View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by DigitalCommons@University of Nebraska

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Virology Papers

Virology, Nebraska Center for

2003

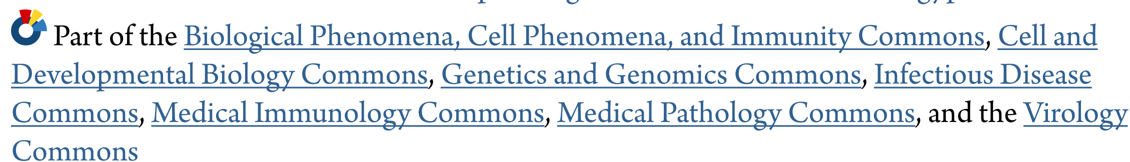
Identification of Novel Domains within Sox-2 and Sox-11 Involved in Autoinhibition of DNA Binding and Partnership Specificity

Matthew S. Wiebe
University of Nebraska Medical Center, mwiebe2@unl.edu

Tamara K. Nowling
University of Nebraska Medical Center

Angie Rizzino
University of Nebraska Medical Center, arizzino@unmc.edu

Follow this and additional works at: https://digitalcommons.unl.edu/virologypub



Part of the Biological Phenomena, Cell Phenomena, and Immunity Commons, Cell and Developmental Biology Commons, Genetics and Genomics Commons, Infectious Disease Commons, Medical Immunology Commons, Medical Pathology Commons, and the Virology Commons

Wiebe, Matthew S.; Nowling, Tamara K.; and Rizzino, Angie, "Identification of Novel Domains within Sox-2 and Sox-11 Involved in Autoinhibition of DNA Binding and Partnership Specificity" (2003). *Virology Papers*. 281.
https://digitalcommons.unl.edu/virologypub/281

This Article is brought to you for free and open access by the Virology, Nebraska Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Virology Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

University of Nebraska Lincoln
digitalcommons.unl.edu

Published in *Journal of Biological Chemistry* 278:20 (May 16, 2003), pp. 17901–17911; doi: 10.1074/jbc.M212211200.
Copyright © 2003 by The American Society for Biochemistry and Molecular Biology, Inc. Used by permission.
Submitted December 2, 2002; revised March 5, 2003; published March 10, 2003.

Identification of Novel Domains within Sox-2 and Sox-11 Involved in Autoinhibition of DNA Binding and Partnership Specificity*

Matthew S. Wiebe,[1,2] Tamara K. Nowling,[1] and Angie Rizzino[1,2]

1. Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805, USA
2. Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805, USA

Corresponding author – Angie Rizzino, Eppley Cancer Institute, University of Nebraska Medical Center, 986805 Nebraska Medical Center, Omaha, NE 68198-6805; telephone 402-559-6338, email arizzino@unmc.edu

Abstract

Sox transcription factors play key regulatory roles throughout development, binding DNA through a consensus (A/T)(A/T)CAA(A/T)G sequence. Although many different Sox proteins bind to this sequence, it has been observed that gene regulatory elements are commonly responsive to only a small subset of the entire family, implying that regulatory mechanisms exist to permit selective DNA binding and/or transactivation by Sox family members. To identify and explore the mechanisms modulating gene activation by Sox proteins further, we compared the function of Sox-2 and Sox-11. This led to the discovery that Sox proteins are regulated differentially at multiple levels, including transactivation, protein partnerships with Pit-Oct-Unc (POU) transcription factors, and DNA binding autoregulation. Specifically, we determined that Sox-11 activates transcription more strongly than Sox-2 and that the transactivation domain of Sox-11 is primarily responsible for this capability. Additionally, we demonstrate that the Sox-11 DNA binding domain is responsible for selective cooperation with the POU factor Brn-2. This requirement cannot be replaced by the DNA binding domain of Sox-2, indicating that the DNA binding domain of Sox proteins is critical for Sox-POU partnerships. Interestingly, we have also determined that a conserved domain of Sox-11 has the novel capability of autoinhibiting its ability to bind DNA *in vitro* and to activate gene expression *in vivo*. Our findings

suggest that the autoinhibitory domain can repress promiscuous binding of Sox-11 to DNA and plays an important role in regulating the recruitment of Sox-11 to specific genes.

The Sox family of transcription factors comprises a diverse group of proteins whose pattern of expression is regulated both spatially and temporally (1). Sox factors are related by the homology (usually >50%) found within their high mobility group (HMG)[1] DNA binding domains. Sox HMG domains have been found to bind to the consensus sequence 5′-(A/T)(A/T)CAA(A/T)G-3′ in the minor groove of DNA (2). As the number of Sox proteins has grown to more than 20 members in multiple species, including both vertebrates and invertebrates, this family has been divided into seven subgroups, the members of which contain high homology within both the HMG domain and flanking regions (1). In some cases, subgroup members have similar patterns of expression during development and perhaps redundant functions. For example, the structurally similar group B members Sox-1, -2, and -3 show overlapping expression in the fetal nervous system, and it has been suggested that they perform similar functions (3, 4). Additionally, the group C members Sox-4 and Sox-11 show high homology within their HMG domain and C-terminal tail, and they may also have complementary roles in the developing nervous system (5).

The observation that the HMG DNA binding domains of many members of the Sox family bind a consensus sequence has prompted the question of whether there are selective mechanisms in place to recruit a specific Sox protein to a gene regulated by an HMG binding site. A review of the literature reveals that genes rely on a diverse array of mechanisms to regulate DNA binding among members of other transcription factor families, which may also apply to Sox proteins. A particularly intriguing example of DNA binding regulation has been found within multiple members of the Ets family. First characterized in Ets-1, the affinity of these proteins for DNA is tightly regulated by autoinhibitory regions flanking the DNA binding domain (6, 7). Interestingly, the ubiquitous HMG-1 and HMG-2 proteins employ a similar regulatory strategy whereby their highly acidic C-terminal tails are capable of negatively affecting the ability of those proteins to bind DNA (8). Importantly, although the autoinhibitory domains of both Ets and HMG proteins repress DNA binding, they can be neutralized by post-translational modification or by interaction with partner proteins, which allows their recruitment to be exquisitely regulated in a gene-specific manner (7, 9–12).

Although specific evidence of DNA binding autoregulation has not yet been observed in Sox proteins, it is clear that cis-regulatory elements are capable of selecting specific Sox transcription factors. Attempts to uncover the mechanisms through which Sox specificity is achieved have revealed that recruitment often requires proper interaction with other transcription factors at adjacent binding sites on the DNA. This is illustrated by one of the cis-regulatory elements that control expression of the fibroblast growth factor 4 (*FGF-4*) gene in embryonal carcinoma cells. *FGF-4* expression depends on a distal enhancer (13, 14), which contains an essential binding site for Sox-2 (15–17). However, Sox-2 does not appear to be capable of substantial activation of this gene on its own, but rather requires cooperation with the POU transcription factor Oct-3 (18, 19). The presence of Oct-3 at a site adjacent to Sox-2 promotes cooperative DNA binding, resulting in a synergistic increase in transactivation (18, 19). The partnership of Sox-2 and Oct-3 is extremely specific for these

two proteins because it requires stereospecific alignment and protein-protein interaction between the HMG and POU DNA binding domains when Sox-2 and Oct-3 are present at their adjacent binding sites (18). Additionally, the role of Oct-3 cannot be fulfilled by Oct-1 for the *FGF-4* enhancer, indicating that only certain Sox-POU combinations can lead to productive regulation of a particular gene (15). Specific partnerships have also been observed among other members of the Sox and POU families. Utilizing the binding sequence found in the *FGF-4* enhancer, it was determined that Sox-10 specifically partners with Oct-6, and Sox-11 cooperates optimally with the POU factors Brn-1 and Brn-2 (20, 21). These observations have led to the hypothesis that a Sox-POU combinatorial code may exist as a mechanism of transcriptional regulation (20).

Although significant progress has been made in understanding the differential function of Sox proteins, both on their own and in combination with other factors, little is known about the mechanisms involved. To elucidate the mechanisms through which Sox proteins may be regulated, we initiated a functional comparison of Sox-2 and Sox-11 using promoter/reporter gene constructs that contain HMG and POU binding sites. Through this study, we determined that Sox-2 performs very differently than Sox-11, being a much weaker transactivator, while at the same time binding more strongly to DNA *in vitro*. To understand the mechanisms involved in the functional differences between Sox-2 and Sox-11, we addressed three questions. First, what role does each of the domains of Sox-2 and Sox-11 play in differential gene activation? Second, how specific are Sox-2 and Sox-11 partnerships with POU proteins, and which regions of Sox-2 and Sox-11 are important for Sox-POU partnerships? Third, which domain(s) are critical for the decreased DNA binding of Sox-11 compared with Sox-2?

We have addressed these questions using Sox deletion mutants and Sox-2/11 chimeric proteins created by interchanging the N terminus, HMG DNA binding domain, or C terminus of both proteins. These studies led to the discovery that the C-terminal transactivation domains play the primary role in the differential capability of Sox-2 and Sox-11 to transactivate. Additionally, we determined that Sox-11, but not Sox-2, can cooperate with Brn-2 and that the Sox-11/Brn-2-specific cooperation requires both adjacent binding of the two factors to DNA and the HMG domain of Sox-11. Therefore, this study further supports the hypothesis that a Sox-POU combinatorial code exists (20) and suggests that this interaction code is dependent on the HMG domain of the Sox protein involved. Finally, we also identified a novel domain within a conserved region of Sox-11 which is capable of negatively regulating both its ability to bind DNA *in vitro* and activate gene expression *in vivo*. Together, this work significantly enhances our understanding of the mechanisms that influence selective recruitment of Sox proteins through partner specificity and DNA binding autoregulation.

Experimental Procedures

Plasmids

Constructs pCATSO3, pCMVSox-2$F_{1\text{–}180}$, pCMVSox-2, pCMVOct-3, pCMVOct-1, and CMV-β-gal have been described previously (19). The promoter/reporter construct pCATDC3 was

constructed similarly to pCATSO3 as follows. Complementary oligonucleotides (synthesized by the Eppley Cancer Institute Molecular Biology Core Facility) containing two sets of the HMG site from the DC5 enhancer site and the POU binding sites (in boldface) from the FGF-4 enhancer region and *Bam*HI sites at both ends (5′-GATCCT**CATTGTT**GTG **ATGCTAAT**GGCT**CATTGTT**GTG**ATGCTAAT**GGAG-3′) were annealed and ligated to generate multimers. Differences in the DC5 and FGF-4 HMG binding sequence are underlined. These multimers were then ligated into a *Bgl*II site upstream from the SV40 promoter of the pCAT vector (Promega) to generate pCATDC3, which contains a total of six tandem copies of the HMG and POU sites. Similarly, pCATS4 contains a total of four tandem copies of the FGF-4 HMG site and was made using identical primers as pCATSO3, with the exception of a scrambled POU site. The upper primer is as follows with the HMG and scrambled POU sites in bold: (5′-GATCT**CTTTGTT**TGG**CGGATCAT** GGCT**CTTTGTT**TGG **CGGATCAT**GGA-3′).

The plasmids pCMVSox-11 and pCMVBrn-2 were gifts from Michael Wegner (20). Sox-11F was constructed using the primers FLAGSox-11 (5′-CGTGCTGGTACCGCCAC-CATGGACTACAAGGACGACGATGATATGGTGCAGCAGGCCGAGAGC-3′) and Sox-11/395 (5′-GCTGGATCCGACCGCCACGACTGCCTCCCG-3′) to amplify the coding region from CMV5-Sox-11. This PCR incorporated the FLAG epitope sequence, preceded by a Kozak sequence and a *Kpn*I site, onto the 5′-end of the coding region of Sox-11 and a *Bam*HI site at the 3′-end. The *Kpn*I/*Bam*HI sites were then digested and the Sox-11F sequence ligated into the *Kpn*I/*Bam*HI sites of the CMV5 expression vector. The Sox-2F expression plasmid referred to in this report was constructed by removing the Sox-2F sequence from the pCEP4 vector (19) using the *Kpn*I/*Bam*HI sites and relegating the Sox-2F sequence into the same sites present in CMV5.

Site-directed mutagenesis or deletions were carried out using a modification of the QuikChange (Stratagene, La Jolla, California) method. All amino acid numbers are based on the wild-type Met being defined as 1, not the Met included with the FLAG tag. Briefly, two primers, complementary to the sequence flanking the targeted region and to each other, were used in PCR amplification of the entire plasmid, thus modifying the indicated template. Unless otherwise noted, each of the primers identified herein represents the sense strand of a perfectly complementary pair of primers used during PCR. To begin the construction of the Sox-2/11 chimeras, unique restriction sites were incorporated at the N-terminal and C-terminal flanks of the HMG box of both the Sox-2F and Sox-11F plasmid. Sites were chosen in such a way that little alteration would be made to the wild-type coding sequence. Mutagenesis to include a *Nhe*I site was performed with either the sox2mut1 (5′-GGCAACCAGAAGAATGCTAGCCCGGAACGTGTCAAGAGGCCC-3′) or sox11mut1 (5′-CCGGACTGGTGCAAGACGGCTAGCGGCCACATCAAACG-3′) primer pair. The *Nhe*I alteration added a Ser following Ala-39 of Sox-2F but did not affect the coding sequence of Sox-11F. The products of these PCR reactions were then used in a second round of mutagenesis to include a *Not*I site using either the sox2mut2 (5′-GTACAC-GCTTCCCGCGGCCGCTTTGCTCGCCCCCGG-3′) or sox11mut2 (5′-GCCAAGCCCAG CGCGGCCGCACAGAGCCCGGACAAGAGC-3′) primer pair. This replaced Gly-131 and Gly-132 of Sox-2F as well as Ala-134 and Gly-135 of Sox-11F with the amino acids Ala-Ala-Ala. This cloning resulted in Sox-2F double mutant and Sox-11F double mutant, each of

which contained a *Kpn*I and *Bam*HI site at their extreme 5′- and 3′-ends, respectively, as well as an *Nhe*I and *Not*I site at the 5′- and 3′- respective ends of their HMG boxes. These sites made it possible to use standard restriction digests and ligations to interchange either the N-terminal, HMG, or C-terminal region of Sox-2F with the corresponding region of Sox-11F, resulting in the creation of the chimeric plasmids Sox-11-2-2F, -11-11-2F, -2-11-2F, -2-11-11F, -2-2-11F, and -11-2-11F.

To create Sox-11FΔTAD, the primer pair sox11Δtad (5′-CTCTACTACAGCTTCAAGTG AGCGGCCGCAAACATCACCAAGCAGCAG-3′) was used to incorporate a stop codon following Lys-282 of Sox-11F. The Sox-11F and Sox-11FΔTAD expression constructs were then used as the templates to create Sox-11FΔAR and Sox-11FΔARΔTAD, respectively, using the primer pairs of sox11Δar (5′-GCCAAGGTGGTCTTCCTGGACGCGCAGCAG-CAACCCCCTCAG-3′) to delete amino acids 190–223. To facilitate *in vitro* translation of the various Sox proteins, the coding regions from each of the above CMV5-derived vectors were cloned into the *Kpn*I/*Bam*HI sites of pCRScript (Stratagene), downstream from the T7 promoter. The identity of all plasmids modified by PCR-based mutagenesis as well as pCATS4 and pCATDC3 was verified by sequence analysis at the Eppley Cancer Institute Molecular Biology Core Facility.

Cell Culture and Transient Transfection

HeLa cells were maintained and transfected using the calcium phosphate precipitation method as described previously (19). All transfections were performed in triplicate with representative experiments shown. Plasmid DNA was isolated and purified using Qiagen tip-500 columns.

Extract Preparation and Western Blotting

HeLa cells were transfected with 5 μg of the relevant plasmids as described above, and extracts were prepared 2 days post-transfection. Whole cell protein extracts were prepared as described previously (19). Nuclear extracts were prepared using the NE-PER kit (Pierce) following the manufacturer's protocol. The isolated nuclei from ~6–7 × 10^6 cells were lysed in 100 μl of nuclear extraction reagent supplemented with various protease and phosphatase inhibitors (22). When preparing Sox-11F, Sox- 11FΔTAD, and Sox-11FΔARΔTAD for the nuclear extracts used in the electrophoretic gel mobility shift assay (EMSA) shown in figure 6c, calpain inhibitor I was added to the growth media 4 h before lysis and was included in the extraction buffer to minimize degradation caused by potential proteasome activity. Extracts were stored at –80°C. Sox proteins produced through *in vitro* transcription and translation were made using the T7 Quick Coupled Rabbit Reticulocyte Lysate (Promega) transcription/translation system.

Western blotting was performed using 10 μl of *in vitro* translated lysate or nuclear lysate from 5 × 10^5 transfected HeLa cells as described (19) using the anti-FLAG M2 antibody (Sigma). Proteins were detected using the enhanced chemifluorescence (ECF) kit (Amersham Biosciences) and scanned on a Storm PhosphorImager (Molecular Dynamics). Quantitation was performed using the ImageQuant 5.0 analysis software (Molecular Dynamics). Western blots were performed in duplicate or triplicate with representative blots shown.

EMSA

Gel mobility shift analysis was based on the method of Fried and Crothers as modified by this laboratory (23). Complementary oligodeoxynucleotide probes were annealed for each probe or competitor, and the resulting double-stranded oligodeoxynucleotide probes were labeled with [α-^{32}P]dCTP by Klenow fill-in reaction of the single-stranded regions at the end of each double-stranded oligodeoxynucleotide probe (lowercase bases). The wild-type Sox probe (hmg1) contained a single HMG binding site (bold), which was based on the sequence present in the *FGF-4* enhancer, whereas in the mutant Sox probe (HMGmut) this sequence is scrambled (underlined). The sequence of hmg1 (sense strand) is 5′-tagaAAA CT**CTTTGTT**TGCCATGTCG-3′, and the sequence of HMGmut (sense strand) is 5′-tagaA AATT<u>AGTCGAA</u>TGCCATGTC-3′. The hmg2 probe was created to mimic the pCATSO3 reporter construct in that it contains multiple HMG binding sites separated by 14bp. The sequence of hmg2 (sense strand) is 5′-tagaT**CTTTGTTT**GGCGGATCATGGCT**CTTTGTT T**GGCGGATCATGGA-3′. For gel mobility shift and competition assays, 1–2 μl of nuclear lysate or 3–4 μl of *in vitro* translated protein was incubated (at room temperature) for 30 min in a 20-μl volume containing 20 mM HEPES pH 7.6, 1 mM EDTA, 2 mM $MgCl_2$, 20% glycerol, 50 mM NaCl, 5 μg of bovine serum albumin, 5 μg of p(dGdC)p(dGdC), and 20,000 cpm of labeled probe (final concentration of 0.5–1 nM) with or without competitor. The exact volume of nuclear lysate used in each assay was determined by first normalizing for differences in Sox protein concentration based on Western blot analysis. In gel mobility supershift assays, reaction mixtures were incubated for 1 h at 4°C with the indicated antibody before addition of the probe, after which the reaction was continued at room temperature for 30 min. Immediately after the completion of binding, the reactions were electrophoresed on a nondenaturing 4% Tris-glycine polyacrylamide gel for 4 h at 150V. The gels were then dried and exposed to a PhosphorImager screen for 1–7 days before scanning on a Storm PhosphorImager. Quantitation of band intensities was performed using the ImageQuant 5.0 analysis software. Experiments using hmg2 (see fig. 6, *b* and *c*) were also performed with hmg1 with the same trend in binding intensity between the Sox-11 mutants. All aspects of each mobility shift assay were repeated and similar results were obtained.

Results

Functional Comparison of Sox-2 and Sox-11

The initial goal of this study was to elucidate the mechanisms permitting selective recruitment and activation by specific Sox proteins at a particular gene. To address this issue, we compared the function of Sox-2 with the related transcription factor Sox-11. These proteins were selected for several reasons. First, the gross domain structure of each has been studied, revealing the location of the DNA binding domain and transactivation domain of each protein (4, 19, 20). Second, to understand the function of a transcription factor *in vivo*, it is best studied in a model system using a known binding sequence from a regulated gene. The binding site for Sox-2 within the enhancer of the *FGF-4* gene has been determined and provides an excellent model system (15, 16). Thus far, target genes of Sox-11 have not been

identified; however, the *FGF-4* enhancer sequence has also been utilized in a partial characterization of Sox-11 (20), indicating that it may provide a relevant system in which to study this transcription factor. Finally, the expression pattern of Sox-2 and Sox-11 within the developing nervous system overlaps, leading to the postulate that there are cell types in which some genes may be regulated differentially depending on whether Sox-2 or Sox-11 binds to a cis-regulatory element (24).

To compare Sox-2 and Sox-11 functionally, we assayed their ability to activate the pCATSO3 promoter/reporter gene construct in HeLa cells, which do not contain any known Sox-like activity (15, 19). The pCATSO3 plasmid contains six tandem repeats of a 24-bp region of the *FGF-4* enhancer upstream from an SV40 promoter driving the expression of a chloramphenicol acetyltransferase (CAT) reporter gene. A promoter/reporter construct containing multiple HMG binding sites was necessary for these studies because gene activation by either Sox-2 or Sox-11 via a single HMG site did not rise above the basal expression of the promoter/reporter gene construct (data not shown). The 24-bp region contains both the HMG binding site and POU binding site found in the *FGF-4* enhancer. The pCATSO3 construct was employed to perform a functional comparison of Sox-2 and Sox-11 because it has been used previously to identify the transactivation domain of Sox-2 and to identify p300 as a potential coactivator of Sox-2 activity (19). Furthermore, a very similar construct, also containing the HMG/POU enhancer sequence of the *FGF-4* gene, was utilized in the initial characterization of Sox-11 (20).

For the functional comparison of Sox-2 and Sox-11, HeLa cells were transiently transfected with pCATSO3 and increasing concentrations of expression plasmid encoding either Sox-2F or Sox-11F, each tagged at their N terminus with a FLAG epitope. This revealed that Sox-11F is a much more potent activator of the pCATSO3 promoter/reporter gene construct than Sox-2F (fig. 1*a*). At all concentrations tested, transfection of Sox-11F led to 40–75-fold more reporter gene activity than Sox-2F. This experiment was also repeated with Sox-2 and Sox-11 expression constructs without FLAG epitopes, and no differences were observed between the ability of the tagged and untagged proteins to transactivate (data not shown). Additionally, to verify that the functional differences observed between Sox-2F and Sox-11F were not caused by variations in protein expression, Western blot analysis was performed on nuclear extracts of transfected HeLa cells using the anti-FLAG M2 antibody. This analysis revealed that there was little difference between Sox-2F and Sox-11F expression in the nucleus (fig. 1*b*). In fact, in multiple experiments, we found that Sox-11F expression was consistently lower than Sox-2F. Thus, the difference in transactivation capability between Sox-2F and Sox-11F (fig. 1*a*) may be an underestimate.

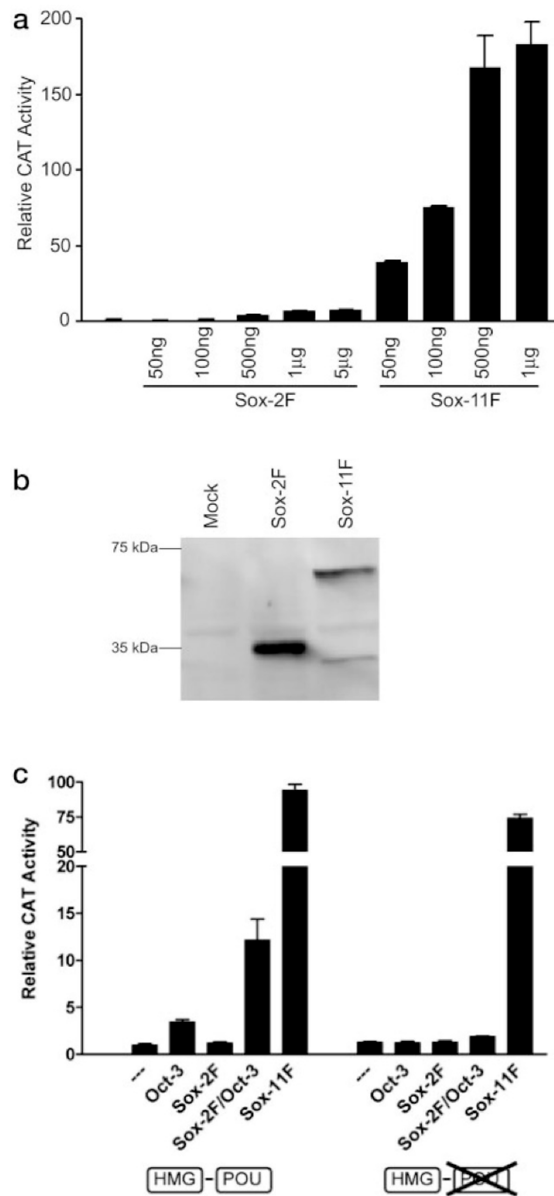


Figure 1. *Transactivation potency and protein expression of Sox-2F and Sox-11F. a* and *c*, duplicate plates of HeLa cells were transfected with 5 μg of pCATSO3 and increasing amounts of plasmid expressing Sox-2F, Sox-11F, or Oct-3, as indicated. Results are shown as the relative CAT expression assayed in each transfection compared with the pCATSO3 reporter alone. The *error bars* represent the S.D. between the CAT activity of the duplicate plates assayed. All transfections included empty CMV5 plasmid to ensure that the final amount of DNA was equivalent. CMV-β-gal was also included to normalize for potential differences in transfection efficiency. Each experiment was repeated at least three times with a representative example shown. *b*, expression of Sox-2F and Sox-11F was determined by Western blot analysis of nuclear extracts obtained from transfected HeLa cells.

The presence of each Sox protein was detected using the anti-FLAG M2 antibody, visualized using ECF, and quantitated using a Storm PhosphorImager and ImageQuant 5.0 software. The molecular masses are shown on the left of the blot.

Our observation that Sox-11F was such a potent activator on its own led us to question whether an endogenous POU factor could be binding at the adjacent site and assisting in Sox-11F-mediated activation. This is an important question because Sox-2 requires the partner Oct-3 for activation. Thus, if Sox-11 can activate significantly without a partner, it would imply that a key functional difference exists between these two proteins. To test this, we constructed the promoter/reporter construct pCATS4, which is similar to pCATSO3 except that the POU site has been scrambled, and there are only four HMG sites in the promoter. When activation of pCATS4 by Sox-2F and Sox-11F is compared with pCATSO3 it is clear that Sox-11F can activate strongly in a partner-independent manner, whereas Sox-2F requires the adjacent POU site and binding of Oct-3 for significant activation (fig. 1*c*).

To understand further the function of Sox-2 and Sox-11, we examined how each of the individual domains of these two proteins contributes to their different abilities to transactivate. For this purpose, we divided Sox-2 and Sox-11 into three regions: the N-terminal region, the HMG DNA binding domain, and the region C-terminal to the DNA binding domain. The N-terminal regions of both proteins contain ~40 amino acids, with no known function in either protein. The HMG DNA binding domains of Sox-2 and Sox-11 are 65% identical and consist of 79 amino acids. The C-terminal regions of both Sox-2 and Sox-11, which are known to contain modular transactivation domains (4, 19, 20), consist of 189 and 265 amino acids, respectively. To determine the role of each of these regions, we created chimeric proteins by interchanging the N-terminal portion, the HMG box (including ~10 amino acids of the C-terminal flanking sequence), or the C-terminal portion of Sox-2 and Sox-11 in all possible combinations. We hypothesized that if one or more of these regions of Sox-11F were responsible for its strong transactivation, placing them into Sox-2F should convert Sox-2F into a strong transactivator. These chimeras were created by first using PCR mutagenesis to insert unique restriction enzyme sites on either side of the plasmid sequence encoding the HMG domain of both proteins. These novel sites were designed to allow for minimal perturbation of the wild-type amino acid sequence, and we determined that they had no effect, either functionally or in protein expression, on Sox-2F or Sox-11F (data not shown). The presence of these unique sites within Sox-2F and Sox-11F allowed us to interchange each domain by restriction digestion and in-frame religation, resulting in six chimeras (fig. 2). The expression level and expected molecular mass of the chimeric proteins were determined by Western blot analysis (fig. 2). Some differences in the level of chimeric protein expression were apparent; for example, the average expression of the constructs containing the Sox-2 C terminus was ~3-fold higher than those containing the Sox-11 C terminus. However, normalizing for these differences during our functional analysis does not affect our conclusions (see below).

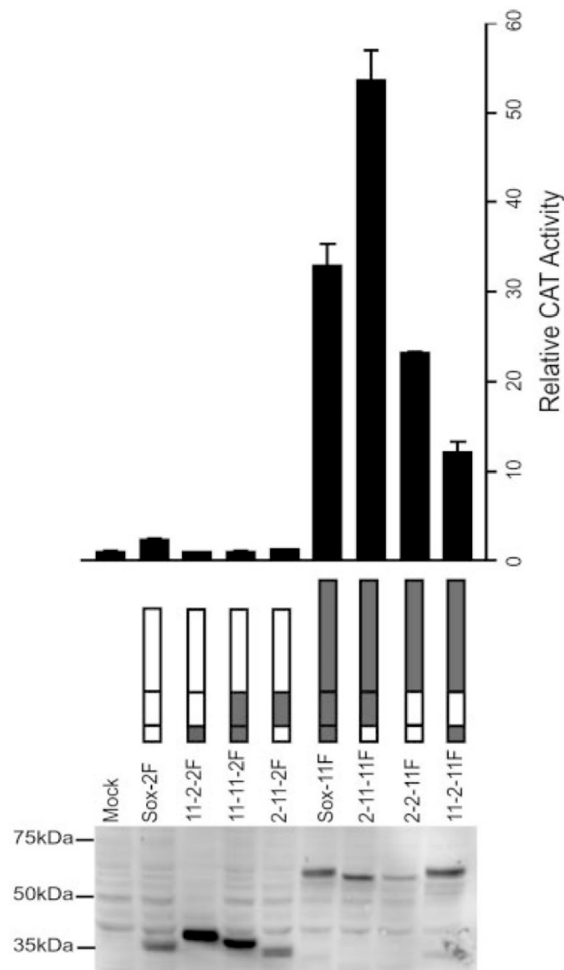


Figure 2. *Transactivation potency and protein expression of Sox- 2/11 chimeras. Top panel*, duplicate plates of HeLa cells were transfected with 5 μg of pCATSO3 and 100 ng of plasmid expressing the indicated Sox chimera. Results are shown as the relative CAT expression assayed in each transfection compared with the pCATSO3 reporter alone. The *error bars* represent the S.D. between the CAT activity of the duplicate plates assayed. *Bottom panel*, expression of chimeras was determined by Western blot analysis of whole cell extracts obtained from transfected HeLa cells. The presence of each Sox protein was detected using the anti-FLAG M2 antibody, visualized using ECF, and quantitated using a Storm PhosphorImager and ImageQuant 5.0 software. The molecular mass is shown on the *left* of the blot. All transfections included empty CMV5 plasmid to ensure that the final amount of DNA was equivalent. CMV-β-gal was also included to normalize for potential differences in transfection efficiency. Each experiment was repeated at least three times with a representative experiment shown.

The ability of each of the chimeras to activate pCATSO3 was then compared with that of Sox-2F and Sox-11F to determine the role of each region in gene activation (fig. 2).

Clearly, there are striking differences in transactivation capabilities between the chimeras containing the C-terminal region of Sox-2F (Sox-11-2-2F, -11-11-2F, and -2-11-2F) and those containing the C-terminal region of Sox-11F (Sox-2-11-11F, -2-2-11F, and -11-2-11F). For example, when the C terminus of Sox-2F is replaced with that of Sox-11F (Sox-2-2-11F) the chimeric protein is capable of activating gene expression 20-fold higher than Sox- 2F. Moreover, in the reciprocal chimera, replacing the C-terminal region of Sox-11F with that of Sox-2F (Sox-11-11-2F) resulted in a 30-fold decrease in transactivation capability. Hence, we conclude that the domain primarily responsible for the functional difference between Sox-2F and Sox-11F lies in their C-terminal regions, which are known to contain their respective transactivation domains (19, 20). As mentioned above, because those chimeras containing the Sox-11 C terminus are less abundant (as determined by Western blot analysis) but exhibit stronger reporter gene activation, normalization to relative protein levels would not alter this conclusion.

Next, we addressed the role of the HMG domain or N terminus in gene activation. When either of these domains is interchanged within those chimeras containing the Sox-2F C terminus, the reporter gene expression is so low that no significant differences can be measured. However, it is apparent that both of these regions play a functional role when the chimeras containing the Sox-11F C terminus are considered (fig. 2). Comparison of either Sox-11F with Sox-11-2-11F or Sox-2-11-11F with Sox-2-2-11F reveals a decrease in activation of ~2.5- fold in each case, indicating that the HMG domain of Sox-11F contributes to stronger potentiation compared with the HMG domain of Sox-2F. A similar consideration of the chimeras containing the N-terminal rearrangements demonstrates that the N terminus of Sox-2F enhances transactivation ~2-fold (compare Sox-11F and Sox-2-11–11F, or Sox-2-2-11F and Sox-11-2-11F). These findings demonstrate that the N terminus and HMG domains of these Sox proteins make a contribution to their ability to activate transcription.

Domain Requirements in Sox-POU Partnerships

Although understanding the properties of a transcription factor on its own is important, it is also critical to understand its ability to work in conjunction with other transcriptional regulators. Indeed, it is widely recognized that many Sox proteins, including Sox-2 and Sox-11, can activate transcription synergistically when bound to DNA adjacent to specific members of the POU family of transcription factors (15, 19, 21, 25). The observation that the HMG and POU domains of Sox-2 and Oct-3 are capable of direct interaction (18) suggests that functional Sox-POU partnerships may depend on a selective interaction between their DNA binding domains. To test this possibility, we compared the ability of Sox-2, Sox-11, and the Sox-2/11 chimeras to cooperate with various POU proteins in transcriptional activation and determined which domains are required for cooperation.

To study the selective nature of Sox-POU cooperation, HeLa cells were transfected with either Sox-2 or Sox-11 and the POU proteins Oct-3 or Brn-2 and then assayed for the ability of each combination to activate the pCATSO3 promoter/reporter gene construct. We first tested the effect of escalating doses of Oct-3 on the ability of Sox-11 to transactivate. Although Sox-11 is known to cooperate with Brn-2 (20), it was unknown how it would respond in the presence of Oct-3. This study demonstrated that Sox-11 is capable of synergistically activating gene expression with Oct-3 (fig. 3*a*), indicating that it cooperates with this POU

factor similarly to Sox-2 (18, 19). To ascertain whether Oct-3 could be cooperating with Sox-11 by up-regulating Sox-11 protein expression, we performed Western blot analysis on duplicate samples of cells transfected with Sox-11F in the presence or absence of Oct-3 and quantitated the expression of Sox-11F (fig. 3*c*). This demonstrated that Oct-3 had little effect on Sox-11 expression. Finally, we examined whether Brn-2 could partner with both Sox-11 and Sox-2. Interestingly, we determined that Sox-11, but not Sox-2, cooperates with Brn-2 to activate transcription (fig. 3*b*). This selective partnership further supports the hypothesis that restrictive mechanisms are in place that allow only specific Sox-POU combinations to cooperate in transcriptional activation.

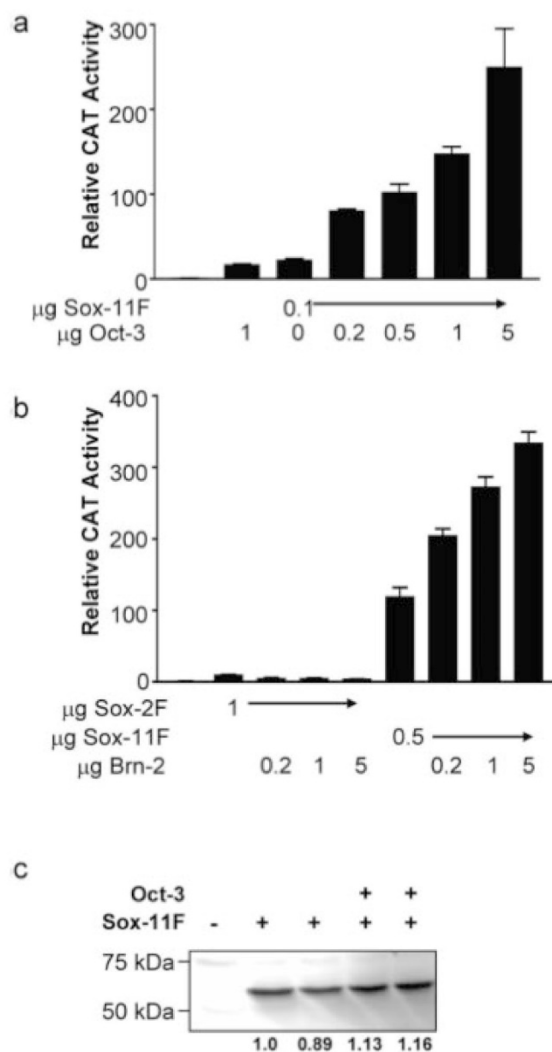


Figure 3. *Sox-2 and -11 cooperate with Oct-3, but only Sox-11 partners with Brn-2.* HeLa cells were transfected with 5 μg of pCATSO3 and the indicated amounts of expression plasmids. *a*, Sox-11F was transfected with increasing doses of Oct-3. *b*, Sox-2F or Sox- 11F was

transfected with increasing doses of Brn-2. Results are shown as the relative CAT expression assayed in each transfection compared with the pCATSO3 reporter alone. The *error bars* represent the S.D. between the CAT activity of the duplicate plates assayed. All transfections included empty CMV5 plasmid to bring the final amount of DNA to 10 μg. CMV-β-gal was also included to normalize for potential differences in transfection efficiency. Each experiment was repeated at least three times with a representative example shown. *c*, expression of Sox-11F was determined by Western blot analysis of whole cell extracts obtained from two separate transfections of HeLa cells with 1 μg of each of the indicated expression plasmids. Sox-11F protein was detected using the anti-FLAG M2 antibody, visualized using ECF, and quantitated using a Storm PhosphorImager and ImageQuant 5.0 software. Molecular mass is shown at the *left* of the blot, and the relative intensities of Sox-11F bands are shown below. All transfections included empty CMV5 plasmid to ensure that the final amount of DNA was equivalent. CMV-β-gal was also included to normalize for potential differences in transfection efficiency. This experiment was repeated three times with a representative example shown.

To begin to elucidate the mechanism underlying the Sox-11/Brn-2 selective partnership, we first examined whether Brn-2 needed a POU binding site adjacent to the HMG site to cooperate with Sox-11 in gene activation. This was addressed by assaying the ability of Sox-11 and Brn-2 to activate pCATS4, in which the POU site is scrambled. In this experiment, we found that removal of the POU site completely abrogates Sox-11/Brn-2 cooperation (fig. 4*a*). This indicates that the enhancement of Sox-11 activation by Brn-2 requires that they be located at adjacent sites on the DNA and is not the result of a general upregulation of Sox-11 protein expression. Next, we examined which domains are needed for these partnerships. Previous studies using Oct-3/1 chimeric proteins have determined that the Oct-3 POU domain is necessary for its cooperation with Sox-2 on the *FGF-4* enhancer sequence. When the POU domain of Oct-3 is exchanged with that of Oct-1, the resulting Oct-3-1-3 protein can no longer synergize with Sox-2 (26). The creation of the Sox-2/11 chimeric proteins provided us with a similar system in which to isolate the domain of Sox-11 which permits selective cooperation with Brn-2. To test this possibility, we compared the ability of Sox-2-11-2F and Sox-11-2-11F to transactivate in the presence of Brn-2 (fig. 4*b*). This study reveals that by interchanging the HMG domain of Sox-2 with that of Sox-11, the chimeric protein Sox-2-11-2F is able to cooperate functionally with Brn-2 in a dose-dependent manner. This is observed in the Brn-2 dependent increase in Sox-11F and Sox-2-11-2F activity above that seen with each protein alone. The importance of the Sox-11 HMG domain is supported further by the fact that when Sox-11-2-11F is transfected in combination with Brn-2 no cooperative activation is detected above the activity of Sox-11-2-11F alone. Thus, we demonstrated that the cooperation of Sox-11 with Brn-2 appears to be dependent on the HMG domain of Sox-11 as well as adjacent binding of the two proteins.

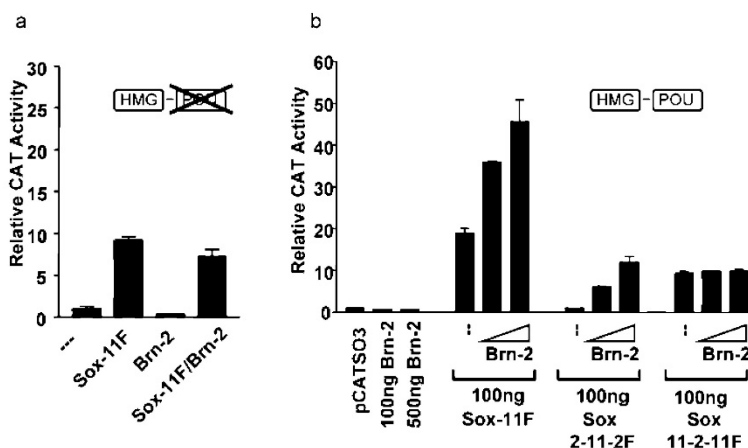


Figure 4. *Adjacent binding sites and the Sox-11 HMG domain are required for cooperation with Brn-2. a*, duplicate plates of HeLa cells were transfected with 5 μg of pCATS4 and 100 ng of Sox-11F, Brn-2, or 100 ng of both as indicated. *b*, duplicate plates of HeLa cells were transfected with 5 μg of pCATSO3 and 100 ng of Sox chimera expression vector with 0, 100, or 500 ng of Brn-2 expression vector. Results are shown as the relative CAT expression assayed in each transfection compared with the pCATS4 or pCATSO3 reporter alone. The *error bars* represent the S.D. between the CAT activity of the duplicate plates assayed. All transfections included empty CMV5 plasmid to bring the final amount of DNA to 10 μg. CMV-β-gal was also included to normalize for potential differences in transfection efficiency. Each experiment was repeated at least three times with a representative experiment shown.

In Vitro DNA Binding of Sox-2 and Sox-11

Binding of Sox transcription factors to DNA has been demonstrated to be an essential and often highly regulated step in their role as transcriptional activators (18, 25, 27–29). With multiple examples of binding regulation within the Sox family and an indication that the HMG DNA binding domains of Sox-2 and Sox-11 have variable effects on selective POU partnerships as shown by our chimeric studies, we examined whether there are differences in the *in vitro* DNA binding capabilities of Sox-2 and Sox-11. To accomplish this goal, HeLa cells were transfected with either Sox-2F or Sox-11F, and nuclear extracts were prepared for use in EMSA to measure the ability of each Sox protein to bind a radiolabeled probe (hmg1) containing a single HMG binding sequence. In each EMSA, the concentration of Sox protein included was first normalized after quantitation by Western blot analysis. When *in vitro* DNA binding of Sox-2F is compared with Sox-11F, a single complex was observed with Sox-2F, whereas little or no binding was detected with Sox-11F (fig. 5*a*). The Sox-2F complex was partially supershifted specifically with the M2 antibody and competed by excess, unlabeled wildtype probe, but not a probe in which the HMG site has been scrambled (fig. 5*a*). In contrast, we detected little or no binding of full-length Sox-11F despite our attempts to use multiple binding buffers, different nonspecific competitors, or lower ionic strength electrophoresis conditions (data not shown). To examine roles of the HMG domains of Sox-2F and Sox-11F in mediating their differences in DNA binding, an

EMSA of the Sox chimeras in which the HMG domains were interchanged was also performed (fig. 5*b*). Surprisingly, Sox-2-11-2F was capable of binding the DNA probe with an intensity similar to that of Sox-2F, whereas DNA binding by Sox-11-2-11F was not detected. Thus, the region(s) responsible for differential DNA binding between Sox-2F and Sox-11F lies outside of the HMG domain.

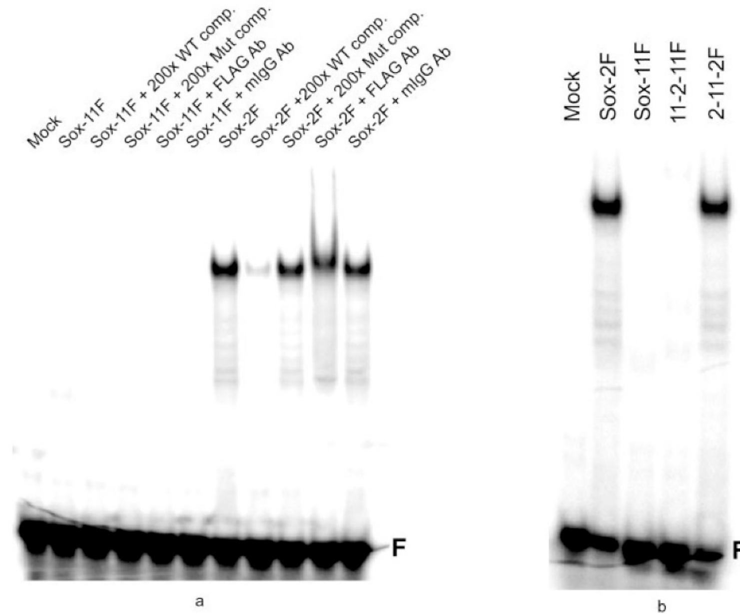


Figure 5. In vitro *DNA binding of Sox proteins. a*, equal amounts of Sox-2F or Sox-11F protein obtained from transfected HeLa cells were studied by EMSA with radiolabeled hmg1 probe. To determine binding specificity, excess unlabeled wild-type or mutant probe was included in the binding reaction as indicated. Supershift analysis was performed by incubating the binding reaction for 1 h at 4°C with either anti-FLAG M2 antibody or an equal amount of nonspecific mouse IgG antibody before adding the probe. *b*, equal amounts of the indicated Sox protein obtained from transfected HeLa cells were studied by EMSA with radiolabeled hmg1 probe. The EMSA gel was exposed to a PhosphorImager cassette for detection and quantification using a Storm PhosphorImager and ImageQuant 5.0 software. Each aspect of these experiments was repeated at least twice using multiple preparations of nuclear extract. The position of the free probe in the gel is indicated by *F*.

Investigation of the DNA Binding Autoregulation of Sox-11

The observation that region(s) of Sox-11 beyond its HMG domains may repress DNA binding led us to hypothesize that a domain capable of autoinhibiting Sox-11 exists in this protein. To determine whether an autoinhibitory region exists within Sox-11, deletion constructs were made and tested for their ability to bind DNA. Our deletion strategy focused on two acidic regions within the C terminus, which were particularly intriguing because of the acidic nature of autoinhibitory regions isolated in other proteins (8, 30–33). The distal region (amino acids 283–395) contains an acidic/hydrophobic region known to

contain the transactivation domain (TAD) (20), while the central region (amino acids 189–224) contains a highly acid-rich (AR) set of amino acids (20/27 consecutive residues being Asp or Glu). The AR region of Sox-11 cannot act as a TAD (20),[2] nor does it have any other known function.

To study each of these regions, three Sox-11F deletion constructs (fig. 6*a*) were placed in vectors allowing protein production either through *in vitro* translation (fig. 6*b*) or in HeLa cells (fig. 6*c*). For these studies, protein production was determined by Western blot analysis (fig. 6*b*, *top*), and a similar amount of each Sox-11F deletion construct was included in the binding reaction (fig. 6*b*, *bottom*). After protein normalization, the DNA binding capability of the Sox-11F deletion mutants was determined by EMSA. This was accomplished using either the hmg1 or the hmg2 probe, which contains two HMG sites separated by 14 bp. The experiments performed with hmg2 revealed the same trend in binding intensity between the Sox-11 mutants as the hmg1 probe (data not shown). Our examination of the DNA binding ability of *in vitro* translated Sox proteins revealed that identical to our observations from HeLa cell extracts, Sox-2F binding was much stronger that Sox-11F (fig. 6*b*, compare *lanes 2* and *3*) despite similar protein expression. Interestingly, when the AR region is removed from Sox-11F, the protein binds well to the radiolabeled probe (*lane 4*). Furthermore, although some binding by Sox-11FΔTAD is observed (*lane 5*), it is also enhanced upon removal of the central AR region (*lane 6*). These results indicate that the AR region of Sox-11 can inhibit DNA binding of *in vitro* translated protein.

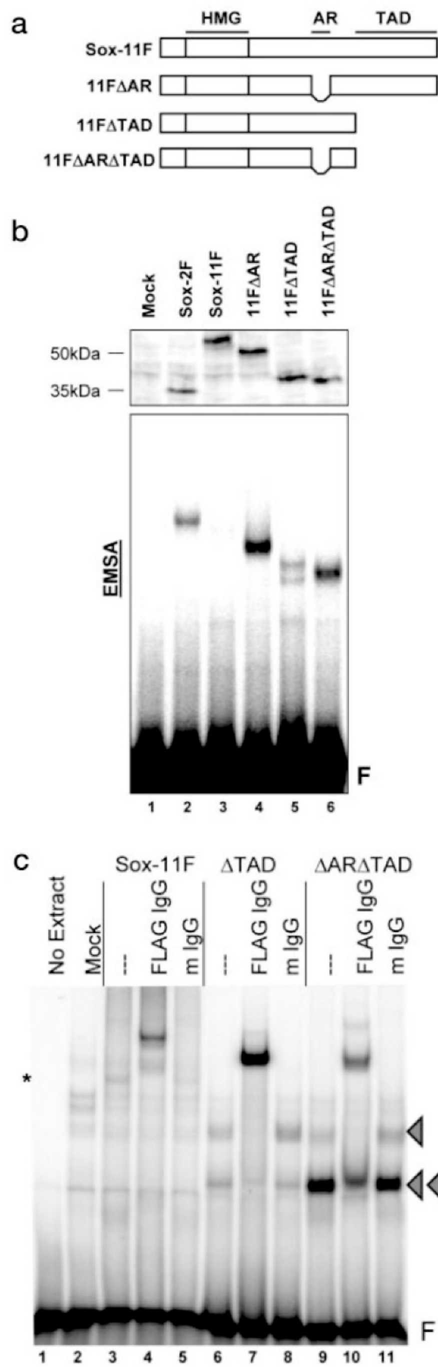


Figure 6. *Identification of autoregulatory regions within Sox-11. a*, diagram of Sox-11F deletion constructs. Sox-11F represents the full-length protein (amino acids 1–395). Amino acids 190–223 were deleted from Sox-11FΔAR. Amino acids 283–395 were deleted from Sox-11FΔTAD. Amino acids 190–223 and 283–395 were deleted from Sox-11FΔARΔTAD. *b, top*, Western blot analysis of the *in vitro* translated Sox proteins. The presence of each Sox

protein was detected using the M2 antibody, visualized using ECF detected by a Storm PhosphorImager, and quantified using ImageQuant 5.0 software. The molecular mass is shown on the *left* of the blot. *Bottom*, equal amounts of *in vitro* translated Sox-2F, Sox-11F protein, or the indicated deletion construct were studied by EMSA with radiolabeled hmg2 probe. *c*, equal amounts of Sox-11F protein or the indicated deletion construct produced in HeLa cells were studied by EMSA with radiolabeled hmg2 probe. Supershift analysis was performed by incubating the binding reaction for 1 h at 4°C with either anti-FLAG M2 antibody or an equal amount of nonspecific mouse IgG antibody before adding the probe. The EMSA gels were exposed to a PhosphorImager cassette for detection using a Storm PhosphorImager and quantified using ImageQuant 5.0 software. Each aspect of these experiments was repeated at least twice using multiple preparations of extract. The position of the free probe in the gel is indicated by *F*.

To determine whether the AR region is able to repress DNA binding of *in vivo* produced protein, we also compared binding of proteins expressed in HeLa cells (fig. 6*c*). This was an important question because Sox-11 may have post-translational modifications *in vivo* which affect DNA binding and are not present on *in vitro* translated protein. As described previously, in this DNA binding analysis, we normalized for any differences in protein expression by including equal amounts in the binding reactions, based on prior Western blot analysis (data not shown). Using extract containing Sox-11F, a faint DNA-protein complex was observed (indicated by *), which could be supershifted specifically by the anti-FLAG M2 antibody (*lanes 4* and *5*). In comparison, Sox-11FΔTAD binding resulted in two complexes (*lane 6*, *single* and *double arrowheads*), each of which was shifted specifically by the M2 antibody (*lanes 7* and *8*). The finding that the intensities of the Sox-11F and Sox-11FΔTAD complexes were relatively similar suggests that the TAD of Sox-11F does little on its own to affect binding to DNA *in vitro*. Next, the role of the AR region of Sox-11 in binding to DNA was examined. Importantly, removal of this region resulted in the formation of an intense, faster migrating complex (*lane 9*, *double arrowhead*) as well as a less intense, slower migrating complex (*lane 9*, *single arrowhead*). Both complexes were shifted specifically by the M2 antibody (*lanes 10* and *11*), indicating they each contain Sox-11FΔARΔTAD, although the supershift of the faster complex resulted in a smaller change in migration. This smaller shift was observed in multiple experiments and may indicate that either this antibody-Sox complex adopts a compact conformation, or perhaps it is destabilized during electrophoresis, resulting in the release of the Sox protein from the antibody. The binding intensity of Sox-11FΔARΔTAD was estimated to be 10-fold greater than Sox-11F (as described under "Experimental Procedures"), indicating that this region acts strongly as a autoinhibitory domain. The significance of multiple DNA-protein complexes with Sox-11FΔARΔTAD and Sox-11FΔTAD is not clear at this time, but they do not alter the conclusion that a strong autoinhibitory region exists in Sox-11F. It is also interesting to note that in the case of both Sox-11F and Sox-11FΔTAD, the addition of the M2 antibody led to an increase in the intensity of the DNA-protein complex (compare *lanes 3* and *4* or *lanes 6* and *7*), indicating that antibody binding may in some way neutralize the autoinhibitory region. We believe this further supports the hypothesis that an autoinhibitory region exists within Sox-11F because antibody binding has been found to release autoinhibition

in both PEA3 and p53 (34, 35). Taken together our findings suggest that a novel autoregulatory domain is present in Sox-11, which we suspect plays an important role in its function as a transcription factor *in vivo*. Furthermore, because autoinhibition occurs with both *in vitro* and *in vivo* produced protein it is most likely an intrinsic property of the protein rather than the result of post-translational modification or interaction with another factor.

Influence of the Autoinhibitory Domain on Sox-11 Gene Activation

Autoinhibitory domains present in other transcription factor families are thought to act as regulatory switches capable of repressing recruitment to a gene unless inhibition is relieved. For example, the Ets family member PEA3 contains two domains that autoinhibit DNA binding to an optimized PEA3 binding site *in vitro*. Evidence that these autoinhibitory domains also regulate *in vivo* recruitment of PEA3 was observed when the ability of full-length PEA3 to activate a reporter gene was compared with a deletion construct in which one autoinhibitory domain was removed. In these studies, removal of the autoinhibitory domain increased activation of the reporter gene ~2-fold, which indicates that the optimal capability of full-length PEA3 to activate gene expression was reduced (34).

To examine whether the AR region can regulate the ability of Sox-11 to activate gene expression from transfected promoter/ reporter gene constructs, we compared Sox-11F and Sox-11FΔAR activation of reporter gene constructs. If the AR region of Sox-11F represses its DNA binding *in vivo*, its removal may enhance gene activation. Alternatively, because Sox-11F can clearly activate pCATSO3 in HeLa cells, autoinhibition may already be relieved *in vivo*, and thus removal of the AR region may have no effect on Sox-11 activation. In these studies we examined the effect of the AR region on Sox-11 activation via two bona fide Sox target sequences, the *FGF-4* HMG site as well as the HMG site present in the δ-crystallin gene, hereafter referred to as the DC5 site. This site was chosen because it is less strongly activated by Sox-11 compared with the FGF-4 HMG site (36).[2] The reporter construct pCATDC3 was made to be identical to pCATSO3 with the exception of 3 bp within each of the six HMG binding sites to conform to the DC5 sequence. When observing the activation of pCATSO3 by Sox-11FΔAR compared with Sox-11F, we found that there was a reproducible 1.4–2-fold increase in activation when the AR region was absent (fig. 7*a*). Furthermore, the activation by Sox-11FΔAR compared with Sox-11F on pCATDC3 revealed that the difference in gene activation was even more pronounced with this construct, increasing in a dose-dependent manner to an almost 5-fold enhancement in gene activation upon removal of the AR region (fig. 7*b*). Finally, because both Sox-11FΔAR and Sox-11F are expressed at almost identical levels in HeLa cells (fig. 7*c*), expression differences do not explain the functional differences we are observing. Therefore, these studies support a model in which the ability of Sox-11 to bind to a gene and regulate its expression *in vivo* may be modulated by a acidic autoinhibitory domain.

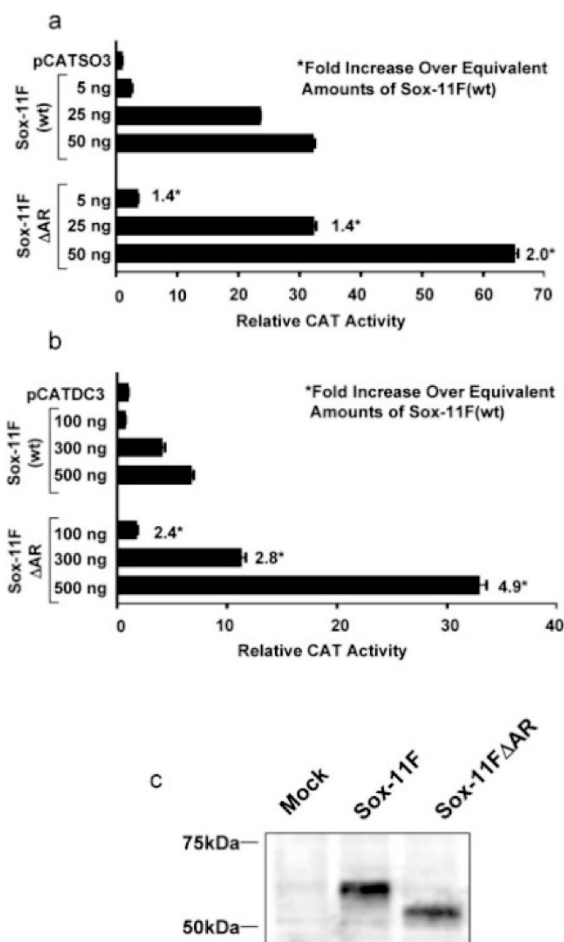


Figure 7. *The Sox-11 autoinhibitory domain represses gene activation* in vivo. *a* and *b*, duplicate plates of HeLa cells were transfected with 5 μg of either pCATSO3 (*a*) or pCATDC3 (*b*) along with the indicated amounts of either Sox-11F or Sox-11FΔAR expression vector. The differences in amounts of Sox plasmid transfected with pCATSO3 compared with pCATDC3 were necessary because of the differing sensitivities of the reporter constructs to Sox-11 and were selected as the minimal amounts needed to achieve activation above the basal reporter gene expression. The *number* to the *right* of each *bar* represents the -fold difference in activation by Sox-11FΔAR compared with full-length Sox-11F at the same amount of transfected expression plasmid. All transfections included the empty CMV5 plasmid to bring the final amount of DNA to 10 μg. CMV-β-gal was also included to normalize for potential differences in transfection efficiency. Each experiment was repeated at least three times with a representative experiment shown. *c*, expression of Sox-11F and Sox-11FΔAR was determined by Western blot analysis of nuclear extracts obtained from transfected HeLa cells. The presence of each Sox protein was detected using the anti-FLAG M2 antibody, visualized using ECF, and quantitated using a Storm PhosphorImager and ImageQuant 5.0 software. The molecular mass is shown on the *left* of the blot.

Discussion

In this study, we compared the ability of Sox-2 and Sox-11 to activate expression of promoter/reporter gene constructs via the HMG binding site found in the *FGF-4* distal enhancer, which is a well-characterized regulatory element for studying the action of Sox proteins (15, 19, 20). This comparison has led to the identification of three Sox regulatory mechanisms, which together can regulate gene expression at multiple levels. First, we demonstrate that the transactivation domains of Sox-2 and Sox-11 are key determinants of their differential effects on transcription. Second, our study of Sox-POU partnerships demonstrates that the selection of Brn-2 by Sox-11 is determined by its HMG domain, highlighting the importance of this domain not only in DNA-protein interactions, but in protein-protein interactions as well. Third, the identification of a novel autoinhibitory region within Sox-11 is an important advance in understanding how its binding to DNA may be regulated. These studies point to a model of Sox function in which DNA binding is not controlled by a single domain, but rather may be modulated by interplay with another region of the protein, leading to finely tuned gene activation.

Function of Sox Proteins Is Highly Dependent on Their TADs

Using a panel of Sox-2/11 chimeric proteins, we demonstrate that Sox-11 is a much more potent transcriptional activator than Sox-2 in the model system used in this study. A functional comparison of these chimeras demonstrates that although the N-terminal and HMG domains modulate gene activation to a small extent, the primary region responsible for the potency of Sox-11 lies within its C-terminal TAD. In several cases, the potency of a TAD has been linked to its primary amino acid sequence. Domains that are rich in acidic and hydrophobic residues function differently in some contexts compared with those composed of glutamine-, serine-, or proline-rich regions (37). We demonstrate that this pattern is also observed in our comparison of Sox-2, which has a serine/proline TAD, and Sox-11, which has a C terminus consisting of acidic and hydrophobic amino acids. Interestingly, Sox-4, Sox-22, and rainbow trout Sox-24 show high homology to the acidic TAD of Sox-11 at their C terminus (38–40). This homology suggests that these proteins are also very strong transcriptional activators. In fact, we have determined that this is true of Sox-4, which we determined to be capable of 8–10-fold greater activation of the pCATSO3 promoter/reporter gene construct than Sox-2.[2] These differences in transactivation potency within the Sox family provide a mechanism for altering gene expression through selective recruitment of one Sox protein over another.

Sox-POU Cooperation Is Selective and Dependent on the HMG Domain

Although the study of a Sox factor on its own yields important insight into protein function, many Sox family members are influenced by other factors that bind DNA at adjacent sites. Specifically, Sox-2 and Sox-11 have been found to cooperate with the POU transcription factors Oct-3 and Brn-2, respectively (15, 19, 20). The observation that partnerships between Sox and POU factors are not promiscuous, but rather are partner-specific has led to the hypothesis that a Sox-POU "code" exists which allows only some Sox-POU pairings to form productive complexes (20). Our studies give further credence to this hypothesis

because Sox-2 is capable of cooperating only with Oct-3 and not Brn-2. Additionally, Sox-11 cooperates with Brn-2 (also shown in previous studies (20)) and Oct-3.

The cooperative nature of Sox proteins has been found to require the HMG domain in multiple cases. Selective HMG-mediated interaction with another transcription factor has been observed on the *FGF-4* enhancer where Sox-2 and Oct-3 have been shown to partner via their DNA binding domains (15, 16). The importance of Sox HMG domains has also been shown in the case of the δ-crystallin (DC5) enhancer, where either Sox-1 or Sox-2 can partner with the paired domain transcription factor Pax6 (29). On the DC5 enhancer, the partnership of Sox-1 or Sox-2 with Pax6 is specific to their HMG domains and cannot be satisfied by the HMG domain of Sox-9 (36). These examples indicate that although HMG domains exhibit the highest homology of any Sox domain, the minor differences that exist in this domain are capable of generating specific partnership interactions. Consistent with these precedents of HMG-mediated selection, we show here that the HMG domain of Sox-11 is responsible for its ability to cooperate with Brn-2 because only Sox-11F and Sox-2-11-2F can partner with Brn-2, whereas Sox-2F and Sox-11-2-11F cannot. Thus, there is growing support for a model in which the HMG domain serves two functions, DNA binding and partner selection, which may permit selective recruitment of Sox proteins only to specific genes.

A Novel Autoinhibitory Region Is Present in Sox-11

To expand our study of Sox-2 and Sox-11, we compared their ability to bind DNA *in vitro*. Using both *in vivo* and *in vitro* produced proteins, we demonstrate the formation of an intense Sox-2·DNA complex, whereas little or no binding of Sox-11 was detected, despite our use of the same amount of protein in the binding reaction. We also observed that Sox-2-11-2F bound DNA with intensity similar to Sox-2, whereas in contrast Sox-11-2-11F was incapable of binding under these conditions. This suggests that the minimal HMG domains of the two proteins are likely to bind DNA to a similar extent. The similar binding of Sox-2F and Sox-2-11-2F indicates that regions of Sox-11 beyond the HMG domain must influence DNA binding *in vitro*. Using a series of deletions constructs, we determined that a highly AR region within Sox-11 is capable of sharply repressing DNA binding by the Sox-11 HMG domain *in vitro*, identifying it as an autoinhibitory region. Importantly, the repression of DNA binding by the AR region is seen with *in vitro* produced protein as well as HeLa-derived proteins. This suggests that inhibition is an intrinsic property of the protein, rather than the result of post-translational modification or interaction with a repressive factor present in HeLa extract. The AR region is particularly interesting because of its homology to Glu/Asp acid-rich stretches present in the HMG-1/2 proteins, which have been found to act as powerful negative regulators of DNA binding in those proteins as well (8, 41). Furthermore, the existence of acidic autoinhibitory regions has been reported in a wide range of transcription factors including RXR, Rfx1, E12, and Nkx6.1, although never before in a Sox protein (30–33). The importance of the acidic region within Sox-11 is highlighted by the fact that long acidic stretches are also present within other Sox family members, such as Sox-22 and Sox-24 (39, 40). Therefore, other Sox proteins may also contain binding autoinhibitory regions.

There are several mechanisms through which binding autoregulatory regions may act. The functional and sequence similarities between the autoinhibitory region of Sox-11 and the extensively characterized acidic tails of chromosomal HMG proteins suggest that the two protein groups may be regulated by related mechanisms involving negative charge. This possibility is supported by a study of the repressive acidic tails of chromosomal HMG-1 and HMG-2. It has been shown that the longer acidic tail of HMG-1 results in decreased ability to bind DNA compared with HMG-2 (41). These negatively charged autoinhibitory domains may be repulsed by the negative charge on the phosphate backbone of DNA or directly contact the positively charged HMG domain, resulting in a decreased affinity of the HMG domain for DNA. Indeed, in the case of HMG-1/2, it has been found that the negative acidic tail can interact directly with its HMG domain (42). It is possible that the autoinhibitory domain of Sox-11 acts similarly, perhaps by directly interacting with other domains of Sox-11 or by inducing conformation changes, resulting in decreased DNA binding as postulated in figure 8.

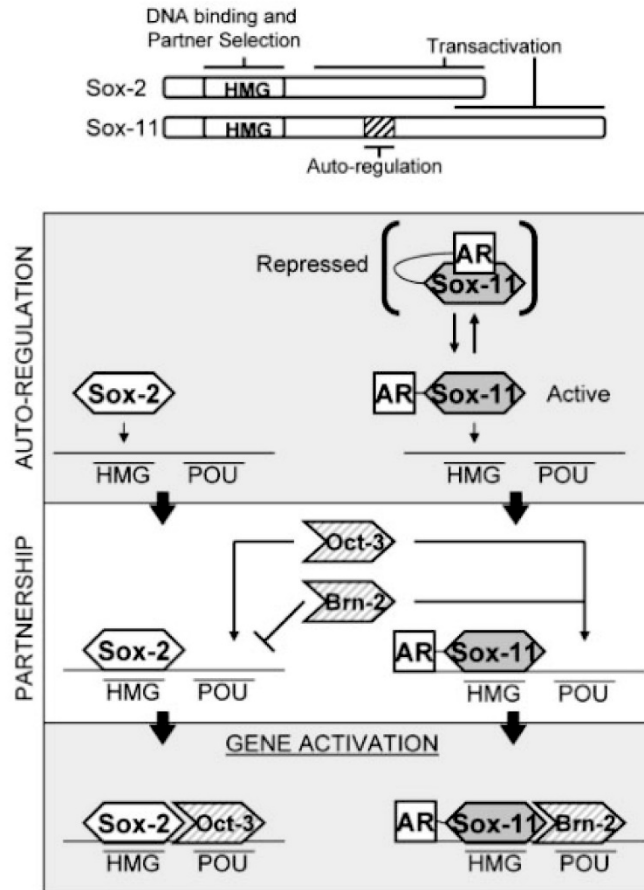


Figure 8. *Multiple levels of regulation impact Sox function.* This model illustrates the regulatory mechanisms examined in this study which may influence Sox protein function. These

> include autoinhibition as a mechanism for controlling DNA binding, which we hypothesize is the result of Sox-11 existing in either an active conformation or a repressed conformation that requires the AR region of Sox-11. Additionally, selective partner interactions through the HMG domains of Sox-2 and Sox-11 can influence which POU proteins cooperate with Sox proteins in gene activation. Finally, if multiple Sox proteins are able to bind the same site, gene activation may be modulated through their differing transactivation capabilities.

To permit DNA binding *in vivo*, autoinhibitory domains must be neutralized. This has been shown to occur through posttranslational modifications and/or direct interactions with partner proteins. Specifically, in the study of cooperative DNA binding between HMG-1 and the TATA-box binding protein, cooperation has been found to require the acidic tail of HMG-1 (11). This is not surprising because the hydrophilic nature of acidic autoinhibitory domains increases the likelihood that they are exposed at the surface of the protein, which places them in an ideal position to interact with other proteins. Additionally, DNA binding of the maize chromosomal HMGB family can be regulated by phosphorylation of the acidic tail, indicating that acidic autoinhibitory regions can be regulated through multiple mechanisms (12). Other examples for such a model of transcription factor cooperation have also been found during the study of the autoinhibitory domains of Ets-1 and PEA3, which are neutralized upon interaction with the transcription factors AML1 and USF-1, respectively, thus enhancing DNA binding (9, 43). Similarly, the transcription factor LEF-1 contains an autoinhibitory domain, which is neutralized upon binding of the coactivator β-catenin, resulting in enhanced DNA binding (44). Interestingly, our promoter/reporter gene activation studies indicate that the Sox-11 autoinhibitory domain may be at least partially neutralized in *vivo* because Sox-11 is able to bind DNA and transactivate when expressed in HeLa cells. Importantly, deletion of the AR inhibitory region leads to an ~2–5-fold boost in gene activation via two different promoter/reporter gene constructs, indicating that its autoinhibitory capacity is present *in vivo* and is capable of regulating gene expression. Based on the conservation of sequence and function between chromosomal HMG proteins and Sox-11, it is plausible that the neutralization of the Sox-11 acidic autoinhibitory domain in the full-length protein may also occur through a number of complex mechanisms, including posttranslational modification or interaction with other proteins.

In conclusion, this study makes several important contributions to the understanding of how Sox-2 and Sox-11 are regulated functionally, both on their own and in cooperation with other transcriptional regulators. We demonstrate that when the two proteins are compared on their own, their transactivation domains play the primary role in determining their functional differences. However, in studying the selective partnership of Sox-11 and Brn-2, the HMG DNA binding domain was found to play a critical role. Building on the example observed in the Sox-2/Oct-3 partnership, this study provides further evidence that the Sox-POU code may depend on the DNA binding domains of these transcription factors. Finally, we have identified a novel autoinhibitory region within Sox-11 capable of significantly influencing both DNA binding *in vitro* and gene expression *in vivo*. This observation places Sox-11 on an expanding list of transcription factors whose DNA binding appears to

be autoregulated by regions of the protein beyond the DNA binding domain. Furthermore, as the autoinhibitory region identified within Sox-11 shows homology to regions of other Sox proteins, this study suggests a similar mechanism may influence the function of several other Sox family members.

Acknowledgments – This work was supported in part by Grant CA 74771 from the NCI, National Institutes of Health, and by Cancer Center Support Grant CA 36727 (to the core facilities of the University of Nebraska Medical Center Eppley Cancer Center). The authors were supported by Training Grant CA 09476 from the NCI, National Institutes of Health. Additionally, Matthew Wiebe was supported by an Emley graduate fellowship.

Notes

1. The abbreviations used are: HMG, high mobility group; AR, acid-rich; *β*-gal, *β*-galactosidase; CAT, chloramphenicol acetyltransferase; CMV, cytomegalovirus; DC5, *δ*-crystallin enhancer; ECF, enhanced chemifluorescence; EMSA, electrophoretic gel mobility shift analysis; FGF-4, fibroblast growth factor 4; POU, Pit-Oct-Unc; TAD, transactivation domain.
2. M. Wiebe and A. Rizzino, unpublished observations.

References

1. Wegner, M. (1999) *Nucleic Acids Res.* 27, 1409–1420.
2. Harley, V. R., Lovell-Badge, R., and Goodfellow, P. N. (1994) *Nucleic Acids Res.* 22, 1500–1501.
3. Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N., and Lovell-Badge, R. (1996) *Development* 122, 509–520.
4. Kamachi, Y., Uchikawa, M., Collignon, J., Lovell-Badge, R., and Kondoh, H. (1998) *Development* 125, 2521–2532.
5. Cheung, M., Abu-Elmagd, M., Clevers, H., and Scotting, P. J. (2000) *Brain Res. Mol. Brain Res.* 79, 180–191.
6. Hagman, J., and Grosschedl, R. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 8889–8893.
7. Petersen, J. M., Skalicky, J. J., Donaldson, L. W., McIntosh, L. P., Alber, T., and Graves, B. J. (1995) *Science* 269, 1866–1869.
8. Isackson, P. J., Beaudoin, J., Hermodson, M. A., and Reeck, G. R. (1983) *Biochim. Biophys. Acta* 748, 436–443.
9. Goetz, T. L., Gu, T. L., Speck, N. A., and Graves, B. J. (2000) *Mol. Cell. Biol.* 20, 81–90.
10. Cowley, D. O., and Graves, B. J. (2000) *Genes Dev.* 14, 366–376.
11. Das, D., and Scovell, W. M. (2001) *J. Biol. Chem.* 276, 32597–32605.
12. Stemmer, C., Schwander, A., Bauw, G., Fojan, P., and Grasser, K. D. (2002) *J. Biol. Chem.* 277, 1092–1098.
13. Curatola, A. M., and Basilico, C. (1990) *Mol. Cell. Biol.* 10, 2475–2484.
14. Ma, Y. G., Rosfjord, E., Huebert, C., Wilder, P., Tiesman, J., Kelly, D., and Rizzino, A. (1992) *Dev. Biol.* 154, 45–54.
15. Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995) *Genes Dev.* 9, 2635–2645.

16. Johnson, L. R., Lamb, K. A., Gao, Q., Nowling, T. K., and Rizzino, A. (1998) *Mol. Reprod. Dev.* 50, 377–386.
17. Nowling, T., Bernadt, C., Johnson, L., Desler, M., and Rizzino, A. (December 17, 2002) *J. Biol. Chem.* 10.1074/jbc.M207567200.
18. Ambrosetti, D. C., Basilico, C., and Dailey, L. (1997) *Mol. Cell. Biol.* 17, 6321–6329.
19. Nowling, T., Johnson, L. R., Wiebe, M., and Rizzino, A. (2000) *J. Biol. Chem.* 275, 3810–3818.
20. Kuhlbrodt, K., Herbarth, B., Sock, E., Enderich, J., Hermans-Borgmeyer, I., and Wegner, M. (1998) *J. Biol. Chem.* 273, 16050–16057.
21. Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I., and Wegner, M. (1998) *J. Neurosci.* 18, 237–250.
22. Scholtz, B., Lamb, K., Rosfjord, E., Kingsley, M., and Rizzino, A. (1996) *Dev. Biol.* 173, 420–427.
23. Kelly, D., Scholtz, B., Orten, D. J., Hinrichs, S. H., and Rizzino, A. (1995) *Mol. Reprod. Dev.* 40, 135–145.
24. Uwanogho, D., Rex, M., Cartwright, E. J., Pearl, G., Healy, C., Scotting, P. J., and Sharpe, P. T. (1995) *Mech. Dev.* 49, 23–36.
25. Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999) *Mol. Cell. Biol.* 19, 5453–5465.
26. Ambrosetti, D. C., Scholer, H. R., Dailey, L., and Basilico, C. (2000) *J. Biol. Chem.* 275, 23387–23397.
27. Lefebvre, V., Li, P., and de Crombrugghe, B. (1998) *EMBO J.* 17, 5718–5733.
28. Peirano, R. I., and Wegner, M. (2000) *Nucleic Acids Res.* 28, 3047–3055.
29. Kamachi, Y., Uchikawa, M., Tanouchi, A., Sekido, R., and Kondoh, H. (2001) *Genes Dev.* 15, 1272–1286.
30. Mirmira, R. G., Watada, H., and German, M. S. (2000) *J. Biol. Chem.* 275, 14743–14751.
31. Sun, X. H., and Baltimore, D. (1991) *Cell* 64, 459–470.
32. Zhao, Q., Chasse, S. A., Devarakonda, S., Sierk, M. L., Ahvazi, B., and Rastinejad, F. (2000) *J. Mol. Biol.* 296, 509–520.
33. Katan-Khaykovich, Y., and Shaul, Y. (2001) *Eur. J. Biochem.* 268, 3108–3116.
34. Bojovic, B. B., and Hassell, J. A. (2001) *J. Biol. Chem.* 276, 4509–4521.
35. Hansen, S., Lane, D. P., and Midgley, C. A. (1998) *J. Mol. Biol.* 275, 575–588.
36. Kamachi, Y., Cheah, K. S. E., and Kondoh, H. (1999) *Mol. Cell. Biol.* 19, 107–120.
37. Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J., and Bentley, D. (1996) *Mol. Cell. Biol.* 16, 2044–2055.
38. van de Wetering, M., Oosterwegel, M., van Norren, K., and Clevers, H. (1993) *EMBO J.* 12, 3847–3854.
39. Jay, P., Sahly, I., Goze, C., Taviaux, S., Poulat, F., Couly, G., Abitbol, M., and Berta, P. (1997) *Hum. Mol. Genet.* 6, 1069–1077.
40. Kanda, H., Kojima, M., Miyamoto, N., Ito, M., Takamatsu, N., Yamashita, S., and Shiba, T. (1998) *Gene* (*Amst*.) 211, 251–257.
41. Lee, K. B., and Thomas, J. O. (2000) *J. Mol. Biol.* 304, 135–149.
42. Ramstein, J., Locker, D., Bianchi, M. E., and Leng, M. (1999) *Eur. J. Biochem.* 260, 692–700.
43. Greenall, A., Willingham, N., Cheung, E., Boam, D. S., and Sharrocks, A. D. (2001) *J. Biol. Chem.* 276, 16207–16215.
44. Tutter, A. V., Fryer, C. J., and Jones, K. A. (2001) *Genes Dev.* 15, 3342–3354.