The murine *Hox-2.4* promoter contains a functional octamer motif

Fried Zwartkruis, Truus Hoeijmakers, Jacqueline Deschamps and Frits Meijlink*
Hubrecht Laboratory, Netherlands Institute for Developmental Biology, Uppsalalaan 8,
3584 CT Utrecht, The Netherlands

Received January 9, 1992; Revised and Accepted March 11, 1992

ABSTRACT

Like other HOX genes the murine *Hox-2.4* gene is thought to be involved in regional specification along the antero-posterior axis. In addition it has been reported to have oncogenic potential. We studied expression *Hox-2.4* in murine EC cells and determined the transcription start site. Studies of DNA-protein interactions in the promoter region showed that the *Hox-2.4* promoter contains a CCAAT box and a perfect octamer motif, which is capable of binding Oct-factors. Cotransfection of Oct expression vectors influences the transcriptional activity of the promoter, suggesting that the Oct-gene family may be involved in regulating HOX genes.

INTRODUCTION

The vertebrate HOX genes (1) constitute a family of related genes that are thought to play a role in regional specification along the anteroposterior (A-P) axis during embryogenesis (2,3). This idea is not only based on the finding that HOX genes are the mouse homologues of the *Drosophila* homeotic genes (4,5), discovered as genes necessary for the specification of the correct segment identity in the fruitfly (6), but is supported by experimental evidence as well. The homology between the murine HOX genes and the *Drosophila* homeotic genes is apparent from the similarities in nucleotide sequence of their DNA-binding domain, the homeobox, and is strengthened by their conserved order in clusters on the chromosome. Consistent with the idea that HOX genes are involved in specification of positional information along the A-P axis, it has been found that each HOX gene has its specific anterior boundary of expression during embryogenesis, which is colinear with the order of the HOX genes on the chromosome (4,5,7). More direct evidence comes from studies in which HOX gene expression has been altered. In *Xenopus* overexpression of *Xhox3* (8) and interference with expression of *Xlhbox1* (9) during embryogenesis leads to morphogenetic changes, that may be explained in terms of alterations in the anteroposterior identity of cells. Moreover, *Xenopus* animal cap cells overexpressing *Xlhbox6* have the capacity to induce a second anteroposterior axis if transplanted into embryos (10).

Overexpression of *Hox-1.1* in the mouse leads to craniofacial abnormalities (11), and to defects in cervical vertebrae that are reminiscent of homeotic variations in *Drosophila* (12). The presence of two defective *Hox-1.5* or *Hox-1.6* genes in mice, obtained after homologous recombination of these genes in embryo derived stem cells leads to more complex, regionally restricted developmental defects. Disruption of *Hox-1.5* leads to the absence of a thymus and parathyroid gland and additional malformations in the throat region (13), while mice having no functional *Hox-1.6* gene miss certain cranial nerves and ganglia and have malformed inner ears and bones of the skull (14).

In *Drosophila* homeotic genes are part of a hierarchical network, that establishes the anterior-posterior axis. Higher in this hierarchy maternal effect, gap and segmentation genes are found, that in concert regulate homeotic genes (15). It is not known which transcription factors control HOX-gene expression in mouse, but Oct proteins (16, reviewed in 17 and 18) are among the candidates as many of them are present during development (19). To test this possibility we have performed a functional analysis of the octamer containing *Hox-2.4* promoter.

The murine *Hox-2.4* gene is located in the 5' region of the HOX-2 cluster on chromosome 11. In addition to specifying positional information, *Hox-2.4* has been shown to have oncogenic properties if overexpressed under certain conditions (20, 21). We have studied the expression of *Hox-2.4* in different embryocarcinoma (EC) cell lines in relation with Oct-factors present. Furthermore we show that Oct-factors binding to the *Hox-2.4* promoter can activate this promoter.

MATERIALS AND METHODS

Cell culture and DNA transfection

C1003 EC, P19 EC and MES-1 cells were grown in DMEM/F12 medium supplemented with 7.5% fetal calf serum in a 7.5% CO₂ incubator as described earlier (22). Cells were transfected by calcium phosphate precipitation (23) as follows: 10 µg of DNA was added to 250 µl Hepes-buffered saline (42 mM Hepes, 275 mM NaCl, 10 mM KCl, 1.4 mM Na₂PO₄ and 10 mM dextrose, pH 7.05). While mixing, an equal volume of 250 mM CaCl₂ was added, the precipitate was allowed to form at room temperature

* To whom correspondence should be addressed

for 25 minutes and was added to the medium of cells cultured in a 6 cm diameter culture dish. The next morning the medium containing the DNA was replaced by fresh medium and cells were cultured for one day before harvesting. In all cases 2.5 μ g of a plasmid in which the *E. coli lacZ* gene is under the control of the promoter from the hydroxy-3-methyl-glutaryl-coenzyme A reductase gene (24), was cotransfected and used as an internal standard for efficiency of transfection. Luciferase and β -Galactosidase activities were determined as described in (25).

For determining *Hox-2.4* promoter activity, 7.5 μ g luciferase reporter constructs was used. In cotransfection experiments 5 μ g reporter construct was combined with various amounts of Oct expression plasmids. The expression plasmids carrying cDNAs of different Oct-factors under the control of the CMV enhancer / TK promoter region have been described previously (26,27). The total amount of DNA was made up to 10 μ g by addition of pEVRF0, the expression plasmid that contains no cDNA. For each cell line expression of Oct-factors after transfection was tested by a band shift assay with oligonucleotides carrying the octamer motif from the *Hox-2.4* promoter and whole cell extracts.

Northern analysis and S1 mapping

Standard recombinant DNA procedures were followed according to (28). *Hox-2.4* sequences used have been published (29) and were derived from the genomic L23 clone (30). Northern analysis was performed as described in (22). A 350-bp *SstI-KpnI* fragment downstream of the coding region was used as a probe for *Hox-2.4*. For S1 analysis a 831 bp *HindIII-BglIII* fragment (−641 to +190), containing some sequences present in an *Hox-2.4* cDNA (29) was cloned into the *HindIII* *BamHI* sites of Bluescript KS+, giving pHBg850. Several DNA fragments were cloned into both M13mp18 and M13mp19. HE300 and HRs550 were made by inserting sequences between the *HindIII* and the *EcoRV* or the *RsaI* site into the *HindIII* and *SmaI* sites of M13. EBg520 was made by inserting an *EcoRV XbaI* fragment from pHBg850 into the same sites of M13. After isolation of single stranded DNA from M13 phages anti-sense as well as sense probes were made by elongation of an M13 universal primer with DNA polymerase I (Klenow fragment) in the presence of [α^{32} P]-dCTP. After restriction digestion of the double stranded product, single stranded probes were eluted from denaturing sequencing gels. S1 analysis was performed as described in (31).

Reporter gene constructs

As a reporter gene we used the firefly luciferase gene (32). All reporter constructs were based on the p19LUC vector (33), which contains two SV40 poly-adenylation sites that prevent readthrough from upstream cryptic promoters in front of the polylinker. Nucleotide coordinates given are relative to the transcription start site. pHBg850Luci was made by inserting sequences present between the *HindIII* site and the *SstI* (−641 to +190) site of pHBg850 (see Northern analysis and S1 mapping) into the same sites of the p19LUC polylinker. pEBg520Luci and pRBg300Luci were made by cloning sequences between the *EcoRV* site (at −313) or the *RsaI* site (at −91) respectively and the *SstI* site into the blunt-ended *SalI* site and the *SstI* site of p19LUC. For 3' deletions pHBg850 was partially digested with *AvaII*, treated with DNA polymerase I (Klenow fragment) and digested with *HindIII*. Insert sequences were cloned into the *HindIII* and *SmaI* site of bluescript KS+ (plasmid pHA700). Subsequently sequences between the *HindIII* site and the *SstI* polylinker site were subcloned into the corresponding sites of p19LUC

(=pHA700Luci). Transferring sequences between the *RsaI* site (at −91) and the *SstI* site from pHA700 resulted in pRA150Luci.

In vitro mutagenesis was done following the method of (34), using *E. coli ung[−] dut[−]* mutants to generate uracil containing M13 DNA which allows for biological selection of mutated DNA. The oligonucleotide used in this construction was AATGGATG-AACATGATGGT. After recloning the mutated M13 insert into Bluescript KS+, sequence analysis showed that no additional mutations were introduced. pHAmutLuci and pRAmutLuci were made in exactly the same way as their unmutated counterparts.

For construction of the SV40 enhancer containing pSD19LUC luciferase vector a 240 bp *HindIII NcoI* fragment was isolated from the pCDX vector (35), made blunt using Klenow fragment and ligated to *BamHI* linkers. After removal of excess *BamHI* linkers, this fragment was inserted in the *BamHI* site downstream of the luciferase gene of p19LUC, giving pSD19LUC. Subsequently *HindIII-SstI* fragments from *Hox-2.4*-luciferase constructs were recloned into the same sites of pSD19LUC.

Band shift assays

Nuclear extracts from cultured cells and dissected embryos were as described in (36). Band shift assays were performed as described previously (37). Sequences of oligonucleotides containing the octamer motif of the *Hox-2.4* promoter are gatc-CCACCATCATTTCATCCATT and its complement gatcA-ATGGATGCAAATGATGGT. Sequences of the *c-abl* oligonucleotides were GGAAGGTCGGTGGAAATGCAAATTCCTCTC and its complementary strand GAGAGGAATTTGCAT-TTCCACCGACCTTCC. CCAAT-boxes containing oligonucleotides used, were based on sequences present in the murine *junD* promoter (gatcCCGTCGGCCAATCGGA and its complementary strand gatcTCCGATTGGCCGACG), the human IGF II gene (gatcCCCATTTGGCGCGGGCGCGAGGA and its complementary strand gatcTCCTCGCGCCCGCGCCAATGGG) and the NF-1 site from the adenovirus origin of replication (ttaTATTGGCTTCAATCCAAAA and its complementary strand taaTTTTGGATTGAAGCCAATA). For whole cell extracts of transfected cells, cells growing in a 20-cm² culture dish were washed with TBS (Tris-buffered saline), scraped in 1 ml TBS and centrifuged for 30 seconds at room temperature. After removal of the supernatant, cells were resuspended in 60 μ l WCE buffer (400 mM KCl, 1 mM EDTA, 10% (v/v) glycerol, 10 mM dithiothreitol, 0.1 mM PMSF, 10 μ g/ml leupeptin, 1 μ M pepstatin, 20 mM Hepes, pH 7.9) and immediately frozen in liquid nitrogen. Just before use cells were disrupted by three freeze-thaw cycles on ice and spun for 5 minutes at 4°C. For band shift assays 1 to 4 μ l was used.

RESULTS

Expression of *Hox-2.4* in embryocarcinoma cells

The murine *Hox-2.4* gene was originally cloned by Hart *et al.* (38) and independently in our lab (30). Analysis of genomic sequences (29,39,40) and cDNAs (29) revealed the presence of two exons of which the 3' one contained an *Antennapedia* type homeobox. During embryogenesis a major 2.2-kb and a minor 1.4-kb *Hox-2.4* transcript can be detected. *In situ* hybridization shows that expression occurs in the spinal cord, part of the *myelencephalon* and in somitic and intermediate mesoderm (5; J. Deschamps *et al.* in preparation).

EC cells are the stem cells of teratocarcinomas, that resemble in many respects the inner cell mass cells of blastocysts. Therefore,

they have been used widely as a model system to study early embryonic events. Because many HOX-genes are known to be expressed in embryocarcinoma cells after treatment with retinoic acid (RA), we tested for expression of *Hox-2.4* in two EC cell lines, that differentially respond to RA treatment. C1003 EC and P19 EC cells were grown in monolayer without or with RA (10^{-6} M) and total RNA from these cells was subjected to Northern analysis. In the presence of RA C1003 EC cells differentiate rapidly (2 days) into mainly neuron-like derivatives (22). RA treatment of P19 EC cells grown in monolayer results in the differentiation to a mixed population of fibroblast-like cells (41) and cells with endodermal characteristics (42). As expected no *Hox-2.4* expression is observed in EC cell lines before RA treatment (Figure 1). High levels of *Hox-2.4* mRNA are found in C1003 EC cells after culturing them in the presence of RA for 24 hours. In P19 EC cells *Hox-2.4* is clearly induced after culturing them for five days in the presence of RA, but the level of expression is much lower as compared to that of C1003 EC cells cultured for one day in the presence of RA. The lower level in P19 EC cells appears to be due to a slower induction of *Hox-2.4*, and not to activation being transient, since induction after two days is even lower. The length of the observed transcripts in EC cells was approximately 2.2 and 1.4 kb, in accordance with the mRNA length reported for murine embryos (38).

Determination of the *Hox-2.4* transcription start site

To determine the transcription start site of *Hox-2.4* we cloned a 520-bp *EcoRV*-*Bgl*III fragment from the upstream region into M13mp18 and M13mp19 (see Figure 2A). The *Bgl*III site was reported to be present in cDNAs of *Hox-2.4* (29). Body-labeled single stranded probes from this M13-construct were isolated from denaturing sequence gels after elongation of the M13 universal primer using Klenow fragment of DNA polymerase I and restriction digestion. Subsequently this probe was used for S1-mapping to analyse RNAs from P19 EC cells cultured in the absence and presence of RA, from embryonic neural tube, embryonic body and from adult kidney (see Figure 2B). No specific protection was seen with mRNA from P19 EC cells cultured in the absence of RA. In all other cases we observed a 206-bp S1-protected fragment with an intensity that correlates with the level of *Hox-2.4* mRNA detected by Northern analysis.

Figure 1. Northern analysis of *Hox-2.4* expression in C1003 and P19 EC cells. Cells were grown in monolayer in the absence of RA or in its presence (10^{-6} M) for the number of days indicated at the top of the figure. In each lane 20 μ g of total RNA was loaded. The probe used to detect *Hox-2.4* transcripts was a 350 bp *Sst*I *Kpn*I fragment located downstream of the *Hox-2.4* coding sequence, that was subcloned from the L23 genomic clone (30; for sequence data, see 29).

The length of this fragment indicates a transcription start site 408 bp upstream of the translation start site. Because no consensus splice is found in the genomic sequence at or near this position, we believe that the protected band is not caused by an intron/exon boundary but reflects a true transcription start site, that is used both in EC cells and in murine tissues. Moreover, the predicted distance between the start and the polyadenylation signal of the mRNA is somewhat more than 2000 bp as would be expected for a 2.2 kb mRNA. Several extra bands at a higher position were also observed, but the intensity of these bands varies strongly between several experiments. To find out whether the higher bands resulted from specific protection of longer transcripts, S1 protection assays with probes containing more upstream sequences (see Figure 2A) were performed. Because no protected fragments were detected with these probes (data not shown) we assume that the higher bands in figure 2B represent partial breakdown products, which are only formed if the EBg520 probe is hybridized to RNA.

Promoter structure of the *Hox-2.4* gene

Sequences upstream of *Hox-2.4* were found to be identical to those published before (29). Sequences surrounding the transcription start site are shown in figure 2C. The transcription

Figure 2. A. Schematic representation of the genomic region encompassing the transcription start site of *Hox-2.4*. Transcribed sequences are shown as a bar. Stripes underneath the scheme indicate the probes used for S1 analysis. B. S1 analysis of mRNA from undifferentiated P19 EC cells (lane 2), P19 EC cells grown in the presence of RA (10^{-6} M) for five days (lane 3), neural tube of 13.5 day murine embryos (lane 4), 13.5 day embryonic body without neural tube (lane 5) and adult kidney (lane 6). In lane 1 only carrier tRNA was present. As a probe antisense EBg520 was used. If a sense probe was used no protection was found (not shown). C. Nucleotide sequence of the region surrounding the transcription start site. Sequences are identical to those published before (29). The transcription start site is indicated by an arrow, the CCAAT-box is underlined and the octamer motif double underlined.

start site is located in between two very pyrimidine-rich regions. The promoter does not contain a typical TATA-box. A CCAAT-box, (reported before (29)) is present 80 bp upstream of the transcription start site, a situation encountered in many other genes. In addition to the CCAAT-box we detected a perfect octamer motif 50 bp upstream of the cap site: ATGCAAATNA. This sequence is a *cis*-acting element in many promoters and enhancers and can be bound by a number of transcription factors (for a review see (17,18)), including the cell ubiquitous factor Oct-1 and tissue specific factors like the products of the *Oct-2* and *Oct-4* genes. The presence of these factors in a cell and the context of the octamer motif within an enhancer or promoter determine the effect of the octamer motif on transcription. We

therefore set out to investigate the possible contribution of this motif to the complex expression pattern of *Hox-2.4*.

Binding of nuclear proteins to the *Hox-2.4* promoter

A 150-bp fragment (called RA150: -91 to +61 nt relative to the transcription start site) containing the CCAAT-box and the octamer motif was used as probe in band shift assays to analyse DNA-protein interactions in the putative promoter region of *Hox-2.4*. Incubation of RA150 with nuclear extracts from undifferentiated P19 EC cells led to the appearance of four bands with intensities depending on the amount of nuclear extract used (Figure 3A, lanes 1 to 3). Band shift assays with nuclear extracts from undifferentiated C1003 or P19 EC cells showed a very similar pattern of retarded bands. Competition experiments with unlabeled probe showed that all four bands were caused by specific binding of proteins (data not shown). To find out which of the four bands resulted from binding by Oct-factors competition experiments were performed with double stranded oligonucleotides containing the octamer motif from the *c-abl* promoter but no additional sequence similarities to RA150 (kindly provided by Dr. D. Meijer, Rotterdam). Addition of these oligonucleotides to the reaction mixture inhibited the formation of the slowest and the fastest migrating DNA-protein complexes but did not affect the other two complexes (Figure 3A, lanes 4 to 6). It is known that Oct-1 and Oct-4 are the most abundant octamer binding proteins in undifferentiated EC cells, which makes it very plausible that the highest and lowest bands in figure 3A represent Oct-1 and Oct-4 binding respectively. When the 25 most 5' located basepairs, harbouring the CCAAT-box, were deleted from the RA150 probe by restriction digestion only two retarded band were observed in band shift assays (3B, lane 2). Therefore, these two bands must be caused by binding of Oct-factors to the intact octamer motif in the 5' deleted probe. We conclude that the middle

Figure 3. A. Band shift assay with the RA150 probe (-90 to +60 relative to the *Hox-2.4* transcription start site). Increasing amounts of P19 EC nuclear extract was used in lanes 1, 2 and 3 (0.5, 2 and 4 μ g respectively). In lanes 4, 5 and 6 (2 μ g P19 EC nuclear extract) addition of a 10-, 25- and 100-fold molar excess of an oligonucleotide containing the octamer motif of *c-abl* inhibits formation of the fastest and the slowest migrating DNA-complexes, showing that these complexes are formed by Oct-factors. Note that complex formation of Oct-1 increases much faster than that of Oct-4. These different kinetics explain the apparent discrepancy between panel 3C (lanes 4 and 5: 2 μ g), where Oct-1 is hardly visible and panel D (lanes 2 and 1: 4 μ g), where it is clearly seen. B. Band shift assay with RA150 (lane 1) and P19 nuclear extract (2 μ g) showing the formation of four complexes. Upon deletion of 29, 5' located bp of the RA150 probe (lane 2) only the highest and lowest persist. The 29, 5' located bp were deleted by BsmAI digestion of the RA150 probe, followed by treatment with DNA polymerase I (Klenow fragment). C. Band shift assay with the RA150 probe and nuclear extracts from undifferentiated and differentiated EC cells (2 μ g): lane 1: no protein; lane 2: P19 EC; lane 3: P19 EC grown in the presence of RA for five days; lane 4: C1003 EC cells; lane 5, 6: C1003 EC cells grown in the presence of RA for 24 and 48 hours respectively; lane 7: C1003 EC cells differentiated by means of serum deprivation (four days). Note that an Oct-factor (marked with an arrowhead) is induced during differentiation of C1003 EC cells by means of RA treatment and by serum deprivation. The identity of this Oct-factor is unknown, but it is most likely one of the neuron-specific Oct-factors of figure 3D. D. Band shift assay with oligonucleotides containing the *Hox-2.4* octamer motif. In all cases 4 μ g nuclear extract is used. Nuclear extract from tissues were made from 13.5 day embryos. Lane 1: C1003 EC cells; lane 2: C1003 treated for 24 hours with RA; lane 3: neural tube; lane 4: brain; lane 5: mesoderm surrounding the neural tube; lane 6: liver.

Figure 4. Transcriptional activity of *Hox-2.4*-luciferase reporter constructs after transfection into MES-1 cells. A schematic representation of the transcripts is given at the left of the figure. The transcription initiation site is indicated with a thin arrow. The open arrow represents the luciferase gene. Positions of nucleotides is given in numbers relative to the transcription start site. Numbers at the right indicate measured luciferase levels, normalized for transfection efficiency. The activity of pHBg850Luci was assumed to be 100%. Each number is the average of three experiments. B. Sequence of the octamer motif before (upper line) and after (lower line) site directed mutagenesis.

two bands in figure 3A are caused by binding of proteins upstream of the octamer motif. Three different CCAAT-box-containing oligos (see Materials and Methods) failed to compete for formation of any of these complexes (data not shown). Therefore, the nature of the proteins giving rise to the middle two bands remains unclear.

As a next step we performed band shift assays with nuclear extracts from EC cells that had undergone the same RA treatment as those used for *Hox-2.4* expression studies. In addition, nuclear extracts of C1003 EC cells differentiated by serum deprivation were tested in this assay. As described previously (22), this latter treatment causes a neuronal differentiation of these cells, but does not induce expression of HOX-genes. In nuclear extracts from P19 EC cells grown in the presence of RA for five days no Oct-4 protein can be detected (Figure 3C, lane 3). Down-regulation of Oct-4 in P19 EC cells starts within several hours after administration of RA and is known to be complete within 48 hours (43 and data not shown). The intensity of the other three retarded bands did not change significantly, except for the Oct-1 complex which became somewhat more pronounced. Treatment of C1003 EC cells with RA for 24 hours also leads to a decrease in the level of Oct-4 and like in P19 EC it cannot be detected anymore after 48 hours (Figure 3C). Moreover, in band shift assays with the RA150 probe one extra band appeared (Figure 3C, lane 5 to 7), which resulted from the induction of an additional Oct-factor as shown by competition experiments with octamer motif containing *c-abl* oligonucleotides (data not shown). A very similar

pattern of retarded bands was found with nuclear extracts of C1003 cells after serum deprivation induced differentiation.

The proteins binding upstream of the octamer sequence were not affected by RA treatment in C1003 EC or P19 EC cells. However, the strong binding of these proteins to the RA150 probe might have obscured the presence of additional Oct-factors. Therefore, band shifts with the 5' deleted probe and with oligonucleotides containing the *Hox-2.4* octamer motif were performed. With these probes no additional binding proteins were found in P19 EC cells. However, two additional Oct-factors not detected with the RA150 probe were found both in RA treated C1003 EC cells and in C1003 EC cells after serum deprivation. As a major fraction of the RA differentiated C1003 EC cells has a neuronal phenotype we questioned whether similar Oct-factors were present in neural tissues of murine embryos. Nuclear extracts were prepared from brain (excluding hindbrain), neural tube, mesoderm directly surrounding the neural tube and liver of 13.5 day mouse embryos and tested in the band shift assay (Figure 3D). In liver and mesoderm only Oct-1 could be detected. In contrast several Oct-factors present both in brain and neural tube bound to the *Hox-2.4* octamer motif. The pattern of retarded bands obtained with nuclear extracts from these neuronal tissues shows striking similarities to that of differentiated C1003 cells. The Oct-factors detected here are likely to correspond to the previously described Oct-2, N-Oct-3 and Oct-7 (19) or N-Oct-2, N-Oct-3 and N-Oct-4 (45). In summary, band shift assays show

Figure 5. Transcriptional activity of *Hox-2.4* luciferase reporter constructs in cotransfection experiments with Oct-expression vectors in MES-1 cells. Luciferase activity was calculated relative to the activity of unmutated luciferase vectors in the absence of Oct-expression vectors. Error bars represent one standard deviation. A. Luciferase activity of pRA150Luci is stimulated by low amounts of cotransfected Oct2A. This stimulation is dependent on an intact octamer motif (compare construct pRAmutLuci). At high levels of the Oct2A expression plasmid transcription of both constructs is inhibited. The results shown represent the outcome of three independent experiments. B. The presence of an SV40 enhancer results in a higher stimulation of the short *Hox-2.4* reporter construct by cotransfected Oct-factors. Again, transcriptional activity of both constructs is inhibited at high concentrations of cotransfected Oct2A, probably as a consequence of squelching. The results shown represent the outcome of three independent experiments. C. Stimulation of luciferase activity by Oct2A is also found with the longer pSDHA700 construct. Stimulation of transcription is not seen in the absence of an intact octamer motif (construct pSDHAMut). The results shown represent the outcome of two independent experiments. D. Cotransfection of Oct-4 expression plasmids results in a two-fold increase in transcription from *Hox-2.4* reporter constructs containing an intact octamer motif. The results shown represent the outcome of two independent experiments.

that after RA treatment different EC cells contain different Oct-factors which resemble those found in tissues that the EC cells phenotypically correspond to. In addition, high levels of Oct-4 are only detected in EC cells that have an undifferentiated phenotype and do not express *Hox-2.4*.

Transcriptional activity of the *Hox-2.4* promoter

The transcriptional activity of the *Hox-2.4* promoter was studied by transient transfections in EC cells and EC cell derived cell lines. An 851 bp fragment containing sequences from -641 to +210 relative to the transcription start site as well as a nested set of deletions of this fragment were cloned in front of the firefly luciferase reporter gene. In all cases the luciferase activity measured was normalized for efficiency of transfection by cotransfection of a plasmid containing the β -galactosidase under the control of a hydroxy-3-methylglutaryl-coenzyme A reductase promoter (24). The results with these reporter constructs in MES-1 (42), a P19 derived cell line having mesodermal characteristics, are shown in figure 4A. The activity of reporter construct pHBg850Luci containing the longest *Hox-2.4* fragment tested was about 50% of the activity of the *Hox-2.3* promoter that we previously characterized (37), and about five times weaker than a minimal thymidine kinase promoter. Deletion of sequences between -641 and -313 (construct pEBg520Luci) causes a nearly two-fold drop in promoter activity. Promoter activity of a DNA fragment containing only 90 bp upstream of the transcription start (construct pRBg300Luci) is three-fold lower as compared to the fragment in pHBg850Luci. Removal of untranslated leader sequences (pRA150Luci) results in luciferase activity slightly higher than in the case of pRBg300Luci. This slight rise is also found when the same untranslated leader sequences are deleted from pHBg850Luci (construct pHA700Luci). In other cell lines similar results were obtained (data not shown). To analyse the role of the octamer motif present in the *Hox-2.4* promoter we used site directed mutagenesis to create a mutation that abolishes binding of Oct-factors (46,27; see Figure 4B). DNA fragments carrying this mutation were inserted into the luciferase vector in exactly the same way as their unmutated counterparts. This yielded two sets of reporter constructs (pHAMutLuci and pHA700Luci; pRAMutLuci and pRA150Luci) that differed only in two point mutations of the octamer motif. We then compared the luciferase activity of the reporter constructs carrying an intact or a mutated octamer motif in MES-1 cells (Figure 4A). Surprisingly, the absence of the octamer motif caused only a minor decrease (20%) in the activity of the *Hox-2.4* promoter. These data show that the octamer motif in the *Hox-2.4* promoter is not important for basal activity.

The *Hox-2.4* promoter can be stimulated by cotransfected Oct-factors

To find out whether *Hox-2.4* promoter activity can be influenced by the presence of Oct-factors cotransfection experiments were performed in MES-1 cells. This cell line contains only Oct-1 as shown by band shift assays. Expression plasmids that were used carry cDNAs encoding Oct-2A, Oct-4 or Oct-6 under the control of the CMV enhancer / TK promoter (26,27). Upon transfection high levels of Oct-2A and Oct-6 protein could easily be demonstrated in whole cell extracts by means of band shift assays, whereas expression of Oct-4 seemed to be much lower (data not shown). The *Hox-2.4* reporter constructs that were used for cotransfection experiments are pHA700Luci, pHAMutLuci, pRA150Luci and pRAMutLuci. As it has been shown in a number

of cases that the presence of a general enhancer like the SV40 enhancer is indispensable for activation of octamer containing promoters (47,48) the same promoter fragments were cloned into the luciferase vector pSD19LUC, which contains an SV40 enhancer downstream of the luciferase gene (constructs pSDHA700, pSDHAMut, pSDRA150 and pSDRAMut). First, reporter constructs containing sequences between -90 and +60 nt relative to the transcription start site (pRA150Luci, pRAMutLuci) were cotransfected with expression vectors containing no Oct-factor cDNA and with increasing amounts of Oct-2A expression vectors. Small amounts (50 to 200 ng) of Oct-2A producing expression vector caused a two- to three-fold stimulation of luciferase activity in cells transfected with pRA150Luci. This weak stimulation was not seen with pRAMutLuci containing a mutated octamer motif. Increasing the amount of Oct-2A expression vector up to 2.5 μ g caused a decrease in luciferase activity with both reporter constructs, probably as a result of squelching (Figure 5A). In cotransfection experiments with pSDRA150 a four-fold stimulation was found. This activation was also completely dependent on an intact octamer motif in the promoter as it was not observed with pSDRAMut (Figure 5B). We then tested longer *Hox-2.4* reporter constructs with Oct-2A expression vectors in cotransfection assays. Essentially the same results were obtained, i.e. a four-fold stimulation of the SV40 enhancer carrying pSDHA600 plasmid, a three-fold induction of pHA600Luci and no induction of reporter constructs without an intact octamer sequence (Figure 5C). Use of Oct-6 instead of Oct-2A expression plasmid resulted in a very similar enhancement of transcription (data not shown).

Finally, the effect of Oct-4 on the same reporter constructs was investigated. Considering the expression pattern of Oct-4 during embryogenesis and its presence in undifferentiated EC cells one might predict that Oct-4 is a repressor of Hox gene expression. However, upon cotransfection of Oct-4 producing expression vectors both the enhancerless pRA150Luci and pHA700Luci constructs as the SV40 enhancer containing pSDRA150 and pSDHA700 construct were stimulated two-fold. No stimulation was found with constructs that carried a mutated octamer motif (Figure 5D). The same experiments were performed in P19 EC cells, that had been differentiated with RA, two days prior to transfection. No essential differences between the two cell lines were observed. These results show that Oct-factors can stimulate the *Hox-2.4* promoter via the octamer motif.

To test the effects of endogenous Oct-factors on the *Hox-2.4* promoter pRA150Luci, pRAMutLuci, pSDHA700 and pSDHAMut were transfected into C1003 EC cells before and after differentiation with RA. In these cells reporter constructs were about twice as active as constructs carrying a mutated octamer motif. Upon addition of RA a 1.9-fold induction was seen for the former, while the latter were activated 1.2 fold. These results suggest that in EC cells too, Oct-4 is not a strong inhibitor of *Hox-2.4* transcription.

DISCUSSION

During development of a multicellular organism many different cell types, expressing their own characteristic set of genes, arise from the zygote. Hence, one of the major challenges of developmental biology is to understand in detail what mechanisms underly the phenomenon of differential gene expression. During the last decade it has become clear that interactions between *cis*-

acting DNA (promoters and enhancers/silencers) and transcription factors are of major importance for gene regulation. One of the best known examples of a DNA sequence responsible for tissue specific expression is the octamer motif. This motif was originally identified as a conserved element in the promoter of B cell-specific immunoglobulin genes (49). Surprisingly, simple artificial promoters in which an octamer motif was placed upstream of a TATA-box appeared to be expressed in a B-cell specific way (46,47). Oligomerization of the octamer motif creates an B cell-specific enhancer (46) and refs therein), although it is also functional in EC and embryonic stem cells (50). The cloning of so-called Oct-factors (e.g. 26,43,44,51–55), transcription factors that bind to the octamer sequence, has led to further insight in tissue specific transcription as it could be shown that these proteins can activate transcription from octamer containing promoters. However, these studies also show the complex nature of gene regulation. For example, Oct-factors appear to be insufficient to stimulate transcription from enhancers (48). In addition, recent studies show that some natural B cell-specific, octamer containing promoters, are not stimulated by cotransfected Oct-factors as was expected on the basis of the behaviour of similar, composite promoters (57). Indeed, many factors including the affinity of a given binding site and the context of this site in a promoter are determining the transcriptional activity of a promoter (57).

HOX genes represent a class of genes that have complex expression patterns during embryogenesis. Expression of HOX-genes starts around day 7.5 to 8.5 p.c. in the allantois and spreads rapidly into an anterior direction. During midgestation a regionally restricted expression is found in neurectodermal and mesodermal tissues (see e.g. 7,5). It is likely that such complex expression patterns are dependent on the combined action of a number of *cis*-acting elements, each binding their specific *trans*-acting factors. In a few cases *cis*-acting elements with a tissue and/or region-specific activity have been shown to exist in HOX clusters (e.g. 25,58–61). Sequence analysis of the *Hox-2.4* promoter shows the presence of binding sites for known transcription factors, like a CCAAT-box and an octamer motif (ATGCAAATNA). This latter motif is especially interesting because of the presence of several Oct-factors during embryogenesis.

One of the Oct-factors that can bind to the octamer motif in the *Hox-2.4* promoter is the germline-specific Oct-4 protein. This factor is highly expressed in embryonic stem and embryo carcinoma cells, while RA treatment of these cells causes a rapid down-regulation (43,44). During early murine embryogenesis Oct-4 is found in the inner cell mass and pluripotent cells of the primitive ectoderm. Expression declines around day 8.5 p.c. and persists only in primordial germ cells (29,54). Both stimulation (43,44) and repression (43,62) of transcription by this Oct-factor have been described. On the basis of its expression Oct-4 seemed to be a candidate for a repressor of *Hox-2.4* expression in EC cells and during early embryogenesis. However, our results show that cotransfection of Oct-4 leads to a weak but significant stimulation instead of a repression of the *Hox-2.4* promoter. Although we cannot exclude the possibility that Oct-4 acts upon *cis*-acting elements in the HOX-2 cluster that are not present in our reporter constructs, our data do not support a role for Oct-4 as a repressor of *Hox-2.4* transcription. Indirect support for this notion comes from the observation that the absence of Oct-4 protein in C1003 EC cells after serum deprivation is not accompanied by detectable level of *Hox-2.4* expression (22; F.Z. and T.H. unpublished data).

We show that apart from Oct-4 several other Oct-factors can bind to the *Hox-2.4* promoter. Some of these Oct-factors appear to be tissue-specific as they can be detected in nuclear extracts from neural tissues of 13.5 day embryos but not in mesoderm or liver extracts. Similarly, they occur in differentiated C1003 EC cells with a neuronal phenotype, but not in P19 EC cells that have differentiated into primitive endoderm and mesodermal cells. These Oct-factors are present both in places where HOX genes are highly expressed like the neural tube, as well as in anterior brain structures where no HOX gene expression is found. This situation is paralleled in C1003 EC cells that have been differentiated by RA or by serum deprivation. Both treatments cause C1003 EC cells to differentiate into neuron-like derivatives, containing similar Oct-factors, but only in RA treated cells HOX gene expression has been found (22). This indicates that these Oct-factors are not main determinants of expression i.e. their presence is not sufficient for *Hox-2.4* expression. However, it should not be expected that complex expression patterns are dependent on a single transcription factor but rather are the result of the simultaneous presence within a cell of a certain combination of transcription factors each having their own unique domains of expression. Here we show that two Oct-factors that have been reported to be expressed in neural tissues, namely Oct-2A (19,63) and Oct-6 (53,55,56) can stimulate the *Hox-2.4* promoter. This stimulation is most clearly seen in the presence of an SV40 enhancer. It is likely that during embryogenesis several enhancing elements from the *Hox-2* cluster determine whether or not the *Hox-2.4* gene is transcribed. In EC cells RA seems to be required for activation of these elements. The role of Oct-factors present in a cell may then be to determine the level of expression. Such a mechanism could explain the higher expression of *Hox-2.4* in the neural tube as compared to the surrounding mesoderm in a 13.5 day embryo or the differences in expression between C1003 EC and P19 EC cells after RA treatment. Apart from *Hox-2.4* several other HOX-genes may be regulated by Oct-factors, as binding sites for Oct-factors also have been found in the promoters of *Hox-2.3* (64) and *Hox-1.3* (cited in 18). Alternative splicing of Oct-gene messengers, post-translational modification of Oct-factors and protein-protein interactions may offer additional levels of regulation.

In summary, the data presented here show for the first time a direct activation of a Hox-promoter by members of the Oct-family. Although a four-fold induction (in the presence of an SV40 enhancer by Oct-2A) may seem relatively low, we believe in its biological significance. It should be borne in mind that the experiments shown here most likely reflect only part of an intricate mechanism underlying control of *Hox-2.4* expression. It is conceivable that full activation of the promoter as it occurs *in vivo* is not seen in the absence of the remaining elements of this mechanism. The recent identification of *cis*-acting elements necessary for tissue and region-specific expression of *Hox-2.3* and *Hox-2.4* (25, and Vogels *et al.* in preparation) will enable us to examine the role of the octamer motif in the *Hox-2.4* promoter *in vivo* by transgenic studies using the β -Galactosidase reporter gene.

ACKNOWLEDGEMENTS

We thank Dr. H.R. Schöler for the CMV-OCT4 expression plasmid, Dr. E. Schreiber and colleagues for the pOEVI(+) Oct2A expression construct and R. de Jong for the TKluc reporter construct. We are grateful to D. Meijer for providing

c-abl oligonucleotides and Oct-6 expression plasmid. We thank M. van Dijk for providing CCAAT-box-containing oligonucleotides from the human IGF II gene and adenovirus, and J. den Hertog for *junD* oligos.

REFERENCES

- Deschamps, J. and Meijlink, F., (1992) *CRC Crit. Rev. Oncogenesis*, **3**, 117–174.
- Gaunt, S.J., Miller, J.R., Powell, D.J. and Duboule, D. (1986) *Nature*, **324**, 662–664.
- Holland, P.W.H. and Hogan, B.L.M. (1988) *Genes Dev.*, **2**, 773–782.
- Duboule, D. and Dollé, P. (1989) *EMBO J.*, **8**, 1497–1505.
- Graham, A., Papalopulu, N. and Krumlauf, R. (1989) *Cell*, **57**, 367–378.
- Gehring, W.J. (1987) *Science*, **236**, 1245–1252.
- Gaunt, S.J., Sharpe, P.T. and Duboule, D. (1988) *Development*, **104** Supplement, 169–179.
- Ruiz i Altaba, A. and Melton, D.A. (1989) *Cell*, **57**, 317–326.
- Wright, C.V.E., Cho, K.W.Y., Hardwicke, J., Collins, R.H. and DeRobertis, E.M. (1989) *Cell*, **59**, 81–93.
- Cho, K.W.Y., Morita, E.A., Wright, C.V.E. and DeRobertis, E.M. (1991) *Cell*, **65**, 55–64.
- Balling, R., Mütter, G., Gruss, P. and Kessel, M. (1989) *Cell*, **58**, 337–347.
- Kessel, M., Balling, R. and Gruss, P. (1990) *Cell*, **61**, 301–308.
- Chisaka, O. and Capecchi, M.R. (1991) *Nature*, **350**, 473–479.
- Lufkin, T., Dierich, A., LeMeur, M., Mark, M. and Chambon, P. (1991) *Cell*, **66**, 1105–1119.
- Ingham, P.W. (1988) *Nature*, **335**, 25–34.
- Herr, W., Sturm, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., Ruvkun, G. and Horvitz, H.R. (1988) *Genes Dev.*, **2**, 1513–1516.
- Rosenfeld, M.G. (1991) *Genes Dev.*, **5**, 897–907.
- Schöler, H.R. (1991) *Trends Genet.*, **7**, 323–329.
- Schöler, H.R., Hatzopoulos, A.K., Balling, R., Suzuki, N. and Gruss, P. (1989) *EMBO J.*, **8**, 2543–2550.
- Perkins, A., Kongsuwan, K., Visvader, J., Adams, J.M. and Cory, S. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8389–8402.
- Aberdam, D., Negreanu, V., Sachs, L. and Blatt, C. (1991) *Mol. Cell. Biol.*, **11**, 554–557.
- Deschamps, J., de Laaf, R., Joosen, L., Meijlink, F. and Destrée, O. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1304–1308.
- Graham, F.L. and Van der Eb, A.J. (1973) *Virology*, **52**, 456–467.
- Gautier, C., Mehtali, M. and Lathe, R. (1989) *Nucleic Acids Res.*, **17**, 8389.
- Kress, C., Vogels, R., de Graaff, W., Bonnerot, C., Meijlink, F., Nicolas, J.-F. and Deschamps, J. (1990) *Development*, **109**, 775–786.
- Müller, M.M., Ruppert, S., Schaffner, W. and Matthias, P. (1988) *Nature*, **336**, 544–551.
- Schöler, H.R., Dressler, G.R., Balling, R., Rohdewold, H. and Gruss, P. (1990) *EMBO J.*, **9**, 2185–2195.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor.
- Kongsuwan, K., Allen, J. and Adams, J.M. (1989) *Nucleic Acids Res.*, **17**, 1881–1891.
- Meijlink, F., de Laaf, R., Verrijzer, P., Destrée, O., Kroezen, V., Hilken, J. and Deschamps, J. (1987) *Nucleic Acids Res.*, **15**, 6773–6786.
- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. eds. (1990) *Current protocols in molecular biology*. Massachusetts General Hospital. Harvard Medical School. Greene Publishing Ass. and Wiley Interscience.
- De Wet, J.R., Wood, K.V., De Luca, M., Helinski, D.R. and Subramani, S. (1987) *Mol. Cell. Biol.*, **7**, 725–737.
- Van Zonneveld, A.-J., Curriden, S.A. and Loskutoff, D.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 5525–5529.
- Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) *Methods Enzymol.*, **154**, 367–382.
- Okayama, H. and Berg, P. (1983) *Mol. Cell. Biol.*, **3**, 280–289.
- Lee, K.A.W., Bindereif, A. and Green, M.R. (1988) *Gene Anal. Techn.*, **5**, 22–31.
- Zwartkruis, F., Hoeijmakers, T., Deschamps, J. and Meijlink, F. (1991) *Mech. Development*, **33**, 179–190.
- Hart, C.P., Awgulewitsch, A., Fainsod, A., McGinnis, W. and Ruddle, F.H. (1985) *Cell*, **43**, 9–18.
- Hart, C.P., Fainsod, A. and Ruddle, F.H. (1987) *Genomics*, **1**, 182–195.
- Blatt, C., Aberdam, D., Schwartz, R. and Sachs, L. (1988) *EMBO J.*, **7**, 4283–4290.
- Roguska, M.A. and Gudas, L.S. (1985) *J. Biol. Chem.*, **260**, 13893–13896.
- Mummary, C.L., Feijen, A., Moolenaar, W.H., van den Brink, C.E. and de Laat, S.W. (1986) *Exp. Cell Res.*, **165**, 229–242.
- Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M. and Hamada, H. (1990) *Cell*, **60**, 461–472.
- Schöler, H.R., Ruppert, S., Suzuki, N., Chowdhury, K. and Gruss, P. (1990) *Nature*, **344**, 435–439.
- Schreiber, E., Harshman, K., Kemler, I., Malipiero, U., Schaffner, W. and Fontana, A. (1990) *Nucleic Acids Res.*, **18**, 5495–5503.
- Gerster, T., Matthias, P., Thali, M., Jiricny, J. and Schaffner, W. (1987) *EMBO J.*, **6**, 1323–1330.
- Dreyfus, M., Doyen, N. and Rougeon, F. (1987) *EMBO J.*, **6**, 1685–1690.
- Müller-Immerglück, M.M., Schaffner, W., and Matthias, P. (1990) *EMBO J.*, **9**, 1625–1634.
- Falkner, F.G. and Zachau, H.G. (1984) *Nature*, **310**, 71–74.
- Schöler, H.R., Balling, R., Hatzopoulos, A.K., Suzuki, N. and Gruss, P. (1989) *EMBO J.*, **8**, 2551–2557.
- Sturm, R.A., Das, G. and Herr, W. (1988) *Genes Dev.*, **2**, 1582–1599.
- Scheidereit, C., Cromlish, J.A., Gerster, T., Kawakami, K., Balmaceda, C.-G., Currie, R.A. and Roeder, R.G. (1988) *Nature*, **336**, 551–557.
- Monuki, E.S., Kuhn, K., Weinmaster, G., Trapp, B.D. and Lemke, G. (1990) *Science*, **249**, 1300–1303.
- Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W.J. and Staudt, L.M. (1990) *Nature*, **345**, 686–692.
- Meijer, D., Graus, A., Kraay, R., Langeveld, A., Mulder, M.P. and Grosveld, G. (1990) *Nucleic Acids Res.*, **18**, 7357–7365.
- Suzuki, N., Rohdewold, H., Neuman, T., Gruss, P. and H.R. Schöler. (1990) *EMBO J.*, **9**, 3723–3732.
- Kemler, I., Bucher, E., Seipel, K., Müller-Immerglück, M.M. and Schaffner, W. (1991) *Nucleic Acids Res.*, **19**, 237–242.
- Zakany, J., Tuggle, C.K., Patel, M.D. and Nguyen-Huu, M.C. (1988) *Neuron*, **1**, 679–691.
- Püschel, A.W., Balling, R. and Gruss, P. (1990) *Development*, **108**, 435–442.
- Bieberich, C.J., Utset, M., Awgulewitsch, A. and Ruddle, F.H. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8462–8466.
- Whiting, J., Marshall, H., Cook, M., Krumlauf, R., Rigby, P.W.J., Stott, D. and Alleman, R.K. (1991) *Genes Dev.*, **5**, 2048–2059.
- Lenardo, M.J., Staudt, L., Robbins, P., Kuang, A., Milligan, R.C. and Baltimore, D. (1989) *Science*, **243**, 544–546.
- Hatzopoulos, A.K., Stoykova, A.S., Erselius, J.R., Goulding, M., Neuman, T. and Gruss, P. (1990) *Development*, **109**, 349–362.
- Verrijzer, C.P., Kal, A.J., and van der Vliet, P.C. (1990) *Genes Dev.*, **4**, 1964–1974.