Conserved POU Binding DNA Sites in the Sox2 Upstream Enhancer Regulate Gene Expression in Embryonic and Neural Stem Cells

Received for publication, May 18, 2004, and in revised form, July 6, 2004
Published, JBC Papers in Press, July 15, 2004, DOI 10.1074/jbc.M405514200

Raffaella Catena‡§, Cecilia Tiveron§%, Antonella Ronchi‡, Silvia Porta‡, Anna Ferri‡, Laura Tatangelo†, Maurizio Cavallaro†, Rebecca Favaro‡, Sergio Ottolenghi†, Rolland Reinbold|, Hans Schöler§§, and Silvia K. Nicolis‡**

From the ‡Department of Biotechnology and Biosciences, University of Milano-Bicocca, piazza della Scienze 2, Milano 20126, Italy, the §Transgenic Mice Service Center, Centro Ricerca Sperimentale, Istituto Regina Elena, via delle Messe d’Oro 156, Roma 00158, Italy, and the ¶Center for Animal Transgenesis and Germ Cell Research, New Bolton Center, Germline Development Group, School of Veterinary Medicine, University of Pennsylvania, Kennett Square, Pennsylvania 19348-1692.

The Sox2 transcription factor is expressed early in the stem cells of the blastocyst inner cell mass and, later, in neural stem cells. We previously identified a Sox2 5′-regulatory region directing transgene expression to the inner cell mass and, later, to neural stem cells and precursors of the forebrain. Here, we identify a core enhancer element able to specify transgene expression in forebrain neural precursors of mouse embryos, and we show that the same core element efficiently activates transcription in inner cell mass-derived embryonic stem (ES) cells. Mutation of POU factor binding sites, able to recognize the neural factors Brn1 and Brn2, shows that these sites contribute to transgene activity in neural cells. The same sites are also essential for activity in ES cells, where they bind different members of the POU family, including Oct4, as shown by gel shift assays and chromatin immunoprecipitation with anti-Oct4 antibodies. Our findings indicate a role for the same POU binding motifs in Sox2 transgene regulation in both ES and neural precursor cells. Oct4 might play a role in the regulation of Sox2 in ES (inner cell mass) cells and, possibly, at the transition between inner cell mass and neural cells, before recruitment of neural POU factors such as Brn1 and Brn2.

Neural stem cells are found throughout the developing embryo, where they are involved in the progressive generation of the mature central nervous system, and also in the adult brain, where neurogenesis persists in selected locations (1–3). In mammalian development, neural stem cells originate, soon after gastrulation, from the early, highly multipotent stem cells of the blastocyst inner cell mass, through sequential development of these cells into primitive ectoderm (epiblast) and neural plate (3, 4). This process, still poorly understood in molecular terms, has many features of a “default” developmental pathway, in embryogenesis (5) as well as in in vitro neural development of embryonic stem (ES) cells (6–8). Together with “general” neural properties (e.g. ability to generate neurons and glia), neural stem cells acquire (following the action of signaling centers) region-specific transcriptional programs (3, 9, 10) that will translate into region-specific characteristics of their differentiated neuronal and glial progeny.

The Sox2 transcription factor is expressed in the ES cells of the inner cell mass and its immediate descendant, the epiblast (11); with the beginning of neurogenesis, Sox2 expression becomes restricted to the neural primordium (neural plate) (10–13). Later, Sox2 is expressed in neural stem cells residing in the ventricular zone of the developing neural tube and in their immediate progeny through embryogenesis (10, 14) and into adulthood (15).

Recent experiments indicate that Sox2 plays critical functional roles within all of these stem/multipotent cell types. Stem cells constituting the early epiblast, which represents the multipotent precursors to all embryonic cell types, are lost in mouse embryos homozygous for a Sox2-null mutation, leading to early embryonic lethality (11); epiblast cells cease to proliferate and self-renew, and some of them instead turn on differentiation markers characteristic of trophoblast cells (11). Indeed, Sox2 was also shown to regulate genes, such as fibroblast growth factor 4 (16), essential for epiblast stem cell maintenance (17). The subsequent role of Sox2 in early neuroectoderm has not yet been addressed in mouse. However, in Xenopus, experiments with dominant-negative forms of Sox2 (interfering with the activity of Sox2 and the related Sox1 and 3 genes) have indicated a requirement for Sox family members in the generation of early neuroectoderm (18). Further, electroporation of Sox2 dominant-negative constructs in the chick embryo neural tube impairs the maintenance of proliferating, multipotent neural precursors and elicits their premature differentiation, whereas conversely, ectopic Sox2 (or Sox1) gene expression maintains the proliferative state of neural precursors and prevents their differentiation (14, 19). Finally, targeted mutations in mice decreasing, but not abolishing, Sox2 expression in neural cells cause an impairment of adult neural precursor cell proliferation and of neurogenesis, together with a depletion of early stem/precursor cells in the hippocampal adult neurogenic...
region, indicating a requirement of Sox2 for adult neural stem/ precursor cells function (15).

Although Sox2 is expressed panneurally, its transcriptional regulation in the central nervous system is guided from its early stages by region-specific regulatory elements (10, 20), suggesting that Sox2 might be differentially regulated by regional signals, and might control general neural properties (e.g., maintenance of a proliferative, self-renewal state) according to the specific developmental needs of the different regions of the nervous system.

In previous work, we identified a 5′-regulatory region of mouse Sox2 which drives expression of a β-gene transgene to neural stem cells and precursors of the developing (and adult) dorsal telencephalon (10). In transgenic mice, this region drives expression of the β-gene reporter already in the blastocyst inner cell mass, followed by restriction of expression to the neural plate and, progressively, to the telencephalon (10).

Here, we have investigated the sequence requirements of this regulation by transgenic analysis and ES cell transformation and, in parallel, the nuclear proteins binding to functionally relevant sequences by gel shift, footprinting, and chromatin immunoprecipitation studies. We find that conserved Pou factor binding sites in the upstream Sox2 enhancer are critical for transgenic Sox2 expression in both embryonic and neural stem/ precursor cells.

EXPERIMENTAL PROCEDURES

DNA Constructs

Modified Fragments of the HSI (Chromatin DNase I-hypersensitive Site 1) Region—Progressively smaller fragments encompassing the HSI region (Fig. 1A) were amplified by PCR from genomic subclones, with primers complementary to the extremities of the different fragments (0.4a, 0.4b, 0.1, Fig. 1, A and B) and carrying a terminal (5′ or 3′) HindIII restriction site; the upstream primer also carried an SalI site (in the upstream primer) and an NotI site (in the downstream primer); sequences in italics are those complementary to relevant sequences by gel shift, footprinting, and chromatin

The 0.1 MUT fragment was obtained by PCR from the 0.4a MUT subclone, with primers carrying external HindIII sites (see above) and subcloned upstream from the Sox2 promoter as done for the 0.4a MUT fragment. Transgenes were isolated as Sall fragments.

The sequences of the primers used are the following: forward primer, 5′-CCTCCTGCACTCTGCGTTGCAACGGACTGAGAAATTTCCACGTAAC′-3′ (POU2 MUTF) and reverse primer T7 (within Bluescript); reverse primer, 5′-GAATTTCCTCAGCTGCAAGTGAGGTCAAGTGGTGCAGC-3′ (POU2 MUTR) and forward primer T3 (within Bluescript); forward primer 5′-GGTCGCGTTTGACTGTTAAAGAAGATTTAGAGAGAAG-3′ (POU1 MUTF) with reverse primer T7; reverse primer 5′-ATCCTTCTATGATGACCCCGACCGCCAGCG-3′ (POU1 MUTR) with forward primer T3.

Transgenic Embryos and X-gal Staining

Transgenic embryos were produced and analyzed by X-gal staining (at E12.5) as described previously (10).

Nuclear Extracts and Electrophoretic Mobility Shift Assay (EMSA)

Nuclear extracts were prepared from ES cells and from dissected embryonic cortex according to the method described in Ref. 21. For ES cell nuclear extracts, one or two subconfluent 10-cm plates were used. For embryonic brain nuclear extracts, 2–3 liters (20–30 embryos) were used (at embryonic day E12.5, or as indicated). The brain was dissection in PBS with protease inhibitors (21) and separated into cortex and into posterior (excluding the spinal cord) plus ventral regions. These were dissociated to single cells by gentle Dounce pestling in PBS with protease inhibitors. After the following steps of cell lysis, nuclear pellet precipitation, and nuclear protein purification, were carried out exactly as described, with the exception that a protease inhibitor mixture (1 mg/ml leupeptin, 1 mg/ml pepstatin, 1 mg/ml chymostatin, 1 mg/ml antipain; all from Sigma) was added to buffers A (1:500) and C (1:300).

EMSA was performed according to Refs. 22 and 23. Briefly, protein binding reactions were performed by combining 32P-labeled double-stranded oligonucleotides (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 mM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C. For "supershift" reactions, 1 μl of the appropriate antibody was added to the complete binding reaction mixture. After the addition of the labeled probe, Unlabeled competitors (15,000–20,000 cpm) and 2–4 μg of nuclear extract with 1.5 μg of poly(dI-dC) and 10 μg of bovine serum albumin in binding buffer (10% glycerol, 12 μM HEPES, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA) and 1 mM dithiothreitol, pH 7.9) and incubating for 20 min at 22°C.
FIG. 1. Sox2 5′-regulatory region. A, mouse Sox2 5′-flanking region. The thick line (from the 5′-end to H) is the DNA region required for telencephalic transgene expression, which comprises the chromatin DNase I-hypersensitive site HSI (vertical arrow) and the PstI-sensitive site P* (10). The thin line (from the H site to the Sox2 gene) is the Sox2 minimal promoter, which is inactive, by itself, in the central nervous system (10). Filled blocks A, B, and C represent sequences highly conserved among mouse, human, and chick. Thin lines underneath represent the fragments used in the functional experiments in this study (see below). Nd, NdeI; Nh, NheI; H, HindIII; P, PstI. Arrow pointing right, transcriptional start site. B, nucleotide sequence of region A, which surrounds DNase I-hypersensitive site HSI, and its alignment among different species. Asterisks mark nucleotides conserved in all three species. Boundaries of fragments used in the functional studies (0.1, 0.4a, 0.4b, and 0.1) are marked by arrows. POU sites 1 and 2 are bound by nuclear POU transcription factors in neural and ES cell nuclear extracts in the experiments shown in Figs. 2, 3, and 5. Underlined nucleotides are mutated in the MUT version of the regulatory fragments (see “Experimental Procedures”). PstI* is the PstI-sensitive site also shown in A.
goat anti-Brd1, goat anti-Brd2, goat anti-Brd3, rabbit anti-Oct (all Santa Cruz Biotechnology), and anti-Act (24).

The following double-stranded oligonucleotides were used (only the top strand is shown): Sox2 POU site 1, 5′-GCCGTGGCATTTCCCAGCTGCCTC-3′; Sox2 POU site 1 MUT, 5′-GCCGGTCGCTTTCCCAGCTGCCTC-3′; Sox2 POU site 2, 5′-TGCTCAAACTCTGCTAATTAGCAATGCTGAGAAATTCCAGTTAACAAGG-3′; Sox2 POU site 2 MUT, 5′-TGCTCAAACTCTGCTAATTAGCAATGCTGAGAAATTCCAGCTAGTTGCTC-3′.

DNAse I in Vitro Footprinting

Two partially overlapping fragments (probes a and b, see Fig. 3), which together encompass the whole 0.4a region (Fig. 3), were amplified in vitro by 20 cycles of PCR using the following primers, one of which had been labeled to high specific activity with [γ-32P]ATP: probe a forward primer, 5′-TACGTTAAGCGGCTACGCGG-3′; probe a reverse primer, 5′-TACCTAATCTCCTTATGG-3′; probe b forward primer, 5′-TGGACCATGGAATATTATCGGAC-3′; probe b reverse primer, 5′-TGGACCATGGAATATTATCGGAC-3′ (labeled); probe β forward primer, 5′-TGCATTCGAAGCTTTGGC-3′; probe β reverse primer, 5′-TGCATTCGAAGCTTTGGC-3′ (labeled). The amplified fragments were run on an 8% 29:1 acrylamide:bisacrylamide gel and eluted. DNAse I footprinting was essentially according to Ref. 25, with some modifications. Each probe (about 15,000 dideoxynucleotides/min) was mixed with 0.6 or 1.8 μg of nuclear extract, 10 μg of bovine serum albumin, 1.5 μg of poly(dI-dC) in footprinting buffer (10% glycerol, 12 μg HEPEs, 4 mM Tris-HCl, 60 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, pH 7.9) in a total of 20 μl for 40 min on ice. In control reactions, the nuclear extract was omitted. At the end of the reaction, 50 μl of footprinting buffer containing 2.5 mM MgCl2 and 2.5 mM CaCl2 was added just before the addition of 6 μl of DNase I (0.4–0.8 μg, RQ1 DNase; Promega). Control reactions without nuclear extracts received 0.2–0.4 μl of DNase I. Digestion proceeded for exactly 2 min at 22 °C before being stopped with 70 μl of stop buffer (0.6 μM NaCl, 20 mM Tris-HCl, 20 mM EDTA, 1% SDS, pH 7.5). After phenol-chloroform extraction and ethanol precipitation, the DNA was resuspended in formamide loading dyes, heated 5 min at 90 °C, and analyzed on a 6% acrylamide:bisacrylamide (19:1) sequencing gel.

ES Cell Transfection (10)

CCE ES cells were electroporated at 10° cells/0.55 ml of PBSCM (PBS + 1 mM CaCl2, 1 mM MgCl2) with a Bio-Rad gene pulse at 220 V, 960 microfarads, with 4 μg of purified transgene DNA. After 5 min on ice, cells were plated on 20 cm-dishes onto STO feeder layers. G418 selection at 300 μg/ml was applied after 24 h. Resistant clones were counted at day 10 of selection.

Purification, Immunoprecipitation, and Decross-linking of Fixed Chromatin Fragments

Cross-linked chromatin complexes were separated from free proteins, DNA, and RNA by CsCl isopycnic centrifugation (essentially as described in Ref. 24). Fractions containing the cross-linked chromatin (~1.38 g/ml) were pooled and dialyzed overnight at 4 °C against 5% glycerol, 1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris-HCl, pH 8.0. Aliquots (200 μg of cross-linked chromatin in 100 μl) were cleared by centrifugation (115,000 rpm at 15,000 rpm, 10 μl of 0.5% 2° TE-EGTA buffer (1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris-CI, pH 8.0), adjusted carefully to 0.1% SDS (w/v) and 0.5 mM NaCl, incubated for 5 min at room temperature, adjusted to 1% Triton X-100 (w/v), 0.1% sodium deoxycholate and 0.1% bovine serum albumin, incubated for 10 min, and cleared again by centrifugation for 15 min at 15,000 rpm. As a preclarifying step, supernatants were incubated with 100 μl of Dynabeads coupled to sheep anti-rabbit IgG for 1 h. Supernatants were removed from the beads by a magnetic particle concentrator (MPC; Dynal), and 10 μg of specific antibodies against Oct4 (or IgG of preimmune serum) was added. Samples were rotated for 3 h at 4 °C, 300 μl of sheep anti-rabbit IgG Dynabeads was added, incubated with rocking for 1 h at 4 °C, and immunocomplexes were pelleted by magnetic field. Pellets were washed (10 min/wash) five times in 1 ml of washing buffer (1% Triton X-100 (w/v), 0.1% sodium deoxycholate (w/v), 0.1% SDS (w/v), 0.1% bovine serum albumin (w/v), 0.5 mM NaCl, 1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris-HCl, pH 8.0), once with 1 ml of LiCl washing buffer (250 mM LiCl, 0.5% Nonidet P-40 (w/v), 0.5% sodium deoxycholate (w/v), 1 mM Na-EDTA, 0.5 mM Na-EGTA, 10 mM Tris HCl, pH 8.0), twice with 1 ml of TE-EGTA buffer, and resuspended in 350 μl of TE-EGTA buffer. 100 μl was kept for protein analysis, and 250 μl was treated 30 min at 37 °C with 50 μg/ml DNase-free RNase A and incubated overnight at 37 °C in 250 μl of proteinase K and 0.25% SDS.

PCR on Immunoprecipitated Cross-linked Chromatin

PCR amplification was performed in a final volume of 50 μl, using 1 ng of genomic ES (E14) DNA or 1 ng of immunoprecipitated cross-linked chromatin as template, 10 pmol of each primer, 2.5 mM dNTPs (Amer- sham Biosciences), 2.5 units of TaqDNA polymerase (Invitrogen), and 1× PCR buffer (Invitrogen). The PCR consisted of 3 min at 94 °C for one cycle and then 33 cycles at 94 °C for 30 s, 54 °C for 45 s, and 72 °C for 45 s, and 1 cycle of 72 °C for 10 min. Upstream and downstream primer pairs were as follows: Sox2p 5′-ATGAAATTTCAAACCTTGTCACCC-3′ and 5′-TTATGTCGCTTTCAAACTCTTCATGTGTGGACAAAAGGCAATAATTAGCAT-3′; Sox2p 5′-GCCGTGGCATTTCCCAGCTGCCTC-3′ and 5′-ATGAAATTTCAAACCTTGTCACCC-3′; Sox2p 5′-GCCGTGGCATTTCCCAGCTGCCTC-3′ and 5′-TGTGTCGCTTTCAAACTCTTATGTGTGGACAAAAGGCAATAATTAGCAT-3′.

RESULTS

A Conserved 400 Base Pair Element 5′ to Sox2 Is Required and Sufficient to Direct Transgene Expression to Brain Neuroepithelial Precursors—In previous experiments (10) we showed that a 5.7-kb region 5′ to mouse Sox2 directs transgene expression to telenchephalic neural stem cells and precursors; we further demonstrated that this expression requires the most 5′ 2.5 kb of this region, which contains a cell type-specific chromatin DNAse I-hypersensitive site, termed Sox2 HSI, which we identified in neural and ES cell nuclei (10) (Fig. 1A). Two more DNAse I-hypersensitive sites (HSII and HSIII) lie in the Sox2 minimal promoter, inactive by itself in the central nervous system (10). The comparison of genomic sequences 5′ to Sox2 in mouse, human, and chick identifies three conserved regions, of about 400, 110, and 90 bp, respectively (Fig. 1; see also Ref. 20), the last corresponding to the transcriptional start site region. The position of the extensive most 5′-region closely matches the DNAse I and Pat1 chromatin-hypersensitive site, HS1 (10) (Fig. 1). This homology thus lies within the region required for transgene expression in the brain.

We asked whether smaller fragments comprising HSI are sufficient to direct telenchephalic expression of the β geo-reporter gene in transgenic assays (Fig. 2). A 400-bp fragment, centered on HSI (and comprising part of the 5′-homology region), was linked to the minimal Sox2 promoter, inactive in the central nervous system (10; Fig. 1A). This construct (0.4-Sox p., Fig. 2A and 2B, images 1–3) is sufficient to direct telenchephalic transgene expression. To assess whether telenchephalic activity requires cooperation with the Sox2 3.3-kb promoter region or can be specified autonomously by the HSI region, we also fused the latter region (as progressively smaller fragments of 3 kb, 800 bp, or 447 bp, the last including the complete homology region) to the heterologous TK minimal promoter, to generate the 3-, 0.8-, 0.4b-TK pr. constructs (Fig. 2A and 2B, images 3–6). The 3-kb and 800-bp fragments, linked to the TK promoter, drive high level telenchephalic activity (Fig. 2B, images 3 and 4). Similarly, the 447-bp fragment drives expression in the telenchephalon, although at a lower level (Fig. 2B, images 5 and 6), together with some ectopic activity (in one embryo, Fig. 2B, image 5). These results indicate that the HSI element is sufficient to specify telenchephalic transgene expression, in the absence of any additional neural regulatory element.

110 BasePairs within the HSI Element Direct Brain-specific Expression—To identify functionally important sequences, we tested smaller fragments within the HSI element. A 110-bp fragment, centered on a site that is accessible to Pat1 digestion in telenchephalic neural cell chromatin (10) (Fig. 1), was linked to the Sox2 minimal promoter (0.1-Sox p. construct, Fig. 2). This construct is still able to drive high level telenchephalic...
expression (Fig. 2A and 2B, image 7). Notably, however, the expression domain is slightly expanded posteriorly, to include the dorsal mesencephalon (Fig. 2B, image 7).

We subsequently tested the same 110-bp sequence linked only to the TK minimal promoter (0.1-TK pr. construct, Fig. 2A and 2B, image 8). Again, expression was detected, although at very low levels, in the mesencephalon and telencephalon, particularly along the medial region (Fig. 2B, image 8). We conclude that the 110-bp element is able to specify brain-specific transgene expression, acting more efficiently when linked to the Sox2 promoter region.

Octamer Sites within the 110-Base Pair Core Region Are Bound by CNS-specific POU Transcription Factors—The 110-bp Sox2 enhancer defined in vivo was mapped in vitro for binding sites recognized by nuclear proteins of the brain. Fig. 3 shows a footprinting assay on DNA probes containing the 110-bp region and spanning the entire 400-bp enhancer. Two regions are protected from DNase I digestion by telencephalic nuclear extracts from E12.5 dorsal telencephalon (Fig. 3A). These protections center on two octamer consensus sequences (Figs. 1 and 3A); these sequences are recognized by members of the POU transcription factor family, some of which are expressed in brain (25). By contrast, no protection is detected using extracts from non-neural cells (CH27 lymphoid cells, Fig. 3A, and data not shown) in this assay.

Two probes encompassing the protected regions were subsequently used in EMSAs (Fig. 4). Both probes form two major complexes with embryonic brain nuclear extracts at E12.5 (Fig. 4, A and B), E11.5 and E10.5 (not shown). The upper complex is more abundant in extracts from dorsal telencephalon, and the lower one in extracts from more posterior central nervous system regions (Fig. 4, A and B; see also Fig. 6A). Non-neural cell extracts (from lymphoid CH27 cells) generate a much slower complex and a band with mobility similar to the upper brain band (Fig. 4, A and B); however, this band is generated by a protein different from the neural factor (see below).

Upon incubation with antibodies specific for the neural POU transcription factors Brn1 and Brn2, a specific decrease of the intensity of the higher complex (with anti-Brn1), or of the lower complex (with anti-Brn2), with concomitant appearance of more slowly migrating bands (Fig. 4, asterisks) was observed (Fig. 4, A and B). This supershift is caused by the formation of a complex between the factor bound to DNA and the antibody; no such supershift was observed with an antibody against Brn3. The mobility of the CH27 band comigrating with the upper telencephalic band was not modified by any of the Brn antibodies, indicating that this band is the result of a different protein.

The intensity of both complexes observed with brain extracts is strongly diminished by competition of the binding of the labeled oligonucleotide probe with nonlabeled oligonucleotides carrying canonical octamer-POU binding sites as well as by a previously characterized POU factor binding site in the neural

---

**Fig. 2.** Sox2 β-geo constructs and their expression in E12.5 transgenic embryos. A, indicated are the different enhancer fragments (black bars, see Fig. 1) that were linked either to the TK (white bars) or the Sox2 (hatched bars) minimal promoters, upstream from the β-geo reporter transgene. White dots, octamer mutations (see Fig. 1B and “Results”). Vertical arrow shows the location of HSI. B, β-galactosidase activity in Sox2 β-geo transgenic embryos at midgestation. Embryos obtained with each of the constructs in A are shown (see “Results”). Image 1 (0.8-Sox pr.), obtained with a construct already reported in Ref. 10, is shown as a reference for comparison.
precursor-specific enhancer of the nestin gene, which is essential for the activity of this enhancer in vivo (25) (Fig. 4C).

Mutations by base substitutions within the octamer binding sequences of POU site 1 or POU site 2 probes (shown in Fig. 1B) abolish binding of both complexes (Fig. 4D).

Octamer Mutations Reduce Neural Transgene Expression in Vivo—To assess the functional importance of the POU binding sites in brain Sox2 transgene expression, point mutations that abolish transcription factor binding (Fig. 4D) to both POU sites 1 and 2 were introduced in the 400-bp, as well as in the 110-bp elements, which were linked to the Sox2 minimal promoter upstream from β-geo and assayed in transgenic embryos (0.4aMUT- and 0.1MUT-Sox pr. constructs; Fig. 2A and 2B, images 9 and 10). As shown in Fig. 2, a reduction in activity is observed in both cases; this is seen both as quantitative expression levels (compare Fig. 2B, images 2 and 9 and images 7 and 10) and number of integrations showing detectable specific expression (Fig. 2A). However, brain-specific expression is not abolished, as can be seen with both mutated constructs (Fig. 2B, images 9 and 10).

HSI Sequences Outside the 110-Base Pair Core Region Are Also Able to Specify Telencephalic Transgene Expression—Because the 110-bp core region had proven able to specify high level brain-specific transgene expression from the minimal Sox2 promoter (0.1-Sox pr. construct; Fig. 2B, image 7), we next asked about the consequences of deleting this region from the full size HSI1 element. In the 0.8Δ0.1-TK pr. construct (Fig. 2A), the 110 bp were deleted from the 0.8-TK pr. construct, which had been shown to drive high level, brain-specific transgene expression (Fig. 2A and 2B, image 4; see above). Surprisingly, the 800-bp HSI enhancer deleted of the 110-bp central region is still able to specify telencephalic transgene expression (Fig. 2B, image 11). Thus, HSI sequences both within the core 110 nucleotides (see above) and in the surrounding region, are able to specify telencephalic transgene expression independently.

The Sox2 HSI 400-Base Pair Element Is Active in ES Cells, and This Activity Depends on the Integrity of the Octamer Binding Sites—We found previously that the full 5.7-kb Sox2 5′-flanking region drives transgene expression already in the blastocyst inner cell mass (10). Further, DNase I-hypersensitive sites are detected in ES cell chromatin, including the HSI region (10). We thus sought to determine whether the HSI element active in neural cells after gastrulation (see above) is also active in ES cells. To this end, the 0.4a-Sox pr. construct, previously found to be active in neural cells in transgenic embryos after gastrulation (Fig. 2), was stably transfected into ES cells. Exploiting the bifunctional nature of the β-geo reporter gene (encoding lacZ and G418 resistance), the number of G418-resistant colonies was taken as a measure of the activity of the linked regulatory sequences. As shown in Fig. 5, the 400-bp element efficiently stimulates activity in ES cells, compared with the promoterless β-geo control and with the minimal 3.3-kb Sox2 promoter cloned upstream from β-geo. However, when the octamer-mutated version of the enhancer was...
assayed (0.4aMUT-Sox pr. construct, see Fig. 2), this activity was reduced to only 17% of the nonmutated control, marginally above the level observed with the 3.3-kb promoter only, and with the promoterless control (Fig. 5). In cells that do not express Sox2, e.g. NIH3T3, both the wild-type and the octamer-mutated constructs were essentially inactive, with the same very low expression levels regardless of octamer integrity (not shown).

Differential POU Transcription Factors Bind to the Sox2 HSI Octamers in ES Cell Nuclear Extracts—EMSAs were done to investigate proteins binding to the HSI octamer sequences in ES cell nuclear extracts (Fig. 6), using the same probes previously used with neural cells (see above, Fig. 4). Two major complexes are seen with both probes: a higher mobility one, observed specifically in ES cells, and a lower mobility one (Fig. 6A). The latter comigrates with the most abundant complex formed by these probes in cells (such as STO fibroblasts, Fig. 6A, or CH27, see Fig. 4A), which do not express Sox2, and where HSI is not accessible. These ES cell complexes differ from those seen with neural cells (Fig. 6A). Both complexes are competed by an unlabeled “canonical” octamer binding site, known to be bound by the POU factors Oct4, Oct1, and Oct6 in ES cells (26), as efficiently as by the Sox2 gene octamer sequence itself (Fig. 6B).

Specific antibodies were used to investigate these complexes further (Fig. 6C). Incubation with an antibody recognizing the ubiquitous Oct1 POU transcription factor results in absence of the slower complex (Fig. 6C). The faster complex is abolished by an antibody specifically recognizing Oct4 (Fig. 6C), a POU transcription factor specifically found in ES cells (and their in vivo counterpart, the inner cell mass and epiblast) and in germ cells.

Chromatin Immunoprecipitation from ES Cells Demonstrates in Vivo Binding of Oct4 to the HSI Element—To assess whether Oct4 does bind to HSI in vivo, we immunoprecipitated ES cell chromatin with an anti-Oct4 antibody (24), and we amplified DNA from the precipitated chromatin with primers for the Sox2 HSI region, flanking the Oct4 binding sites, and

---

**Fig. 3.** Neural cell nuclear proteins generate DNase I footprints on the 0.4-kb regulatory element. A, foot-printing assay with nuclear extracts from neural (E12.5 dorsal telencephalon; dors. tel.) and non-neural (CH27) cells, using probe α (left panel) and β (right panel), which span the 0.4-kb element (fragment 0.4a in Fig. 1). Increasing amounts of nuclear extracts (0.6 and 1.8 µg) were added where indicated. DNase I digestion was at 0.6 µg/reaction in the presence of nuclear extracts and at 0.2 µg/reaction (lanes 4, 7, 8, and 11), or 0.4 µg/reaction (lane 1) in the absence. Regions of protection are marked on the left side of panels; numeric coordinates refer to the sequence in Fig. 1B. The most strongly footprinted regions cover two octamer sequences (POU1 and POU2 sites in Fig. 1B); asterisks mark nucleotides within these sequences whose mutated version has been functionally assayed (see text). B, diagram showing probes α and β and the footprinted regions (filled bars) seen with neural extracts in A, with respect to the 0.4-kb enhancer element. Round dots mark the labeled end of the probes used in A. In probe α, the forward strand was labeled; in probe β, the reverse strand. Sequence numbering is as in Fig. 1B. Filled bars represent the strongest footprints, dotted lines connecting them show a region that is also protected, although more weakly (see A: probe α also shows some protection upstream from position 387 (left panel), and the footprint seen with probe β appears to continue beyond the PstI site (position 368, right panel)).

---

By guest on July 25, 2018
Fig. 4. The two POU sites in the Sox2 enhancer are recognized by neural POU transcription factors. A, EMSA of a probe encompassing POU site 1. With E12.5 dorsal telencephalic nuclear extracts (dors. tel.) or E12.5 extracts of brain devoid of dorsal telencephalon (res. brain), two major DNA-protein complexes are seen (black and white arrows at the left of the figure), of which the upper one is stronger with dorsal telencephalic extracts, and the lower one with residual brain extracts. EMSA with CH27 lymphoid cells shows two different complexes (black and white arrowheads at the right of the figure). The addition of antibody against Brn1 disrupts the upper complex in neural cells (lanes 2 and 7) and produces a supershift, stronger in lane 2 (asterisks); no change is observed in CH27 (lane 12). Brn2 antibody disrupts the lower band in neural cells (lanes 3 and 8), generating a supershift (asterisk; stronger in lane 8) and has no effect in CH27 (lane 13). Anti-Brn3 antibody does not affect any of the gel shifts (lanes 4, 9, and 14). B, EMSA of a probe encompassing POU site 2 with E12.5 neural extracts (lanes 1–4) or with non-neural CH27 (lane 6). Two complexes are seen with neural extracts (black and white arrows on the left of the panel) of the same mobility as those seen with POU site 1 probe (compare lanes 3, 4, and 5). The addition of Brn1 or Brn2 antibodies supershifts (asterisks) the upper complex (lane 1) or the lower complex (lane 2), respectively. Non-neural CH27 cells (lane 6) shows two different complexes of the same mobility as those seen in A with POU site 1 (black and white arrowheads). C, binding of neural proteins to POU sites 1 and 2 is competed efficiently by an octamer site in the nestin gene neural precursor-specific enhancer (25) (lanes 5–7 and 16–18). The competition efficiency is similar to that of the self competitors (lanes 2–4 and 13–15). A nonrelated oligonucleotide (GATA-1-binding oligonucleotide), by contrast, does not compete in the same concentration range (lanes 9–11 and 20–22). D, point mutations in POU sites 1 and 2 abolish nuclear protein binding. Oligonucleotides carrying nucleotide substitutions in the octamer sequences (see Fig. 1B) no longer form the protein-DNA complexes (compare lane 4 with lane 1, and lane 10 with lane 7) and fail to compete for binding to the nonmutated probes (lanes 2 and 3 and lanes 8 and 9). wt, wild-type.
with primers for a control gene (actin), not regulated by Oct4 (Fig. 7). The HSI region was amplified efficiently from the immunoprecipitated chromatin, whereas the actin region was not amplified (Fig. 7B). As a further control, preimmune serum was unable to immunoprecipitate both the HSI and actin regions (Fig. 7C). These results indicate that Oct4 is bound to the HSI element in vivo in ES cells.

**DISCUSSION**

In previous work, we identified a regulatory region of the Sox2 gene which drives reporter gene expression in transgenic mice to neural stem and precursor cells within a specific subset of the panneural Sox2 expression domain, the telencephalon (10). In this paper, we have restricted the critical region for Sox2 telencephalic expression and identified a sequence encompassing a cluster of POU factor binding sites. Mutation of these sites considerably decreases transgene expression. Interestingly, the same mutation affects expression of the transgene in ES cells. These results highlight the role of POU factors in Sox2 regulation, both in totipotent cells and in early neural precursors.

**Sox2 Regulatory Sequences Involved in Telencephalic Expression**—In vivo work in mouse, Xenopus, and chicken (10, 11, 14, 15, 18, 19) has suggested important functional roles for Sox2 in stem/precursor cells, ranging from the highly multipotent early epiblast cells to neural stem cells of the adult brain. In the nervous system, Sox2 is expressed during embryogenesis along the entire neural tube. We showed previously (10) that a 5.7-kb fragment, encompassing the Sox2 promoter, drives reporter gene expression in transgenic embryos exclusively in the telencephalon; this suggested that Sox2 regulatory sequences may respond in the telencephalon to local signals. In this work, we sought to identify the shortest Sox2 regulatory regions required for the specification of telencephalic expression.

By linking progressively shorter fragments of DNA surrounding the HSI site (0.8-, 0.4-, 0.1-Sox pr. constructs, Figs. 1 and 2A) to the Sox2 minimal promoter (10), we have shown that 110 bp (centered on the PsI site shown to be hypersensitive in chromatin; Ref. 10) drive, in conjunction with the Sox2 promoter, strong reporter expression to the telencephalic region (Fig. 2B, images 1, 2, and 7). Because the Sox2 3.3-kb promoter alone is not able to direct telencephalic expression (10), these results suggest that the HSI sequences are sufficient to specify telencephalic expression. However, we cannot rule out that sequences within the minimal Sox2 promoter may be required to cooperate with HSI and potentiate its activity. To test this point, we linked HSI fragments to a TK minimal promoter (inactive in the telencephalon, 25). The HSI fragments (3, 0.8, and 0.4 kb) still drive clear (Fig. 2B, images 3–6) reporter gene expression to the telencephalon. This proves that the HSI fragment contains sequences that are sufficient by themselves to specify regional expression in neural cells. The 110-bp fragment is also active in the telencephalon, when linked to the TK promoter, although to a much lesser extent than when linked to the Sox2 promoter (Fig. 2B, image 8). These results further indicate that, although the 110-bp sequence is sufficient to specify telencephalic expression, cooperation with promoter sequences (Fig. 2B, image 7) or with other sequences within the HSI region (Fig. 2B, images 4 and 5) increases the overall efficiency.

An interesting and unexpected observation is that restriction to telencephalic expression requires more than the 110-bp fragment, as indicated by the relatively strong mesencephalic activity detected with the constructs including the 110-bp element (Fig. 2B, images 7, 8, and 10; see also below) and the absence of such expression with all other constructs. These results indicate that the HSI region contains sequences negatively regulating the activity of the core HSI 110-bp element.

The 110- and 400-bp HSI elements are, to our present knowledge, the only short enhancer sequences able to drive specific expression to the telencephalon.

**POU Binding Sequences in the HSI 110-Base Pair Region Contribute to Transgene Expression**—Within the core 400-bp HSI region, in vitro footprinting (Fig. 3) and subsequent in vitro EMSAs (Fig. 4) reveal prominent binding of neural specific POU factors, such as Brn1 and Brn2, to conserved binding sites, here named POU1 and POU2 sites. Similar sites, bound by the same factors in neural cells extracts, are known to contribute to the level of expression of the nestin and the BFAP genes in neural regions of transgenic mice (25). Mutation of POU sites reduces both the number of expressing embryos and the level of reporter gene activity, within the 0.4- and 0.1-Sox pr. constructs (Fig. 2A and 2B, compare images 2 and 9, and images 7 and 10). Thus, POU factors binding to the 110 bp are likely responsible for at least part of the HSI activity. It should be noted that other potential POU-binding elements, although not detected as in vitro footprints (not shown), lie both 5’ and 3’ to the core region, within the 0.8-kb fragment, and might contribute to the overall activity of the HSI region.

**A Modular Organization of the Sox2 Regulatory Region?**—Using relatively large Sox2 constructs, consisting of 5.7–12 kb of promoter-enhancer, including HSI, the majority of integrations within transgenic mice was specifically active in the telencephalon, with little or no ectopic expression (10). By dissecting the region to fragments as small as 0.1 kb, telencephalic
expression is maintained, even with the smallest fragment (0.1 kb) (see above), although the proportion of active integrations, and/or the level of expression, may be decreased, and some ectopic expression is observed. These effects may be the result of the smaller size of the constructs, allowing stronger position dependence. An additional explanation might be that the Sox2 promoter-enhancer region may consist of several, partially redundant modules, able to specify telencephalic expression independently (see above) and to interact to promote optimal activity. Indeed, the 110-bp sequence that is sufficient for telencephalic expression (Fig. 2B, images 7 and 8) is clearly not strictly required because its deletion from the 0.8-kb region is still compatible with specific telencephalic expression (Fig. 2B, images 11-13).

A number of data in the literature suggest a modular organization of the Sox2 regulatory sequences, in which multiple partially redundant elements are spread along the gene and control expression level and region-specific activity. Although Sox2 is expressed from the endogenous gene along the entire neural tube, the HSI enhancer is active specifically in the telencephalon in transgenic mice (Ref. 10 and present paper). Other elements may participate in regulating telencephalic expression because a targeted deletion of the HSI enhancer from the endogenous mouse Sox2 gene does not completely abolish Sox2 telencephalic expression, but rather reduces it by about 50% in neural stem/precursor cells from adult and embryonic mouse brains (15). In keeping with this, Uchikawa et al. (20) reported several distinct regulatory elements along the chicken Sox2 gene, active at distant and/or overlapping neural tube regions during development, and Tomioka et al. (27) and Miyagi et al. (28) demonstrated, by transfection and by retroviral transduction assays, a Sox2 enhancer 3′ to the mouse Sox2 gene itself, active in neural cells. Our present dissection of the HSI region points to further functional heterogeneity even within a short DNA sequence active within a restricted brain region.

**HSI POU Factor Binding Sites Are Essential for Sox2 Transgene Expression in ES Cells**—ES cells represent an “in vitro” counterpart of early epiblast stem cells, the highly multipotent precursors to all embryonic cell types (including neural cells). Sox2 is essential, in mouse, for growth and maintenance of these cells (11), and the phenotype of Sox2−/− mutants has some resemblance to that of Oct4−/− embryos (29). The regulation of the expression of Sox2 is still poorly understood; however, we reported previously that the same transgenic con-
struct that is highly expressed in neural stem cells is also active, prior to that stage, in the blastocyst inner cell mass and epiblast (10; see also 27).

Starting from our observation that POU-binding elements are important in neural expression of Sox2, we asked whether the same sites are recognized by Oct4, a POU factor that is expressed exclusively in inner cell mass and epiblast, germ cells, and very early embryonic tissues (29).

EMSA experiments (Fig. 6, A and B) with ES cell nuclear extracts, in combination with the use of specific antibodies (Fig. 6C), show that at least two different proteins, Oct4 and Oct1, bind to the POUS binding sites with neural nuclear extracts (Figs. 4 and 6A; lanes 9 and 10). Oct4 is specifically expressed in the blastocyst inner cell mass (and in in vitro derived ES cells) and in germ cells (30–32). Sox2 is highly expressed in the inner cell mass and in ES cells (11). The same POUS site mutations that affect neural expression of the Sox2 constructs both abolish Oct4 and Oct1 in vitro binding and drastically decrease (by more than 80%) the activity of the 0.4 Sox pr. construct in stable transfections in ES cells (Fig. 5).

These data suggest that Oct4 (and/or Oct1) are important for Sox2 expression in ES cells.

Transfection experiments in ES cells previously showed a dependence on the cooperation between Sox2 itself and Oct4 for autoregulation of a different Sox2 enhancer, located downstream from the gene (27). To evaluate the in vivo relevance of Oct4 to Sox2 expression, we performed chromatin immunoprecipitation with anti-Oct4 antibodies; indeed, Oct4 was found to be present on the HSI region of the endogenous Sox2 gene in ES cells nuclei, thus implying a functional in vivo role (Fig. 7). The importance of Oct4 for optimal Sox2 expression is further suggested by the observation that Oct4-deficient ES cells have a lower level of Sox2 mRNA than wild-type cells (33). The presence of Oct4 in the endogenous Sox2 promoter region was previously demonstrated in ES cells by chromatin immunoprecipitation with anti-Oct4 antibodies; indeed, Oct4 was found to be present on the HSI region of the endogenous Sox2 gene in ES cells nuclei, thus implying a functional in vivo role (Fig. 7).

The importance of Oct4 for optimal Sox2 expression is further suggested by the observation that Oct4-deficient ES cells have a lower level of Sox2 mRNA than wild-type cells (33). The same POU sites control expression of the Sox2 constructs both abolish Oct4 and Oct1 in vitro binding and drastically decrease (by more than 80%) the activity of the 0.4 Sox pr. construct in stable transfections in ES cells (Fig. 5).

The importance of Oct4 for optimal Sox2 expression is further suggested by the observation that Oct4-deficient ES cells have a lower level of Sox2 mRNA than wild-type cells (33). The same POU sites control expression of the Sox2 constructs both abolish Oct4 and Oct1 in vitro binding and drastically decrease (by more than 80%) the activity of the 0.4 Sox pr. construct in stable transfections in ES cells (Fig. 5).

The same POU sites mutations that affect neural expression of the Sox2 constructs both abolish Oct4 and Oct1 in vitro binding and drastically decrease (by more than 80%) the activity of the 0.4 Sox pr. construct in stable transfections in ES cells (Fig. 5).

The importance of Oct4 for optimal Sox2 expression is further suggested by the observation that Oct4-deficient ES cells have a lower level of Sox2 mRNA than wild-type cells (33). The same POU sites control expression of the Sox2 constructs both abolish Oct4 and Oct1 in vitro binding and drastically decrease (by more than 80%) the activity of the 0.4 Sox pr. construct in stable transfections in ES cells (Fig. 5).

The importance of Oct4 for optimal Sox2 expression is further suggested by the observation that Oct4-deficient ES cells have a lower level of Sox2 mRNA than wild-type cells (33). The same POU sites control expression of the Sox2 constructs both abolish Oct4 and Oct1 in vitro binding and drastically decrease (by more than 80%) the activity of the 0.4 Sox pr. construct in stable transfections in ES cells (Fig. 5).
34. Fuhrmann, G., Chung, A. C., Jackson, K. J., Hummelke, G., Baniahmad, A., Sutter, J., Sylvester, I., Schöler, H. R., and Cooney, A. J. (2001) Dev. Cell 1, 377–387
35. Burgess, S., Reim, G., Chen, W., Hopkins, N., and Brand, M. (2002) Development 129, 905–916
36. Schier, A. F., Neuhauss, S. C., Harvey, M., Malicki, J., Solnica-Krezel, L., Stainier, D. Y., Zwartkruis, F., Abdelilah, S., Stemple, D. L., Rangini, Z., Yang, H., and Driever, W. (1996) Development 123, 165–178
37. Pfeffer, P. L., Payer, B., Reim, G., di Magliano, M. P., and Busslinger, M. (2002) Development 129, 307–318
38. Schonemann, M. D., Ryan, A. K., McEvilly, R. J., O’Connell, S. M., Arias, C. A., Kalla, K. A., Li, P., Sawchenko, P. E., and Rosenfeld, M. G. (1995) Genes Dev. 9, 3122–3135
39. Alvarez-Bolado, G., Rosenfeld, M. G., and Swanson, L. W. (1995) J. Comp Neurol. 355, 297–299
40. Katsuoka, F., Motohashi, H., Onodera, K., Suwabe, N., Engel, J. D., and Yamamoto, M. (2000) EMBO J. 19, 2980–2991
41. Yamagishi, H., Maeda, J., Hu, T., McNally, J., Conway, S. J., Kume, T., Meyers, E. N., Yamagishi, C., and Srivastava, D. (2003) Genes Dev. 17, 269–281
Conserved POU Binding DNA Sites in the Sox2 Upstream Enhancer Regulate Gene Expression in Embryonic and Neural Stem Cells

Raffaella Catena, Cecilia Tiveron, Antonella Ronchi, Silvia Porta, Anna Ferri, Laura Tatangelo, Maurizio Cavallaro, Rebecca Favaro, Sergio Ottolenghi, Rolland Reinbold, Hans Schöler and Silvia K. Nicolis

J. Biol. Chem. 2004, 279:41846-41857.
doi: 10.1074/jbc.M405514200 originally published online July 15, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M405514200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 41 references, 18 of which can be accessed free at http://www.jbc.org/content/279/40/41846.full.html#ref-list-1