Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3

Huabing Yuan,[1] Nicoletta Corbi,[2] Claudio Basilico,[3] and Lisa Dailey[3,4]

Department of Microbiology and Kaplan Cancer Center, New York University School of Medicine, New York, New York 10016 USA; [4]The Department of Molecular Biology, The Rockefeller University, New York, New York 10021 USA

Fibroblast growth factor 4 (FGF-4) has been shown to be a signaling molecule whose expression is essential for postimplantation mouse development and, at later embryonic stages, for limb patterning and growth. The FGF-4 gene is expressed in the blastocyst inner cell mass and later in distinct embryonic tissues but is transcriptionally silent in the adult. In tissue culture FGF-4 expression is restricted to undifferentiated embryonic stem (ES) cells and embryonal carcinoma (EC) cell lines. Previously, we determined that EC cell-specific transcriptional activation of the FGF-4 gene depends on a synergistic interaction between octamer-binding proteins and an EC-specific factor, Fx, that bind adjacent sites on the FGF-4 enhancer. Through the cloning and characterization of an F9 cell cDNA we now show that the latter activity is Sox2, a member of the Sry-related Sox factors family. Sox2 can form a ternary complex with either the ubiquitous Oct-1 or the embryonic-specific Oct-3 protein on FGF-4 enhancer DNA sequences. However, only the Sox2/Oct-3 complex is able to promote transcriptional activation. These findings identify FGF-4 as the first known embryonic target gene for Oct-3 and for any of the Sox factors, and offer insights into the mechanisms of selective gene activation by Sox and octamer-binding proteins during embryogenesis.

[*Key Words*: FGF-4 expression; Oct proteins; Sox factors; mouse embryogenesis]

Received July 5, 1995; revised version accepted September 13, 1995.

The complex intercellular interactions that take place during development are directed by the orchestrated activities of extracellular signaling molecules, whose precise temporal and spatial expression is fundamental to proper embryogenesis and is largely determined at the transcriptional level. Thus, insight into the transcriptional regulation of the genes for these signaling peptides is prerequisite to understanding the process of development. Although the identification of many developmental-specific transcription factors and their cognate recognition elements have been reported, in most cases the actual mechanisms by which these interactions generate the complex patterns of gene activity taking place during embryogenesis still remains to be clarified.

We have been studying the regulatory elements controlling expression of fibroblast growth factor 4 (FGF-4), a member of the FGF family originally identified as an oncogene (Delli Bovi et al. 1987; Taira et al. 1987). FGF-4 recently has been shown to provide several important signaling functions in embryogenesis. FGF-4 is essential for survival of the postimplantation mouse embryo (Feldman et al. 1995) and, at later embryonic stages, is an essential component of a signaling network required for growth and patterning of the developing limb (Niswander et al. 1993; Laufer et al. 1994). The FGF-4 gene is expressed in the blastocyst inner cell mass and later in distinct embryonic tissues but is transcriptionally silent in the adult (Niswander and Martin 1992; for review, see Basilico and Moscatelli 1992). Accordingly, FGF-4 expression is observed in embryonic stem (ES) cells and embryonal carcinoma (EC) tissue culture cell lines such as F9 and P19, but not in their differentiated counterparts or in HeLa or 3T3 cells (Velcich et al. 1989; Schoorlemmer and Kruijer 1991). Extensive mutational analysis has demonstrated that EC cell-specific expression is conferred by the combinatorial interaction of factors that bind to multiple *cis*-acting elements within the FGF-4 enhancer (Curatola and Basilico 1990; Dailey et al. 1994). One of these elements, OCTA2, is recognized by both the ubiquitous octamer-binding protein Oct-1 and the embryonic-specific Oct-3 protein. However, mutation of a separate element that binds a second, EC-specific factor that we designated as Fx, also abolishes enhancer activity. The Fx-binding site lies immediately upstream of the octamer motif, and we established that enhancer activity correlates with the ability of the FGF-4 enhancer DNA to form ternary complexes (Oct*) composed of Fx

**Present addresses: [1]Brigham & Women's Hospital, Hematology–Oncology Division, Boston, Massachusetts 02115; [2]Dipartimento di Biopatologia Umana, Universita di Roma La Sapienza, 00161 Rome, Italy.
[3]Corresponding authors.**

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

and octamer-binding protein (Dailey et al. 1994). Together these observations indicate that EC cell-specific transcriptional activation is determined by a synergistic interaction between Fx and octamer-binding proteins on the FGF-4 enhancer.

Several properties of Fx (Dailey et al. 1994) suggested that it might be a member of the Sry-related Sox factor family. Mutations within certain Sox factor genes have been found to underlie several developmental abnormalities including sex reversal, campomelic dysplasia, and Borjeson–Forssman–Lehmann syndrome, suggesting that Sox factors play an essential role in the execution of specific developmental programs (Gubbay et al. 1990; Sinclair et al. 1990; Stevanovic et al. 1993; Tommerup et al. 1993; Wright et al. 1993; Foster et al. 1994; Wagner et al. 1994). However, none of the target genes regulated by Sox factors have been identified to date. Sox factors are expressed in distinct embryonic tissues and are related by a highly conserved 80-amino-acid high mobility group (HMG) domain that mediates factor binding within the minor groove of the DNA sequence TCTTTGTT (Gubbay et al. 1990; Sinclair et al. 1990; Nasrin et al. 1991; van de Wetering and Clevers 1992). These DNA-binding properties also appear to be exhibited by Fx. In addition, Fx activity is tissue restricted and is observed only in cells of embryonic origin (Dailey et al. 1994). We therefore searched for a complementary DNA clone encoding an EC cell-specific Sox factor to compare its binding and transcriptional activation properties with those of Fx. As we report here, these efforts have led to the isolation of a full-length cDNA clone for murine Sox2. We show that the DNA-binding and antigenic properties of Sox2 are indistinguishable from those of Fx, indicating that these are identical proteins. We also demonstrate that Sox2 acts as a transcriptional activator of the FGF-4 enhancer. Remarkably, however, *trans*-activation of FGF-4 enhancer chloramphenicol acetyltransferase (CAT) constructs by Sox2 requires the coexpression of Oct-3 and was not observed with Oct-1. These findings thus identify FGF-4 as a target gene specifically activated by an Oct-3/Sox2 complex in EC cells. The specificity of Sox2 activation with Oct-3 also offers insight into the mechanism of selective gene activation by individual members of these multifactor families and the transcriptional regulation of the FGF-4 expression during embryogenesis.

```
GTGTTTGCAAAAAGGGAAAAGTACTTTGCTGCCTCTTTAAGACTAGGGCTGGGAGAAAGA        60
AGAGGAGAGAGAAAGAAAGGAGAGAAGTTTGGAGCCCGAGGCTTAAGCCTTTCCAAAAAC       120
TAATCACAACAATCGCGGCGGCCCGAGGAGGAGAGTCTCCCTGTTTTTTCATCCCAATTG       180
CACTTCGCCCGTCTCGAGCTCCGCTTCCCCCCAACTATTCTCCGCCAGATCTCCGCGCAG       240
GGCCGTGCACGCCGAGGCCCCCGCCCGCGGCCCCTGCATCCCGGCCCCCGAGCGCGGCCC       300
CCACAGTCCCGGCCGGGCCGAGGGTTGGCGGCCGCCGGCGGGCCGCCCGCCCAGCGCCCG       360
CATGTATAACATGATGGAGACGGAGCTGAAGCCGCCGGGCCCGCAGCAAGCTTCGGGGGG       420
 M  Y  N  M  M  E  T  E  L  K  P  P  G  P  Q  Q  A  S  G  G    20
CGGCGGCGGAGGAGGCAAGCCCACGGCGGCGGCGACCGGCGGCAACCAGAAGAACAGCCC       480
 G  G  G  G  G  K  P  T  A  A  A  T  G  G  N  Q  K  N  S  P    40
GGACCGCGTCAAGAGGCCCATGAACGCCTTCATGGTATGGTCCCGGGGGCAGCGGCGTAA       540
 D  R  V  K  R  P  M  N  A  F  M  V  W  S  R  G  Q  R  R  K    60
GATGGCCCAGGAGAACCCCAAGATGCACAACTCGGAGATCAGCAAGCGCCTGGGCGCGGA       600
 M  A  Q  E  N  P  K  M  H  N  S  E  I  S  K  R  L  G  A  E    80
GTGGAAACTTTTGTCCGAGACCGAGAAGCGGCCATTCATCGACGAGGCCAAGCGGCTGCG       660
 W  K  L  L  S  E  T  E  K  R  P  F  I  D  E  A  K  R  L  R   100
CGCTCTGCACATGAAGGAGCACCCCGATTATAAATACCGGCCGCGGCGGAAAACCAAGAC       720
 A  L  H  M  K  E  H  P  D  Y  K  Y  R  P  R  R  K  T  K  T   120
GCTCATGAAGAAGGATAAGTACACGCTTCCCGGAGCTTGCTGGCCCCCGCGGAACAGCAT       780
 L  M  K  K  D  K  Y  T  L  P  G  A  C  W  P  P  R  N  S  M   140
GGCGAGCGGGGTTGGGGTGGGCGCCGGCCTGGGTGGCGGGCTGAACCAGCGCATGGACAG       840
 A  S  G  V  G  V  G  A  G  L  G  G  G  L  N  Q  R  M  D  S   160
CTACGCGCACATGAACGGCTGGAGCAACGGCAGCTACAGCATGATGCAGGAGCAGCTGGG       900
 Y  A  H  M  N  G  W  S  N  G  S  Y  S  M  M  Q  E  Q  L  G   180
CTACCCGCAGCACCCGGGCCTCAACGCTCACGGCGCGGCACAGATGCAACCGATGCACCG       960
 Y  P  Q  H  P  G  L  N  A  H  G  A  A  Q  M  Q  P  M  H  R   200
CTACGTCGTCAGCGCCCTGCAGTACAACTCCATGACCAGCTCGCAGACCTACATGAACGG      1020
 Y  V  V  S  A  L  Q  Y  N  S  M  T  S  S  Q  T  Y  M  N  G   220
CTCGCCCACCTACAGCATGTCCATCTCGCAGCAGGGCACCCCCGGTATGGCGCTGGGCTC      1080
 S  P  T  Y  S  M  S  I  S  Q  Q  G  T  P  G  M  A  L  G  S   240
CATGGGCTCTGTGGTCAAGTCCGAGGCCAGCTCCAGCCCCCCCGTGGTTACCTCTTCCTC      1140
 M  G  S  V  V  K  S  E  A  S  S  S  P  P  V  V  T  S  S  S   260
CCACTCCAGGGCGCCCTGCCAGGCCGGGGACCTCCGGGACATGATCAGCATGTACCTCCC      1200
 H  S  R  A  P  C  Q  A  G  D  L  R  D  M  I  S  M  Y  L  P   280
CGGCGCCGAGGTGCCGGAGCCCGCTGCGCCCAGTCGACTGCACATGGCCCAGCTCATCCT      1260
 G  A  E  V  P  E  P  A  A  P  S  R  L  H  M  A  Q  L  I  L   300
GTGCGGCCCGGTGCCCGGCACGGCCAAATACGGCACACTGCCCCTGTCGCACATGTGAGG      1320
 C  G  P  V  P  G  T  A  K  Y  G  T  L  P  L  S  H  M  *     318
GCTGGACTGCGAACTGGAGAAGGGGAGAGATTTTCAAAGAGATACAAGAGAATTGGGAGG      1380
GGTGCAAAAAGAGGAGAGTAGGAAAAATCTGATAATGCYCAAAAGGAAAAAAACCACCAA      1440
TCCCATCCAAATTAACGCTTTTTCGGTGATGCCGACTAGAAAACTTTTATGAGAGATCTG      1500
GAGGAAAAAAACTACGCAAAACTTTTTTTAAAGTTCTAGTGGTACGTTAGGCGCTTCGC       1560
AGGGAGTTCGCAAAAGTCTTTACCAGTAATATTTAGAGCTAGACTCCGGGCGATGAAAAA      1620
AAAGTTTTAATATTTGCAAGCAACTTTTGTACAGTATTTATCGAGATAAACATGGCAATC      1680
AAATGTCCATTGTTTATAAGCTGAGAATTTGCCAATATTTTTCGAGGAAAGGGAACAAGC      1740
TGGGAAAAGATTCTGCAGTTAAATTTAGGACCGTTACAAACAAGGAAGGAGTTTATTCGG      1800
ATTTGAACATTTTAGTTTTAAAATTGTACAAAAGGAAAACATGAGAGCAAGTACTGGCAA      1860
GACCGTTTTCGTGGTCTTGTTTAAGGCAAACGTTCTAGATTGTAGTAAATTTTTAACTTA      1920
CTGTTAAAGGCAAAAAAAAAATGCCCATGCAGGTTGATATCGTTGGTAATTTATAATAGC      1980
TTTTGTTCAATCCTACCCTTTCATTTTGTTCACATAAAAAATATGGAATTACTGTGTTTG      2040
AAATATTTTCTTATGGTTTGTAATATTTCTGTAAATTGTGATATTTTAAGGTTTTTCCCC      2100
CCTTTTATTTTCCGTAGTTGTATTTTAAAAGATTCGGCTCTCTTATTGGAATCAGGCTGG      2160
CGAGAATCCATGTATATATTTGAACTAATACCATCCTTATAACAGCTACATTTTCGGCTT      2220
AAGTTTTTACTCCATTATGCACAGTTTGAGATAAATAAATTTTTGAAATATGGACACTGA      2280
AAAAAAAAAAAAAAAAAAAAAAA                                           2303
```

Figure 1. DNA and predicted amino acid sequence of murine Sox2. Degenerate primers 5′-AAGGCCGGATCCATGAA(TC)GC(TCA)TT(TC)AT(TGA)GT(TGCA)TGG-3′ and 5′-AAGGCCGGATCC(AGCT)GG(TC)TT(AG)TA(TC)TT(AG)TA(AG)TC(AGCT)GG-3′, corresponding to the highly conserved Sox family HMG-box, were used to amplify HMG domain fragments from EC cell cDNA by PCR as described in Materials and methods. PCR products were subcloned into the *Bam*HI site of the Bluescript SK plasmid (Stratagene) and sequenced. The 200-bp Sox2 HMG box DNA was radiolabeled and used as a probe to screen an F9 cell cDNA library. The sequence of the largest of the positive clones is shown with the amino acid sequence encoded by the longest ORF depicted below the nucleotide sequence. The HMG box is underlined whereas the regions complementary to the PCR primers are boxed. There is an in-frame stop codon at position 104 of the nucleotide sequence and a polyadenylation signal at position 2255. The sequence has been submitted to EMBL/GenBank (accession no. U31967).

Results

Isolation of murine Sox2 cDNA

To search for an EC cell-specific Sox factor cDNA, we exploited the fact that all Sox factors contain a highly conserved 80-amino-acid Sry-related HMG domain (van de Wetering and Clevers 1992) and designed oligonucleotide primers complementary to the most conserved, least degenerate regions within the 3′ and 5′-most borders of this HMG box (Fig. 1). PCR amplification was performed using either P19 or F9 cDNA, and amplified products of ~200 bp were gel purified and subcloned. DNA sequence analysis revealed that PCR products obtained using either EC cell cDNA contained an open reading frame (ORF) for a peptide corresponding to the amino acid sequence within the HMG domain of Sox2 (Wright et al. 1993). Because the isolation of Sox2 cDNA sequences other than those within the HMG domain had not been reported, we used the Sox2 HMG box DNA fragment to screen an F9 cDNA library and obtain a full-length Sox2 cDNA. The largest of the positive isolates

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

contains an ORF predicted to encode an ~35-kD protein composed of 318 amino acids as shown in Figure 1. Amino acid sequences comprising the Sox2 HMG domain span residues 41–120. No other Sox factor cDNA clones were obtained from the library screening, even when low stringency conditions were employed.

The expression of Sox2 mRNA in mouse tissue was determined by Northern analysis using a radiolabeled Sox2 DNA probe derived from the 3'-untranslated region of the Sox2 cDNA. The ~2.3-kb Sox2 transcript was detected in brain and retina and at very low levels in lung and stomach, whereas no expression was seen in any of the other tissues examined (Fig. 2A; data not shown). Significantly, Sox2 mRNA was prominent in RNA prepared from F9 and P19 (data not shown) EC cell lines as well as from ES cells (Fig. 2B). Differentiation of the F9 cells using retinoic acid (RA) resulted in a decrease in the level of Sox2 mRNA that was not observed for the GAPDH mRNA control (Fig. 2B). The predominant expression of Sox2 transcripts in ES and undifferentiated EC cells indicated that Sox2 expression was consistent with the distribution of Fx binding activity (Dailey et al. 1994).

Comparison of DNA-binding and antigenic properties of Sox2 and Fx

To determine the relationship between Sox2 and Fx, antibodies (αHSox2) were raised against purified recombinant Sox2 HMG domain. The αHSox2 antibodies specifically immunoprecipitated a single protein of ~35 kD from [^{35}S]methionine-labeled F9 (Fig. 3A, lane 2) and P19 (data not shown) cell extracts. This species, which was down-regulated in RA-treated F9 cells (Fig. 3A, lane 4) and was not detected in HeLa cell extracts (Fig. 3A, lane 6), is of the same molecular mass predicted for Sox2 (Fig. 1). Thus, although the αHSox2 antiserum was generated against an HMG domain that is common to all Sox family members, only one polypeptide, most likely Sox 2, is recognized by these antibodies in EC cells. Significantly, the αHSox2 antiserum specifically inhibited the interaction of Fx with its DNA site in electrophoretic mobility shift assays (EMSAs) performed with partially purified Fx from F9 cells and an FGF-4 enhancer oligonucleotide DNA probe (Fig. 3B). In contrast, neither the preimmune nor αOct-3 antibodies had any effect on Fx DNA binding. This result demonstrates that Fx must possess a Sox HMG domain that mediates its binding to DNA. Therefore, the finding that the αHSox2 antibody recognizes a single species in F9 extracts (Fig. 3A) and inhibits Fx DNA-binding activity strongly suggests that Fx is Sox2.

To further establish the identity of Fx as Sox2, we assessed the binding properties of Sox2 and its ability to form the Oct* ternary complexes. To this end, a construct containing Sox2 cDNA-coding sequences downstream of a cytomegalovirus (CMV) promoter was made and transfected into HeLa cells. Nuclear extracts prepared from the transfected cells were tested by EMSA using an oligonucleotide DNA probe containing the Fx and octamer-binding sites (Fig. 4). This analysis revealed

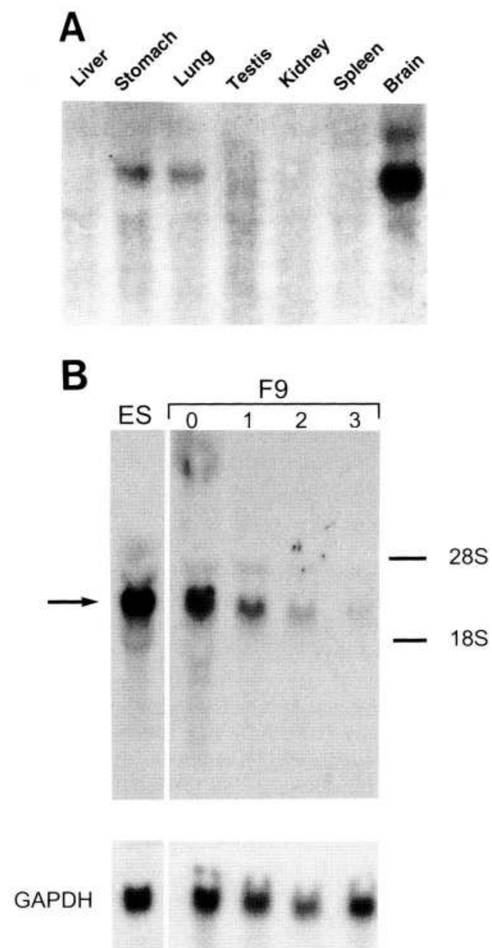


Figure 2. Expression of Sox2 mRNA. (*A*) Northern analysis of Sox2 mRNA in adult mouse tissues. Total RNAs (10 μg) were subjected to Northern blotting and hybridized to a radiolabeled 500-bp fragment corresponding to the 3' end of the untranslated region of Sox2 cDNA (nucleotides 1757–2303). Negative tissues not shown included ovary, uterus, and muscle. (*B*) Expression of Sox2 mRNA transcript in ES (CCE) cells or undifferentiated and differentiated F9 cells. Total RNA was prepared from F9 cells that were treated with 10^{-6} M all *trans*-RA for 0, 1, 2, or 3 days. RNAs (10 μg) were subjected to Northern blotting and hybridized with the same 500-bp radiolabeled Sox2 DNA fragment described in *A*. Hybridization of the same blot with a rat glyceraldehyde 3-phosphate dehydrogenase (GAPDH) probe was used to normalize for RNA quantity loaded in each lane.

that the expressed Sox2 protein generated a DNA–protein complex of the same mobility as that produced by Fx (Fig. 4A). Sox2 binds specifically to the Fx-binding site, as a mutant oligonucleotide containing base substitutions within this motif was unable to compete for Sox2 binding (Fig. 4A; lanes 4,5). Expression of the Sox2 protein induced formation of the ternary Oct-1* complex in the transfected HeLa cell nuclear extracts (Fig. 4A, lanes 3,5) comparable to that produced by Fx using F9 nuclear extracts (Fig. 4A, lane 1). In addition, extracts prepared from HeLa cells cotransfected with CMV–Oct-3 and CMV–Sox2 expression plasmids generated an

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

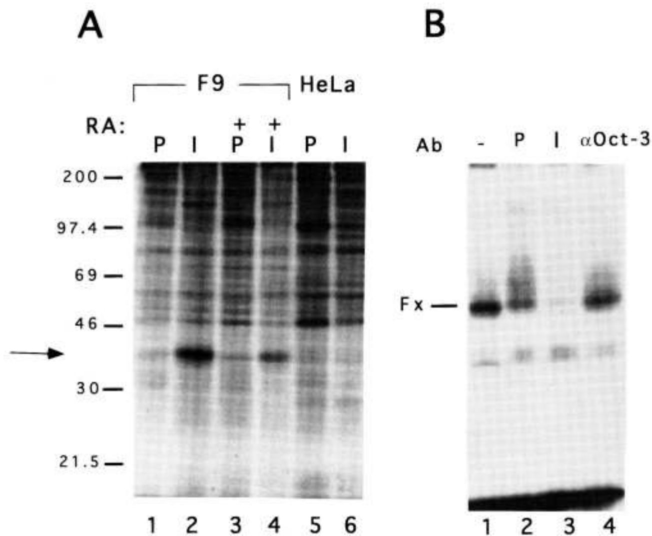


Figure 3. Fx is Sox2. (*A*) Immunoprecipitation of cell extracts using αHSox2 antibody. [^{35}S]Methionine-labeled protein lysates from undifferentiated or RA-treated F9, or HeLa cells were incubated with 3 μl of preimmune (P) or αHSox2 (I) IgG and the complexes precipitated using protein A–agarose beads as described in Materials and methods. Immunoprecipitated proteins were resolved by SDS-PAGE in a 10% polyacrylamide gel along with protein standards (Amersham) and visualized by autoradiography. (*B*) Inhibition of Fx DNA binding by αHSox2 antibodies. EMSA using radiolabeled FGF-4 enhancer oligonucleotide (see Fig. 4) and partially purified Fx protein from F9 cells was performed in the absence or presence of preimmune (P), αHSox2 (I), or anti-Oct-3 (αOct-3, 17) antisera as described in Materials and methods.

additional new complex that corresponds to the mobility of the Oct-3* complex observed in F9 cells (Fig. 4B). Taken together, these data demonstrate that Sox2 displays all of the binding properties of Fx and is consistent with the notion that Fx is Sox2.

Transcriptional activation by Sox2 and octamer-binding proteins

The ability of Sox2 and the individual octamer-binding proteins to *trans*-activate the FGF-4 enhancer was assessed using transient transfection into HeLa cells. Initial experiments utilized the fgfpML–CAT reporter construct that contains a single copy of a 115-bp segment derived from the FGF-4 enhancer (base pairs 93–207, Dailey et al. 1994) inserted upstream of the adenovirus major late TATA box and initiator region within the CAT reporter plasmid pML-1 (Du et al. 1993). This enhancer segment contains all of the *cis*-acting elements shown previously to be required for efficient enhancer function in F9 cells (Dailey et al. 1994). As expected, CAT gene expression was specifically activated in undifferentiated F9 cells, whereas that of the enhancerless pML-1CAT construct was very low (data not shown). No activation by the FGF-4 enhancer DNA sequences was observed in transfected HeLa cells.

To directly assess the roles of Sox2 and the individual octamer-binding proteins in modulating F9-specific gene activation, CMV promoter-driven expression plasmids of Oct-1 (pCGOct-1, Tanaka and Herr 1990), Oct-3, or Sox2 were compared for their ability to activate expression of the fgfpML–CAT plasmid after transfection into HeLa cells. No activation of CAT gene expression was observed upon cotransfection with any of the individual octamer or Sox factor expression plasmids (Fig. 5). Likewise, cotransfection of CMV–Sox2 with Oct-1 expression plasmid failed to elicit CAT gene expression. In contrast, cotransfection of CMV–Sox2 with CMV–Oct-3 resulted in an ~11-fold induction of CAT activity (Fig. 5). To determine whether other Sox factors, which would recognize the same consensus sequence of the FGF-4 enhancer DNA, were capable of *trans*-activation in combination with Oct-3, we tested a CMV promoter-driven plasmid expressing Sox5 (Denny et al. 1992). Sox5 was incapable of activating expression of the fgfpML–CAT plasmid in the HeLa cells alone or in conjunction with Oct-3 (Fig. 5). These results therefore indicate that only a complex formed by the combination of the tissue-specific Oct-3 and Sox2 factors can elicit transcriptional activity from the FGF-4 enhancer.

The results presented above along with our previous demonstration (Dailey et al. 1994) that FGF-4 enhancer activity required the Sox and octamer motifs strongly supported the notion that EC-specific transcription of the FGF-4 gene is primarily dictated by these two DNA-binding motifs. To test this directly, multimers of the 22-bp FGF-4 enhancer segment containing only these two binding sites (enhancer positions 118–139, Dailey et al. 1994) were inserted either 130 bp upstream [pSVCat(S/O)5′] or 1710 bp downstream [pSVCat(S/O)3′] of the transcription start site within an SV40 promoter–CAT gene plasmid (Fig. 6A). Analysis of CAT gene expression from these constructs in differentiated and undifferentiated F9 or HeLa cells revealed that the Sox and octamer motifs mimicked the behavior of the FGF-4 enhancer, as they were both necessary and sufficient to act as an F9-specific, position-independent enhancer element (Fig. 6A). Transfection of HeLa cells with these reporter plasmids and CMV–Sox2, CMV–Oct-1, or CMV–Oct-3 (Fig. 6B) demonstrated that the simple multimers of octamer- and Sox factor-binding motifs responded to the expression of these *trans*-acting proteins in a manner analogous to the natural enhancer. Only cotransfection of pSVCat(S/O)5′ or pSVCat(S/O)3′ with Sox2 and Oct-3 expression plasmids led to transcriptional activation (an approximately eightfold induction) whereas Sox2, Oct-1, Oct-3, or Sox2 plus Oct-1 was ineffective (Fig. 6B). Importantly, specific activation by Oct-3 and Sox2 was also observed using the pSVCat(S/O)3′ reporter construct, indicating that activation by the Sox2/Oct-3 complex is at least partially independent of location (Fig. 6B, but see Discussion). These results demonstrate that the octamer- and Sox factor-binding motifs can together act as a minimal EC cell-specific enhancer by virtue of their ability to recruit both Oct-3 and Sox2. The ability of Oct-3, but not Oct-1, and of Sox2, but not

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

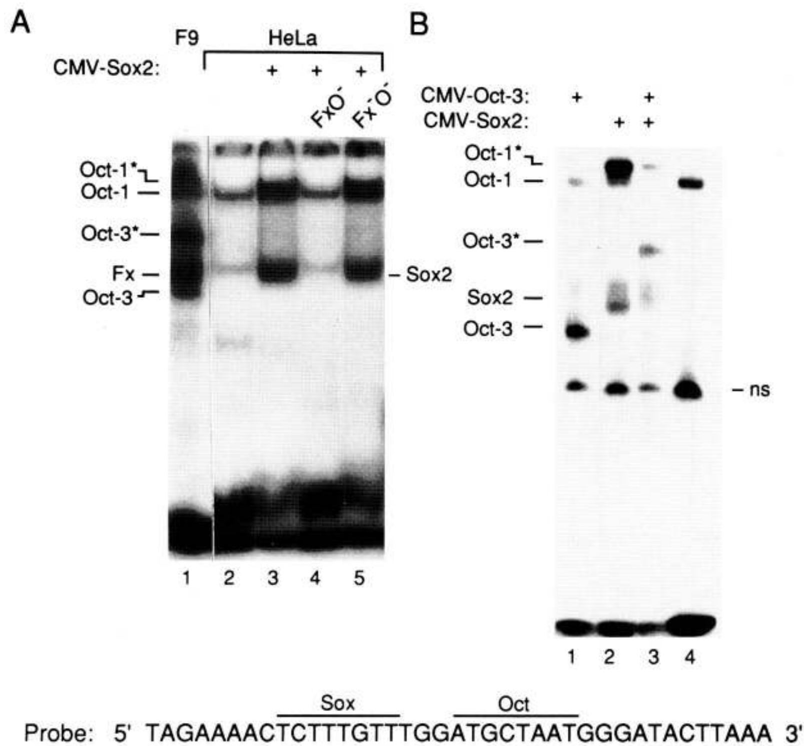


Figure 4. Comparison of Sox2 and Fx interaction with FGF-4 enhancer DNA sequences. (*A*) Five microliters of nuclear extracts prepared from CMV–Sox2-transfected (lanes *3*–*5*), or untransfected (lane *2*) HeLa cells, or 0.5 λ of F9 nuclear extract (10 mg/ml, lane *1*), was incubated with a radiolabeled FGF-4 enhancer DNA sequence oligonucleotide that spans both the Fx- and octamer-binding sites as shown. The sample in lane *4* also contains 2 ng of unlabeled FGF-4 enhancer oligonucleotide containing mutation of the octamer-binding site (FxO⁻; Dailey et al. 1994), whereas that of lane *5* includes a competitor oligonucleotide with mutations of both the Fx and octamer-binding sites (Fx⁻O⁻). Samples were resolved by electrophoresis in a 5% native polyacrylamide gel as described. (*B*) Sox2 induces formation of Oct3*. HeLa cells were transfected with CMV–Oct3 (lane *1*), CMV–Sox2 (lane *2*), both CMV–Oct-3 and CMV–Sox2 (lane *3*), or with the CMV vector alone (lane *4*). Nuclear extracts prepared from the transfected cells were analyzed by EMSA as in *A*. (N.S.) A nonspecific DNA-binding activity observed in some extracts.

Sox5, to activate CAT gene transcription from these reporter constructs demonstrates that a specific combination of unique domains within Sox2 and Oct-3 generates a new, higher order complex with distinct activation properties and that assembly of this activation complex is directed by the octamer- and Sox factor-binding motifs present within the FGF-4 enhancer. This suggests that FGF-4 gene expression in F9 cells requires the activities of the Sox2–Oct-3 complex bound at the enhancer and that FGF-4 is an embryonic target gene for these factors.

The Sox2 HMG domain is not sufficient for activation of the FGF-4 enhancer

The HMG domains of Sox2 and Sox5 share 53% amino acid homology (Denny et al. 1992), whereas the amino acid composition of these two proteins is completely divergent outside of this domain. One possibility raised by

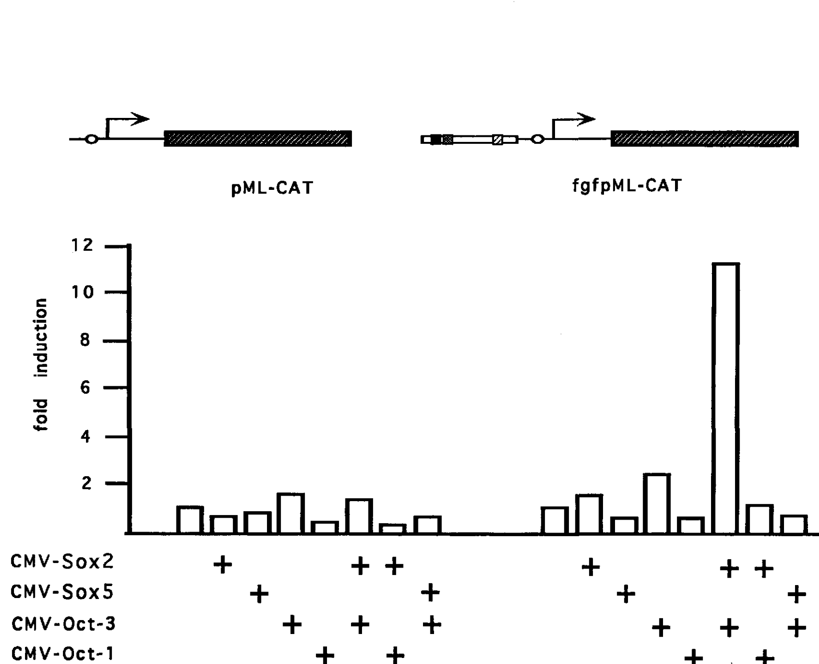


Figure 5. Activation of the FGF-4 enhancer requires Sox2 and Oct-3. Schematic representations of the pML–CAT and fgfpML–CAT reporter constructs are shown at the *top*. The line indicates sequences (−45 to +65) of the adenovirus major late basal promoter; the position of the TATA box is shown as an open circle, and the transcriptional start site is indicated by the arrow. The 115-bp FGF-4 enhancer fragment is shown by the open rectangle containing the binding sites for Sox2 (solid box), octamer-binding proteins (shaded hatched box), and Sp1 (open hatched box). CAT-coding sequences (not drawn to scale) are indicated by the shaded hatched rectangle. HeLa cells were transfected with 5 μg of either reporter construct in the presence of various combinations of 1 μg of each of the CMV expression plasmids as indicated by the + below the histogram. CAT activity was determined ~60 hr posttransfection and quantitated by PhosphorImager (Molecular Dynamics). CAT activity generated by each transfected reporter construct alone was give the value of 1. The values shown represent the average of at least three experiments.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

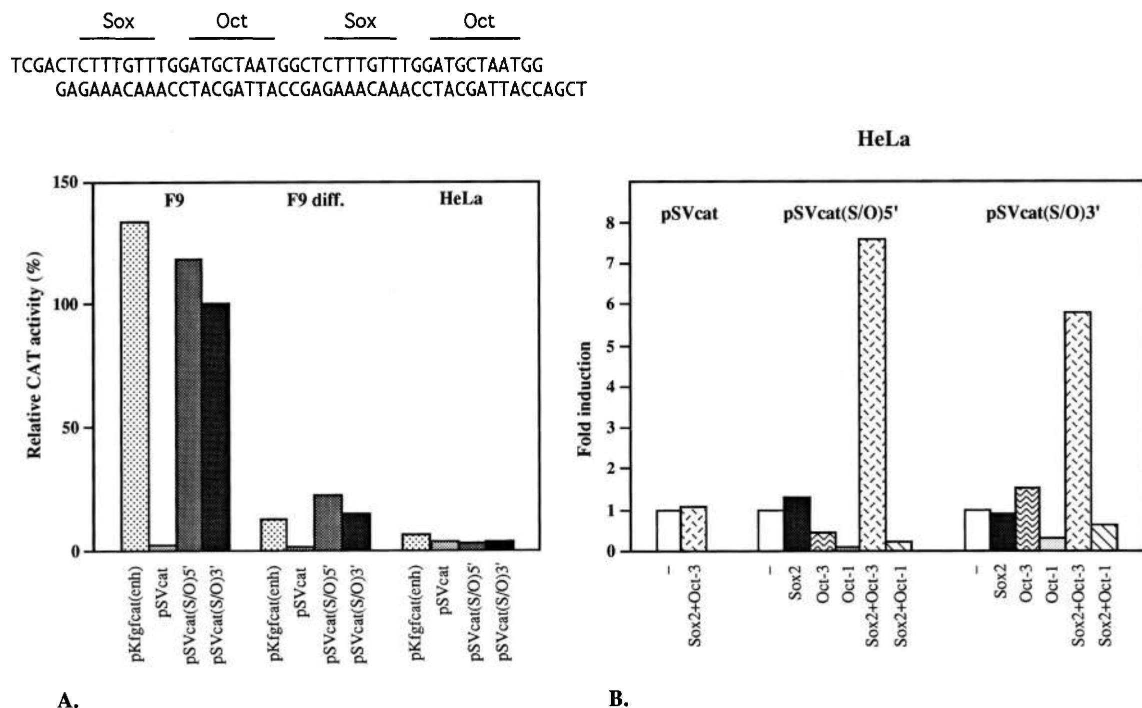


Figure 6. Expression and *trans*-activation of CAT plasmids under the control of a synthetic FGF-4 enhancer in F9 and HeLa cells. (*A*) Cells were transfected with 5 μg of each of the indicated plasmids using the calcium phosphate method as described previously (Curatola and Basilico 1989). The total amount of DNA transfected was brought to 25 μg by the addition of salmon sperm or mouse genomic DNA. CAT activity is reported as percent conversion relative to pRSVCAT. pKfgfcat(enh) contains the human FGF-4 promoter upstream and the 380-bp murine FGF-4 enhancer downstream of the CAT gene (Curatola and Basilico 1990). pSVCat (Promega) contains the SV40 promoter upstream of the CAT gene (no enhancer). pSVcat(S/O)5′ contains 12 copies of the FGF-4 murine enhancer sequences containing the Sox- and Oct-binding sites cloned 5′ to the SV40 promoter of pSVCat. pSVCat(S/O)3′ contains six copies of the Sox/Oct binding sequences cloned 3′ of the CAT gene of pSVCat as described in Materials and methods. The sequence shown at the *top* represents the repeated FGF-4 enhancer unit multimers of the Sox/Oct binding sites present in the pSVCat(S/O) plasmids. (*B*) Five micrograms of pSVCat, pSVCat(S/O)5′, or pSVCat(S/O)3′ plasmid DNA was transfected into HeLa cells alone or together with CMV plasmids expressing Sox2, Oct-1, or Oct-3 as indicated. Results are presented as the ratio between the CAT activity obtained from extracts of cells cotransfected with each CAT reporter plasmid and the indicated Sox2, Oct-1, or Oct-3 expression plasmids and the CAT activity from extracts of cells transfected with the reported constructs alone. Each of the plasmids expressing Sox2, Oct-1, or Oct-3 (5 μg) was used in cotransfections with pSVCat(S/O)5′, whereas 1 μg of each expression plasmid was used for pSVCat(S/O)3′. The results shown are the average of three independent experiments. Sox2, Oct-3, and Oct-1 represent plasmids in which the expression of Sox2, Oct-3, and Oct-1, respectively, is under the control of a CMV promoter.

the observation that Sox5 is unable to activate fgfCAT gene expression (Fig. 5) is that a unique regions of Sox2 other than the HMG domain is required for this activity. A deletion mutant of Sox2, NP2, in which the carboxy-terminal portion of Sox2 downstream of amino acid 179 is removed but the HMG domain remains intact, was tested for its ability to activate fgfpML–CAT gene expression in HeLa cells with cotransfected Oct-3. As shown in Figure 7, the NP2- truncated protein was localized to the nucleus, produced in amounts comparable to wild-type Sox2, and bound the FGF-4 enhancer oligonucleotide probe in a manner comparable to wild-type Sox2. However, the NP2 protein was unable to activate CAT gene expression from the fgfpML–CAT reporter (Fig. 7C). These results indicate that the HMG domain of Sox2 is not sufficient for activation of the FGF-4 enhancer and suggest the contribution of a separate, non-DNA-binding domain within Sox2 to its function as an activator.

Discussion

In this report we have described the full-length cDNA cloning and activation properties of Sox2. Although the amino acid sequence of HMG domains among the many Sox factors are generally >60% homologous (for review, see Laudet et al. 1993), we find no significant similarities with Sry, Sox4, or Sox5 (Sinclair et al. 1990; Denny et al. 1992; van de Wetering and Clevers 1992) outside of this domain. The expression pattern of Sox2 was found to be highly tissue restricted. Sox2 displays all of the DNA-binding properties of the activity we identified previously as Fx and interacts with an FGF-4 enhancer DNA element that is essential to enhancer function. Furthermore, Sox2 can form ternary complexes with Oct-1 or Oct-3 but acts as a transcriptional activator of FGF-4 enhancer-containing reporter constructs only in conjunction with Oct-3. One prediction of these results was that overexpression of Oct-1 would compete with Oct-3

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

array of unique functional domains. Furthermore, the enhancer itself provides a specific arrangement of binding sites for these factors that leads to assembly and tethering of a transcriptionally productive complex.

The mechanism of activation by Sox factors and the identity and role of different domains in this activation still remain to be clarified. The HMG domain of the Sox factors, as well as that of the closely related protein LEF-1, induces a strong bend upon binding to the DNA (Ferrari et al. 1992; Giese et al. 1992); thus, the role of these factors may be primarily architectural and facilitate protein–protein interactions at the promoter or enhancer. A central role for the DNA-binding and bending functions of Sry has been underscored by the observation that there is little evolutionary conservation of amino acid sequence outside of the Sry HMG domain and all mutations leading to sex reversal map within the HMG domain of Sry (Berta et al. 1990). Additionally, previous analyses of Sry and Sox5 demonstrated that these proteins are unable to activate transcription from reporter constructs containing multimers of a Sox DNA-binding site, supporting the notion that transcriptional activation mediated by Sry and Sox5 requires multiple protein–protein interactions (for review, see Grosschedl et al. 1994) . Likewise, LEF-1 is unable to activate transcription from the T-cell receptor α (TCRα) enhancer by itself (Waterman and Jones 1990; Travis et al. 1991). In this report we have shown that transcriptional activation by Sox2 along with Oct-3 on the FGF-4 enhancer requires a domain located outside of the HMG domain of Sox2 and, thus, that the DNA-binding functions of this protein are by themselves not sufficient for full factor activity (Fig. 7). A similar result was observed for LEF-1 (Carlsson et al. 1993; Giese and Grosschedl 1993). Although more precise mapping of this domain within Sox2 is in progress, it is intriguing to note that the region deleted in the NP2 mutant (amino acids 179–318) includes a sequence (Fig. 1) that is 32% serine (amino acids 210–262) and harbors a subregion (amino acids 217–255) sharing significant homology with the activation domain of LEF-1 (Kamachi et al. 1995). In this regard it is noteworthy that Sox5, which is unable to activate fgfCAT gene expression (Fig. 5), also contains a serine-rich region, but its amino acid composition is completely unrelated to that of Sox2 (Denny et al. 1992; Fig. 1). The possible role of this Sox2 domain in protein-protein interactions, stabilization of DNA binding, or *trans*-activation function is currently under investigation.

A recent report by Zwilling et al. (1995) has demonstrated that HMG2, a distantly related ancestor of the Sox factors, can activate transcription mediated by Oct-1 and Oct-2. In contrast to Sox proteins, HMG2 binds DNA in a relatively non-sequence-specific manner and apparently potentiates octamer protein-mediated transcriptional activation by direct protein–protein interactions occurring between the HMG and POU domains. Analogous interactions were not observed using Sry or LEF-1 and therefore do not reflect a general feature common to all HMG domains. It is tempting to speculate that the productive interaction of the proteins that we have described may, in part, represent the evolutionary refinement of a more general, ancient interaction between HMG and POU domain proteins.

One issue raised by this study is the molecular basis of the selective action of Sox2 with Oct-3 rather than Oct-1. The highly divergent sequences of Oct-3 outside of the POU-specific and POU homeo domains could play a role in establishing a direct interaction with Sox2 or other transcription factors necessary to activate the basal transcription machinery or, alternatively, may contribute a *trans*-activation domain that is functional only when juxtaposed to complementary domains within Sox2. A proline-rich region noted previously within Oct-3 and not present in Oct-1 is one candidate for this activity (Okamoto et al. 1990; Rosner et al. 1990; Schöler et al. 1990b). Alternatively, or additionally, differences within the conserved POU domain could be important. This latter possibility is supported by studies on the activation of herpes simplex virus immediate early gene transcription by the specific interaction of the viral factor VP16 with Oct-1 but not Oct-2. The basis of this selective interaction has been attributed to a single amino acid difference within the POU domains of Oct-1 and Oct-2 (Lai et al. 1992; Pomerantz et al. 1992). It is also possible that the DNA-bending property of Sox2 (L. Dailey, unpubl.) is affected differently by, or elicits different conformational consequences on, Oct-3 and Oct-1.

The observation that Sox2 acts specifically with Oct-3 but not Oct-1 illustrates one mechanism by which selective gene activation can be achieved by individual members within these families. Other studies, particularly those concerning the B cell-specific expression of immunoglobulin genes, have also revealed that transcriptional activation by octamer-binding proteins requires additional, cell-specific activities and that some of these can discriminate among the different octamer-binding factors. For example, the B-cell factor originally identified as OCA-B, serves as a coactivator of either Oct-1 or Oct-2 and is required for transcriptional activation of reporter genes containing an octamer element located proximal to the TATA box (Luo et al. 1992; Staiger et al. 1995; Strubin et al. 1995). In this case, it is OCA-B, rather than Oct-2, that confers tissue-specific promoter activity through the octamer motif. However, separate studies have shown that transcriptional activation of reporter constructs containing octamer elements located at a remote position from the TATA box require a tissue-specific activator that functions selectively with a specific octamer-binding protein. Such an activity, observed in B cells, is believed to be distinct from OCA-B and functions with Oct-2 but not Oct-1 (Pfisterer et al. 1994). Likewise, the adenovirus E1A protein will activate reporter gene transcription through remote octamer motifs preferentially in conjunction with Oct-3 and is believed to be analogous to a cellular factor present in F9 cells (Schöler et al. 1991). These two still uncharacterized activities are thus functionally similar to Sox2, both with respect to their role as cell type-specific transcriptional activators and their selective action with particular octamer-binding proteins.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

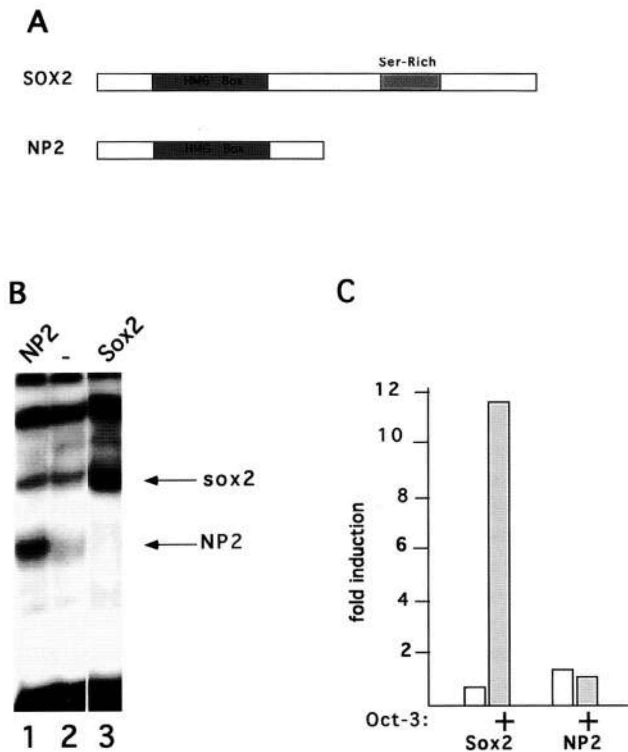


Figure 7. The Sox2 HMG domain is not sufficient for *trans*-activation. (*A*) Schematic representation of Sox2 and the NP2 mutant cDNAs. Shown are the relative location of the HMG and serine-rich regions. (*B*) EMSA. Nuclear extracts prepared from HeLa cells (lane *2*) or HeLa cells transfected with CMV-Sox2 (lane *3*) or NP2 (lane *1*) were analyzed by EMSA using the FGF-4 enhancer oligonucleotide probe shown in Fig. 4. (*C*) The fgfpML–CAT reporter plasmid was transfected into HeLa cells with either CMV–Sox2 or NP2 DNA in the presence (+) or absence of CMVOct-3 as indicated. CAT activity was assessed as described in Materials and methods.

for octamer site binding on the FGF-4 enhancer and interfere with enhancer activation. Preliminary results (data not shown) indicate that this is the case. We were able to observe transcriptional activation of reporter constructs that contained a single, proximally located FGF-4 enhancer as well as reporters that contained simple multimers of the octamer- and Sox factor-binding motifs located either proximal or distal to the transcriptional start site. However, attempts to activate transcription from constructs containing the natural enhancer at a distal 3′ position were not successful (data not shown). This may reflect an inappropriate factor stoichiometry in the transfected HeLa cells or the requirement for additional, perhaps F9-specific, factors that in some way mediate activation of this construct by the Sox2/Oct-3 complex located at a distance. Nonetheless, these findings, together with our previous mutagenesis studies (Dailey et al. 1994), demonstrate that EC-specific transcriptional activation of the FGF-4 gene is determined by a synergistic interaction between Oct-3 and Sox2 at the FGF-4 enhancer.

The identity of the activity that we previously designated Fx as Sox2 relies on several observations. (1) Sox2 binds the same DNA sequences recognized by Fx; (2) The functional ability of Sox2 to *trans*-activate the FGF-4 enhancer is clearly demonstrated by cotransfection assays using HeLa cells; (3) Antiserum (αHSox2) raised against the HMG domain of Sox2 also recognizes Fx, establishing that Fx is a Sox factor; (4) αHSox2 antiserum immunoprecipitates what is apparently a single protein from F9 nuclear extracts that corresponds in molecular mass (35 kD) to that predicted by the cDNA sequence of Sox2 (which suggests that only one Sox factor is present in F9 nuclear extracts and/or only this one species is recognized by αHSox2); (5) Sox2 protein–DNA complexes exhibit the same mobility as Fx–DNA complexes in nondenaturing gels, suggesting that the molecular mass of Fx corresponds to that of Sox2. Although it is formally possible that Fx represents an additional 35-kD Sox factor present in F9 nuclear extracts, this is not the published molecular mass of murine Sox4, Sox5, and Sox9 and human Sox9 or chicken Sox11 (Denny et al. 1992; van DeWetering et al. 1993; Tommerup et al. 1993; Wright et al. 1993, 1995; Uwanogho et al. 1995). One exception might be Sox3 because the recently published chicken Sox3 cDNA sequence encodes a protein of 315 amino acids (Uwanogho et al. 1995). However, the sequence of the murine Sox3 homolog is not yet available and it is also not known whether Sox3 is expressed in EC or ES cells. In this regard, in spite of extensive efforts to isolate additional Sox factor cDNAs from our F9 cDNA library, we consistently only obtained Sox2. Therefore, although we cannot completely exclude the possibility that Sox3 contributes to the activity we have named Fx, the simplest interpretation of our results is that Fx is Sox2.

The specific activation by Oct-3 and Sox2 described in this report represents the first demonstration of a corequirement for factors from these two important developmental transcription factor families. The Sox- and octamer-binding factor families are each defined by the presence of highly conserved DNA-binding domains (Sox–HMG box and POU homeo domain, respectively (Herr et al. 1988; Gubbay et al. 1990; Sinclair et al. 1990), and members within each group bind to the same DNA target elements with comparable affinity. For this reason it is unlikely that gene-specific activation by an individual Sox or octamer-binding factor is primarily determined by differential DNA binding. Rather, because transcriptional activation generally results from the combinatorial interaction of multiple factors (Tjian and Maniatis 1994), unique activator functions for an individual Sox- or octamer-binding protein are likely to be revealed only in the presence of specific factor partners through direct or indirect, protein–protein interactions. A fundamental aspect of this mechanism is the central role played by enhancer DNA sequences to recruit and assemble these biologically relevant transcription complexes. Thus, the tissue-specific activation of the FGF-4 enhancer by Oct-3 and Sox2 reflects the modular nature of the factors and that each is composed of a particular

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

Finally, our results contain important implications for the expression of the FGF-4 gene during embryogenesis. The enhancer element that we have analyzed is active in F9 cells that are comparable to cells of the blastocyst inner cell mass. At this stage, FGF-4, Sox2 (ES cell RNA, Fig. 2), and Oct-3 are all expressed (Rosner et al. 1990; Schöler et al. 1990a; Niswander and Martin 1992), consistent with the notion that Oct-3 and Sox2 determine expression of FGF-4 in the early embryo through interactions at the enhancer that we have defined. However, it is not clear at the moment whether FGF-4, Oct-3, and Sox2 are coexpressed at later embryonic stages (e.g., the limb bud), and some data (Fig. 2; Rosner et al. 1990; Schöler et al. 1990a; Niswander and Martin 1992) actually suggest that this may not be the case. Thus, it is quite possible that FGF-4 gene activation at later embryonic stages may be accomplished by different Sox and octamer partners or possibly even by distinct enhancer elements. Future efforts will be directed toward the elucidation of these novel factors and/or enhancer elements. Thus, we believe that the FGF-4 enhancer analysis using EC cells has defined an element that is specific to expression of this gene in early embryogenesis by virtue of the specific functions provided by the combination of Oct-3 and Sox2.

Materials and methods

Plasmids, cell culture, and transfections

Cell culturing of F9 and HeLa cells was performed as described elsewhere (Dailey et al. 1994). $CaPO_4$ transfections, performed essentially as described in Curatola and Basilico (1990), generally utilized 5 μg of reporter plasmid and 1 or 5 μg of expression plasmid in a total of 25 μg of DNA (from salmon sperm or mouse) as indicated for individual experiments. Transfections additionally utilized the CMV–βGal plasmid for normalization of transfection variations.

The pRSVCAT plasmid and pKfgfcat(enh), that contains the human FGF-4 promoter upstream and the 380-bp murine FGF-4 enhancer downstream of the CAT gene have been described elsewhere (Curatola and Basilico 1990). The Sox2 expression plasmid CMV–Sox2 contains the 2-kb *Not*I–*Xho*I fragment (nucleotides 329–2303) from Bluescript–Sox2 plasmid inserted downstream of the CMV promoter within plasmid pCEP4 (Invitrogen). The cloned fragment contains the entire Sox2-coding region as well as some 3′- and 5′-untranslated sequences. To generate the Sox2 mutant NP2, pCEP4 Sox2 (CMV–Sox2) was digested at the *Bam*HI site within the plasmid polylinker just 3′ to the Sox2 cDNA insert and then subjected to partial digestion with *Pvu*II. The ends were filled in using the Klenow fragment of DNA polymerase I, and the reaction products were ligated using T4 DNA ligase under conditions favoring recircularization of the digested DNA. After transformation, plasmid DNAs were enzymatically screened for deletion of the DNA segment between the *Pvu*II site at nucleotide 895 within the Sox2 cDNA sequence (Fig. 1) and the *Bam*HI site downstream of the cDNA insert. This Sox2 variant contains Sox2 cDNA sequences between the *Not*I and *Pvu*II sites and thus is designated as mutant NP2. The entire Oct-3-coding region located within a 1.4-kb *Bam*HI–*Xho*I fragment (nucleotides 1–1339 of the published Oct-3 nucleotide sequence (Okamoto et al. 1990) of Oct-3 cDNA was inserted downstream of the CMV promoter in plasmid pCEP4 (Invitrogen) to create CMV–Oct-3. The Oct-1 expression plasmid pCGOct-1 and the Sox5 expression plasmid pd740 were generously provided by Winship Herr and Alan Ashworth, respectively (Tanaka and Herr 1990; Denny et al. 1992). To construct the fgfpML–CAT reporter plasmid, oligonucleotide primers, including positions 118 (5′ primer) and 139 (3′ primer) of the FGF-4 enhancer (Dailey et al. 1994), were used as primers for PCR. The 5′ primer contained the recognition site for *Hin*dIII, whereas the 3′ primer contained the recognition site for *Sal*I. After *Hin*dIII and *Sal*I digestion, the gel-purified 116-bp PCR product was inserted into the *Hin*dIII and *Sal*I sites of the pMLCAT plasmid (Du et al. 1993) that contains the basal promoter elements of the adenovirus major late promoter upstream of CAT-coding sequences. To construct the FGF-4 enhancer–CAT reporter plasmids pSVcat(S/O)5′ and pSVcat(S/O)3′, oligonucleotides (shown in Fig. 6A) consisting of two copies of the 22-bp FGF-4 enhancer segment encompassing the Sox factor- and octamer-binding motifs and *Sal*I recognition sites on each end were synthesized using the Applied Biosystem 392 DNA/RNA synthesizer. These oligonucleotides were annealed, phosphorylated, and multimerized by self-ligation. The resulting multimers were cloned upstream of the SV40 promoter at the *Bgl*II site of pSVCat (Promega) to create pSVcat(S/O)5′. Alternatively, the ends of the self-ligated multimers were filled in with Klenow and inserted at the *Sal*I site (that had also been filled in after digestion) downstream of the CAT-coding sequences within pSVCat to create pSVcat(S/O)3′. The number of oligonucleotide repeats were determined, by DNA sequencing (Sanger et al. 1977), to be 12 within pSVcat(S/O)5′ and 6 within pSVcat(S/O)3′.

Cloning of Sox2 cDNA

Twenty-five micrograms of reverse-transcribed RNA from P19 cells or phage supernatant from a λZAP (Stratagene) F9 cDNA library were amplified by PCR using 100 pmoles of the degenerate Sox HMG box oligonucleotide primers described in Figure 1. The 5′ primer contained the recognition site for *Bam*HI, whereas the 3′ primer contained recognition sites for either *Bam*HI or *Eco*RI. The PCR consisted of 25 cycles of 1 min at 94°C, 2 min at 37°C, and 2 min at 72°C. Amplification products (200–250 bp) were gel purified after electrophoresis and eluted from a native 5% polyacrylamide gel, digested with *Bam*HI, inserted into the Bluescript SK vector, and sequenced. All but 1 of 30 inserts that were found to contain HMG box sequences encoded a peptide corresponding to the published DNA sequence of the Sox2 HMG box (Wright et al. 1993); one clone contained a mutant variant of the Sox2 HMG box that was probably generated during the PCR. The Sox2 HMG box DNA sequences were radiolabeled by random priming and used as a probe to screen a λZAP cDNA library derived from undifferentiated F9 cell mRNA. Fifteen positive clones were isolated. Insert DNAs larger than 2 kb were excised from 10 positive clones according to the manufacturer's protocol (Stratagene) and subjected to dideoxy sequencing using a series of oligonucleotide primers.

RNA analysis

RNA isolation, electrophoresis, and Northern blotting were as described previously (Curatola and Basilico 1990).

Antibody production and immunoprecipitation analysis

Sox2 HMG box DNA sequences generated by PCR amplification using primers containing a recognition site for *Bam*HI (5′ primer) or *Eco*RI (3′ primer), as described above, were digested with *Bam*HI and *Eco*RI and inserted, in-frame, into the *Bam*HI and *Eco*RI sites of the pGEX-2T expression plasmid (Promega).

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

Overexpressed GST–Sox2 HMG fusion protein was purified according to the manufacturer's instructions. Thirty micrograms of purified protein was coupled to aluminum hydroxide (AluS-gel, SERVA) and injected into a rabbit. Boosts were performed at 1-month intervals, and sera collected 10–14 days after each injection. IgG from the preimmune and immune antisera was selected by protein A–agarose chromatography.

Immunoprecipitation

Undifferentiated or RA-treated F9, or HeLa cells were incubated in the presence of [^{35}S]methionine (0.5 mCi/ml) for 3 hr, and lysates were generated using RIPA buffer (0.15 M NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 0.05 M Tris-HCl at pH 7.5). Extracts were normalized to equivalent cell number and incubated with 3 μl of preimmune or αHSox2 IgG and the complexes precipitated using protein A–agarose beads. Immunoprecipitated proteins were resolved by SDS-PAGE on a 10% polyacrylamide gel along with protein standards (Amersham) and visualized by autoradiography.

Nuclear extracts preparation, chromatography, and DNA-binding assays

To prepare small-scale extracts from transfected cells, the plates were washed with PBS ~60 hr after transfection and collected (in 1 ml of PBS) into an Eppendorf tube. After brief centrifugation, the cell pellet was resuspended in hypotonic buffer (Dignam et al. 1983), incubated on ice for 10 min, Dounce homogenized, and spun in a microcentrifuge for 5 min. The nuclear pellet was resuspended in 50 μl of BC350N (20 mM HEPES at pH 7.9, 350 mM KCl, 20% glycerol, 0.01% NP-40, 0.2 mM PMSF, 0.5 mM DTT) and incubated for 30 min at 4°C with rotation. After a 5-min centrifugation, the supernatant was withdrawn and used for EMSA. Large-scale preparations of nuclear extracts were performed as described previously (Dailey et al. 1994). Partial purification of Fx was achieved by fractionation of F9 nuclear extracts using ammonium sulfate (to 30% saturation). The precipitated proteins were resuspended in BC50N buffer and applied to a desalting column (p-6DG resin, BioRad) equilibrated previously in BC50N. This fraction was subjected to chromatography on an S-Sepharose column, and Fx activity was eluted using 0.5 M KCl in BCN buffer. Fx activity was monitored by EMSA during the course of the fractionation.

EMSAs were performed as described previously (Dailey et al. 1994) either in the presence or absence of oligonucleotide competitor DNAs or antisera described for individual figures.

CAT assays

Extracts prepared for CAT assay were normalized either by β-galactosidase activity or protein concentration (BioRad). CAT assays were performed essentially as described in Curatola and Basilico (1990).

Acknowledgments

We thank Winship Herr for kindly providing the pCG Oct-1 plasmid, A. Ashworth for the gift of plasmid pd740, Beth Moorefield, David Levy, and Alka Mansukhani for helpful comments on the manuscript, R. Trautmann and D. Ambrosetti for discussions, and Vera Levytska for excellent technical assistance. Some of the oligonucleotides used in this study were synthesized by the Biopolymer facility of the Rockefeller University. This investigation was supported by U.S. Public Health Service grant CA42568 (C.B.) and CA49472 (L.D.).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Basilico, C. and D. Moscatelli. 1992. The FGF family of growth factors and oncogenes. *Adv. Cancer Res.* **59:** 115–163.

Berta, P., J.R. Hawkins, A.H. Sinclair, A. Taylor, B.L. Griffiths, P.N. Goodfellow, and M. Fellous. 1990. Genetic evidence equating SRY and the testis-determining factor. *Nature* **348:** 448–450.

Carlsson, P., M.L. Waterman, and K.A. Jones. 1993. The hLEF/TCF-1α HMG protein contains a context-dependent transcriptional activation domain that induces the TCRα enhancer in T cells. *Genes & Dev.* **7:** 2418–2430.

Curatola, A.M. and C. Basilico. 1990. Expression of the K-fgf proto-oncogene is controlled by 3′ regulatory elements which are specific for embryonal carcinoma cells. *Mol. Cell. Biol.* **10:** 2475–2484.

Dailey, L., H. Yuan, and C. Basilico. 1994. Interaction between a novel F9-specific factor and octamer-binding proteins is required for cell-type-restricted activity of the fibroblast growth factor 4 enhancer. *Mol. Cell. Biol.* **14:** 7758–7769.

Delli Bovi, P., A.M. Curatola, F.G. Kern, A. Greco, M. Ittmann, and C. Basilico. 1987. An oncogene isolated by transfection of Kaposi's sarcoma DNA encodes a growth factor that is a member of the FGF family. *Cell* **50:** 729–737.

Denny, P., S. Swift, F. Connor, and A. Ashworth. 1992. An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA binding protein. *EMBO J.* **11:** 3705–3712.

Dignam, J.D., R.M. Lebovitz, and R.G. Roeder. 1983. Accurate transcription initiation by polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11:** 1475–1489.

Du, H., A.L. Roy, and R.G. Roeder. 1993. Human transcription factor USF stimulates transcription through the initiator elements of the HIV-1 and the Ad-ML promoters. *EMBO J.* **12:** 501–511.

Feldman, B., W. Poueymirou, V.E. Papaioannou, T.M. DeChiara, and M. Goldfarb. 1995. Requirement of FGF-4 for postimplantation mouse development. *Science* **267:** 246–250.

Ferrari, S., V.R. Harley, A. Pontiggia, P.N. Goodfellow, R. Lovell-Badge, and M.E. Bianchi. 1992. SRY, like HMG 1, recognizes sharp angles in DNA. *EMBO J.* **11:** 4497–4506.

Foster, J.W., M.A. Dominguez-Steglich, S. Guioli, G. Kowk, and P.A. Weller. 1994. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* **372:** 525–530.

Giese, K. and R. Grosschedl. 1993. LEF-1 contains an activation domain that stimulates transcription only in a specific context of factor-binding sites. *EMBO J.* **12:** 4667–4676.

Giese, K., J. Cox, and R. Grosschedl. 1992. The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* **69:** 185–195.

Grosschedl, R., K. Giese, and J. Pagel. 1994. HMG domain proteins: Architectural elements in the assembly of nucleoprotein structures. *Trends Genet.* **10:** 94–100.

Gubbay, J., J. Collignon, P. Koopman, B. Capel, A. Economou, A. Munsterberg, N. Vivian, P.N. Goodfellow, and R. Lovell-Badge. 1990. A gene mapping to the sex-determining region of the mouse Y-chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346:** 245–250.

Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, P.A. Sharp, H.A. Ingraham, M.G. Rosenfeld, M. Finney, G.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

Ruvkin, and H.R. Horvitz. 1988. The POU domain: A large conserved region in the mammalian Pit-1, Oct-1, Oct-2, and *Caenorhabditis elegans unc-86* gene products. *Genes & Dev.* **2:** 1513-1516.

Kamachi, Y., S. Sockanathan, Q. Liu, M. Breitman, R. Lovell-Badge, and H. Kondoh. 1995. Involvement of Sox proteins in lens-specific activation of crystallin genes. *EMBO J.* **14:** 3510–3519.

Lai, J.-S., M.A. Cleary, and W. Herr. 1992. A single amino acid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeo domain. *Genes & Dev.* **6:** 2058–2065.

Laudet, V., D. Stehelin, and H. Clevers. 1993. Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res.* **21:** 2493–2501.

Laufer, E., C.E. Nelson, R.L. Johnson, B.A. Morgan, and C. Tabin. 1994. Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. *Cell* **79:** 993–1003.

Luo, Y., H. Fujii, T. Gerster, and R.G. Roeder. 1992. A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. *Cell* **71:** 231–241.

Nasrin, N., C. Buggs, X.F. Kong, J. Carnazza, M. Goebl, and M. Alexander-Bridges. 1991. DNA-binding properties of the product of the testis-determining gene and a related protein. *Nature* **354:** 317–320.

Niswander, L. and G.M. Martin. 1992. Fgf-4 expression during gastrulation, myogenesis, limb and tooth development in the mouse. *Development* **114:** 755–768.

Niswander, L., C. Tickel, A. Vogel, I. Booth, and G.R. Martin. 1993. FGF-4 replaces the apical ectodermal ridge and directs outgrowth and patterning of the limb. *Cell* **75:** 579–587.

Okamoto, K., H. Okazawa, A. Okuda, M. Sakai, M. Muramatsu, and H. Hamada. 1990. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60:** 461–472.

Pfisterer, P., A. Annweiler, C. Ullmer, L.M. Corcoran, and T. Wirth. 1994. Differential transactivation potential of Oct1 and Oct2 is determined by additional B cell-specific activities. *EMBO J.* **13:** 1655–1663.

Pomerantz, J.L., T.M. Kristie, and P.A. Sharp. 1992. Recognition of the surface of a homeo domain protein. *Genes & Dev.* **6:** 2047–2057.

Rosner, M.H., M.A. Vigano, K. Ozato, P.M. Timmons, F. Poirer, P.W.J. Rigby, and L.M. Staudt. 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345:** 686–692.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.

Schöler, H.R., G.R. Dressler, R. Balling, H. Rohdewold, and P. Gruss. 1990a. Oct-4: A germline specific transcription factor mapping to the mouse t-complex. *EMBO J.* **9:** 2185–2195.

Schöler, H.R., R. Siegfried, N. Suzuki, K. Chowdhury, and P. Gruss. 1990b. New type of POU domain in germ line-specific protein Oct-4. *Nature* **344:** 435–439.

Schöler, H.R., T. Ciesiolca, and P. Gruss. 1991. A nexus between Oct-4 and E1A: Implications for gene regulation in embryonic stem cells. *Cell* **66:** 291–304.

Schoorlemmer, J. and W. Kruijer. 1991. Octamer-dependent regulation of the kFGF gene in embryonal carcinoma and embryonic stem cells. *Mech. Dev.* **36:** 75–86.

Sinclair, A.H., P. Berta, M.S. Palmer, J.R. Hawkins, B.L. Griffiths, D.J. Smith, J.W. Foster, A. Frischaux, R. Lovell-Badge, and P.N. Goodfellow. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346:** 240–244.

Staiger, M.G., L. Knoopfel, O. Georgiev, W. Schaffner, and C.M. Hovens. 1995. A B-cell coactivator of octamer-binding transcription factors. *Nature* **373:** 360–362.

Stevanovic, M., R. Lovell-Badge, J. Collignon, and P.N. Goodfellow. 1993. SOX3 is an X-linked gene related to SRY. *Hum. Mol. Genet.* **2:** 2013–2018.

Strubin, M., J.W. Newell, and P. Matthias. 1995. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* **80:** 497–506.

Taira, M., T. Yoshida, K. Miyagawa, H. Sakamoto, M. Terada, and T. Sugimura. 1987. cDNA sequence of a human transforming gene hst and identification of the coding sequence required for transforming activity. *Proc. Natl. Acad. Sci.* **84:** 2980–2984.

Tanaka, M. and W. Herr. 1990. Differential transcriptional activation by Oct-1 and Oct-2: Interdependent activation domains induce Oct-2 phosphoryolation. *Cell* **60:** 375–386.

Tjian, R. and T. Maniatis. 1994. Transcriptional activation: A complex puzzle with few easy pieces. *Cell* **77:** 5–8.

Tommerup, N., W. Schempp, P. Meinecke, S. Pedersen, L. Bolund, C. Brandt, C. Goodpasture, P. Guldberg, K.R. Held, H. Reinwein et al. 1993. Assignment of an autosomal sex reversal locus (SRA1) and campomelic dysplasia (CMPD1) to 17q24.3-q25.1. *Nature Genet.* **4:** 170–174.

Travis, A., A. Amsterdam, C. Belanger, and R. Grosschedl. 1991. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor α enhancer function. *Genes & Dev.* **5:** 880–894.

Uwanogho, D., M. Rex, E.J. Cartwright, G. Pearl, C. Healy, P.J. Scotting, and P.T. Sharpe. 1995. Embryonic expression of the chicken *Sox2*, *Sox3* and *Sox11* genes suggests an interactive role in neuronal development. *Mech. Dev.* **49:** 23–26.

van de Wetering, M. and H. Clevers. 1992. Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *EMBO J.* **11:** 3039–3044.

van de Wetering, M., M. Oosterwegel, K. van Norren, and H. Clevers. 1993. Sox-4 an Sry-like HMG protein, is a transcriptional activator in lymphocytes. *EMBO J.* **12:** 1247–3854.

Velcich, A., P. Delli-Bovi, A. Mansukhani, E.B. Ziff, and C. Basilico. 1989. Expression of the K-fgf protooncogene is repressed during differentiation of F9 cells. *Oncogene Res.* **5:** 31–37.

Wagner, T., J. Wirth, J. Meyer, B. Zabel, M. Held, J. Zimmer, J. Pasantes, F.D. Bricarelli, J. Keutel, E. Hustert, U. Wolf, N. Tommerup, W. Schemp, and G. Scherer. 1994. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* **79:** 1111–1120.

Waterman, M. and K. Jones. 1990. Purification of TCF-1α, a T-cell-specific transcription factor that activates the T-cell receptor cα gene enhancer in a context-dependent manner. *New Biol.* **2:** 621–636.

Wright, E.M., B. Snopek, and P. Koopman. 1993. Seven new members of the Sox gene family expressed during mouse development. *Nucleic Acids Res.* **21:** 744.

Wright, E., M.R. Hargrave, J. Christiansen, L. Cooper, J. Kun, T. Evans, U. Gangadharan, A. Greenfield, and P. Koopman. 1995. The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nature Genet.* **9:** 15–20.

Zwilling, S., H. König, and T. Wirth. 1995. High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. *EMBO J.* **14:** 1198–1208.

Downloaded from genesdev.cshlp.org on August 22, 2022 - Published by Cold Spring Harbor Laboratory Press

Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3.

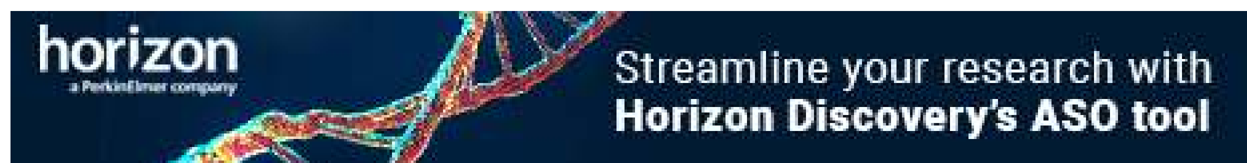
H Yuan, N Corbi, C Basilico, et al.

Genes Dev. 1995, **9:**
Access the most recent version at doi:10.1101/gad.9.21.2635

References This article cites 52 articles, 11 of which can be accessed free at:
http://genesdev.cshlp.org/content/9/21/2635.full.html#ref-list-1

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**



Copyright © Cold Spring Harbor Laboratory Press