of total cellular RNA was loaded in each lane. Blots were probed with a ^{32}P-labeled 1.3 kb rHox cDNA and exposed to X-ray film for 1 week. The position of 28S and 18S rRNAs are indicated by *arrows*. *Lanes 1* and *2* show RNA from untreated ROS 17/2.8 and UMR 201 cells, respectively. *Lanes 3–5* show RNA from UMR 201 cells treated for 24 h with the following agents: lane 3, retinoic acid (10^{-6} M); lane 4, TGF-β1 (1 ng/ml); lane 5, retinoic acid and TGF-β1 simultaneously. The results show a major 4.7 kb (**A**) transcript in ROS 17/2.8 cells which is absent from UMR 201 cells. UMR 201 cells had transcripts of 3.8 kb (**B**) and 3.2 kb (**C**), which were not regulated by treatment with either retinoic acid or TGF-β1, either separately or together.

Probe	Engrailed oligonucleotide						Collbox	rOC box
Fusion protein Induction (IPTG)	–	–	–	+	+	+	+	+
Sense			+			+	+	+
Antisense		+			+			
Empty vector	+			+				

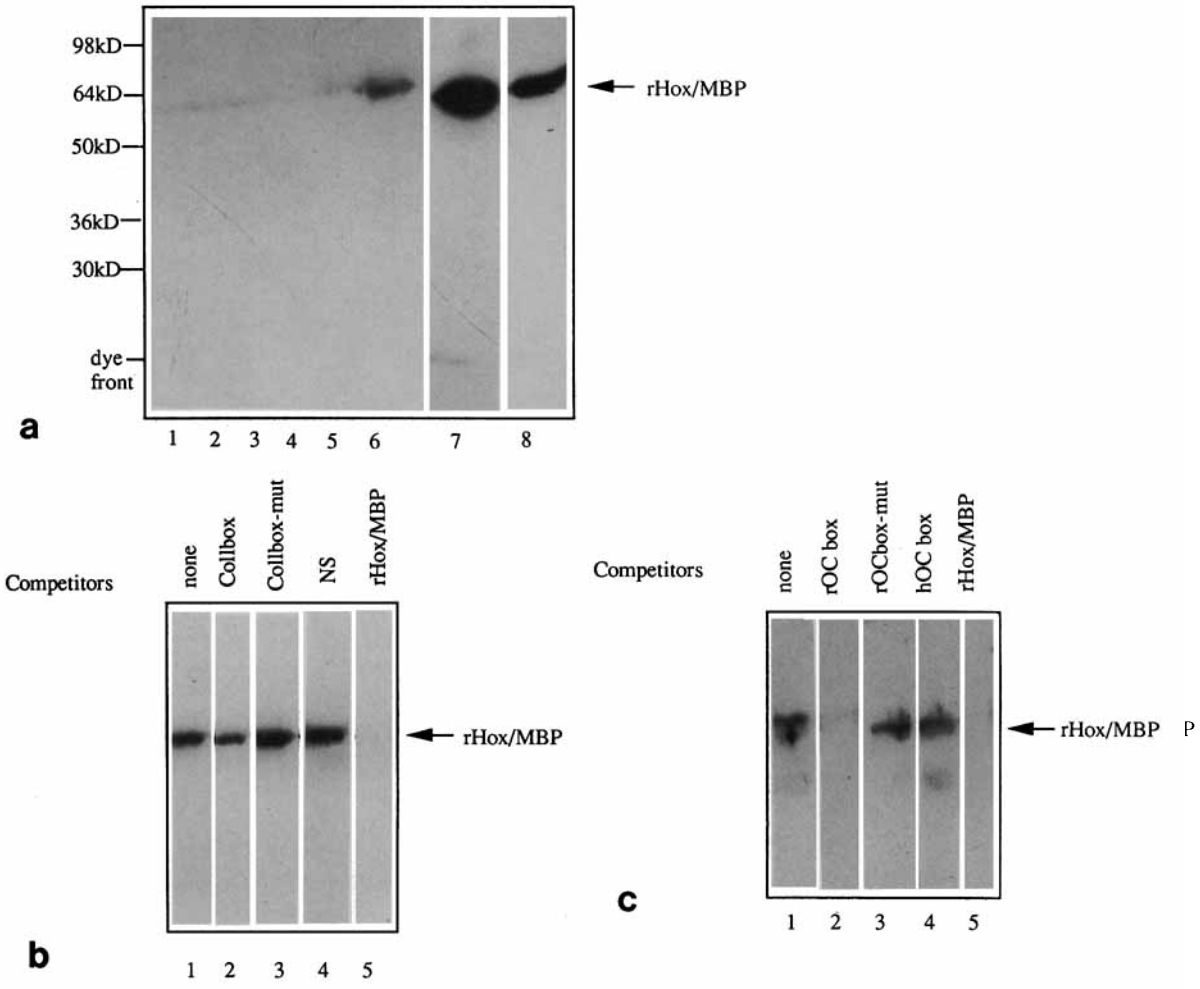

Fig. 4. Southwestern analysis. The rHox–MBP fusion protein was expressed in *E. coli* using the pMAL-c system. The crude protein extracts were separated by SDS-PAGE in 10% acrylamide and electroblotted onto nitrocellulose membranes. The oligonucleotide probes were end labeled, annealed and concatamerized. **a:** The rHox–MBP was probed with the homeodomain protein binding site of the engrailed promoter (En), the rat collagen I α1 promoter (−1,763/−1,742 bp) (Collbox), and the rat osteocalcin promoter (−100/−83 bp) (rOC box) (*lanes 6–8*). The negative controls (uninduced empty vector, sense and antisense; induced empty vector and antisense) show no binding (*lanes 1–5*, respectively) to the engrailed promoter element (En). The same results were observed in experiments with the collagen I α1 promoter (Collbox) and rOC box probe (controls not shown). **b:** Competition experiment. Unlabeled Collbox, Collbox-mut, NS oligonucleotide, or rHox–MBP fusion protein were coincubated with the labeled Collbox at a 200-fold excess to compete for binding of the rHox–MBP fusion protein on the blot. **c:** Competition experiment. Unlabelled rOC box, rOCbox-mut, hOC box oligonucleotides, and the rHox–MBP fusion protein were added at a 200-fold excess to the labeled rOC box probe to compete for binding to the rHox–MBP fusion protein on the blot. The migration position of the rHox–MBP fusion protein is indicated by an arrow. The protein–DNA complexes were visualized by autoradiography.

confirm this finding, we carried out cotransfection assays of pRcCMV–rHox with pSV40CAT, which contains the SV40 early promoter; pOSCAT2, which consists of 2 kb of osteocalcin promoter; and pOSCAT17, which contains a minimal promoter containing 41 bp upstream of the osteocalcin transcription start site. pOSCAT2 contains the human osteocalcin box and a substantial number of upstream regulatory elements, while pOSCAT17 contains only the TATA box of the osteocalcin promoter. These three constructs were strongly repressed by pRcCMV–rHox transfection (Fig. 6b). The pRcCMV–rHox antisense orientation vector had little effect on CAT activity, except in one transfection experiment (prOCbox–TKCAT) which probably is an outlier effect (Fig. 6a).

In a total of 17 different transfections in NIH 3T3 fibroblasts, the pRcCMV–rHox expression vector demonstrated repression of target promoter activity (6.6 ± 3.2-fold, $P = 0.007$). Repression of promoter activity was not dependent on the promoter examined, and even the minimal pOSCAT17 was equally repressed, suggesting that the rHox repression relates either to an effect on TATA box function or to a global effect on cell activity. In ROS 17/2.8 cells the rHox expression vector repressed pTKCAT in a similar manner (data not shown). In analysis of Msx-1, Catron et al. [1995] showed that repression by Msx-1 did not require a homeodomain response element in the target gene sequence.

Since it is possible that rHox could have a biphasic effect on target promoters we examined the dose–response by cotransfection of 0.15–16.5 μg pRcCMV–rHox (sense construct) into NIH 3T3 cells with prOCbox–TKCAT and pOSCAT2, respectively. Controls in this experiment were antisense constructs. The results with both promoter constructs revealed that the CAT activity decreased as the amount of rHox expression vector increased (Fig. 6c). Under the conditions used no evidence for biphasic responses was observed. Repression was evident when data were plotted as CAT activity (Fig. 6c) or ex-

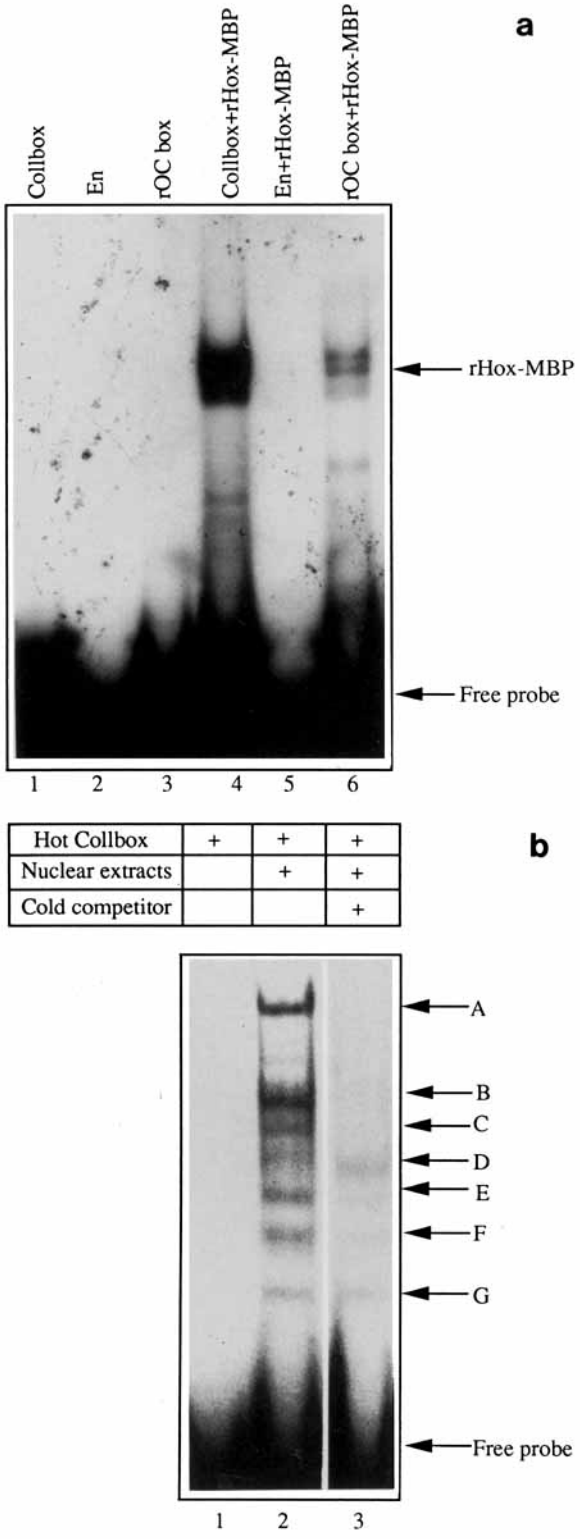

Fig. 5. **a:** Gel mobility shift analysis demonstrating binding of the rHox–MBP fusion protein to oligonucleotide probes. In this experiment the relative binding affinities of the rHox–MBP fusion protein for the Collbox, the engrailed element (En), and the rOC box are compared. *Lanes 1–3* are Collbox, En, and rOC box free probe, respectively. *Lanes 4–6* show the interaction between the rHox–MBP fusion protein and the labeled Collbox, En, and rOC box oligonucleotides, respectively. Clearly, the rHox–MBP fusion protein binds most strongly to the Collbox sequence. Binding to the En sequence was not observed. The DNA–rHox fusion protein complexes are indicated by an arrow. **b:** Nuclear proteins from ROS 17/2.8 osteoblastic cells bind to the collagen I α1 promoter (Collbox) sequence. Seven DNA–protein complexes are indicated by **A–G.** Competition (lane 3) was with 200-fold excess of unlabeled Collbox.

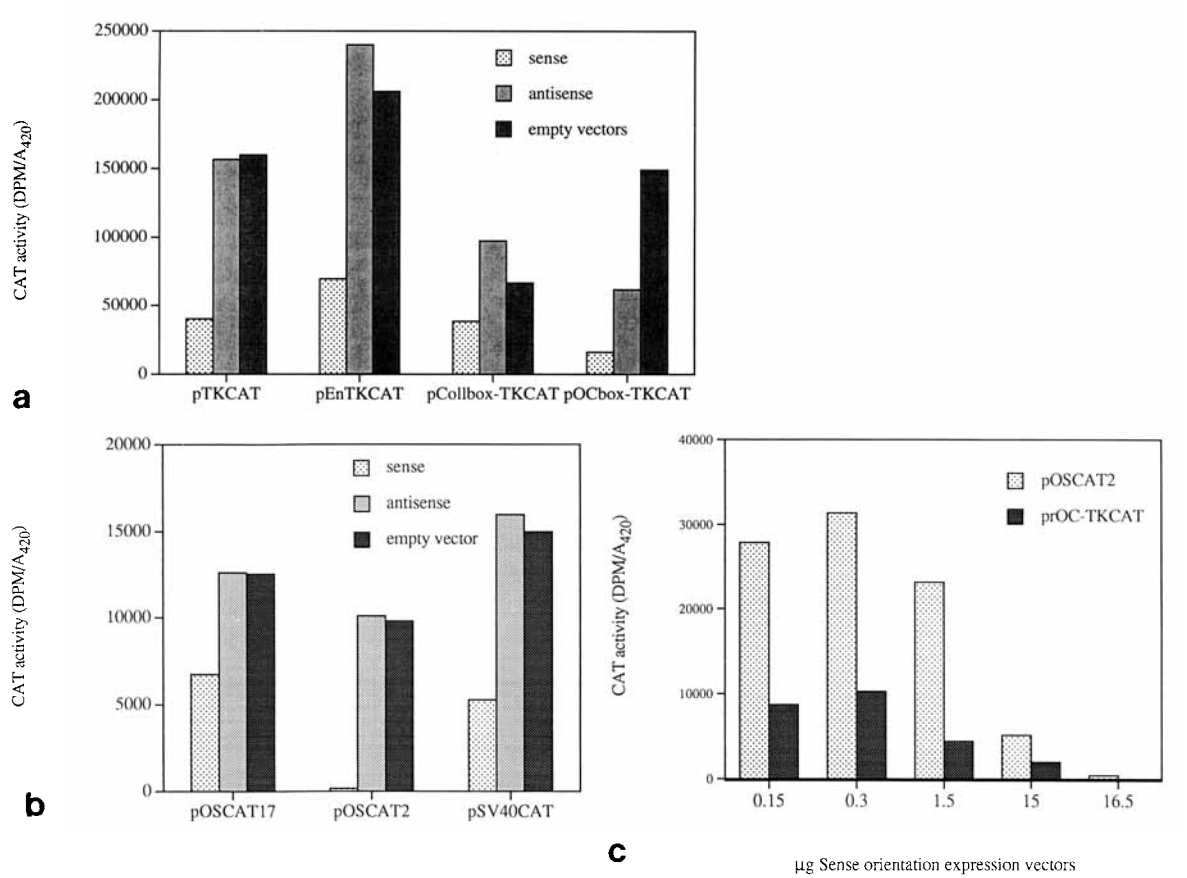

Fig. 6. Regulation of target gene activity by rHox determined by transfection of expression vectors. pRcCMV–rHox was transfected into NIH 3T3 cells in both the sense and antisense orientations with (**a**) pTKCAT, pEnTKCAT, pCollbox–TKCAT, and prOCbox–TKCAT, or (**b**) pOSCAT17, pOSCAT2, or pSV40CAT. Cells were assayed for CAT activity 36 h after transfection. The results are presented as the average of triplicate measurements in the same experiment, corrected for transfection efficiency. The results showed that the pRcCMV–rHox sense orientation vector repressed promoter–CAT activity compared to the antisense orientation or the empty vector controls. **c**: Dose–response experiments were performed by cotransfection of 0.15–16.5 μg pRcCMV–rHox (sense construct) with pOSCAT2 or prOCbox–TKCAT into NIH 3T3 cells. The results showed that CAT activity decreased as the amount of pRcCMV–rHox (sense construct) increased.

pressed as a ratio of the antisense control (not shown).

DISCUSSION

We have described the isolation of a homeobox gene, termed rHox, using a homeodomain protein binding site to probe a rat osteosarcoma cell line (ROS 17/2.8) expression cDNA library. Sequence analysis indicates that rHox is a member of the Pmx/Prx family, which belongs to the paired subfamily. Members of the paired gene family encode a paired box, a DNA binding motif with or without a homeodomain [Treisman et al., 1991]. The major difference between rHox/Pmx/Prx and a typical paired family member is a glutamine substitution for serine at position 50 in the homeodomain, a position that may contribute to DNA binding specificity between the classes of homeobox genes [Treisman et al., 1989]. In the antennapedia class of homeodomain proteins, glutamine is at this position; therefore, the binding specificity of rHox/Pmx may be similar to that of antennapedia. Sequence comparison also suggests that rHox is most likely the rat homolog of mouse Pmx.

Studies with the engrailed promoter indicated that homeodomain proteins can bind to an A-T-rich consensus sequence which contains an ATTA core motif essential for DNA binding activity [Hanes et al., 1991]. At least four murine homeodomain proteins—Hox 7.1, Hox 1.5, En-1 [Catron et al., 1993], and Pmx, in addition to some *Drosophila* homeodomain proteins, such as Antennapedia and Bicoid [Hanes et al., 1991]—can recognize this A-T-rich binding site. The particular significance of homeodomain proteins for bone development is indicated by the fact that a mutation of the human Msx-2 gene

results in craniosynostosis [Jab et al., 1993]. However, few target genes of the vertebrate homeodomain proteins have been identified. Towler et al. [1994] and Hoffmann et al. [1994] have identified an A-T-rich element or homeodomain protein binding site in the rat OC box region of the osteocalcin promoter. When this homeodomain binding site was mutated, the osteocalcin promoter basal activity in ROS 17/2.8 cells was decreased [Towler et al., 1994; Hoffmann et al., 1994]. Similarly, in the collagen I α1 promoter, a sequence similar to a homeodomain protein binding site in the upstream region (−2,003/−1,668 bp) contributes to collagen I α1 expression in vitro and in vivo [Kronenberg et al., 1993; Bogkanovic et al., 1994]. Our data show that rHox binds to both the osteocalcin and collagen I α1 promoters, suggesting that these genes may be targets for transcriptional regulation by rHox in bone. Consistent with this conclusion is the fact that rHox, collagen I α1, and osteocalcin are all expressed in the osteoblastic cell line ROS 17/2.8. The 4.7 kb transcript of rHox is expressed in the osteoblastic ROS 17/2.8, but not in the preosteoblastic UMR 201 cells (which do not express osteocalcin). Expression of the 4.7 kb transcript in mature osteoblastic cells may be due to differential splicing of rHox mRNA.

The apparent higher affinity of the rHox–MBP fusion protein for the Collbox collagen promoter sequence compared to the rOC box suggests that the rHox–DNA interaction is dependent not only on the ATTA core motif, but is also influenced by sequences surrounding that core sequence. This result agrees with Hanes and Brent [1991], who found that the homeodomain binding site is composed of two subsites: a common core element and specificity-determining bases that lie 3′ to it.

The binding of homeodomain proteins to a promoter motif suggests direct transcriptional regulation. Towler et al. [1994] reported that transfected Msx-2 was capable of repressing osteocalcin promoter activity in MC3T3E1 osteoblastic cells. We observed similar repression of pOSCAT2 when rHox was cotransfected into NIH 3T3 cells (Fig. 6b). However, the data on rHox activity in NIH 3T3 is not consistent with a specific, selective repression of those promoters carrying ATTA motifs. We tested nonbone-related promoters in the same assay (pSV40CAT, etc.) and noted a generalized repression of transfected promoter activity under rHox cotransfection. The quantities of expression vector used were below the level where squelching is a reasonable explanation. rHox may therefore have a global effect on cellular function, leading to decreased promoter activity. Elucidating the molecular basis of this intriguing activity requires further research. However, since Msx-2 also represses promoter activity, it may be a feature of homeodomain genes to negatively control osteocalcin gene expression. Catron et al. [1995] recently suggested that the mechanism of Msx-1 repression may be mediated through interactions with one component of the general transcription complex. Examples of homeodomain proteins interacting with other transcriptional mechanisms include human Phox-1 [Grueneberg et al., 1992], which is involved in formation of the serum response factor complex, and Prx-1, which may be an intracellular selector determining cell fate in response to growth and differentiation factors [Nohno et al., 1993]. In our study, the results of transient transfection show that rHox repressed all the promoter–CAT constructs used, suggesting that rHox may be a nonspecific repressor interacting with general transcription mechanisms.

Further studies are required to understand the regulatory role of rHox. In all, the demonstration in this study that the rHox protein binds to homeodomain binding sites in the rat collagen I α1 and osteocalcin promoters suggests that it may play a role in controlling bone gene expression. At least seven DNA–protein complexes were observed in gel shift analysis between the collagen I α1 promoter element (Collbox) and ROS 17/2.8 osteoblastic cell nuclear proteins. This suggests that greater complexity of interactions are to be expected as this system is investigated further.

ACKNOWLEDGMENTS

Y. Hu was supported by the Garvan Research Foundation. We thank Ms. Elizabeth Buttrose for preparing and maintaining cell lines. This research was supported by the Australian National Health and Medical Research Council. We thank Dr. Bill Kalionis and Mr. Joe Badolato for advice on screening methods.

REFERENCES

Bogdanovic Z, Bedalov A, Krebsbach PH, Pavlin D, Woody CO, Clark SH, Thomas HF, Rowe DW, Kream BE, Lichtler AC (1994): Upstream regulatory elements necessary for expression of the rat COL 1A1 promoter in transgenic mice. J Bone Miner Res 9:285–292.

Castrillo JL, Theill LE, Karin M (1991): Function of the homeodomain protein GHF1 in pituitary cell proliferation. Science 253:197–199.

Catron KM, Iler N, Abate C (1993): Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. Mol Cell Biol 13:2354–2365.

Catron KM, Zhang H, Marshall SC, Inostroza JA, Wilson JM, Abate C (1995): Transcriptional repression by Msx-1 does not require homeodomain DNA-binding sites. Mol Cell Biol 15:861–871.

Chenu C, Colucci S, Grano M, Zigrino P, Barattolo R, Zambonin G, Baldini N, Vergnaud P, Delmas PD, Zallone AZ (1994): Osteocalcin induces chemotaxis, secretion of matrix proteins, and calcium-mediated intracellular signalling in human osteoclast-like cells. J Cell Biol 127:1149–1158.

Cserjesi P, Lilly B, Bryson L, Wang Y, Sassoon DA, Olson EN (1992): MHox: A mesodermally restricted homeodomain protein that binds an essential site in the muscle creatine kinase enhancer. Development 115:1087–1101.

De Robertis EM, Morita EA, Cho KW (1991): Gradient fields and homeobox genes. Development 112:669–678.

Duboule D, Dolle P (1989): The structural and functional organization of the murine HOX gene family resembles that of *Drosophila* homeotic genes. EMBO J 8:1497–1505.

Frain M, Swart G, Monaci P, Nicosia A, Stampfli S, Frank R, Cortese R (1989): The liver-specific transcription factor LF-B1 contains a highly diverged homeobox DNA binding domain. Cell 59:145–157.

Gaunt, SJ (1991): Expression patterns of mouse Hox genes: Clues to an understanding of developmental and evolutionary strategies. Bioessays 13:505–513.

Gehring WJ (1987): Homeo boxes in the study of development. Science 236:1245–1252.

Glowacki J, Rey C, Glimcher MJ, Cox KA, Lian JB (1991): A role for osteocalcin in osteoclast differentiation. J Cell Biochem 45:292–302.

Gorman CM, Merlino GT, Willingham MC, Pastan I, Howard BH (1982): The Rous sarcoma virus long terminal repeat is a strong promoter when introduced into a variety of eukaryotic cells by DNA-mediated transfection. Proc Natl Acad Sci USA 79:6777–6781.

Grueneberg DA, Natesan S, Alexandre C, Gilman MZ (1992): Human and *Drosophila* homeodomain proteins that enhance the DNA-binding activities of serum response factor. Science 257:1089–1095.

Guazzi S, Lonigro R, Pintonello L, Boncinelli E, Di Lauro R, Mavilio F (1994): The thyroid transcription factor-1 gene is a candidate target for regulation by Hox proteins. EMBO J 13:3339–3347.

Hanes SD, Brent R (1991): A genetic model for interaction of the homeodomain recognition helix with DNA. Science 251:426–430.

Hoffmann HM, Catron KM, van Wijnen AJ, McCabe LR, Lian JB, Stein GS, Stein JL (1994): Transcriptional control of the tissue-specific, developmentally regulated osteocalcin gene requires a binding motif for the Msx family of homeodomain proteins. Proc Natl Acad Sci USA 91:12887–12891.

Ingham PW (1988): The molecular genetics of embryonic pattern formation in *Drosophila*. Nature 335:25–34.

Jabs EW, Muller U, Li X, Ma L, Luo W, Haworth IS, Klisak I, Sparkes R, Warman ML, Mulliken JB, Snead ML, Maxson R (1993): A mutation in the homeodomain of the human MSX2 gene in a family affected with autosomal dominant craniosynostosis. Cell 75:443–450.

Kern MJ, Argao EA, Birkenmeier EH, Rowe LB, Potter SS (1994): Genomic organization and chromosome localization of the murine homeobox gene Pmx. Genomics 19:334–340.

Kessel M, Gruss P (1990): Murine developmental control genes. Science 249:374–379.

Kozak M (1984): Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acid Res. 12:857–872.

Kronenberg MS, Bedalov A, Salvatori R, Lichtler AC (1993): Detection of a differentially expressed DNA binding activity in the rat α1 (I) collagen promoter. J Bone Miner Res 8:163.

Krumlauf R (1992): Evolution of the vertebrate Hox homeobox genes. Bioessays 14:245–252.

Laughon A (1991): DNA binding specificity of homeodomains. Biochemistry 30:11357–11367.

Martin TJ, Findlay DM, Heath JK, Ng KW (1993): Osteoblasts: differentiation and function. "Physiology and pharmacology of bone" Mundy GR, Martin TT, eds. Springer-Verlag publish. pp 149–173.

McGinnis W, Krumlauf R (1992): Homeobox genes and axial patterning. Cell 68:1497–1505.

Morrison NA, Eisman JA (1992): Role of the negative glucocorticoid regulatory element in glucocorticoid repression of the human osteocalcin gene. J Bone Miner Res 8:969–975.

Morrison NA, Shine J, Fragonas J, Verkest V, McMenemy LM, Eisman JA (1989): 1,25-Dihydroxyvitamin D-responsive element and glucocorticoid repression in the osteocalcin gene. Science 246:1158–1161.

Mundy GR, Poser JW (1983): Chemotactic activity of the gamma-carboxyglutamic acid containing protein in bone. Calcif Tissue Int 35:164–168.

Nohno T, Koyama E, Myokai F, Taniguchi S, Ohuchi H, Saito T, Noji S (1993): A chicken homeobox gene related to Drosophila paired is predominantly expressed in the developing limb. Dev Biol 158:254–264.

Pastoureau P, Vergnaud P, Meunier PJ, Delmas PD (1993): Osteopenia and bone remodeling abnormalities in warfarin-treated lambs. J Bone Miner Res 8:1417–1425.

Price PA, Williamson MK, Haba T, Dell RB, Jee WS (1982): Excessive mineralization with growth plate closure in rats on chronic warfarin treatment. Proc Natl Acad Sci USA 79:7734–7738.

Sambrook J, Fritsch EF, Maniatis T (1989): "Molecular cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbour, NY: Cold Spring Harbor Laboratory.

Sanger F, Nicklen S, Coulson AR (1977): DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467.

Scott MP, Tamkun JW, Hartzell GW (1989): The structure and function of the homeodomain. Biochim Biophys Acta 989:25–48.

Stein GS, Lian JB (1993): Molecular mechanisms mediating proliferation–differentiation interrelationships during progressive development of the osteoblast phenotype. Endocr Rev 14:424–442.

Stein GS, Lian JB (1995): Molecular mechanisms mediating proliferation–differentiation interrelationships during progressive development of the osteoblast phenotype: Update 1995. Endocr Rev Monographs 4:290–297.

Towler DA, Bennett CD, Rodan GA (1994): Activity of the rat osteocalcin basal promoter in osteoblastic cells is dependent upon homeodomain and CP 1 binding motifs. Mol Endocrinol 8:614–624.

Treisman J, Gonezy P, Vashishtha M, Harris E, Desplan C (1989): A single amino acid can determine the DNA binding specificity of homeodomain proteins. Cell 59:553–562.

Treisman J, Harris E, Desplan C (1991): The paired box encodes a second DNA-binding domain in the paired homeodomain protein. Genes Dev 5:594–604.

Vinson CR, LaMarco KL, Johnson PF, Landschuloz WH, McKnight SL (1988): In situ deletion of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev 2:801–806.

Yokouchi Y, Ohsugi K, Sasaki H, Kuroiwa A (1991): Chicken homeobox gene Msx-1: Structure, expression in limb buds and effect of retinoic acid. Development 113:431–444.

Zhou H, Hammonds RG Jr, Findlay DM, Fuller PJ, Martin TJ, Ng KW (1991): Retinoic acid modulation of mRNA levels in malignant, nontransformed, and immortalized osteoblasts. J Bone Miner Res 6:767–777.

Zhou H, Hammonds RG Jr, Findlay DM, Martin TJ, Ng KW (1993): Differential effects of TGFβ1 and BMP4 on gene expression and differentiated function of preosteoblasts. J Cell Physiol 155:112–119.

Zhou H, Manji SS, Findlay MD, Martin TJ, Heath JK, Ng KW (1994): Novel action of retinoic acid. Stabilization of newly synthesized alkaline phosphatase transcripts. J Biol Chem 269:22433–22439.