Nucleic Acids Research, 2003, Vol. 31, No. 5 **1541–1553**
DOI: 10.1093/nar/gkg230

Adjacent DNA sequences modulate Sox9 transcriptional activation at paired Sox sites in three chondrocyte-specific enhancer elements

Laura C. Bridgewater*, Marlan D. Walker, Gwen C. Miller, Trevor A. Ellison, L. Daniel Holsinger, Jennifer L. Potter, Todd L. Jackson, Reuben K. Chen, Vicki L. Winkel, Zhaoping Zhang[1], Sandra McKinney[1] and Benoit de Crombrugghe[1]

Department of Microbiology and Molecular Biology, Brigham Young University, 591 WIDB, Provo, UT 84602, USA and [1]Department of Molecular Genetics, the University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Received October 29, 2002; Revised and Accepted December 16, 2002

ABSTRACT

Expression of the type XI collagen gene *Col11a2* is directed to cartilage by at least three chondrocyte-specific enhancer elements, two in the 5′ region and one in the first intron of the gene. The three enhancers each contain two heptameric sites with homology to the Sox protein-binding consensus sequence. The two sites are separated by 3 or 4 bp and arranged in opposite orientation to each other. Targeted mutational analyses of these three enhancers showed that in the intronic enhancer, as in the other two enhancers, both Sox sites in a pair are essential for enhancer activity. The transcription factor Sox9 binds as a dimer at the paired sites, and the introduction of insertion mutations between the sites demonstrated that physical interactions between the adjacently bound proteins are essential for enhancer activity. Additional mutational analyses demonstrated that although Sox9 binding at the paired Sox sites is necessary for enhancer activity, it alone is not sufficient. Adjacent DNA sequences in each enhancer are also required, and mutation of those sequences can eliminate enhancer activity without preventing Sox9 binding. The data suggest a new model in which adjacently bound proteins affect the DNA bend angle produced by Sox9, which in turn determines whether an active transcriptional enhancer complex is assembled.

INTRODUCTION

The fibrillar collagen type XI plays a critical role in cartilage formation. It is therefore essential for proper skeletal morphogenesis, because much of the skeletal bone forms from a cartilage template through the process of endochondral ossification (1,2). Type XI collagen is also critical for the proper development and maintenance of articular cartilage (3,4). A mature type XI collagen molecule is a triple-helical trimer composed of α1, α2 and α3 protein subunits, products of the *Col11a1*, *Col11a2* and *Col2a1* genes, respectively. Assembly of type XI pro-collagen takes place in the rough endoplasmic reticulum, where the three subunits associate at the C-terminus and fold in zipper-like fashion toward the N-terminus. The immature pro-collagen is then secreted by the chondrocyte. Final cleavage of the N- and C-termini takes place extracellularly, and the mature type XI collagen molecules are incorporated along with type II and type IX collagen into cartilage collagen fibrils. Accumulating evidence suggests that type XI collagen is incorporated into the interior of the fibril and there functions to limit fibril diameter (5–7). This conclusion is supported by observations of unusually thick collagen fibrils in the cartilage extracellular matrix of both mice and humans that lack normal type XI collagen (2,4).

The *Col11a2* gene is expressed almost exclusively in cartilage, making it a good model for the study of chondrocyte-specific gene expression. Information gained from the study of this gene will be relevant to the eventual development of gene-based treatments for cartilage disorders such a chondrodysplasias and osteoarthritis (8–10). Accumulating evidence suggests that tissue-specific gene expression is often controlled by the cooperative functions of multiple independent, modular enhancer elements. For example, the mouse *Pax6* gene contains at least three independent enhancer elements that mediate gene expression in different tissues of the developing eye and pancreas. These enhancers are highly conserved in organisms as diverse as pufferfish, human and quail (11,12). The *Drosophila FMRFα* gene also contains at least three separate enhancer elements distributed over 8 kb of DNA that drive expression in about 17 different cell types in the *Drosophila* central nervous system. (13). Finally, the mouse pro-α1(I) collagen gene (*Col1a1*) contains at least three separate enhancer elements that independently direct gene expression to dermal fibroblasts, osteoblasts and odontoblasts, and tendon and fascia fibroblasts (14,15). Each of these independent, modular enhancer elements functions by binding multiple proteins that activate or repress transcription.

*To whom correspondence should be addressed. Tel: +1 801 422 2434; Fax: +1 801 422 0519; Email: laura_bridgewater@byu.edu

 © Oxford University Press 2003; all rights reserved

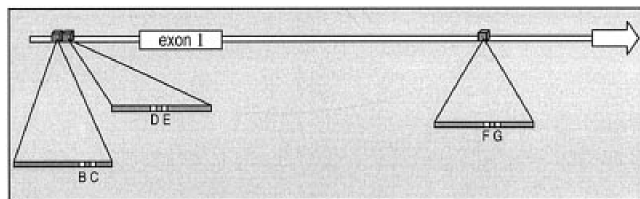

Figure 1. Schematic diagram of the 5′ region of mouse *Col11a2*. The B/C and D/E enhancers in the upstream promoter region and the F/G enhancer in the first intron are represented as boxes, with expanded views of each enhancer showing approximate locations of the two heptameric Sox-binding consensus sequences in each enhancer.

A thorough understanding of a particular gene's regulatory mechanisms must begin with the delineation of critical protein-binding sequences in the enhancer elements, identification of the proteins that bind at those sites and determination of how those proteins interact with one another.

A protein family whose members are known to play key regulatory roles in a variety of critical developmental processes such as germ layer formation, organ development and cell type specification is the Sox protein family (16). Sox proteins form a subgroup of the HMG box superfamily of DNA-binding proteins. Members of the HMG superfamily all contain an HMG box DNA-binding domain consisting of three α-helices and an N-terminal extension, folded into a twisted L-shape. This L-shaped DNA-binding domain binds in the minor groove of DNA and induces a bend in the DNA. Some HMG proteins bind DNA non-specifically, simply by recognizing distorted DNA structures, but others recognize specific DNA sequences (17). Sox proteins fall into the latter category. The consensus binding sequence for Sox proteins has been defined as the heptameric sequence (A/T)(A/T)-CAA(A/T)G (18).

One Sox protein in particular, Sox9, has been shown to play a critical role in cartilage development and to activate expression of the type II collagen gene *Col2a1*, a classic marker of chondrogenesis (19–22). It is co-expressed with *Col2a1* in all cartilage tissue (23). This protein activates *Col2a1* expression by binding to heptameric Sox-binding sequences in a chondrocyte-specific enhancer element in the first intron of the gene.

By examining similar heptameric sequences in the regulatory regions of the also chondrocyte-specific type XI collagen gene *Col11a2*, we identified two independent modular enhancer elements in its 5′ region, herein referred to as B/C and D/E (Fig. 1) (24). These enhancers each contain two sites with homology to the heptameric Sox protein-binding consensus sequence. When multimerized as four tandem copies upstream of a minimal human β-globin promoter, each of these enhancers was found to direct reporter gene expression specifically to cartilage in developing transgenic mice. The enhancers were both activated in transient transfection experiments by SOX9, and mutating the Sox sites eliminated enhancer activity (24). Liu *et al*. subsequently published work documenting the presence of a third chondrocyte-specific enhancer element in the first intron of the mouse *Col11a2* gene (25). This element, which we call F/G, contains a 7 bp sequence that binds Sox9 *in vitro*. We present evidence herein demonstrating that a novel intronic Sox site, 3 bp downstream from the first 7 bp sequence identified by Liu *et al*., is also critical for the transcriptional activity of the F/G enhancer and participates in binding of Sox9.

The structural similarities between the three *Col11a2* chondrocyte-specific enhancer elements is striking. All three contain pairs of heptameric Sox protein-binding consensus sequences arranged in opposite orientation to each other and 3–4 bp apart. We herein demonstrate that correct spacing between the paired Sox sites is critical for enhancer activity. We also show that the third intronic enhancer can be activated *in vivo* by SOX9, that the three modular enhancers serve non-redundant functions and that other protein-binding sites besides the Sox sites participate in the activation of each enhancer element. The information gained about each of the three *Col11a2* chondrocyte-specific enhancer elements has opened the way for the identification of the proteins that cooperate with Sox9 in the activation of *Col11a2* gene expression.

MATERIALS AND METHODS

Plasmid constructions

The three *Col11a2* enhancer elements were designated B/C, D/E and F/G, for clarity in referring to the different elements (Fig. 1). Each letter represents a site with homology to the heptameric Sox protein-binding consensus sequence (A/T)-(A/T)CAA(A/T)G (18). A site called A is found towards the 5′ end of the B/C element, but it was shown previously to play no role in enhancer activity (24). Site D is a perfect match to the heptameric Sox protein-binding consensus sequence, sites C, E and F are 1 bp mismatches, and sites B and G are 2 bp mismatches.

Luciferase reporter plasmids. The following experimental enhancer elements were synthesized as complementary oligonucleotides: the wild-type B/C, D/E and F/G elements, the Sox site mutant versions of these (Bm/C, B/Cm, Bm/Cm, Dm/E, D/Em, Dm/Em, Fm/G, F/Gm and Fm/Gm), the B/C, D/E and F/G elements containing insertion (B/C+3, D/E+3, F/G+3), deletion (D/E–3) and substitution (B*C, D*E, F*G) mutations between the paired Sox sites, the elements from *Col9a2* and *Aggrecan*, and the series of 3 bp mutations in the B/C, D/E and F/G elements (Life Technologies, Inc.). (See Table 1 and Figs 8–10 for sequences.) The complementary oligonucleotides were purified by polyacrylamide gel electrophoresis, annealed, and cloned into the *Bam*HI–*Bgl*II site of the p89Col2a1Bs plasmid, which was a kind gift of Veronique Lefebvre of the Cleveland Clinic, Cleveland, OH (26). The enhancer elements were multimerized to four tandem copies in this plasmid and then transferred along with a minimal promoter (p89) into the luciferase reporter vector pLuc4 (26). The final version of each construct therefore contained four tandem repeats of each element, cloned directly upstream of the minimal promoter and the luciferase reporter, as was described previously for the B/C and D/E enhancer constructs (24).

pCXIwt2/3 constructs. A 1122 bp fragment between –742 and +380 of mouse *Col11a2* was amplified by PCR and cloned into the *Eco*RI–*Xho*I polylinker region in the expression vector pNASSβ as described by Tsumaki *et al*. (27). A 684 bp fragment from +1468 to +2151 representing the middle

Table 1. Sequences of elements tested for enhancer activity

Element	Sequence
B/C	tgggcctacagagccggtgcagggtggggaCACAAGGcgtgCTTTGTCgggagg
Bm/C	------------------------------CAC**GC**GG-----------------
B/Cm	---C**GGC**GTC------
Bm/Cm	------------------------------CAC**GC**GGcgtgC**GGC**GTC------
B*C	------------------------------CACAAGGc**t**tgCTTTGTC------
B/C+3	------------------------------CACAAGGcgtg**cgt**CTTTGTC------
D/E	gcggctgcttTTCAAAGgcgCCTTGTTtgccggtgcccttccagcgccag
Dm/E	----------T**C**C**GGC**G---------------------------------
D/Em	--------------------CC**GC**G**G**T-----------------------
Dm/Em	----------T**C**C**GGC**GgcgCC**GC**G**G**T-----------------------
D*E	----------TTCAAAGg**a**gCCTTGTT-----------------------
D/E+3	----------TTCAAAGgcg**gac**CCTTGTT-----------------------
D/E-3	----------TTCAAAGCCTTGTT-----------------------
F/G	cggtttcctcagctcctggaCTCAAAGggcCTTTTCTctcctgcctgccc
Fm/G	--------------------C**G**CA**CC**G-----------------------
F/Gm	------------------------------CT**GG**TCT-------------
Fm/Gm	--------------------C**G**CA**CC**GggcCT**GG**TCT-------------
F*G	--------------------CTCAAAGg**t**cCTGGTCT-------------
F/G+3	--------------------CTCAAAGggc**gac**CTTTTCT-------------
Col9a2	gtgtgactaTGCATTGtggtGTTTGATaGTCTAAGggaCACTTTTcatagaacaa
Aggrecan	aactttccaacagtgttcTATGAAGattCTTTCAAatcgcattcccagag

Sites with homology to the HMG protein-binding consensus sequence are shown in upper case. Substituted and inserted nucleotides are shown in bold. The sequence of each wild-type element is given in full, but only the sequence of the HMG sites and intervening sequences are shown in mutant elements.

one-third of the mouse *Col11a2* first intron was also amplified by PCR and cloned into the *Sal*I–*Pst*I polylinker region in the same pNASSβ vector, downstream of the β-galactosidase reporter gene to create pCXIwt2/3. A cosmid containing the entire mouse *Col11a2* gene, kindly provided by Dr Darwin Prockop (Center for Gene Therapy, Tulane University Health Science Center, New Orleans, LA), served as the PCR template.

ΔA, ΔB, ΔC, ΔD, ΔE, ΔF and ΔG constructs were built as described above for pCXIwt2/3, except that a targeted mutation was introduced by PCR into the indicated Sox site before the *Col11a2* regions were cloned into the pNASSβ vector. Each of the mutant sites contained a total of four mismatches with the heptameric Sox protein-binding consensus sequence [(A/T)(A/T)CAA(A/T)G], except for ΔD, which contained two mismatches. The following mutations were introduced: ΔA, TACAGAG→TGTGGAG; ΔB, CACAAGG→CATGAGG; ΔC, CTTTGTC→CTGCATG; ΔD, TTCAAAG→TTTGAAG; ΔE, CCTTGTT→CCGCATT; ΔF, CTCAAAG→CGCACCG; ΔG, CTTTTCT→CTGGTCT.

Tissue culture and transfections

The rat chondrosarcoma (RCS) cell line was obtained as described (28). BALB-3T3 cells and RCS cells were cultured at 37°C under 5% CO_2 in Dulbecco's modified Eagle's medium supplemented with penicillin (50 U/ml), streptomycin (50 μg/ml), L-glutamine (2 mM) and 10% fetal calf serum. Transfections were performed using LipofectAMINE reagent (Life Technologies, Inc.) according to the manufacturer's instructions. A total of 2 μg of DNA was used to transfect each 10 cm^2 dish. Cellular extracts were prepared, and luciferase and β-galactosidase activities were assayed as described (28).

Electrophoretic mobility shift assays (EMSAs)

Probes were radiolabeled by end-filling with Klenow fragment. Protein–DNA binding reactions were performed as described (26). Nuclear extracts were prepared as described (29), and SOX9 was synthesized by *in vitro* trancription–translation from a previously described SOX9-pcDNA-5′-UT expression vector using the STP3 T7 kit from Novagen, Inc. (Madison, WI) (22). Assays were performed using 9 μg of nuclear extract with 1.0 μg of poly(dG–dC)·poly(dG–dC) as non-specific competitor, or 1 μl of *in vitro* synthesized SOX9 with 0.1 μg of poly(dG–dC)·poly(dG–dC) as non-specific competitor. Electrophoretic gel separations were run for 2–3 h at 140 V. When anti-Sox9 antibody was included, it was combined first with the SOX9 protein and incubated in binding buffer at room temperature for 30 min before radiolabeled probe was added.

Generation and analysis of transgenic mice

The plasmid pCXIwt2/3 was digested with *Eco*RI and *Pst*I to remove extraneous plasmid sequences prior to microinjections to make transgenic mice. Microinjections were performed as described previously (30). Founder embryos were sacrificed at 14.5 days post-coitum, and the presence of the transgene was detected by Southern analysis of placental DNA. Transgenic embryos were fixed and then stained with X-gal, photographed as whole mounts, and then embedded and sectioned for histological analysis as described previously (30).

RESULTS

Each of the two heptameric Sox consensus sequences in the intronic F/G enhancer element is necessary for transcriptional activity and for protein binding in EMSA

Previous work by others demonstrated that the intronic element that we call F/G has chondrocyte-specific enhancer activity in transgenic mice (25). That work also showed that the Sox site we call F, which is a 1 bp mismatch to the Sox

consensus binding sequence, binds Sox9 *in vitro*. In this study, we examined the importance of that heptameric site for enhancer activity. We also tested an adjacent site, G, which is a 2 bp mismatch to the Sox consensus binding sequence. The G site lies 3 bp downstream from the end of the F site and in reverse orientation to it, after the pattern of the B and C sites and the D and E sites found in the two upstream enhancers (24).

The importance of the F and G sites for enhancer activity was tested by constructing mutant versions of the F/G enhancer element containing mutations in the F site (Fm/G), the G site (F/Gm) or both sites (Fm/Gm). Mutations were designed to produce a total of four mismatches to the heptameric Sox consensus sequence, with As and Ts being replaced by Gs or Cs, and vice versa (Table 1). These mutant elements were multimerized as four tandem copies upstream of a minimal promoter and luciferase reporter. Earlier experiments had demonstrated that four copies of the B/C and D/E enhancers was the least number of copies that produced easily detectable enhancer activity, and so the same number of copies was used with F/G for consistency. The mutant enhancer constructs were transiently transfected into RCS cells to determine the effect of each mutation on enhancer activity. Mutation of only the F site decreased enhancer activity dramatically relative to the control. Mutation of only the G site had a similar effect, demonstrating that this site likewise is necessary for enhancer activity (Fig. 2A).

The wild-type and mutant versions of the F/G element were used as probes in EMSA experiments under the same conditions in which the *Col11a2* B/C and D/E elements and the *Col2a1* 48 bp chondrocyte-specific enhancer have been shown to form unique DNA–protein complexes with proteins found in chondrocyte but not fibroblast nuclear extracts (24,31). The complex with the *Col2a1* 48 bp enhancer previously was designated CSEP, for chondrocyte-specific enhancer-binding protein (Fig. 2B, first panel, large arrow) (31). Antibodies against Sox9 supershifted a portion of the CSEP complex (Fig. 2B, first panel, small arrow) (31). Under the same conditions, the F/G element formed a complex that had the same mobility as CSEP and that was also partially supershifted by antibodies against Sox9 (Fig. 2B, second

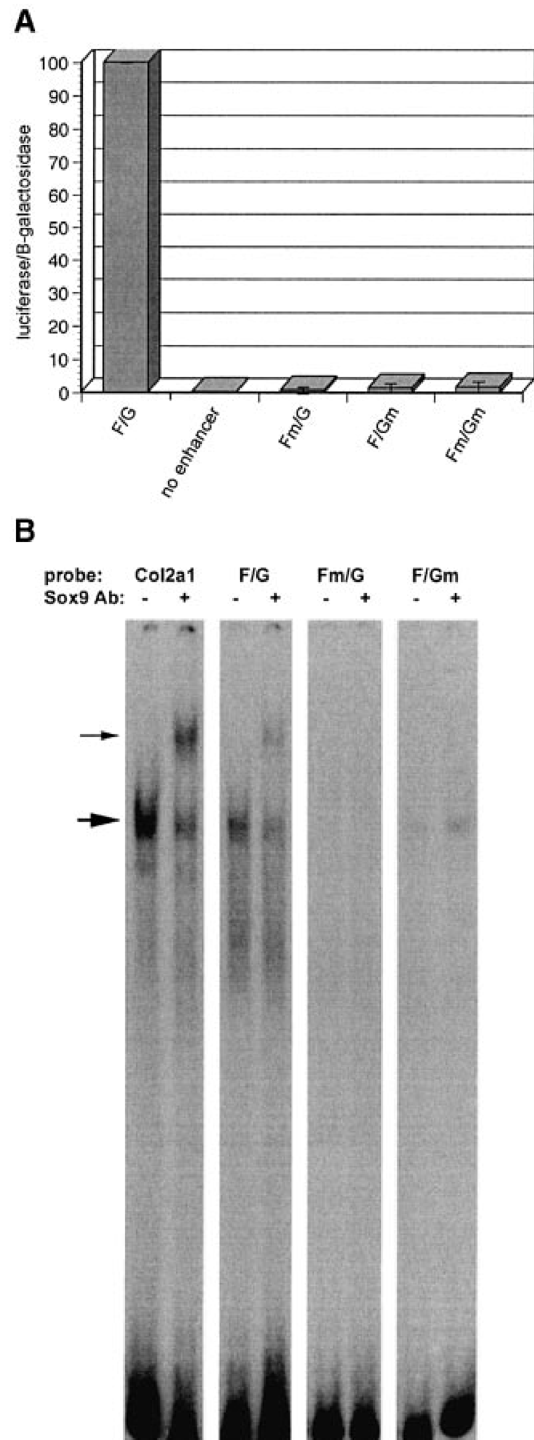

Figure 2. Site F and site G are both required for the transcriptional activity of the F/G enhancer and for the formation of a unique CSEP-like DNA–protein complex that includes Sox9. (**A**) Four copies of the wild-type F/G enhancer element (F/G), or mutant elements containing point mutations in either the F site (Fm/G), the G site (F/Gm) or both sites (Fm/Gm) were cloned upstream of a minimal promoter and a luciferase reporter gene (see Table 1 for sequences). The ability of each enhancer element to activate expression of the reporter gene was tested by transient transfection in RCS cells. Each experiment included 1.5 µg of reporter plasmid and 0.5 µg of an internal control for transfection efficiency, the pSV-β-galactosidase vector (Promega, Madison, WI). Results are presented as luciferase units per β-galactosidase unit ± SD, and are normalized to 100% activity for the wild-type F/G enhancer. Experiments were performed in duplicate, and the results of three independent repeats are shown. (**B**) EMSA was performed using nuclear extracts from chondrocytic RCS cells and four different DNA probes (named at the top of the figure). The probe in the first panel is the mouse *Col2a1* 48 bp chondrocyte-specific enhancer, which has been shown to form a unique chondrocyte-specific DNA–protein complex with proteins from RCS but not fibroblast nuclear extracts (31). This complex previously was named CSEP, and is indicated by the large arrow. Antibodies against Sox9 supershift a portion of the CSEP complex (small arrow). The probe in the second panel is the wild-type F/G enhancer element. It forms a DNA–protein complex that has similar mobility to CSEP and is also partially supershifted by antibodies against Sox9. The third and fourth panels contain mutant enhancer probes Fm/G and F/Gm as indicated. Mutating either the F site or the G site greatly reduced the CSEP-like complex, and no supershift by Sox9 antibodies is detectable with either mutant probe, indicating that the F and G sites are both needed for Sox9 binding. The four panels shown are all from the same EMSA experiment, but intervening lanes have been removed to bring the lanes of interest into closer proximity with each other.

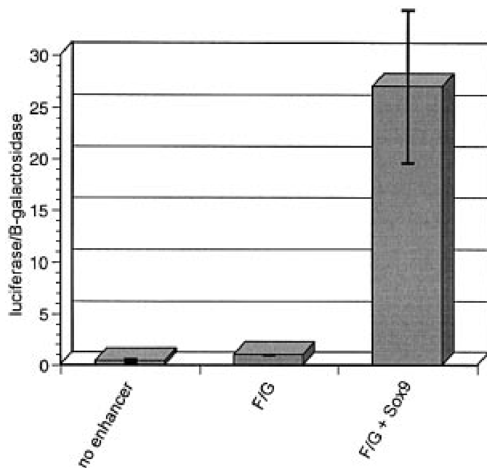

Figure 3. The F/G enhancer is strongly activated by ectopic expression of SOX9 in fibroblasts. The same F/G enhancer construct as described in Figure 2A was transiently transfected into BALB-3T3 fibroblasts, either with or without co-transfection of a SOX9 expression plasmid (SOX9-pcDNA-5′-UT). Each experiment included 1 µg of F/G reporter plasmid, 0.3 µg of an internal control for transfection efficiency, the pSV-β-galactosidase vector (Promega, Madison, WI) and 0.35 µg of the SOX9 expression plasmid or an empty expression plasmid control. Results are presented as luciferase units per β-galactosidase unit ± SD, and are normalized to the activity of the F/G enhancer without SOX9. The graph reflects the results of three independent experiments, each performed in duplicate.

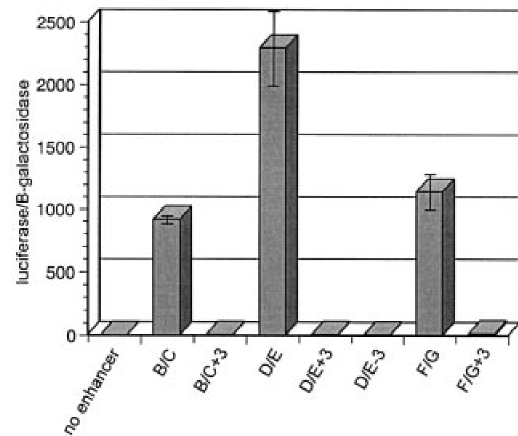

Figure 4. The spacing between the paired heptameric Sox consensus sequences is critical for enhancer activity. Insertion mutations (and one deletion mutation) were introduced between paired heptameric Sox-binding consensus sequences in the enhancers B/C, D/E and F/G to determine whether proximity of the sites to each other is necessary for their activity (see Table 1 for sequences). Mutated enhancers were multimerized to four tandem copies upstream of a minimal promoter and a luciferase reporter. The effect of each mutation was tested by transiently transfecting plasmids into RCS cells and measuring reporter gene expression induced by each mutant enhancer compared with its wild-type counterpart. Each experiment included 1.5 µg of experimental plasmid and 0.5 µg of an internal control for transfection efficiency, pSV-β-galactosidase (Promega, Madison, WI). Results are presented as luciferase units $\times 10^4$ per β-galactosidase unit ± SEM, and include two independent experiments, each performed in duplicate.

panel). Mutation of the F site or the G site in the DNA probe virtually eliminated this complex, and antibodies against Sox9 no longer produced any supershift (Fig. 2B, third and fourth panels). These results are consistent with those shown in Figure 2A, in which the same mutations abolished activity of the F/G enhancer in transient transfection experiments (Fig. 2A), indicating that the proteins that bind these Sox sites in EMSA probably play a critical role in transcriptional activation.

The F/G enhancer can be activated by ectopic expression of SOX9 in non-chondrocytic cells

The wild-type F/G reporter construct described above was tested in transient transfections of BALB-3T3 fibroblasts and found to be inactive, as expected for a chondrocyte-specific enhancer. Co-transfection with the previously described expression plasmid SOX9-pcDNA-5′-UT containing the coding sequences for human SOX9, however, produced 25-fold increased activity in fibroblasts (Fig. 3) (22). Previous experiments have shown that the B/C and D/E enhancers are also activated by SOX9 under similar conditions (24). This result, together with the demonstration of SOX9 binding to the F/G element in EMSA, suggests that Sox9 participates in transcriptional activation of *Col11a2* through the F/G enhancer.

The spacing between the paired Sox sites is critical for enhancer activity

The heptameric Sox consensus sites that have been shown to be important for transcriptional activity in all three *Col11a2* enhancers are arranged in pairs separated by 3 or 4 bp. If this spacing was necessary to allow protein–protein interactions between factors bound at the paired sites, and if those interactions were in turn needed for transcriptional activation, then changing the spacing between the sites would be expected to reduce transcriptional activity. Accordingly, mutant versions of each of the enhancers were constructed. Three additional base pairs were inserted between B and C in the B/C element (B/C+3), between D and E in the D/E element (D/E+3) and between F and G in the F/G element (F/G+3). In addition, the D/E element was altered to remove all three of the base pairs from between the two Sox sites, positioning D and E immediately adjacent to each other (D/E–3) (Table 1). All four of these spacing mutants were found to be completely inactive in transient transfections of RCS cells (Fig. 4). This result indicates that the arrangement of the Sox sites in correctly spaced pairs is necessary to allow productive interactions between adjacently bound proteins, which interactions are required for transcriptional activation.

The sequence of the nucleotides between the paired Sox sites is not critical for enhancer activity

In order to determine whether the sequence or simply the size of the DNA between the paired Sox sites is important, substitution mutations were introduced between the two consensus Sox sites in each enhancer element. The resulting elements were named B*C, D*E and F*G (Table 1). Mutant enhancer elements were multimerized to four tandem copies upstream of a minimal promoter and a luciferase reporter, and tested in transient transfections of RCS cells. The substitution mutations were found to have little effect on transcriptional activation, with the exception of F*G, which was more active than the wild-type F/G enhancer (Fig. 5). In no case did the substitution mutations between the Sox sites abolish enhancer

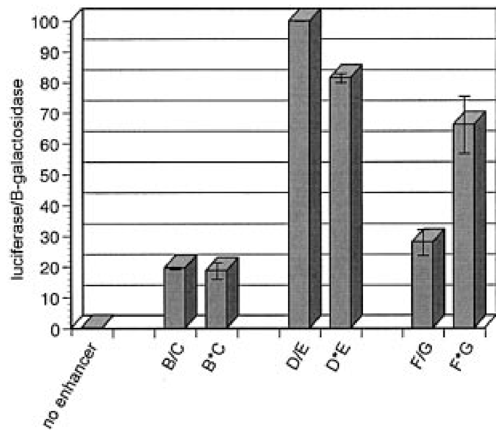

Figure 5. The sequence of the nucleotides between the paired heptameric Sox sites is not critical for enhancer activity. Single base pair substitution mutations were introduced between the paired Sox sites in each of the three enhancers to determine whether those intervening sequences participate in protein binding (see Table 1 for sequences). Mutated enhancers were multimerized to four tandem copies upstream of a minimal promoter and a luciferase reporter. The effect of each mutation was tested by transiently transfecting plasmids into RCS cells and measuring reporter gene expression induced by each mutant enhancer compared with its wild-type counterpart. Each experiment included 1.5 µg of experimental plasmid and 0.5 µg of an internal control for transfection efficiency, pSV-β-galactosidase (Promega, Madison, WI). Results are presented as luciferase units per β-galactosidase unit ± SEM, and include three independent experiments, each performed in triplicate and normalized to the activity of the D/E enhancer.

activity, in striking contrast to the effects of insertion and deletion mutations between the Sox sites. These results indicate that the spatial separation between the two binding sites is critical for productive protein interactions, but the exact sequence of the nucleotides separating the protein-binding sites is not.

The Sox sites each contribute to transcriptional activation in the context of the larger *Col11a2* regulatory region

The B, C, D, E, F and G sites have all been shown to be necessary for the activity of the short modular enhancer elements within which they reside (Fig. 2) (24). In order to determine whether these sites are also important within the context of the entire *Col11a2* regulatory region, they were mutated individually using PCR to create the series of 'Δ' constructs (containing Sox site substitution mutations) in the pCXIwt2/3 vector. This plasmid contains 1122 bp of the 5′ region and 684 bp of the first intron of *Col11a2* driving expression of a β-galactosidase reporter gene.

The pcXIwt2/3 plasmid was developed initially by combining the 1122 bp 5′ region that previously had been shown to activate transcription in chondrocytes (albeit with an irregular distribution of expression) with subsections of the first intron, which had been shown to increase and equalize expression levels in cartilage structures when combined with the 5′ region (24,27). Constructs containing the 5′ region plus the first third, the middle third or the last third of the intron were tested in transient tranfection experiments. The construct containing the middle third (which is 684 bp long and includes the F/G enhancer) produced more reporter gene expression than either the first or the last thirds, and was 30% more active than the construct containing the entire first intron (data not shown). This plasmid was named pCXIwt2/3 and was tested as a transgene in developing mouse embryos to confirm that it was an appropriate positive control for the subsequent Sox site targeted mutational analysis. In three out of three separate transgenic founder mice obtained, this transgene directed β-galactosidase reporter gene expression specifically to cartilage. No mice were found that contained the transgene (as determined by Southern analysis) but did not express the β-galactosidase reporter. The cartilage-specific expression pattern was documented both in whole-mount embryos stained with lacZ and in histological sections of the stained embryos (Fig. 6). This transgene experiment confirmed that pCXIwt2/3 contains the regulatory elements necessary to produce chondrocyte-specific gene expression *in vivo*, and therefore established this plasmid as an appropriate model in which to test the effects of targeted Sox-binding site mutations on gene expression.

In the context of the pCXIwt2/3 plasmid, mutations were introduced into each of the heptameric Sox protein-binding

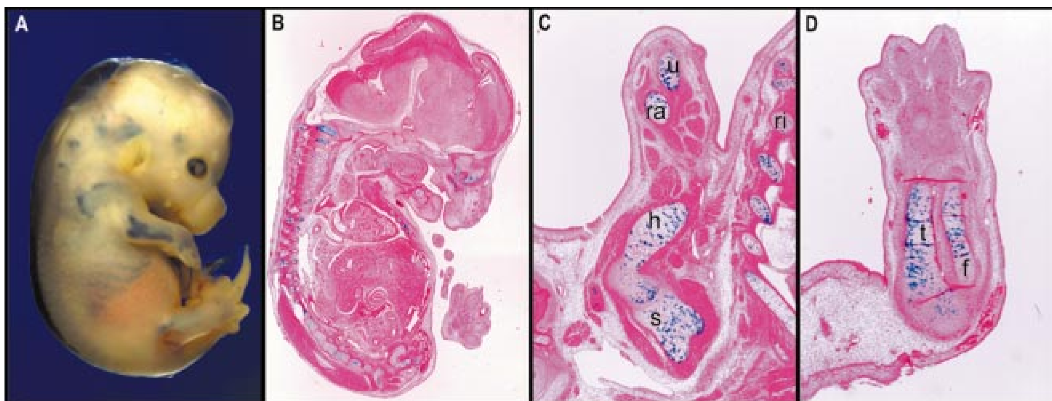

Figure 6. A total of 1122 bp from the 5′ region and 684 bp from the first intron of *Col11a2* direct transgene expression in a pattern that is cartilage specific. Transgenic mouse embryos generated using the vector pCXIwt2/3 were stained with X-gal at day 14.5 of embryonic development (when the skeleton is largely composed of cartilage) and were examined in whole mount and after histological sectioning. (**A**) Whole mount showing transgene expression in scapula, radius, ulna, tibia, fibula, digits, ribs, vertebrae and portions of the cranium. (**B**) Low magnification sagittal section. (**C**) Forelimb region showing transgene expression in scapula (s), humerus (h), radius (ra), ulna (u) and ribs (ri). (**D**) Hindlimb section showing transgene expression in the tibia (t) and fibula (f).

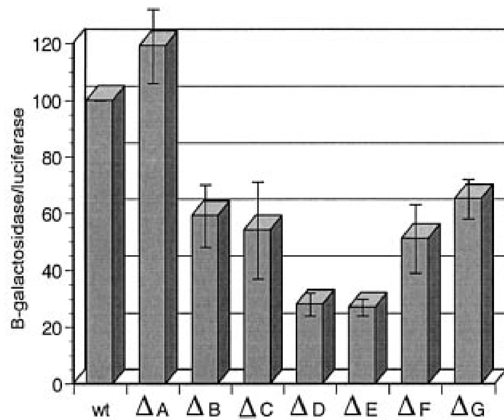

Figure 7. Sites B, C, D, E, F and G are each important in the context of the larger *Col11a2* regulatory region. Targeted substitution mutations were introduced by PCR into each of the heptameric Sox-binding sites, one by one, in the pCXIwt2/3 plasmid (labeled 'wt' in this figure) to create the mutant plasmid series ΔA, ΔB, ΔC, ΔD, ΔE, ΔF and ΔG. The mutant plasmids were transiently transfected into RCS cells, and the reporter β-galactosidase gene expression from each mutant plasmid was measured and compared with the amount of reporter gene expression from the pCXIwt2/3 control plasmid. Experiments included 1.5 µg of experimental plasmid and 0.5 µg of an internal control for transfection efficiency, pGL-3, which expresses luciferase (Promega, Madison, WI). Results are presented as β-galactosidase units per luciferase unit ± SEM, and are normalized to 100% activity for the pCXIwt2/3 wild-type plasmid. The figure includes data from seven independent experiments, all normalized to 100% activity for the pCXIwt2/3 positive control and each performed in duplicate.

sites individually, using PCR. Each of the resulting Δ plasmids was therefore identical to the pCXIwt2/3 plasmid, except for the introduction of substitution mutations within the indicated heptameric site (see Materials and Methods for sequences.) Each plasmid was transiently transfected into RCS cells to measure the effect of the mutation. These experiments showed that mutation of any of the paired Sox sites B, C, D, E, F or G decreased the activity of the pCXIwt2/3 construct (Fig. 7), suggesting that each of these sites plays a significant role in transcriptional activation in the context of the entire *Col11a2* regulatory region. Mutation of the A site, in contrast, did not decrease transcriptional activity of the β-galactosidase reporter gene. The A site lies 18 bp upstream of the B site in the B/C element and has no paired partner. This result confirms our earlier conclusion that the A site does not bind proteins that are necessary to activate transcription (24).

The results of this experiment are consistent with our prior observation that in individual enhancer elements, mutation of either site in a pair prevented protein binding and eliminated transcription (Fig. 2) (24). Mutation of either site in a pair in the context of the entire regulatory region produced an almost equal reduction in reporter gene expression, suggesting once again that the paired sites act together, with mutation of either site inactivating the pair.

Interestingly, the ΔF and ΔG constructs containing mutations that are known to completely inactivate the F/G enhancer element produced reporter gene expression that was still twice as high as the pCXIwt plasmid (data not shown). Since the pCXIwt plasmid is completely lacking intronic DNA, this result suggests that the F/G enhancer is not the only positive element present in the middle third of the *Col11a2* first intron.

Paired Sox sites are not alone sufficient to produce chondrocyte-specific enhancer activity

The discovery that all three *Col11a2* enhancers contain paired Sox sites suggested that this feature might define a chondrocyte-specific enhancer element. To test this idea, computer searches of the mouse *aggrecan* gene and the mouse *Col9a2* gene (both largely chondrocyte specific in expression pattern) were used to identify regions containing paired Sox sites (Table 1). These regions were multimerized to four tandem copies upstream of the same minimal promoter and luciferase reporter as the *Col11a2* enhancers, and were tested in transient transfections of RCS cells to see if they contained enhancer activity. These elements both failed to produce any more transcriptional activity than the 'no enhancer' control, demonstrating that the presence of paired Sox sites is not alone sufficient to produce enhancer activity (data not shown).

Each *Col11a2* chondrocyte-specific enhancer element contains other important sequences besides the heptameric Sox-binding sites

Because paired heptameric Sox-binding sites alone were not sufficient to produce chondrocyte-specific enhancer activity, thorough mutational analyses of the B/C, D/E and F/G enhancers were performed to determine what other sequences in each of these enhancers were necessary for activity. A mutation series in which substitutions were introduced 3 bp at a time, every 6 bp along the length of each enhancer, was constructed (Figs 8A, 9A and 10A). The mutated enhancer elements were multimerized to four tandem copies and cloned upstream of the same minimal promoter and luciferase reporter as before. The effects of each mutation were tested in transient transfections of RCS cells, and mutations that decreased enhancer activity by 50% or more were interpreted as having prevented the binding of important proteins. The selection of 50% reduction as the critical level was somewhat arbitrary, but served the purpose of providing a starting point for later protein identification experiments.

The mutational analysis identified an important region immediately upstream of the B site in the B/C enhancer element, spanning mutations 3, 4 and 5 (Fig. 8). Mutations in this region did not change either the B or C Sox sites, and yet they decreased enhancer activity dramatically, indicating that the region probably binds one or more proteins that are critical to the activity of the enhancer. Mutations 6–9, which did change the B and C sites, also decreased activity as expected. Mutations 1 and 2 reduced enhancer activity to ~40 and 50% of wild-type, respectively, and so the region of these mutations might also merit further examination.

The D/E element mutations 3 and 5, which disrupted the heptameric Sox consensus sites, markedly decreased enhancer activity as expected (Fig. 9). Mutation 4, which introduced one additional mismatch into the E site, also reduced activity somewhat. In addition, mutation 1 decreased activity to ~40% of wild-type. The region of mutation 1, therefore, is the most likely location in the D/E enhancer for the binding of an additional activating protein.

The F/G element mutations 5–7 all disrupted Sox consensus sites and decreased activity as expected (Fig. 10). Mutation 4, interestingly, also reduced enhancer activity even though it converted the F site to a perfect match of the heptameric

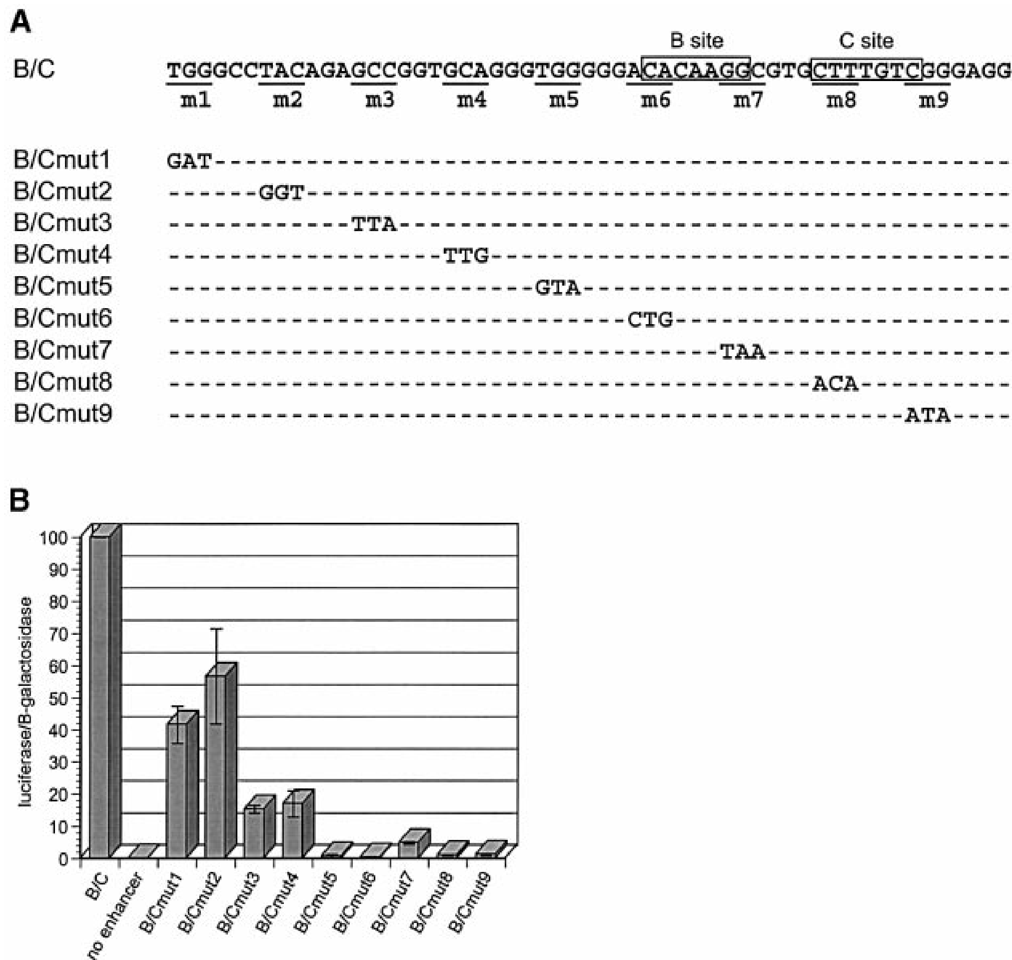

Figure 8. The B/C enhancer contains sequences outside the B and C sites that are also necessary for enhancer activity. (**A**) To locate protein-binding sites in the B/C enhancer, 3 bp mutations were introduced along the length of the enhancer as shown. (**B**) Mutated enhancers were multimerized to four tandem copies upstream of a minimal promoter and a luciferase reporter. The effect of each mutation was tested by transiently transfecting the plasmids into RCS cells and measuring reporter gene expression induced by each mutant enhancer compared with the wild-type B/C element. Each transfection included 1.5 µg of experimental plasmid and 0.5 µg of an internal control for transfection efficiency, pSV-β-galactosidase (Promega, Madison, WI). Results are presented as luciferase units per β-galactosidase unit ± SEM, and include three independent experiments, each performed in duplicate or triplicate and normalized to 100% activity for the B/C enhancer.

Sox-binding consensus sequence. This result supports the hypothesis that Sox9 binding is not alone sufficient to activate the F/G enhancer, because F/Gmut4 would still be expected to bind Sox9. Mutation 8 did not affect a Sox site and yet it virtually eliminated enhancer activity. Mutation 3, likewise outside of the Sox-binding sites, reduced activity to ~40% of wild type. The regions around and including mutations 3, 4 and 8 therefore probably bind activating proteins.

Mutations that prevent enhancer activity do not necessarily inhibit SOX9 binding

Sox9 has been shown to activate each of the three *Col11a2* enhancer elements. The mutational analysis described above, however, demonstrated that mutations outside of the heptameric Sox consensus binding sites that would not be expected to disrupt Sox9 binding directly can still reduce enhancer activity. EMSA experiments were performed to determine whether these non-Sox mutations might actually be preventing Sox9 binding even though they lie outside the predicted binding sites.

When the B/C enhancer mutation series of probes was tested, only mutations 6 and 8 prevented binding of *in vitro* transcribed and translated SOX9 in EMSA (Fig. 11A). These mutations fell within the recognized B and C Sox-binding sites (see Fig. 8A). Mutations 7 and 9 also affected the two Sox sites, but these mutations produced new Sox sites that were equal or better matches to the heptameric Sox consensus binding sequence. They did not prevent SOX9 binding in EMSA experiments. They did, however, abolish the transcriptional activity of the B/C enhancer (Fig. 8). Likewise, mutations 3, 4 and 5 did not prevent SOX9 binding and yet they greatly reduced or eliminated transcriptional activity. These results demonstrate that mutations can inhibit transcription from the B/C enhancer without affecting SOX9 binding *in vitro*, suggesting once again that transcriptional activation from the B/C enhancer requires one or more proteins in addition to Sox9.

The D/E enhancer mutation series was also tested in EMSA. Mutants 3 and 5, which contained severe disruptions to the D and E HMG sites and eliminated enhancer activity, prevented

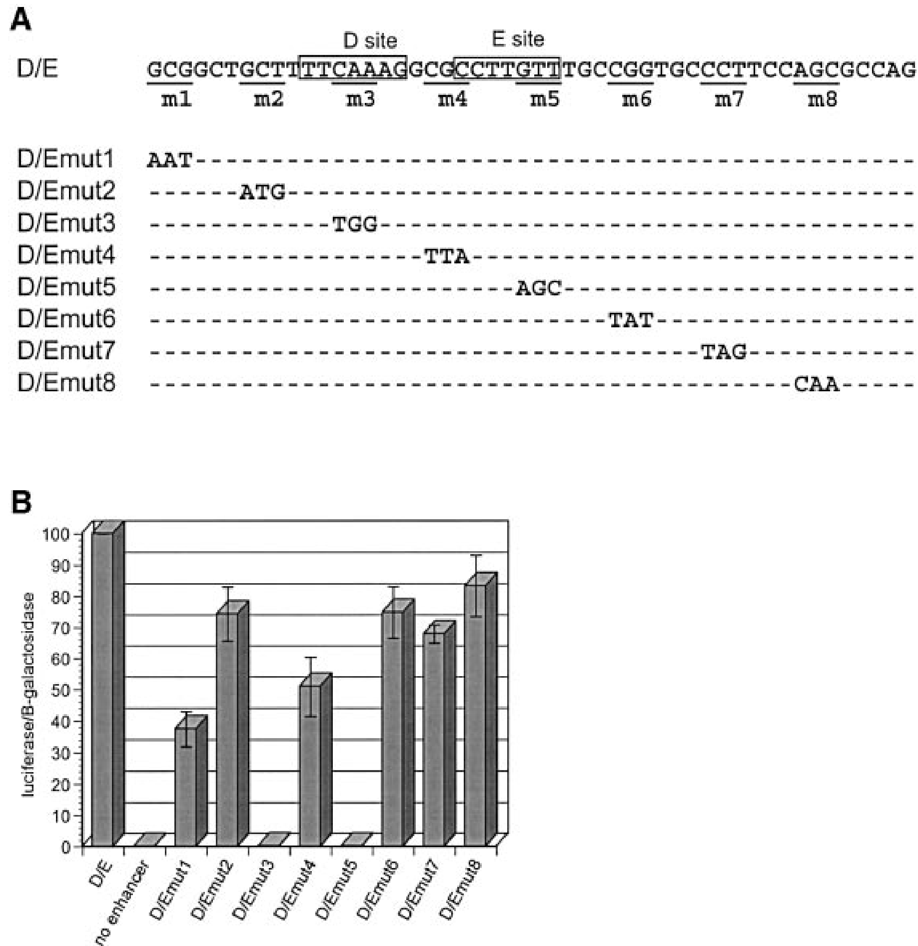

Figure 9. The D/E enhancer contains sequences outside the D and E sites that are also necessary for enhancer activity. (**A**) To locate protein-binding sites in the D/E enhancer, 3 bp mutations were introduced along the length of the enhancer as shown. (**B**) Mutated enhancers were multimerized to four tandem copies upstream of a minimal promoter and a luciferase reporter. The effect of each mutation was tested by transiently transfecting the plasmids into RCS cells and measuring reporter gene expression induced by each mutant enhancer compared with the wild-type D/E element. Each transfection included 1.5 µg of experimental plasmid and 0.5 µg of an internal control for transfection efficiency, pSV-β-galactosidase (Promega, Madison, WI). Results are presented as luciferase units per β-galactosidase unit ± SEM, and include three independent experiments, each performed in duplicate or triplicate and normalized to 100% activity for the D/E enhancer.

SOX9 binding (compare Figs 11B and 9). Mutation 1, however, which reduced transcriptional activity to 40%, had no effect on SOX9 binding. This result suggests that another protein besides Sox9 helps activate the D/E enhancer through interactions in the area of mutation 1.

When the F/G enhancer mutation series was tested by EMSA with *in vitro* transcribed and translated SOX9, mutations 5, 6 and 7, which all reduced transcriptional activity, also reduced or prevented SOX9 binding. However, mutations 3, 4 and 8, which also reduced transcriptional activity, did not affect SOX9 binding (compare Figs 11C and 10). The F/Gmut4, which mutation created a perfect Sox consensus sequence at the F site but also changed 2 bp upstream of the F site, is of particular interest. As predicted, this mutant probe binds SOX9 as well or better than the wild-type probe *in vitro*. Transcriptionally, though, it is almost inactive. This result again supports the concept that other proteins that bind adjacent to Sox9 are essential for enhancer activation.

Together, these EMSA experiments support the hypothesis that while Sox9 binding is critical for *Col11a2* enhancer activity, it alone is not sufficient to produce that activity. All the mutations that prevented SOX9 binding in EMSA also reduced or eliminated transcriptional activity, but several mutations that did not affect SOX9 binding were also able to inhibit transcriptional activity. Those mutations probably interfered with the binding of other essential proteins.

DISCUSSION

The primary goals of the work presented here were to assess the importance of the structural features held in common by the three chondrocyte-specific *Col11a2* enhancer elements, to examine binding of Sox9 to the heptameric Sox sites in the three enhancers and to identify any other protein-binding sites that were also needed for enhancer activity. This work has paved the way for the identification of all the additional

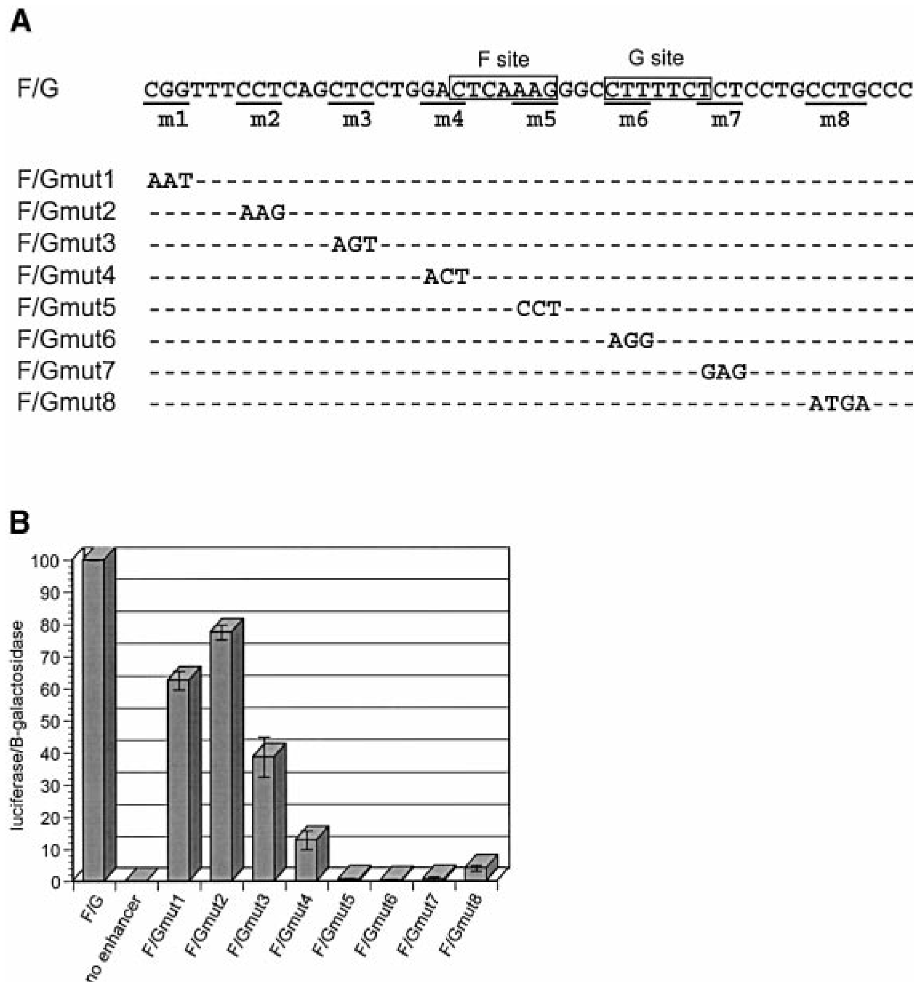

Figure 10. The F/G enhancer contains sequences outside the F and G sites that are also necessary for enhancer activity. (**A**) To locate protein-binding sites in the F/G enhancer, 3 bp mutations were introduced along the length of the enhancer as shown. (**B**) Mutated enhancers were multimerized to four tandem copies upstream of a minimal promoter and a luciferase reporter. The effect of each mutation was tested by transiently transfecting the plasmids into RCS cells and measuring reporter gene expression induced by each mutant enhancer compared with the wild-type F/G element. Each transfection included 1.5 µg of experimental plasmid and 0.5 µg of an internal control for transfection efficiency, pSV-β-galactosidase (Promega, Madison, WI). Results are presented as luciferase units per β-galactosidase unit ± SEM, and include three independent experiments, each performed in duplicate and normalized to 100% activity for the F/G enhancer.

proteins that bind the *Col11a2* enhancers and cooperate with Sox9 to activate transcription.

The B/C and D/E elements were already known to contain two critical Sox-binding consensus sequences each, which can bind SOX9 *in vitro* (24). We have shown herein that the F/G enhancer element also contains two Sox sites, both of which participate cooperatively in SOX9 binding *in vitro*. We have also shown that SOX9 can activate the chondrocyte-specific F/G enhancer in fibroblasts, where it normally is not active. This study, then, has demonstrated a consistent structural pattern in all three of the known *Col11a2* chondrocyte-specific enhancer elements. All three contain two Sox sites, which are separated from each other by 3 or 4 bp and arranged in opposite orientation to each other. SOX9 binding to these sites was prevented by mutating either one of the sites in each pair. Mutation of either site in a pair also eliminated enhancer activity, as did changing the spacing between the paired sites. These results suggest a scenario in which Sox9 binds as a dimer to the paired heptameric Sox sites in each enhancer, and from there activates transcription.

The importance of paired Sox sites for Sox9 binding was clearly demonstrated by the mutational analysis. One might expect, however, that even if Sox9 normally binds the *Col11a2* enhancers as a dimer, mutation of one of the Sox sites would leave monomeric Sox9, as was previously observed for Sox10 bound to paired Sox consensus sequences (32). Our EMSA experiments using *in vitro* transcribed and translated SOX9 with each of the three enhancer elements, however, showed no evidence of separate monomeric and dimeric Sox9–DNA complexes. Only one complex was supershifted by the addition of antibodies against Sox9, and mutation of either one of the Sox-binding sites in each pair abolished that complex formation (Fig. 11). It appears that Sox9 binds the *Col11a2* enhancers as a dimer or not at all.

These data could be interpreted in two different ways. The first possible interpretation is that Sox9 binds the enhancers as

a heterodimer with another protein such as L-Sox5 or Sox6, both of which have been shown to play a role in cartilage development, and that mutation of either the Sox9-binding site or the other heptameric site prevents formation of the entire complex (33,34). This interpretation is unlikely, however, because dimerization was clearly present in our EMSA experiments with *in vitro* transcribed and translated SOX9, where the only other proteins present besides SOX9 were those from the *in vitro* transcription/translation mix. Still, it is possible that L-Sox5 and Sox6 could compete with Sox9 *in vivo* for binding at either or both of the Sox sites in a pair, perhaps exerting a negative regulatory effect. The second possible interpretation is that Sox9 binds to the *Col11a2* enhancers as a homodimer. Sox10, the Sox protein most closely related to Sox9, is known to form homodimers as it binds to paired Sox sites in a regulatory element of the P_0 gene (32). The 36 amino acid Sox10 homodimerization domain is 78% conserved in Sox9, consistent with the idea that Sox9 also binds the *Col11a2* enhancers as a homodimer (16,32). Neither Sox10 nor Sox9, however, is capable of forming homodimers in solution (32,34). Their dimerization domains instead appear to be configured upon DNA binding so that they only function on appropriate DNA-binding sites (32).

It is interesting that Sox9 appears capable of binding to six different heptameric sites in the three *Col11a2* enhancer elements, and yet none of those sites has exactly the same binding sequence. Similarly, the Sox protein family includes at least 20 members that all share affinity for the same heptameric consensus sequence (A/T)(A/T)CAA(A/T)G (18). In spite of sharing the same consensus binding sequence preference, the different Sox proteins precisely regulate their own unique target genes and no others. The question of how such similar DNA-binding domains achieve target site specificity *in vivo* represents a critically important aspect of our understanding of gene regulation. It appears that Sox proteins achieve binding site specificity in many cases by interacting with other proteins that bind at adjacent sites (32,35–37).

Experiments by others in two different genetic systems yielded relevant information about interactions between Sox proteins and proteins bound at adjacent binding sites. The DC5 enhancer element from the δ-*crystallin* gene contains binding sites for Sox2 and δEF3 separated by 3 bp. Increasing the spacing between these sites by an additional 4 bp eliminated enhancer activity (35). In EMSA experiments, however, Sox2 was still able to bind the mutant enhancer. Separate experiments examined a regulatory element from the fibroblast growth factor-4 (FGF-4) gene that contains paired binding sites for Sox2 and Oct-3 separated by 3 bp. In this case, increasing the separation to ⩾6 bp eliminated transcriptional activity but still allowed binding of the individual proteins in EMSA (36). These experiments demonstrated the necessity for protein–protein interactions between Sox proteins and adjacently bound proteins in two different genes.

Besides providing evidence of interactions between Sox proteins and adjacently bound proteins, these results also

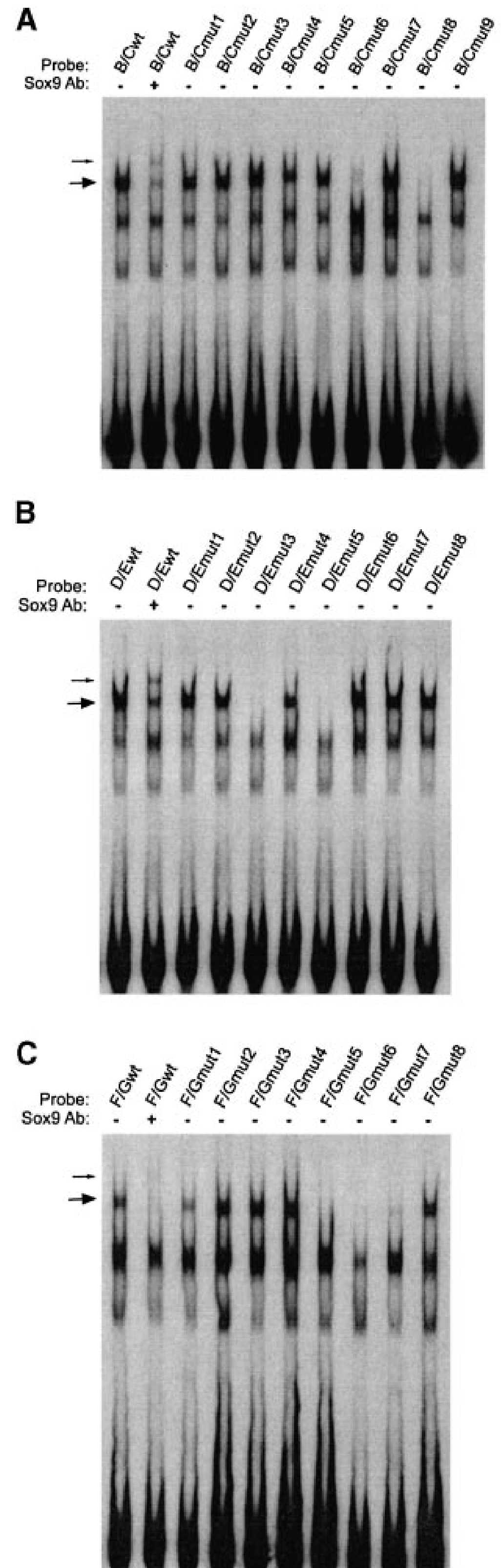

Figure 11. Mutations that reduce transcriptional activity do not necessarily prevent SOX9 binding. The mutant enhancer elements shown in Figures 8–10 were used as probes in EMSA experiments with *in vitro* transcribed and translated SOX9 to determine whether SOX9 binding correlates with transcriptional activity. Antibody against Sox9 was included with each wild-type probe to determine which DNA–protein complex contained SOX9. Complexes containing SOX9 are marked by large arrows, and small arrows indicate the supershifted SOX9 antibody complex. (**A**) B/C mutant probe series; (**B**) D/C mutant probe series; (**C**) F/G mutant probe series.

parallel our observations in EMSA experiments that the ability of an enhancer to bind SOX9 at critical Sox-binding sites *in vitro* does not guarantee that the element will have transcriptional activity *in vivo*. Even when both Sox sites were intact and correctly spaced, and SOX9 binding to the *Col11a2* enhancer elements appeared robust in EMSA experiments, we identified several mutant enhancers that had little or no transcriptional activity in transient transfection experiments. These mutants demonstrated that although the paired Sox sites were both necessary and sufficient to allow SOX9 binding to the *Col11a2* enhancers *in vitro*, other adjacent protein-binding sites were also needed to accomplish transcriptional activation *in vivo*.

Of particular interest were three mutations, B/Cmut7, B/Cmut9 and F/Gmut4, which altered Sox sites and decreased transcriptional activity, but did not affect SOX9 binding in EMSA. In every case, these mutations produced new Sox sites that were equal or better matches to the heptameric Sox consensus sequences than were the original Sox sites, accounting for their ability to bind SOX9 *in vitro*. The B/Cmut9 and F/Gmut4 mutations each altered two additional base pairs immediately outside of the paired Sox sites, probably disrupting protein binding at those adjacent sites to prevent transcription. The B/Cmut7 mutation, however, altered only one additional base pair between the paired Sox sites. Our experiments (Figs 4 and 5) indicated that the number but not the sequence of the nucleotides between the paired sites is important for enhancer activity, and so it was surprising that this mutant was transcriptionally inactive. It is possible that the 2 bp mutation in the Sox-binding site of B/Cmut7 allowed SOX9 binding and DNA-dependent dimerization *in vitro*, but that *in vivo* the sequence alteration prevented formation of the correctly configured protein–DNA complex required to activate transcription.

Sox9 has an intrinsic transactivation domain, and so it is somewhat surprising that additional adjacently bound proteins would be required for transcriptional activation. Recent conformational studies on HMG domains may shed some light on this issue. Nuclear magnetic resonance (NMR) and crystallographic analysis of HMG boxes from Sox proteins have revealed a unique feature. HMG boxes in general consist of three α-helices and an N-terminal extension that together fold into a twisted L-shape. The 'L' consists of two wings that together form the DNA-binding and bending domain. Evidence suggests that one of the two wings has a disordered structure in solution, and only upon binding to DNA does it assume a rigid shape and bend angle (17). The final shape would be unique to the DNA-binding sequence and could enhance binding of other proteins to adjacent DNA sites. Likewise, the prior presence of other proteins at adjacent DNA sites could alter the manner in which the disordered HMG wing binds to its consensus sequence. Binding of Sox9 to sites in different enhancers with different adjacently bound proteins would probably result in slightly different protein conformations and DNA bend angles. This in turn would be expected to attract different accessory proteins such as heat shock protein 70 (HSP70), which recently has been shown to interact with Sox9 in multiprotein complexes during transcription (38). In this way, two enhancer-binding complexes that both contained Sox9 nevertheless could have different and distinct effects on transcription.

In contradiction to the apparent requirement for proteins that bind outside the Sox sites to activate the *Col11a2* enhancers, ectopic expression of only SOX9 is sufficient to activate the F/G enhancer in fibroblasts where it is not normally active (Fig. 3). Ectopic expression of SOX9 can also activate the B/C and D/E chondrocyte-specific enhancers in fibroblasts (24). This result suggests that the other proteins that cooperate with Sox9 may be more ubiquitous than Sox9, and may already be present in fibroblasts. In this case, Sox9 would provide the tissue specificity for gene expression and the ubiquitous proteins would be recruited in unique combinations and utilized to activate gene expression only in the presence of Sox9.

In summary, the experiments presented here suggest the following model for protein binding to the *Col11a2* enhancer elements. Ubiquitous proteins bind to sequences adjacent to the Sox sites in each enhancer element. Sox9 then binds to the Sox consensus sequences in each element, but it can only form stable interactions with the DNA if it binds as a dimer with another Sox9 protein bound at the adjacent Sox site. The Sox9 dimerization domain is configured upon DNA binding, and so dimerization occurs simultaneously with or very soon after the initial Sox9–DNA binding. Another change also takes place at the time of Sox9 binding. The disordered wing of the Sox9 HMG box domain takes a stable and ordered conformation that is dependent on the precise DNA sequence it has bound and upon the identity of the proteins bound at the adjacent non-Sox sites. This conformation determines the angle of the bend that is induced in the DNA, which in turn produces a unique DNA and protein complex configuration that may attract additional protein cofactors. Only when the correct cofactors are all assembled in a properly configured enhancer complex can Sox9 transactivate gene expression.

Through the analyses presented here, the important protein-binding sites in each of the *Col11a2* enhancers have been delineated, and the role of Sox9 as an activator of all three enhancers is firmly established. The way is now open for the identification of all the other proteins that bind these enhancers.

ACKNOWLEDGEMENTS

We thank James H. Kimura of the Henry Ford Hospital in Detroit, MI, for RCS cells; Darwin J. Prockop of the Center for Gene Therapy, Tulane University Health Science Center in New Orleans, LA, for the cosmid containing the mouse *Col11a2* gene; and Veronique Lefebvre of the Cleveland Clinic Foundation in Cleveland, OH, for the *Col2a1* enhancer probe and the p89Col2a1Bs plasmid. We also thank Gerald J. Pinero of the University of Texas Dental Branch at Houston for histological work. This work was supported by National Institutes of Health Grant AR46362 (to L.C.B.), by National Institutes of Health Grant AR42919 (to B.d.C.) and by the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation Fellowship DRG-1464 (to L.C.B.).

REFERENCES

1. Melkoniemi,M., Brunner,H.G., Manouvrier,S., Hennekam,R., Superti-Furga,A., Kaariainen,H., Pauli,R.M., van Essen,T., Warman,M.L., Bonaventure,J., Miny,P. and Ala-Kokko,L. (2000)

Autosomal recessive disorder otospondylomegaepiphyseal dysplasia is associated with loss-of-function mutations in the COL11A2 gene. *Am. J. Hum. Genet.*, **66**, 368–377.

2. Li,Y., Lacerda,D.A., Warman,M.L., Beier,D.R., Yoshioka,H., Ninomiya,Y., Oxford,J.T., Morris,N.P., Andrikopoulos,K., Ramirez,F. *et al*. (1995) A fibrillar collagen gene, Col11a1, is essential for skeletal morphogenesis. *Cell*, **80**, 423–430.
3. Vikkula,M., Mariman,E.C., Lui,V.C., Zhidkova,N.I., Tiller,G.E., Goldring,M.B., van Beersum,S.E., de Waal Malefijt,M.C., van den Hoogen,F.H., Ropers,H.H. *et al*. (1995) Autosomal dominant and recessive osteochondrodysplasias associated with the COL11A2 locus. *Cell*, **80**, 431–437.
4. van Steensel,M.A., Buma,P., de Waal Malefijt,M.C., van den Hoogen,F.H. and Brunner,H.G. (1997) Oto-spondylo-megaepiphyseal dysplasia (OSMED): clinical description of three patients homozygous for a missense mutation in the COL11A2 gene. *Am. J. Med. Genet.*, **70**, 315–323.
5. Gregory,K.E., Oxford,J.T., Chen,Y., Gambee,J.E., Gygi,S.P., Aebersold,R., Neame,P.J., Mechling,D.E., Bachinger,H.P. and Morris,N.P. (2000) Structural organization of distinct domains within the non-collagenous N-terminal region of collagen type XI. *J. Biol. Chem.*, **275**, 11498–11506.
6. Bos,K.J., Holmes,D.F., Kadler,K.E., McLeod,D., Morris,N.P. and Bishop,P.N. (2001) Axial structure of the heterotypic collagen fibrils of vitreous humour and cartilage. *J. Mol. Biol.*, **306**, 1011–1022.
7. Blaschke,U.K., Eikenberry,E.F., Hulmes,D.J., Galla,H.J. and Bruckner,P. (2000) Collagen XI nucleates self-assembly and limits lateral growth of cartilage fibrils. *J. Biol. Chem.*, **275**, 10370–10378.
8. Prockop,D.J. (1999) Heritable osteoarthritis. Diagnosis and possible modes of cell and gene therapy. *Osteoarthritis Cartilage*, **7**, 364–366.
9. Prockop,D.J. (1999) Hopkins Memorial Medal lecture. Pleasant surprises en route from the biochemistry of collagen to attempts at gene therapy. *Biochem. Soc. Trans*, **27**, 15–31.
10. Sandell,L.J. (2000) Genes and gene expression. *Clin. Orthop.*, S9–S16.
11. Kammandel,B., Chowdhury,K., Stoykova,A., Aparicio,S., Brenner,S. and Gruss,P. (1999) Distinct *cis*-essential modules direct the time–space pattern of the Pax6 gene activity. *Dev. Biol.*, **205**, 79–97.
12. Plaza,S., Saule,S. and Dozier,C. (1999) High conservation of *cis*-regulatory elements between quail and human for the Pax-6 gene. *Dev. Genes Evol.*, **209**, 165–173.
13. Benveniste,R.J. and Taghert,P.H. (1999) Cell type-specific regulatory sequences control expression of the *Drosophila* FMRF-NH2 neuropeptide gene. *J. Neurobiol.*, **38**, 507–520.
14. Rossert,J., Eberspaecher,H. and de Crombrugghe,B. (1995) Separate *cis*-acting DNA elements of the mouse pro-alpha 1(I) collagen promoter direct expression of reporter genes to different type I collagen-producing cells in transgenic mice. *J. Cell Biol.*, **129**, 1421–1432.
15. Rossert,J.A., Chen,S.S., Eberspaecher,H., Smith,C.N. and de Crombrugghe,B. (1996) Identification of a minimal sequence of the mouse pro-alpha 1(I) collagen promoter that confers high-level osteoblast expression in transgenic mice and that binds a protein selectively present in osteoblasts. *Proc. Natl Acad. Sci. USA*, **93**, 1027–1031.
16. Wegner,M. (1999) From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.*, **27**, 1409–1420.
17. Weiss,M.A. (2001) Floppy SOX: mutual induced fit in hmg (high-mobility group) box–DNA recognition. *Mol. Endocrinol.*, **15**, 353–362.
18. Harley,V.R., Lovell-Badge,R. and Goodfellow,P.N. (1994) Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res.*, **22**, 1500–1501.
19. Bell,D.M., Leung,K.K., Wheatley,S.C., Ng,L.J., Zhou,S., Ling,K.W., Sham,M.H., Koopman,P., Tam,P.P. and Cheah,K.S. (1997) SOX9 directly regulates the type-II collagen gene. *Nature Genet.*, **16**, 174–178.
20. Bi,W., Deng,J.M., Zhang,Z., Behringer,R.R. and de Crombrugghe,B. (1999) Sox9 is required for cartilage formation. *Nature Genet.*, **22**, 85–89.
21. Bi,W., Huang,W., Whitworth,D.J., Deng,J.M., Zhang,Z., Behringer,R.R. and de Crombrugghe,B. (2001) Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. *Proc. Natl Acad. Sci. USA*, **98**, 6698–6703.
22. Lefebvre,V., Huang,W., Harley,V.R., Goodfellow,P.N. and de Crombrugghe,B. (1997) SOX9 is a potent activator of the chondrocyte-specific enhancer of the pro alpha1(II) collagen gene. *Mol. Cell. Biol.*, **17**, 2336–2346.
23. Zhao,Q., Eberspaecher,H., Lefebvre,V. and De Crombrugghe,B. (1997) Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis. *Dev. Dyn.*, **209**, 377–386.
24. Bridgewater,L.C., Lefebvre,V. and de Crombrugghe,B. (1998) Chondrocyte-specific enhancer elements in the Col11a2 gene resemble the Col2a1 tissue-specific enhancer. *J. Biol. Chem.*, **273**, 14998–15006.
25. Liu,Y., Li,H., Tanaka,K., Tsumaki,N. and Yamada,Y. (2000) Identification of an enhancer sequence within the first intron required for cartilage-specific transcription of the alpha2(XI) collagen gene. *J. Biol. Chem.*, **275**, 12712–12718.
26. Lefebvre,V., Zhou,G., Mukhopadhyay,K., Smith,C.N., Zhang,Z., Eberspaecher,H., Zhou,X., Sinha,S., Maity,S.N. and de Crombrugghe,B. (1996) An 18-base-pair sequence in the mouse proalpha1(II) collagen gene is sufficient for expression in cartilage and binds nuclear proteins that are selectively expressed in chondrocytes. *Mol. Cell. Biol.*, **16**, 4512–4523.
27. Tsumaki,N., Kimura,T., Matsui,Y., Nakata,K. and Ochi,T. (1996) Separable *cis*-regulatory elements that contribute to tissue- and site-specific alpha 2(XI) collagen gene expression in the embryonic mouse cartilage. *J. Cell Biol.*, **134**, 1573–1582.
28. Mukhopadhyay,K., Lefebvre,V., Zhou,G., Garofalo,S., Kimura,J.H. and de Crombrugghe,B. (1995) Use of a new rat chondrosarcoma cell line to delineate a 119-base pair chondrocyte-specific enhancer element and to define active promoter segments in the mouse pro-alpha 1(II) collagen gene. *J. Biol. Chem.*, **270**, 27711–27719.
29. Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.
30. Zhou,G., Garofalo,S., Mukhopadhyay,K., Lefebvre,V., Smith,C.N., Eberspaecher,H. and de Crombrugghe,B. (1995) A 182 bp fragment of the mouse pro alpha 1(II) collagen gene is sufficient to direct chondrocyte expression in transgenic mice. *J. Cell Sci.*, **108**, 3677–3684.
31. Zhou,G., Lefebvre,V., Zhang,Z., Eberspaecher,H. and de Crombrugghe,B. (1998) Three high mobility group-like sequences within a 48-base pair enhancer of the Col2a1 gene are required for cartilage-specific expression *in vivo*. *J. Biol. Chem.*, **273**, 14989–14997.
32. Peirano,R.I. and Wegner,M. (2000) The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences. *Nucleic Acids Res.*, **28**, 3047–3055.
33. Lefebvre,V., Behringer,R.R. and de Crombrugghe,B. (2001) L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. *Osteoarthritis Cartilage*, **9** (Suppl. A), S69–S75.
34. Lefebvre,V., Li,P. and de Crombrugghe,B. (1998) A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *EMBO J.*, **17**, 5718–5733.
35. Kamachi,Y., Cheah,K.S. and Kondoh,H. (1999) Mechanism of regulatory target selection by the SOX high-mobility-group domain proteins as revealed by comparison of SOX1/2/3 and SOX9. *Mol. Cell. Biol.*, **19**, 107–120.
36. Ambrosetti,D.C., Basilico,C. and Dailey,L. (1997) Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein–protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.*, **17**, 6321–6329.
37. Kamachi,Y., Uchikawa,M. and Kondoh,H. (2000) Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet.*, **16**, 182–187.
38. Marshall,O.J. and Harley,V.R. (2001) Identification of an interaction between SOX9 and HSP70. *FEBS Lett.*, **496**, 75–80.