THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 50, Issue of December 14, pp. 47296–47302, 2001
Printed in U.S.A.

Spatially Precise DNA Bending Is an Essential Activity of the Sox2 Transcription Factor*

Received for publication, August 9, 2001, and in revised form, September 20, 2001
Published, JBC Papers in Press, October 2, 2001, DOI 10.1074/jbc.M107619200

Paola Scaffidi and Marco E. Bianchi‡

From the San Raffaele Scientific Institute, via Olgettina 58, 20132 Milan, Italy and Faculty of Medicine, Universitá Vita-Salute San Raffaele, via Olgettina 58, 20132 Milan, Italy

Sox proteins, a subclass of high mobility group box proteins, govern cell fate decisions by acting both as classical transcription factors and architectural components of chromatin. We aimed to demonstrate that the DNA bending activity of Sox proteins is essential to regulate gene expression. We focused on mouse Sox2, which participates in the transactivation of the *Fgf4* (fibroblast growth factor 4) gene in the inner cell mass of the blastocyst. We generated six substitutions in the high mobility group box of Sox2. One mutant showed a reduced DNA bending activity on the *Fgf4* enhancer (46° instead of 80°), which resulted in more powerful transactivation compared with the wild type protein. We then selected two single-base mutations in the *Fgf4* enhancer that make the DNA less bendable by the Sox2 protein. Again, a different DNA bend (0° and 42° instead of 80°) resulted in a different activation of transcription, but in this case reduced bending corresponded to decreased transcription. We found that the opposite effect on transcription of similar DNA bending angles is due to a 20° difference in the relative orientation of the DNA bends, proving that a correct three-dimensional geometry of enhanceosome complexes is necessary to promote transcription.

Transcriptional activation of eukaryotic genes requires the assembly of multiprotein complexes on promoter and enhancer sequences. The right combination of factors binding to specific DNA sequences ensures specificity, and the contacts between the different proteins increase the stability of the complex. However, the various transcription factors are often bound to DNA at nonadjacent sites, and direct contacts between them are possible only if they are juxtaposed through the deformation of the DNA helix. Long DNA fragments have sufficient flexibility and can form loops, bringing in close contact proteins bound to distant sites. On the contrary, fragments smaller than 500 bp[1] are rather stiff (1), and the formation of multiprotein complexes can require specific proteins that introduce bends in DNA. Proteins of the HMG box family have DNA bending activity and are thus expected to act as architectural factors.

The HMG box family is composed of chromatin proteins and transcription factors that contain the DNA binding domain of the same name (2, 3). All of these proteins share the ability to recognize distorted DNA structures and introduce bends in DNA. However, although chromatin proteins HMGB1 and 2 bind to DNA in a sequence-independent manner (4), transcriptional regulators such as the sex-determining factor SRY (5) are sequence-specific.

SRY is one of the best characterized DNA-bending proteins. It is encoded on the Y chromosome and promotes the development of male gonads. Through its HMG box, it recognizes sites similar to the sequence A/TAACAAA/T (6), and, upon binding DNA from the minor groove, it bends it at an 83° angle (7). A mutation that significantly reduces the DNA bending activity of human SRY causes sex reversal (8). However, a direct correlation between the bending and transactivational activities of SRY could never be definitively established, because so far no SRY target gene has been identified with certainty, precluding direct experiments on transcriptional control.

To prove beyond doubt that the bending activity of HMG box proteins has a functional role in the regulation of transcription, we focused our attention on Sox2. Sox (SRY-related HMG box) proteins are stage- and tissue-specific transcription factors that specify cell fate during development (9). Sox2 contains an HMG box that recognizes and bends the same DNA sites as SRY and has a transactivation domain at its C terminus (10). Sox2 starts to be expressed at the blastocyst stage, in the inner cell mass; there, in collaboration with the octamer-binding protein Oct-3 (11), it binds to the enhancer of the *Fgf4* (fibroblast growth factor 4) gene and determines its expression (12, 13).

We generated six different mutants in the HMG box of Sox2. We found one mutant with reduced bending activity; this mutant activates *Fgf4* transcription *more* than the wild type protein. We also mutated the Sox2 binding site and selected for single-nucleotide substitutions that render the DNA molecule stiffer toward the action of wild type Sox2; again we found that different DNA bending resulted in a different level of transcription. In this case, reduced bending *decreased* transcription. When we compared the complexes obtained using the mutated protein and the mutated binding site, we found a 20° difference in the relative orientation of the DNA bends. Taken together, these data prove that a correct three-dimensional geometry of nucleoprotein complexes is crucial for transcriptional activation by HMG box proteins.

EXPERIMENTAL PROCEDURES

Bacterial Expression Vectors and Protein Purification—Plasmid pGex-Sox2.HMG was constructed by cloning into the *Eco*RI-*Sal*I sites of plasmid pGex-4T1 (Amersham Pharmacia Biotech), a 387-bp fragment of mouse Sox2 cDNA (nucleotides 449–835; Ref. 10) coding for the Sox2 HMG box plus 26 amino acids upstream and 24 amino acids downstream. The fragment was obtained by PCR from mouse genomic DNA using primers Sox2for (5′-GGGAATTCCAGCAAGCTTCGGGGGG-3′) and Sox2rev (5′-AAGTCGACTAGCTCGCCATGCTGTTCCC-3′). To in-

* This work was supported by grants from Associazione Italiana Ricerca Cancro and the European Union program (to M. E. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Fax: 39-0226434861; E-mail: bianchi.marco@hsr.it.

[1] The abbreviations used are: bp, base pair(s); aa, amino acid; CPA, circular permutation analysis; GST, glutathione *S*-transferase; HMG, high mobility group; PCR, polymerase chain reaction; wt, wild type; EMSA, electrophoretic mobility shift assay.

troduce mutations in the Sox2 HMG box, pGex-Sox2.HMG was used as template in two-step PCR mutagenesis, using Sox2for and Sox2rev as external primers and six pairs of internal mutagenic primers (substitutions in bold type): M47Ifor: 5′-GAGGCCC**ATA**AACGCCTTC-3′, and M47Irev: 5′-GAAGGCGTT**TAT**GGGCCTC-3′; N48Qfor: 5′-GAGGCCC-ATG**CAA**GCCTTC-3′, and N48Qrev: 5′-GAAGGC**TTG**CATGGGCCTC-3′; R58Kfor: 5′-GGGGCAG**AAG**CGTAAGATGG-3′, and R58Krev: 5′-C-CATCTTACG**CTT**CTGCCCC-3′; K89Ifor: 5′-GAGACCGAG**ATA**CGG-CCGT-3′, and K89Irev: 5′-ACGGCCG**TAT**CTCGGTCTC-3′; F92Sfor: 5′-CGGCCG**TCC**ATCGACGAG-3′, and F92Srev: 5′-CTCGTCGAT**GG-A**CGGCCG-3′, and Y112Ffor: 5′-CGGATTATAAA**TTC**CGGCCGC-3′, and Y112Frev: 5′-GCGGCCG**GAA**TTTATAATCCG-3′. The final PCR products were cloned into the *Eco*RI-*Sal*I sites of pGex-4T1 to obtain the mutant pGex-Sox2.HMG plasmids. Sequences were confirmed by double-stranded sequencing. Plasmid pGex-Oct-3.POU was kindly provided by Hans Schöler (University of Pennsylvania).

Glutathione *S*-transferase (GST) fusion proteins were expressed in *Escherichia coli* and purified on glutathione-Sepharose (Amersham Pharmacia Biotech) according to the manufacturer's protocol. The GST moiety was separated from Sox2.HMG by incubating the fusion protein bound to glutathione-Sepharose beads for 2 h at room temperature with thrombin (final concentration, 2 units/μl in phosphate-buffered saline). The Sox2 HMG boxes (wt and mutant) released in the buffer were quantified by Coomassie staining in comparison with a known marker and stored at −80 °C in phosphate-buffered saline containing 0.1 mM dithiothreitol and 0.05% Nonidet P-40.

Electrophoretic Mobility Shift Assay—The γ-^{32}P-labeled Fgf probe was made of annealed oligonucleotides Fgf-probe.for (5′-ctagAACTCTT-TGTTTGGATGCTAATGGGA-3′) and Fgf-probe.rev (5′-tcgaTCCCATT-AGCATCCAAACAAAGAGTT-3′). The Fgf probe reproduces the sequence of the murine *Fgf4* enhancer (14) and contains adjacent Sox2 and Oct-3 binding sites. Purified polypeptides were incorporated in the indicated amounts, in a final volume of 10 μl. Binding buffers, electrophoresis, and autoradiography were as desribed (7). The Oct-3 POU domain alone has a low solubility after removal of the GST moiety; thus, the binding reactions with the Oct-3 POU domain were performed by incubating, first, the Sox2 HMG box with ~500 ng of GST-OCT-3.POU bound to glutathione-Sepharose in 8 μl of binding buffer, allowing the two proteins to interact. After 45 min at 4 °C, the probe (0.1 nM) was added together with thrombin (0.01 unit) to remove the GST, and after 40 min at room temperature the samples were applied to the gel.

Determination of the DNA-Protein Complex Dissociation Constants—The amount of active protein in the preparations was determined by competitive EMSA, adding increasing quantities (0–400 nM) of unlabeled probe to binding mixture containing a small amount of labeled Fgf probe (0.1 nM) and 150 nM Sox2 HMG box. Incubation and electrophoresis were carried out as described in the previous paragraph. The radioactivity present in the bands was measured by exposing the dried gel to PhosphorImager screens (Molecular Dynamics). The concentration of active protein was deduced from the concentration of unlabeled probe that starts to displace the labeled probe.

To determine the dissociation constants of the protein-DNA complexes, wt and mutant Sox2 HMG boxes were titrated into binding mixtures containing a fixed amount (0.1 nM) of probe. The samples were electrophoresed, and the radioactivity was measured as described in the previous paragraph. Under conditions of protein excess, the dissociation constant is equivalent to the concentration of polypeptide that binds 50% of the probe.

Calculation of DNA Bend Parameters—For circular permutation analysis, plasmids containing wt and mutant *Fgf4* enhancers (pBend2-Fgf4, pBend2-Fgf4.mut4 and pBend2-Fgf4.mut11) were prepared by insertion of annealed synthetic oligonucleotides (Bend-Fgf4: 5′-CTAG-AACTCTTTGTTTGGATGCTAATGGGA-3′ and 5′-TCGATCCCATTA-GCATCCAAACAAAGAGTT-3′; Bend-Fgf4.mut11: 5′-CTAGAACTCTT-TATTTGGATGCTAATGGGA-3′ and 5′-TCGATCCCATTAGCATCCA-AATAAAGAGTT-3′; and Ben2-Fgf4.mut4: 5′-CTAGAACTCTTTGGTT-GGATGCTAATGGGA-3′ and 5′-TCGATCCCATTAGCATCCAACCAA-AGAGTT-3′) between *Xba*I and *Sal*I restriction sites in plasmid pBend2. Preparation of probes, electrophoresis, and data analysis were carried out as previously reported (7).

The series of phasing vectors was constructed by cloning the oligonucleotide pairs described above into the *Eco*RI-*Sal*I sites of pSB10 81–86 vectors (15). Probes were prepared by end-labeling *Rsa*I-*Pvu*II fragments with ^{32}P. Each probe was incubated in binding mixtures containing wt or N48Q Sox2 HMG box in concentrations equal to the K_D of the complexes. After 10 min of incubation in ice, the samples were electrophoresed at 4 °C in 10% polyacrylamide gels at 10 V/cm for 18–20 h. The relative mobilities (complex mobility divided by probe mobility) were plotted as a function of the length of the spacer between the Sox2 binding site and the intrinsic DNA bend, and the best fit to a cosine function (phasing function) was determined (Prism software). The relative orientations of DNA bends were determined from the minima of the phasing functions. The helical periodicity was assumed to be 10.5 bp/turn of the DNA helix.

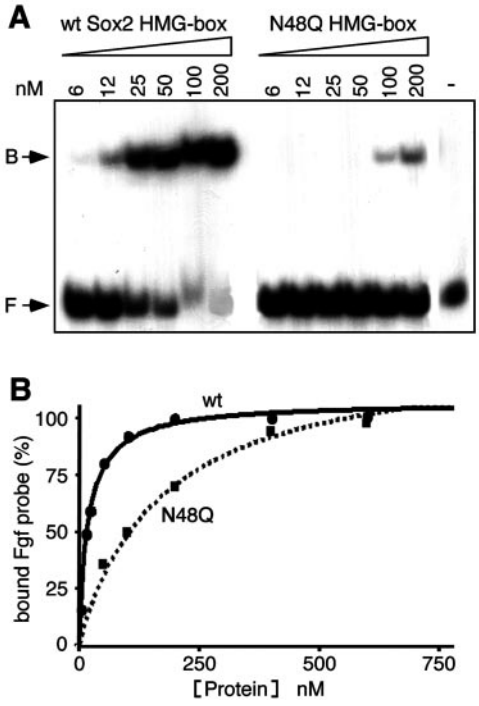


FIG. 1. **Determination of the affinity of wt and mutant Sox2 HMG boxes to the *Fgf4* enhancer.** *A*, increasing amounts of each polypeptide, indicated at the *tops* of the *lanes*, were incubated in 10 μl of binding buffer with a fixed amount of Fgf probe (0.1 nM). EMSA and quantification of the bands are described under "Experimental Procedures." *Arrow F*, free Fgf probe; *arrow B*, wt and mutant HMG box/DNA complex. *B*, fraction of bound Fgf probe as a function of active protein concentration. The K_D corresponds to the protein concentration that binds 50% of the probe (mean of at least three independent experiments with each polypeptide).

Selection of Mutant Sox2 Binding Sites—Sequences recognized by the wt Sox2 HMG box, but intrinsically less bendable than the wt sequence, were selected from a pool of 131-bp DNA molecules containing mutant Sox2-binding sites in the middle of the fragment (D probe). Plasmid pBend2-Fgf4 was used as template in two-step PCR mutagenesis using as external primers bend.for (5′-AGATATCCAGCTGCC-C-3′) and bend.rev (5′-GGATATCTTTAAACTCGAG-3′) and a pair of degenerated primers, where one nucleotide of the Sox2 binding site was individually substituted by the other three nucleotides: fgf1, CTAGA-ACTC**V**TTGTTTGG and CCAAACAA**B**GAGTTCTAG; fgf2, CTAGAAC-TCT**V**TGTTTGG and CCAAACA**B**AGAGTTCTAG; fgf3, GAACTCTT-**V**GTTTGGATG and CCAAAC**B**AAGAGTTCTAG; fgf4, GAACTCTTT-**H**TTTGGATG and CATCCAAA**D**AAAGAGTTC; fgf5, GAACTCTTTG-**V**TTGGATG and CATCCAA**B**CAAAGAGTTC; fgf6, AACTCTTTGT**V**T-GGATGC and CATCCA**B**ACAAAGAGTTC; and fgf7, CTTTGTT**V**GGA-TGCTAATG and CATTAGCATCC**B**AACAAAG, where **H** = A/C/T, **B** = C/G/T, **V** = A/C/G, and **D** = A/T/G.

The seven PCR products were mixed together, obtaining a pool of 21 different DNA molecules containing mutant Sox2 binding sites in the middle. The pool was end-labeled with ^{32}P and used as probe in a band shift assay with wt Sox2 HMG box in the conditions described above. After autoradiography, shifted bands corresponding to lower mobility complexes were cut out from the gel, and DNA was electroeluted and reamplified by PCR. The PCR product was end-labeled with ^{32}P and used again as probe in a band shift assay. Four rounds of selection/amplification were performed to enrich the pool of less bendable molecules, and finally the molecules were cloned into the *Eco*RV restriction site of pBluescript (Stratagene) and sequenced. Two of the mutations (mut4 and mut11) were inserted into pBend2-Fgf4 and tested in circular permutation assays.

Transient Transfection Assays—Plasmid pT81-Fgf4 was constructed

by cloning into the *Bam*HI-*Kpn*I sites of pT81-luc (ATCC) a 286-bp fragment (nucleotides 2939–3224) containing the *Fgf4* enhancer (14), obtained by PCR from mouse genomic DNA with primers enhFgf.for (5′-GCGGATCCTTAGCTCGCTTCAGG-3′) and enhFgf.rev: (5′-GCGGTACCGAGCCACCAGACAGAAAG-3′). To introduce in the *Fgf4* enhancer the nucleotide substitutions that decrease DNA flexibility, pT81-Fgf4 was used as template in two-step PCR mutagenesis using enhFgf.for and enhFgf.rev together with the two oligonucleotide pairs mut4 (5′-GAAAACTCTTTGGTTGGATGC-3′ and 5′-CTTTTGAGAAACCAACCTACG-3′) and mut11 (5′-GAAAACTCTTTATTTGGATGC-3′ and 5′-CTTTTGAGAAATAAACCTACG-3′). The final PCR products were cloned into *Bam*HI-*Kpn*I sites of pT81-luc, obtaining the pT81-Fgf4mut4 and pT81-Fgf4mut11.

Plasmid pSG5-Sox2 was constructed by cloning into the *Bam*HI site of pSG5 (Stratagene) a 977-bp fragment (nucleotides 406–1383) corresponding to the Sox2 coding region (10). To construct variants of pSG5-Sox2, *Hin*dIII-*Sac*II fragments obtained from each mutant pGex-Sox2.HMG were cloned into the same sites of wt pSG5-Sox2, substituting the wt sequence.

HeLa and 3T3 cells ($\sim 1.5 \times 10^5$ cells in a 6-cm dish) were transfected by the calcium phosphate method with various combinations of reporter plasmids (pT81-Fgf4, pT81-Fgf4mut4, and pT81-Fgf4mut11), wt and mutant Sox2 expression plasmids, and pCMV-OCT3. PGK-β-gal was used as normalizing vector. The total amount of transfected DNA was adjusted to 9 μg with empty pBluescript. The amounts of plasmids used in each experiment are reported in the legends of the figures.

RESULTS

Mutagenesis of the Sox2 HMG Box—The DNA binding and bending activities of HMG box proteins are completely specified by their HMG box. We therefore studied the interaction between the Sox2 HMG box and its cognate DNA target site.

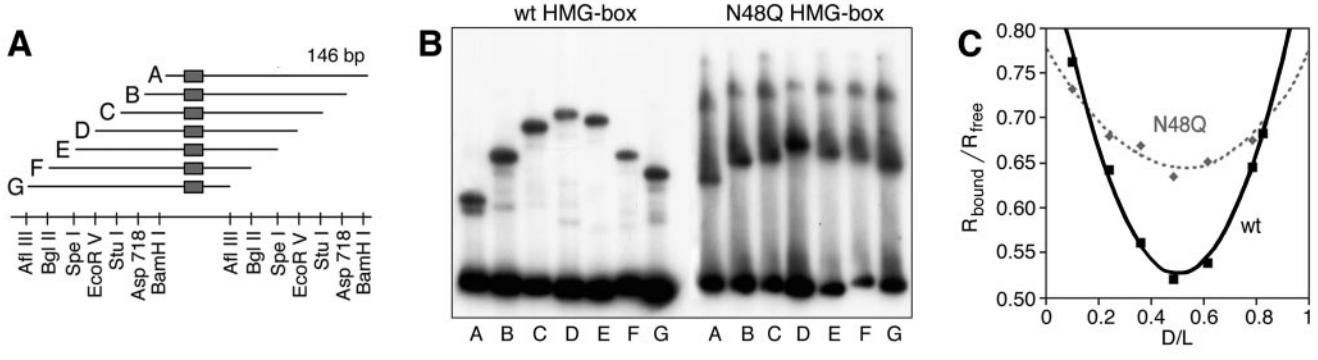
We generated six point mutations within the HMG box to find mutant proteins that bend DNA with different angles. The sequences of Sox2 and SRY HMG boxes are very similar (about 85% identity); we then used the structure of the human SRY HMG box-DNA complex (16) as a model to identify the Sox2 residues involved in contacts to DNA. In SRY, five amino acids (Met-64, Asn-65, Phe-67, Ile-68, and Trp-98) form a T-shaped wedge in direct contact with the central base pairs of the DNA binding site. They mediate the widening of the minor groove and the bending of the helix toward the major groove. Mutation of Met-64 alters DNA bending and causes sex reversal (8). We generated in Sox2 the mutations M47I, K89I, and F92S, which correspond to SRY mutations that cause sex reversal; we also selected in Sox2 Asn-48, adjacent to Met-47; Tyr-112, corresponding to a SRY residue that contacts the 5′ part of the binding site; and Arg-58, which corresponds to a SRY residue that forms salt bridges to the phosphate backbone. These residues were replaced by conservative substitutions (Asn to Gln, Arg to Lys, and Tyr to Phe) to avoid radical modifications of the protein-DNA interaction surfaces. We produced wt and mutant HMG boxes in *E. coli* (see "Experimental Procedures"), and we tested *in vitro* their DNA binding and bending activities.

TABLE I
DNA affinity and bending activity of Sox2 mutants

Sox2 HMG box	K_D	Angle
	M	degrees
Wild type	1.5×10^{-8}	80
M47I	1.8×10^{-8}	80
N48Q	12×10^{-8}	46
R58K	18×10^{-8}	80
K89I	3×10^{-8}	80
F92S	0.9×10^{-8}	80
Y112F	1.2×10^{-8}	80

Analysis of DNA Binding Activity of wt and Mutant Sox2—We then determined the dissociation constant (K_D) of the DNA-protein complexes formed by the wt and mutant Sox2 HMG boxes with the *Fgf4* enhancer (only the concentration of active protein was considered). The seven polypeptides were titrated on a limiting amount of probe; K_D is equal to the concentration of active protein that binds 50% of the probe. Fig. 1 shows titrations with wt HMG box and mutant N48Q; the K_D values of all polypeptides are summarized in Table I. The K_D determined for wt Sox2 HMG box (1.5×10^{-8} M) is similar to those estimated for SRY and other Sox proteins (8, 17). The DNA affinities of mutants M47I, K89I, F92S, and Y112F were indistinguishable from that of the wt, whereas mutants N48Q and R58K showed 8- and 12-fold reductions, respectively. Thus, all of the mutants bind the *Fgf4* target site, and only two have slightly reduced affinity.

DNA Bending Activity of wt and Mutant HMG Boxes—The effect of the mutations on the DNA bending activity was assayed by circular permutation analysis (CPA). This is a standard method to investigate distortions in DNA, based on the observation that a fragment containing a bend near the end migrates faster than a fragment with a bend in the middle (18). The *Fgf4* binding site was cloned into the pBend2 vector be-



Scaffidi Fig. 2

FIG. 2. **Mutation N48Q reduces Sox2 DNA bending activity.** *A*, DNA fragments used as probes in CPA. Plasmid pBend2-Fgf4, containing the wt *Fgf4* enhancer (*black box*) flanked by tandemly repeated DNA sequences, was cleaved at the restriction sites indicated in the map. The DNA fragments (designated *A–G*) all contain circular permutations of the same sequence of 146 bp. *B*, circular permutation analysis of DNA bending induced by wt and mutant Sox2 HMG boxes. 3 fmol of labeled DNA probes were incubated in 10 μl of binding buffer with wt and mutant HMG boxes (wt, 25 nM; N48Q, 300 nM). The retarded bands running behind the main protein-DNA complex band are complexes containing more than one polypeptide molecule/DNA molecule. *C*, calculation of DNA bending angles. The mobility of the protein-DNA complexes (R_{bound}) was normalized to the mobility of the corresponding free probe (R_{free}). The distance of the center of the Sox2 binding site from the 5′ end of the fragment was divided by the total length of the probe (flexure displacement, *D/L*). The plotted points were interpolated with quadratic functions: $y = 1.508x^2 - 1.516x + 0.907$ ($r^2 = 0.994$) for wt Sox2, and $y = 0.470x^2 - 0.500x + 0.777$ ($r^2 = 0.952$) for N48Q. The first and second order parameters of the equation are in close agreement and yield an estimate of flexure angles of 80° for wt Sox2 and 46° for N48Q.

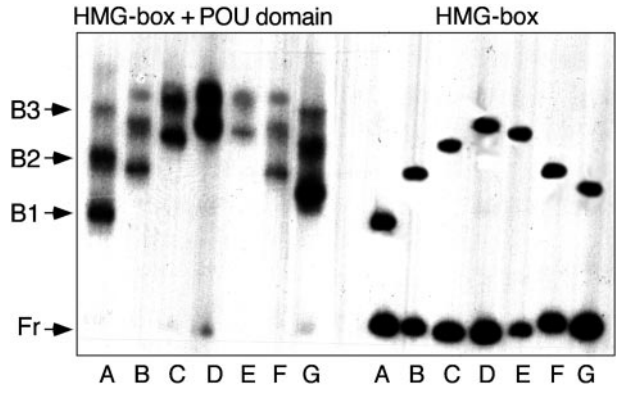


FIG. 3. **Oct-3 does not affect Sox2 DNA bending activity.** Wild type Sox2 HMG box (25 nM) was incubated at 4 °C with or without 1 μM of Oct-3 POU domain and circular permutation probes (0.3 nM) (see "Experimental Procedures" for details). *Arrow Fr*, free DNA; *arrow B1*, HMG box-DNA complex; *arrow B2*; HMG box-POU domain-DNA ternary complex; *arrow B3*, complex with two HMG boxes binding to the same DNA molecule. The bending angles induced by the Sox2 HMG box alone (B1 complexes) and by the HMG box in cooperation with the Oct-3 POU domain (B2 complexes) are not significantly different.

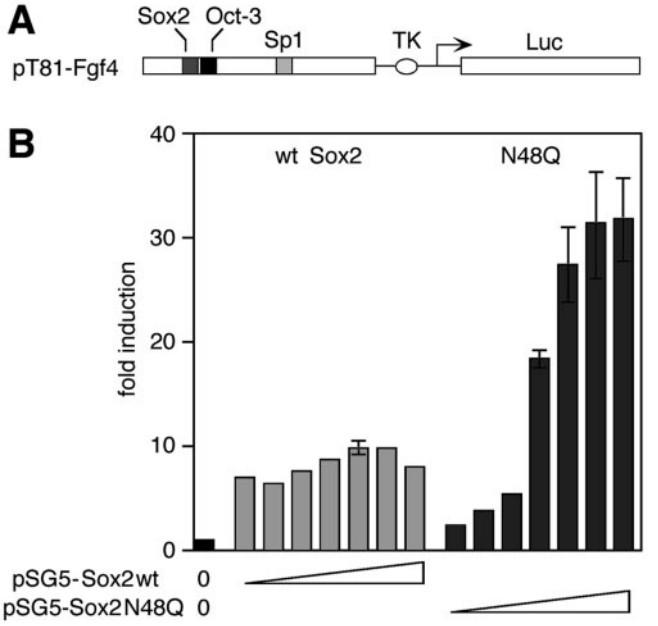


FIG. 4. **Mutation N48Q increases the transactivational potential of Sox2 protein.** *A*, diagram representing the reporter plasmid pT81-Fgf4. The *Fgf4* enhancer, containing Sox2, Oct-3, and Sp1 binding sites, was cloned in front of the TK promoter and luciferase coding region. *B*, transcriptional activities of wt and N48Q Sox2 proteins. 3T3 fibroblasts were transiently transfected with Sox2 expression plasmids (0, 0.2, 0.5, 1, 2.5, 3, 3.5, and 4 μg) and pT81-Fgf4 reporter plasmid (2 μg). PGK-β-gal plasmid expression was used for internal normalization. All transfections were carried out in triplicate, in three independent experiments.

tween two repeated 121-bp units, and seven circularly permutated fragments were obtained by restriction enzyme digestion (Fig. 2*A*) and used in EMSA as probes with wt and mutant HMG boxes. On the basis of three independent experiments we estimated that wt Sox2 induces an 80° deviation of the DNA axis from linearity. This value is close to that determined for SRY (83°; Ref. 7). Mutants M47I, R58K, K89I, F92S, and Y112F distort DNA exactly as the wt protein, whereas N48Q is significantly different (46°; Fig. 2*B*). A summary of the biochemical properties of the seven polypeptides is presented in Table I.

Effect of Oct-3 on Sox2 DNA Binding and Bending Activities—The transcriptional activation of *Fgf4* gene is mediated by the synergistic action of Sox2 and Oct-3 (12, 13). The two proteins bind contiguous target sites in the *Fgf4* enhancer and interact with each other through the POU domain and the HMG box. We then checked the influence of Oct-3 on the DNA binding and bending activities of the Sox2 HMG box. Titration and circular permutation experiments were performed again in the presence of the Oct-3 DNA-binding region, composed by the POU-specific domain and the POU homeodomain. In these conditions, a ternary complex is formed by the Sox2 HMG box, the Oct-3 POU domain, and DNA. The apparent DNA binding affinity of both wt and N48Q HMG box is increased by the presence of the Oct-3 POU domain (data not shown). On the other hand, Oct-3 has little effect on the DNA bending activity of Sox2 (Fig. 3). Two additional sets of bands appear. The middle one corresponds to HMG box-POU domain-DNA complexes, and the upper one corresponds to two HMG boxes bound to DNA. The bending angle calculated for the ternary complex (77°) is very close to that estimated for the HMG box-DNA complex (80°). We conclude that the bending activity of Sox2 is not significantly affected by Oct-3.

The N48Q Mutation Enhances the Transcriptional Activity of Sox2—We then correlated the N48Q bending activity with its transcriptional activity. 3T3 and HeLa cells were transiently co-transfected with fixed amounts of Oct-3 expressing vector and a reporter plasmid driven by the *Fgf4* enhancer (Fig. 4*A*), together with increasing amounts of expression vectors for either wt or N48Q Sox2 (full-length proteins). Wt Sox2 determines, at the maximum level, a 7-fold induction of luciferase expression (Fig. 4*B*), whereas the mutant N48Q reaches a 21-fold induction. At low concentrations, N48Q is actually less effective than the wt protein, and this is probably due to its reduced DNA binding affinity. At high concentrations, all DNA target sites are presumably saturated even by the mutant, and differences in the binding affinities become irrelevant, whereas the transcriptional activities must depend exclusively on the bending activity of Sox2. We conclude that the bending activity of Sox2 affects transcriptional control directly.

DNA Mutations That Affect DNA Bending by Sox2—To further confirm the correlation between DNA bending and transcription, we used a complementary approach. We generated a population of mutated Sox2 binding sites (Fig. 5*A*) by substituting individually every nucleotide to find DNA molecules that are intrinsically stiffer and less bendable by wild type Sox2 protein. By PCR we obtained a mixed population of 131-bp fragments containing the mutated Sox2 binding sites in the middle, and we used it as probe in EMSA with wt Sox2. Because of the position of the site, the migration of DNA-protein complexes is very sensitive to differences in the bending angle. We recovered from the gel the complexes with higher mobility than the wt sequence, containing less distorted DNA fragments. After several cycles of selection and amplification we ended up with fragments with reduced bendability that were cloned and sequenced. The mutant sequences were inserted in the pBend-Fgf4 plasmid and compared in CPA. Mutant 11 substitutes the fourth nucleotide of the binding site (A instead of G), and this change abolishes distortion by Sox2 (Fig. 5*C*). In mutant 4 the fifth nucleotide is G instead of a T, and this reduces the bending induced by wt Sox2 to 46° (Fig. 5*B*).

By itself, this result confirms that the DNA sequence determines the general architecture of nucleoprotein complexes, and even slight variations can affect severely the overall geometry. We then tested whether a different DNA bend (obtained through modification of the binding site rather than the mutation of the binding protein) could result in a difference in transcription.

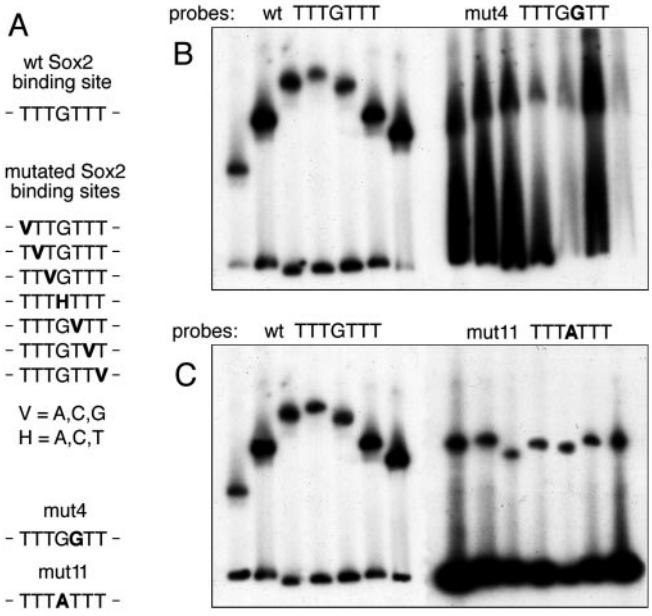


FIG. 5. **Single-nucleotide substitutions in the Sox2 binding site reduce DNA flexibility.** *A*, single nucleotides in the Sox2 binding site of the *Fgf4* enhancer were individually substituted, obtaining a pool of 21 mutants. From these, mut4 and mut11 were selected as having reduced flexibility. *B* and *C*, circular permutation analysis of DNA bending induced by wt Sox2 HMG box in mut4 (*B*) and mut11 (*C*). 3 fmol of circularly permutated DNA probes (*A–G*) containing either mutation were incubated in 10 μl of binding buffer with wt protein (25 nM). Bending angles were calculated as described in the legend to Fig. 3.

Reduced DNA Flexibility Affects Transactivation Mediated by Sox2—The nucleotide substitutions reducing DNA flexibility (mut4 and mut11) were inserted in the reporter plasmid pT81-Fgf4 and tested in transient co-transfections of wt Sox2 and Oct-3. Fig. 6 shows that a different DNA bend determines a different activation of transcription. In this case, a smaller distortion leads to a reduced activation rather than an increase as seen with N48Q Sox2 protein. The planar angles formed by wt Sox2 protein on the mut4 site and by N48Q Sox2 protein on the wt site are similar, as measured by CPA. To unravel this apparent paradox we made the hypothesis that the three-dimensional geometries of the complexes N48Q/*Fgf4* and wtSox2/*Fgf4*mut4 were different. CPA is a good tool to detect general protein-induced perturbations of DNA, including static bending, increased DNA flexibility, and aberrant protein structures; however, it gives no information concerning the direction of the DNA bend. A more specific method for the study of protein-directed bends is phasing analysis, which allows delineation of the direction of a DNA distortion in space (19).

Phasing Analysis—Phasing analysis utilizes as probes in EMSA DNA fragments having an intrinsic bend at different distances from the site of induced bending. The spacing between the intrinsic and the induced bends is incrementally varied over one turn of the DNA helix (Fig. 7*A*); thus, there is a spacing where the two bends are in phase and form a complementary larger angle and a spacing where the two bends are out of phase and have the effect of straightening the DNA fragment. The mobility of the complexes in a polyacrylamide gel is slowest when the two bends add up (*cis*-isomer) and highest when the two bends counteract each other (*trans*-isomer).

We compared by phasing analysis the complexes wtSox2/*Fgf4*, N48Q/*Fgf4*, and wtSox2/*Fgf4*mut4 (Fig. 7*B*). We plotted the relative mobilities of the complexes as a function of the linker length (Fig. 7*C*), and we determined for each complex the spacer length that gives the lowest complex mobility. The min-

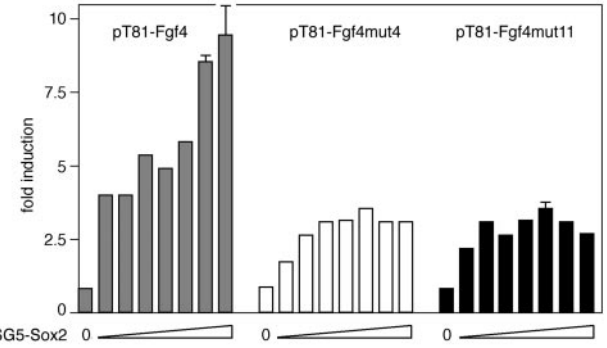


FIG. 6. **Mutations mut4 and mut11 decrease Sox2-mediated transactivation of the *Fgf4* enhancer.** 3T3 fibroblasts were transfected with 2 μg of either wt pT81-Fgf4, pT81-Fgf4mut4, or pT81-Fgf4mut11, and increasing amounts of wt pSG5-Sox2 (0, 0.2, 0.5, 1, 2.5, 3, 3.5, and 4 μg). All transfections were carried out in triplicate, in three independent experiments, as described in the legend to Fig. 4.

ima for N48Q/*Fgf4* and for wtSox2/*Fgf4* are indistinguishable, indicating that the bends point in the same direction. On the contrary, a significant change was observed in the case of wtSox2/*Fgf4*mut4, the spacer length giving the *cis* isomer differs by 0.6 bp in comparison with wtSox2/*Fgf4*. Assuming an average of 10.5 bp/turn of the DNA helix, this corresponds to a 20° difference in the relative orientations of the DNA bends. These data prove that the mutations in Sox2 protein (N48Q) and the *Fgf4* enhancer (mut4) determine different alterations in the architecture of the corresponding nucleoprotein complexes, compared with the wt complex.

DISCUSSION

In this study we investigated the role of the bending activity of the Sox2 transcription factor in the transcriptional activation of the *Fgf4* gene. We first quantified the affinity of the Sox2 HMG box to its target site. The estimated $K_D = 1.5 \times 10^{-8}$ M indicates a quite strong DNA binding activity, comparable with that observed for other HMG box transcription factors, such as SRY, LEF1, and Sox5 (8, 20, 21). We then calculated that, upon binding, the HMG box induces an 80° distortion of the DNA double helix. We also proved that Oct-3, which cooperates with Sox2 in *Fgf4* regulation, does not affect the DNA bending activity of Sox2.

These results indicate that the interactions with DNA of SRY and Sox2 are similar. Taking advantage of this similarity, we introduced mutations into the Sox2 HMG box that were presumed to alter its DNA bending activity, using SRY as a guide. The most dramatic effect was obtained by substituting Asn-48 with Gln (N48Q mutant); the bending activity was strongly reduced. The substituted Asn corresponds to one of the five amino acids of SRY that enter into the minor groove and induce the distortion of the DNA helix (16). The side chain of this Asn is involved in hydrogen bonding and electrostatic interactions with the fourth and fifth base pairs of the binding site. The Asn → Gln substitution adds a methylene group to the side chain, increases its bulk, and thus affects the induced DNA bending. The DNA binding affinity is also reduced 8-fold.

An unexpected result came from mutant M47I. In SRY, the corresponding substitution causes a reduction in the bending activity of the protein, which phenotypically results in sex reversal (8). The mutation in Sox2, on the contrary, has no effect. This is even more surprising if we consider that the adjacent Asn is indeed involved in DNA bending. Obviously, the fine details of the interactions of SRY and Sox with DNA are not completely identical.

Our goal was to prove that the ability of Sox2 to activate

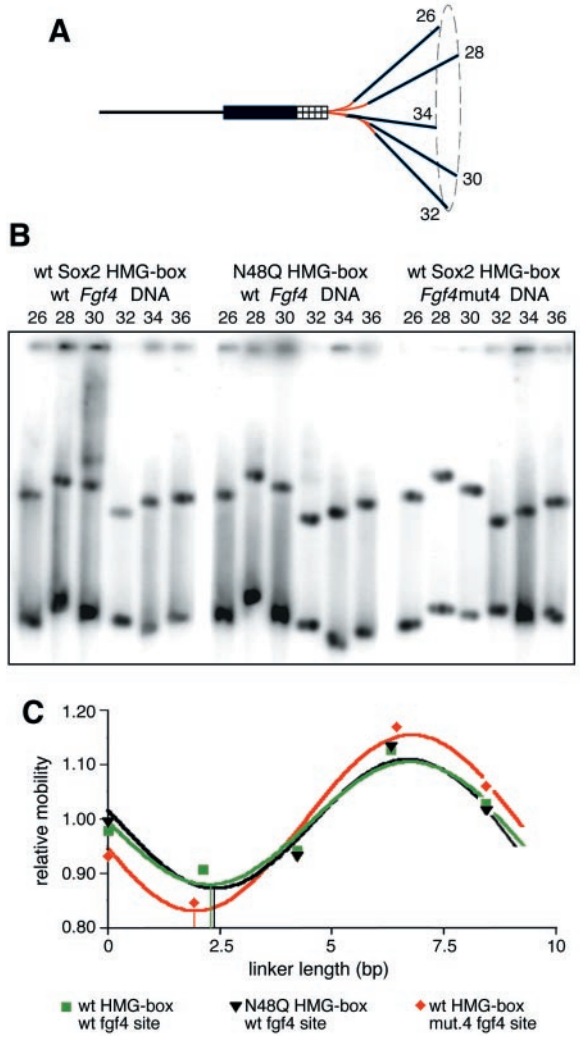


FIG. 7. **Different orientation of the DNA bends in N48Q/*Fgf4* and wtSox2/*Fgf4*mut4 complexes.** *A*, DNA fragments used as probes in phasing analysis contain either the wt or mut4 *Fgf4* enhancer (*black box*) separated from the $(AT)_6$ intrinsic bend (*red line*) by a spacer of variable length (*open boxes*). The drawing represents the orientation in space of the fixed bends. The *numbers* indicate the distance in base pairs between the centers of the Sox2 binding site and the $(AT)_6$ intrinsic bend. The total length of the fragments varies from 310 to 320 bp. *B*, phasing analysis. 3 fmol of labeled DNA probes were incubated in 10 μl of binding buffer with either wt Sox2 or mutant N48Q HMG boxes (wt, 25 nM; N48Q, 300 nM). *C*, calculation of the relative orientations of the directed bends. The normalized complex mobilities (from at least five independent experiments) were plotted as a function of the length of the linker spacer, and the best fit to a cosine function was determined. The minima of the phasing functions for N48Q/*Fgf4* and wtSox2/*Fgf4*mut4 differs by 0.6 bp, which corresponds to a 20° difference in the relative orientation of the directed bends, assuming a helical DNA periodicity of 10.5 bp. The minima for N48Q/*Fgf4* and wtSox2/*Fgf4* coincide.

transcription is dependent on its ability to bend DNA. Indeed we found that the reduction of the bending activity in mutant N48Q determines a 3-fold increase of the protein transactivation potential, indicating that at least one of the mechanisms used by Sox2 to promote transcription is based on its capacity to manipulate DNA geometry. At first glance, the finding that reduced bending corresponds to increased transcription seems surprising. A similar result, however, was found in the case of IHF, an *E. coli* protein that activates transcription at several promoters by inducing a 180° DNA bend, thus promoting DNA-RNA polymerase interactions (22). In a random mutation screening, mutants of IHF that maximize transcriptional activation were identified; a reduction in the bending activity corresponded to stronger activation (23). Three-dimensional models showed that wt IHF bends DNA more sharply than is necessary for DNA back looping to occur, and the optimal bending angle is much less than the 180° turn imposed by wt IHF. Even though the molecular mechanisms underlying IHF- and Sox2-mediated transactivation are different, our data can have a similar explanation; the distortion induced by wt Sox2 might be larger than optimal to promote contacts among the proteins involved in *Fgf4* transactivation.

One might ask why evolution selected a protein that is non-optimal in its capacity for transcriptional activation. Sox2 regulates expression of γ- and δ-crystallin in the lens (24), osteopontin during hypoblast formation (25), and probably several other genes. The expression of different genes might have different architectural demands, and the entity of the distortion imposed by wt Sox2 is possibly a compromise that allows a sufficient expression of all its target genes. A different explanation is also possible; an increased expression of *Fgf4* is not necessarily an advantage for the organism. Fgf4 is a secreted protein that acts within complex regulatory pathways, in which intracellular and extracellular signals cooperate to define the body plan of the embryo (26). Increased *Fgf4* expression could disrupt the delicate equilibrium that regulates intercellular communication.

In the second part of our work, we confirmed the relation between DNA bending and transcriptional activity by modulating DNA flexibility. The sequence of the Sox2 binding site affects the flexibility of the DNA molecule; single nucleotide substitutions reduce (mut4) or completely abolish (mut11) the ability of wt Sox2 protein to distort DNA. The mut4 sequence is bent by wt Sox2 by an angle similar to the one produced by the mutant N48Q protein on the wild type sequence, as measured by CPA. Unexpectedly, the substitutions in the protein and in the DNA site had opposite effects on transcription. We explained this apparent paradox by means of phasing analysis; we found that the mutations in the protein and in the binding sites affect the three-dimensional geometry of the DNA-protein complex in different ways. In particular, the orientation of the directed DNA bend induced by N48Q is essentially identical to that induced by wt Sox2. On the contrary, when mut4 is bound by the wt protein, the bend points 20° away from the direction of the wild type bend, most likely preventing the correct assembly of the nucleoprotein complex on the *Fgf4* enhancer.

In conclusion, in this work we have directly confirmed the hypothesis that an essential component of the action of Sox proteins is DNA bending. Furthermore, such bending must be precise both in amplitude and in direction to provide a correct framework for the rest of the transcriptional machinery to operate.

Acknowledgments—We thank Dr. H. R. Schöler for providing reagents. We are very grateful to Dr. S. Guazzi for preliminary experiments.

REFERENCES

1. Travers, A. A., Ner, S. S., and Churchill, M. E. A. (1994) *Cell* **77,** 167–169
2. Bianchi, M. E., and Beltrame, M. (1998) *Am. J. Hum. Genet.* **63,** 1573–1577
3. Grosschedl, R., Giese, K., and Pagel, J. (1994) *Trends Genet.* **10,** 94–100
4. Thomas, J. O., and Travers, A. A. (2001) *Trends Biochem. Sci.* **26,** 167–174
5. Goodfellow, P. N., and Lovell-Badge, R. (1993) *Annu. Rev. Genet.* **27,** 71–92
6. Harley, V. R., Lovell-Badge, R., and Goodfellow, P. N. (1994) *Nucleic Acids Res.* **22,** 1500–1501
7. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1992) *EMBO J.* **11,** 4497–4506
8. Pontiggia, A., Rimini, R., Harley, V. R., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1994) *EMBO J.* **13,** 6115–6124
9. Pevny, L. H., and Lovell-Badge, R. (1997) *Curr. Opin. Genet. Dev.* **7,** 338–344
10. 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
11. Schöler, H. R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990) *Nature* **344,** 435–439

12. Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995) *Genes Dev.* **9,** 2635–2645
13. Ambrosetti, D.-C., Basilico, C., and Dailey, L. (1997) *Mol. Cell. Biol.* **17,** 6321–6329
14. Brookes, S., Smith, R., Thurlow, J., Dickson, C., and Peters, G. (1989) *Nucleic Acids Res.* **17,** 4037–4045
15. Drak, J., and Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 3074–3078
16. Werner, M. H., Huth, J. R., Gronenborn, A., and Clore, G. M. (1995) *Cell* **81,** 705–714
17. Connor, F., Wright, E., Denny, P., Koopman, P., and Ashworth, A. (1995) *Nucleic Acids Res.* **17,** 3365–3372
18. Wu, H.-M., and Crothers, D. M. (1984) *Nature* **308,** 509–513
19. Zinkel, S. S., and Crothers, D. M. (1987) *Nature* **328,** 178–181
20. Giese, K., Amsterdam, A., and Grosschedl, R. (1991) *Genes Dev.* **5,** 2567–2578
21. Connor, F., Cary, P. D., Read, C. M., Preston, N. S., Driscoll, P. C., Denny, P., Crane-Robinson, C., and Ashworth, A. (1994) *Nucleic Acids Res.* **22,** 3339–3346
22. Rice, P. A., Yang, S., Mizuuchi, K., and Nash, H. N. (1996) *Cell* **87,** 1295–1306
23. Engelhorn, M., and Geiselmann, J. (1998) *Mol. Microbiol.* **30,** 431–441
24. Kamachi, Y., Uchikawa, M., Collignon, J., Lovell-Badge, R., and Kondoh, H. (1998) *Development* **125,** 2521–2535
25. Botquin, V., Hess, H., Fuhrmann, G., Anastassidis, C., Gross, M. K., Vriend, G., and Schöler, H. R. (1998) *Genes Dev.* **12,** 2073–2090
26. Basilico, C., and Moscatelli, D. (1992) *Adv. Cancer Res.* **59,** 115–165

Spatially Precise DNA Bending Is an Essential Activity of the Sox2 Transcription Factor

Paola Scaffidi and Marco E. Bianchi

J. Biol. Chem. 2001, 276:47296-47302.
doi: 10.1074/jbc.M107619200 originally published online October 2, 2001
