⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 40, pp. 30496–30505, October 1, 2010

Transcriptional Regulation of an Axonemal Central Apparatus Gene, Sperm-associated Antigen 6, by a SRY-related High Mobility Group Transcription Factor, S-SOX5*[S]

Received for publication, March 9, 2010, and in revised form, July 15, 2010 Published, JBC Papers in Press, July 28, 2010, DOI 10.1074/jbc.M110.121590

Elizabeth Anne Kiselak[‡1], Xuening Shen[‡], Jingmei Song[‡], David Roberto Gude[‡2], Jiannan Wang[‡], Steven L. Brody[§], Jerome F. Strauss III[‡], and Zhibing Zhang[‡3]

From the [‡]Department of Obstetrics and Gynecology, Virginia Commonwealth University, Richmond, Virginia 23298 and the [§]Department of Internal Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

SOX5 is a transcription factor with homology to the high mobility group box region of the testis-determining factor, SRY. Both the mouse and human *SOX5* genes encode a 48-kDa SOX5 protein (S-SOX5) that is only present in tissues containing cells with motile cilia/flagella. The mammalian sperm-associated antigen 6 gene (*SPAG6*) encodes an axoneme central apparatus protein. Because human and mouse *SPAG6* gene promoters contain multiple potential binding sites for SOX5, *SPAG6* gene regulation by S-SOX5 was investigated in BEAS-2B cells, a line derived from human bronchial cells. Like FOXJ1, a transcription factor known to be essential for motile ciliogenesis, S-SOX5 stimulated mouse and human *SPAG6* promoter function in BEAS-2B cells, but the effect was abrogated when the SOX5 binding sites were mutated or deleted. S-SOX5 and FOXJ1 functioned cooperatively in stimulating *SPAG6* promoter activity. The *SPAG6* message was up-regulated when S-SOX5 was overexpressed in BEAS-2B cells, and silencing of *S-SOX5* by RNA interference down-regulated *SPAG6* transcripts. Chromatin immunoprecipitation and EMSA experiments demonstrated that S-SOX5 associates with the *SPAG6* promoter directly. The present study demonstrates that *SPAG6* is a S-SOX5 target gene, indicating a key role for S-SOX5 in the formation and function of motile cilia.

SOX proteins are transcription factors with a high motility group box DNA binding domain similar to that of the sex-determining region (SRY) protein (1, 2). Based on phylogenetic analysis of their high motility group domains, *SOX* genes can be separated into 10 groups, A to J (3). Many genes within each subgroup also share conserved regions outside the high motility group domain. SOX proteins activate or repress target genes by binding to specific DNA sequences (4, 5). They regulate diverse developmental processes, including lens, muscle, blood vessel, hair follicle, gut, B cell, and cartilage development (6–13). *SOX* genes are expressed in many tissues and are implicated in the etiology of many diseases (14).

SOX5 is a member of the *SOXD* group, which includes three genes, *SOX5*, *SOX6*, and *SOX13* (15). Mouse *Sox5* is expressed as a short transcript (2 kb) in adult testis (16) and a longer transcript (6 kb) in other tissues (17). The short transcript encodes a 48-kDa protein isoform that lacks the N-terminal half of the larger protein encoded by the 6-kb transcript. The former protein was the first to be discovered and was named SOX5 (called S-SOX5 in this paper). The longer SOX5 isoform was originally named L-SOX5 (18), but most authors refer to this isoform as SOX5. L-SOX5 is highly expressed in chondrocytes and striated muscles, indicating a likely role in human cartilage and muscle development (19, 20). The phenotype of *L-Sox5* knock-out mice demonstrated that L-SOX5 is essential for cartilage formation and notochord extracellular matrix sheath formation, notochord cell survival, and development of the nucleus pulposus of intervertebral discs (21–23).

The different tissue distribution of S-SOX5 suggests that it plays roles that are distinct from those of L-SOX5. It has been reported that S-SOX5 regulates genes expressed in the testis including the testicular isoform of hormone-sensitive lipase, IκBβ, and ZNF230 (24–26). Besides the testis, S-SOX5 has been reported to be highly expressed in human brain (27). The high abundance of S-SOX5 in testis and brain, tissues that contain motile cilia/flagella, suggests that S-SOX5 might play a role in ciliogenesis and/or spermatogenesis. To test this idea, we selected a target gene, sperm-associated antigen 6 (*SPAG6*), that is representative of "9 + 2" cilia, encoding a protein in the axoneme central pair.

Mammalian sperm-associated antigen 6 is the orthologue of *Chlamydomonas* PF16. *SPAG6* was first identified in a human testis cDNA expression library (28). Mouse *Spag6* was subsequently cloned from a library generated from mixed germ cells (29). Both human and mouse *SPAG6* genes each encode 1.8- and 2.8-kb mRNAs highly expressed in testis. In sperm, SPAG6 protein is present in the sperm tail, and immunoelectron microscopy demonstrated that SPAG6 is localized to the central apparatus of the axoneme, consistent with its localization in the *Chlamydomonas* flagella (30). In addition to testis and sperm, SPAG6 is also expressed in tissues that contain cells bearing 9 + 2 motile cilia, such as brain and lung (28, 31–33). As

* This work was supported, in whole or in part, by National Institutes of Health Grant HD37416 (to J. F. S. III).

⌘ *Author's Choice*—Final version full access.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–6 and Table 1.

[1] This research was submitted in fulfillment of a Master's degree research project at Virginia Commonwealth University.

[2] Supported by National Institutes of Health F31 Predoctoral Fellowship IF3IHD0G2314-01 through the NICHD.

[3] To whom correspondence should be addressed: 11028A Sanger Hall, 1101 E. Marshall St., Richmond, VA 23298. Fax: 804-928-0573; E-mail: zzhang4@vcu.edu.

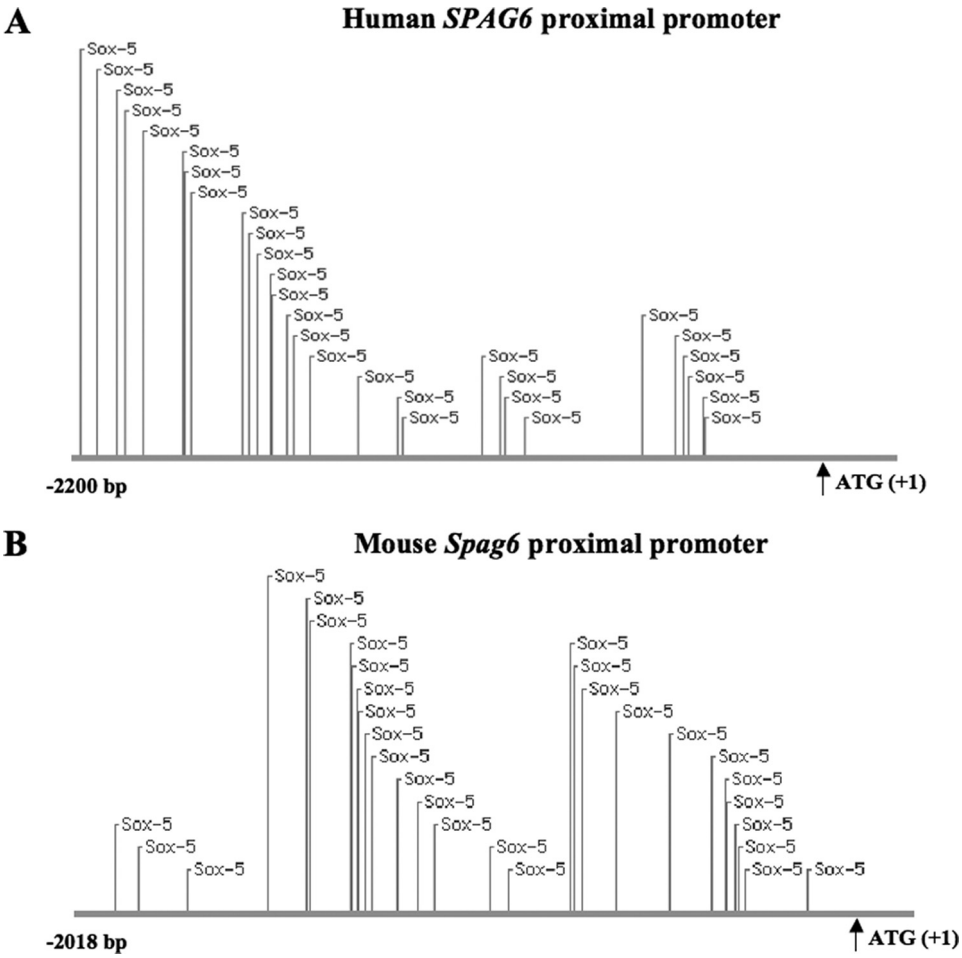

FIGURE 1. **Analysis of potential SOX5 binding sites in the human and mouse *SPAG6* gene proximal promoter regions.** 2-kb human *SPAG6* (*A*) and mouse *Spag6* (*B*) proximal promoter regions were scanned with ConSite for potential SOX5 binding sites. Translational start sites are indicated by *arrows*.

in *Chlamydomonas* and *Trypanosome brucei* (34–37), SPAG6 is essential for flagellar motility and maintenance of the structure of the axoneme of mature mammalian sperm (30).

SPAG6 protein contains eight contiguous armadillo repeats, which places it in a family of proteins known to mediate protein-protein interactions. It has been shown that SPAG6 associates with SPAG16L and SPAG17, two other proteins localized to the axoneme central apparatus (32, 39, 40). In *Spag6* mutant mice, both SPAG16L and SPAG17 were missing from sperm (32, 39, 40). These data imply that SPAG6 is a critical protein in either the assembly or structural integrity of the sperm tail axoneme.

Using a bioinformatics approach, we identified potential transcription factors that might bind to the mouse and human *SPAG6* proximal promoter regions and regulate *SPAG6* gene expression. This analysis revealed multiple potential binding sites for SOX5, a transcription factor that has not been previously implicated in the regulation of genes encoding ciliary/flagellar proteins. We therefore investigated the role of S-SOX5 in the regulation *SPAG6* gene expression.

EXPERIMENTAL PROCEDURES

5′-Rapid Amplification of cDNA Ends

5′-Rapid amplification of cDNA ends was conducted to define the 5′-untranslated region (UTR) of *SPAG6* mRNA and transcriptional start sites using a Human Marathon cDNA Amplification kit (Clontech) according to the manufacturer's instructions.

DNA Constructs

Human SPAG6 Promoter Constructs—1.5 kb of the human *SPAG6* promoter that contains multiple putative SOX5 binding sites was amplified and cloned into KpnI/NheI sites of the PGL3 basic vector. A 500-bp *SPAG6* promoter sequence containing one putative SOX5 binding site was also cloned into KpnI/NheI sites of the PGL3 basic vector.

Mouse Spag6 Promoter Constructs—A 1.5-kb mouse *Spag6* promoter fragment was cloned in PGL3 basic vector as previously described (31). A 250-bp mouse *Spag6* promoter sequence containing two putative SOX5 binding sites, including the transcriptional start site, was cloned into KpnI/XhoI sites of the PGL3 basic vector.

Site-directed Mutagenesis of SOX5 Binding Sites in SPAG6 Promoter Constructs—SOX5 binding sites in the 500-bp human and the 250-bp mouse *SPAG6* promoter constructs were mutated or deleted using a QuikChange XL Site-directed Mutagenesis kit (Stratagene, La Jolla, CA), and the mutations/deletions were confirmed by DNA sequencing.

Mouse and Human S-SOX5 Expression Constructs—Full-length mouse *S-Sox5* cDNA was amplified and cloned into the BamHI/SalI site of the pTarget (pT) vector (Promega). *S*-Sox5/pTarget plasmid was digested with KpnI and XhoI restriction enzymes, and the *S-Sox5* cDNA was subcloned to the KpnI/XhoI site of the pcDNA3 mammalian expression vector (Invitrogen). pcDNA3 carries a neomycin resistance cassette, which was used for G418 selection to generate stable BEAS-2B cell lines for use in subsequent experiments. A mouse *S-Sox5* expression plasmid with a FLAG tag fusion in the N terminus was also generated. The *S-Sox5* cDNA was cloned into the SalI/KpnI sites of the pRK-7 vector (kindly provided by Dr. Wenhui Hu, Department of Physiology, Virginia Commonwealth University). A Foxj1 expression plasmid in pcDNA3 was previously described (41).

To create a human *S-SOX5*/pTarget construct, full-length human *S-SOX5* cDNA was amplified from human testis cDNA (Clontech) and subcloned into the BamHI/NotI sites of pTarget vector.

Human SOX5 RNAi Constructs—RNAi constructs targeting human *SOX5* transcript were generated with the iLenti^{TM} siRNA Expression System (Capital Biosciences, Inc., Washington, D. C.).

The target sequences were identified through BLOCK-iT™ RNAi Designer (Invitrogen). Three sets of oligonucleotides were designed according to the user's manual. Double-stranded oligonucleotides (see supplementary Table 1) were generated, phosphorylated, and then ligated to linearized iLenti™ vector. The ligation products were used for transformation, and positive clones were identified by restriction digestion and sequencing.

Western Blot Analysis

Equal amounts of protein (50 μg/lane) were heated to 95 °C for 10 min in 4× sample buffer, loaded onto 10% SDS-polyacrylamide gels, electrophoretically separated, and transferred to polyvinylidene difluoride membranes. The membranes were blocked and then incubated with antibodies against SOX5 (Aviva Systems Biology, San Diego, CA) or GAPDH (Santa Cruz Biotechnology, Santa Cruz, CA) overnight at 4 °C. After washing, the blots were incubated with an anti-rabbit immunoglobulin conjugated to horseradish peroxidase for 1 h at room temperature. SOX5 or GAPDH protein was detected with the Super Signal Pico Chemiluminescent or Femto Maximum Sensitive system (Pierce).

Immunofluorescence Staining

Human bronchial BEAS-2B cells (American Type Culture Collection (ATCC), Rockville, MD) were seeded into two-chamber slides (BD Biosciences) and transfected with pTarget (control) and S-SOX5/pTarget. Two days after transfection, cells were washed with PBS and fixed in 100% methanol at −80 °C for 7 min. Cells were blocked in PBS containing 10% goat serum for 30 min at room temperature followed by overnight incubation with the polyclonal anti-rabbit SOX5 primary antibody at 4 °C. After washing cells three times with PBS, the secondary antibody, Alexa Fluor 488 anti-rabbit IgG (Invitrogen), was applied in 10% goat serum for 1 h at room temperature. After washing three more times with PBS, the slides were mounted and sealed with nail polish. SOX5 staining was visualized using a Nikon Eclipse TS100 phase contrast/fluorescent microscope; images were obtained with a Nikon Coolpix 4500 digital camera.

Generation of S-SOX5 Stable Cell Lines

BEAS-2B cells were seeded at 1.5 × 10^6 cells/100-mm dish and the next day transfected with 6 μg/dish pcDNA3 or *S-SOX5*/pcDNA3 plasmids. 24 h after transfection, the cells were supplied with fresh medium containing 1.5 mg/ml G418. Resistant clones were selected within 15–20 days and expanded. The G418 concentration was maintained at 1 mg/ml in the cell culture medium (42).

Generation of Adenovirus-expressing S-SOX5

An adenovirus (Ad)[4]-expressing mouse S-SOX5 (AdS-SOX5) was generated with AdEasy™ Adenoviral Vector system (catalogue no. 240010; Stratagene) following the instruction manual. Briefly, mouse *S-Sox5* cDNA was subcloned into adenovirus shuttle vector, pShuttle-CMV, and the cDNA was transferred into the pAdEasy-1 virus genome by homologous recombination in Ad packaging HEK-293 cells. The control virus was made with empty pShuttle-CMV plasmid. The expression of S-SOX5 was tested by Western blot analysis of protein from COS-1 cells and BEAS-2B cells infected with AdS-SOX5 or control Ad virus. The AdS-SOX5 and control Ad virus were amplified in the Virus Facility of Virginia Commonwealth University.

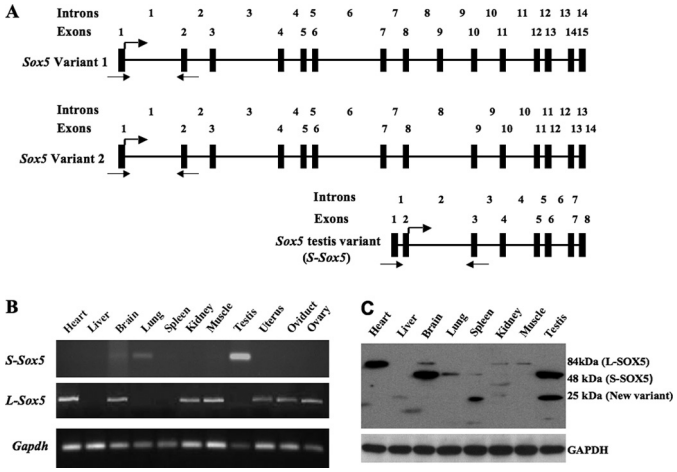

FIGURE 2. **Structure of mouse *Sox5* gene and expression pattern of SOX5 *in vivo*.** *A*, structures of *Sox5* transcripts. The primer sets for amplifying individual *Sox5* transcripts and translational start sites are indicated. *B*, analyses of tissue distribution of indicated *Sox5* transcripts by RT-PCR using primer sets shown in *A*. *C*, analysis of SOX5 protein expression in the indicated tissues. The antibody recognizes both the 48-kDa and the 84-kDa SOX5 proteins.

RT-PCR and Real-time PCR

Total RNA was isolated from cultured cells and indicated tissues with TRIzol (Invitrogen), the RNA was reversed transcribed, and the cDNAs were used for RT-PCR or real-time PCR. For real-time PCR, the cDNAs from BEAS-2B cells transfected with pcDNA3 or *S-Sox5*/pcDNA3 plasmids, or *SOX5* RNAi plasmid DNA, infected with AdS-SOX5 or control Ad virus, were utilized for PCR. 18 S rRNA was amplified simultaneously to normalize the expression level of target genes. The primers were synthesized according to the published sequences (43). Real-time PCRs were carried out using 2× SYBR Green master mix (Bio-Rad).

Transient Transfection and Promoter Activity Assays

The promoter activity was assessed using BEAS-2B human bronchial epithelial cells and MDCK cells (ATCC). Cells were maintained in BEBM (Cambrex, for BEAS-2B cells) or DMEM with 10% fetal calf serum (for MDCK cells). 24 h prior to transfection, 3 × 10^4 cells were plated in 12-well plates. Triplicate wells were transfected with indicated plasmids using FuGENE 6 (Roche Diagnostics). In addition, 25 ng of the *Renilla* luciferase plasmid was transfected into each well to evaluate transfection efficiency. 48 h after transfection, the cells were harvested, and the promoter activity was measured with the Dual Luciferase Reporter Assay system (Promega).

Chromatin Immunoprecipitation (ChIP) Assay

ChIP assay was performed using a ChIP assay kit (Upstate Millipore) following the manufacturer's instructions. Briefly,

[4] The abbreviations used are: Ad, adenovirus; ChIP, chromatin immunoprecipitation; MDCK, Madin-Darby canine kidney; RT, reverse transcription.

the BEAS-2B cells were infected with AdS-SOX5. 48 h after infection, the proteins bound to DNA were cross-linked with 1% formaldehyde; cells were subsequently resuspended in the lysis buffer, followed by sonication. After preclearing with salmon sperm DNA/protein A-agarose (50% slurry), the protein extracts were subjected to immunoprecipitation with antibody against SOX5 or normal rabbit IgG. Immunocomplexes were extensively washed, and the DNA was recovered and resuspended in 50 μl of 10 mM Tris-HCl (pH 8.5). The DNA was used as template for PCR amplification with three sets of primers flanking SOX5 binding sites. As a negative control, PCR was carried out with a primer set 3.5 kb upstream of the transcription start site where no SOX5 binding site is nearby. The experiment was repeated three times. The primer sequences are seen in the supplementary Table 1.

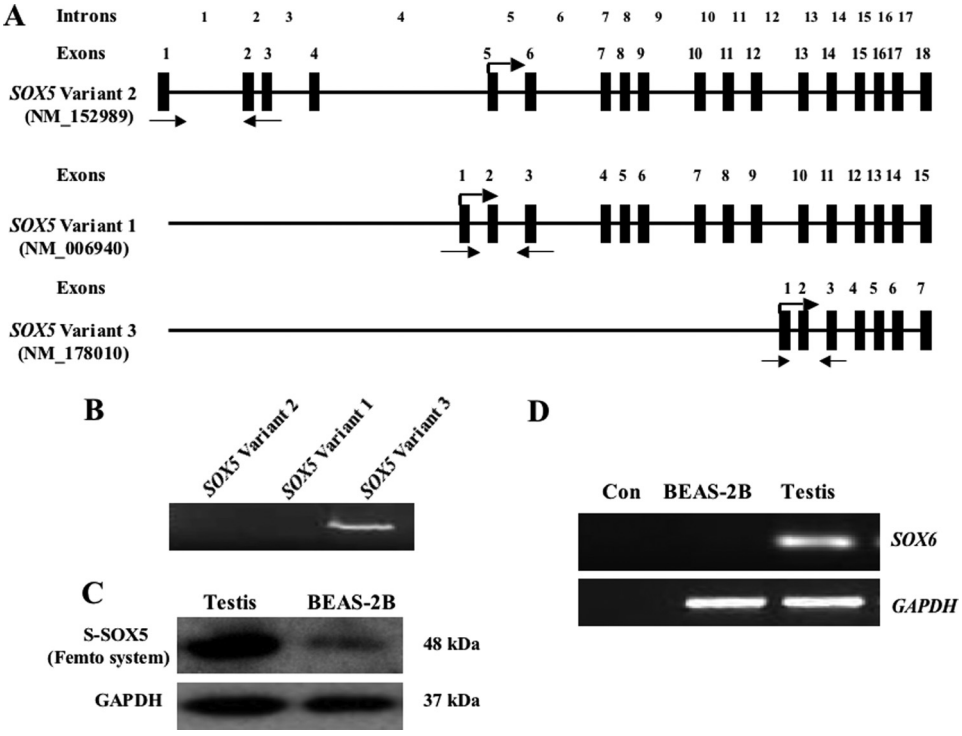

FIGURE 3. **Structures of human *SOX5* variants and analysis of *SOX5* expression in BEAS-2B cells.** *A*, structures of human *SOX5* variants. The translational start sites and primers for amplifying individual transcript are indicated. *B*, RT-PCR amplification of the three *SOX5* transcripts with specific primers shown in *A* with RNA isolated from BEAS-2B cells. *C*, Western blot analysis of S-SOX5 protein expression in mouse testis and BEAS-2B cells. Note that only 48-kDa S-SOX5 protein was detected when the Femto system was used. *D*, RT-PCR amplification of *SOX6* transcripts in BEAS-2B cells and human testis. *Con*, no template control.

Purification of FLAG-tagged S-SOX5 Protein and Electrophoretic Mobility Shift Assay (EMSA)

To purify FLAG-tagged S-SOX5 protein, the plasmid containing FLAG-S-SOX5 cDNA was transfected into 293T cells, 48 h after transfection, the protein was purified with EZviewRed Anti-FLAG M2 Affinity Gel system (Sigma-Aldrich) following the manufacturer's manual. For EMSA experiments, six double-stranded synthetic oligonucleotides were made; three contained putative SOX5 binding sites, and the other three had the same sequences as the previous ones except that five nucleotides in the SOX5 binding sites were mutated. The double-stranded synthetic oligonucleotides were labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP. The EMSA binding reaction was mixed in 2× binding buffer (Promega) with up to 4 μg of purified S-SOX5 protein and 1 × 10^5 cpm of ^{32}P-labeled double-stranded oligonucleotide probes (1 ng). Reaction mixtures were incu-

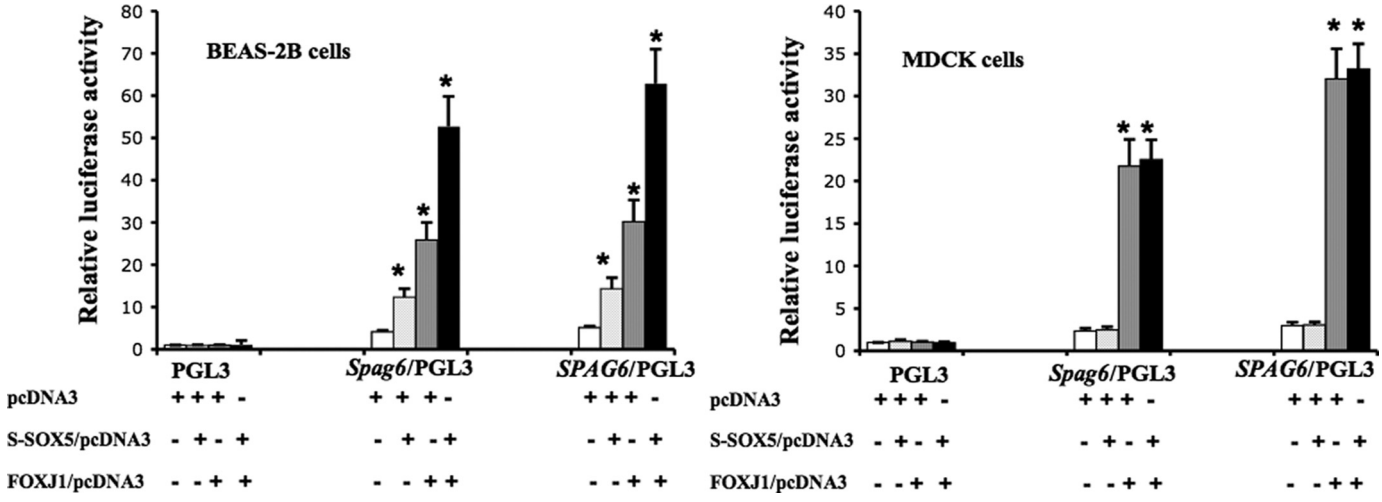

FIGURE 4. **Effect of SOX5 and FOXJ1 on human and mouse *SPAG6* promoter function in BEAS-2B and MDCK cells.** 1.5-kb human and mouse *SPAG6* promoter constructs were transfected to BEAS-2B or MDCK cells. The cells were also transfected with S-SOX5/pcDNA3, or FOXJ1/pcDNA3, or a combination of the two plasmids simultaneously. 48 h after transfection, relative luciferase activity was analyzed. Values given are means ± S.E. (*error bars*) from three separate experiments. *, $p < 0.05$ compared with controls.

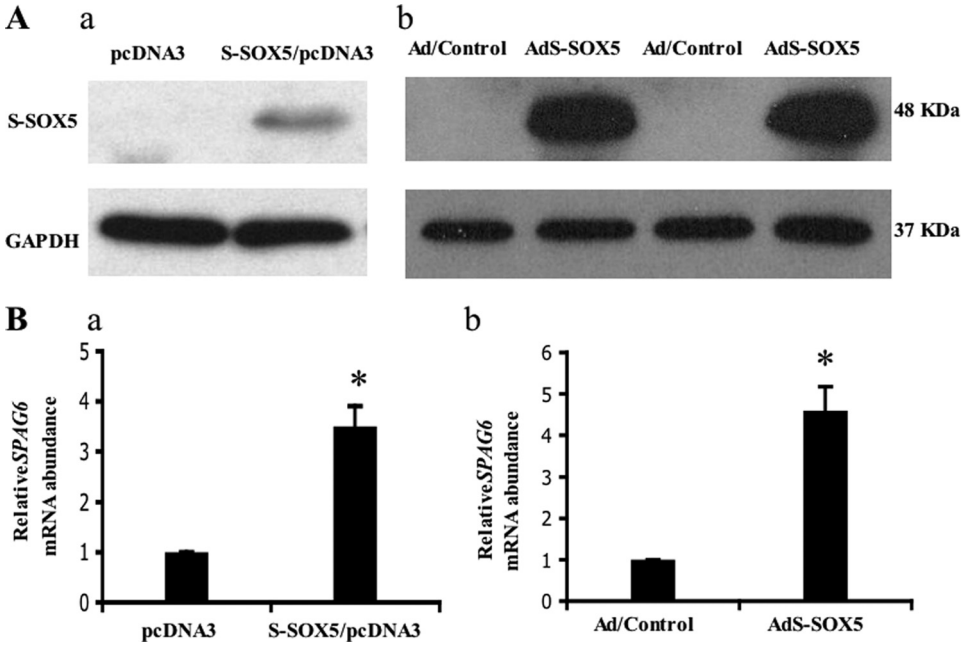

FIGURE 5. **Exogenous S-SOX5 up-regulates *SPAG6* mRNA expression in BEAS-2B cells.** *A*, Western blot analysis of S-SOX5 protein expression in BEAS-2B cells stably expressing S-SOX5 (*a*) or infected with AdS-SOX5 (*b*) detected with the Pico system. *B*, analysis of *SPAG6* mRNA expression by real-time PCR. Total RNA was isolated from BEAS-2B cells stably expressing S-SOX5 or control cells transfected with pcDNA3 only (*a*) or infected with AdS-SOX5 or Ad/control (*b*). Real-time PCR was performed with a primer set specifically targeting *SPAG6* message. 18 S rRNA was also amplified simultaneously to normalize the results. Values are mean ± S.E. (*error bars*). *, $p < 0.05$ compared with controls.

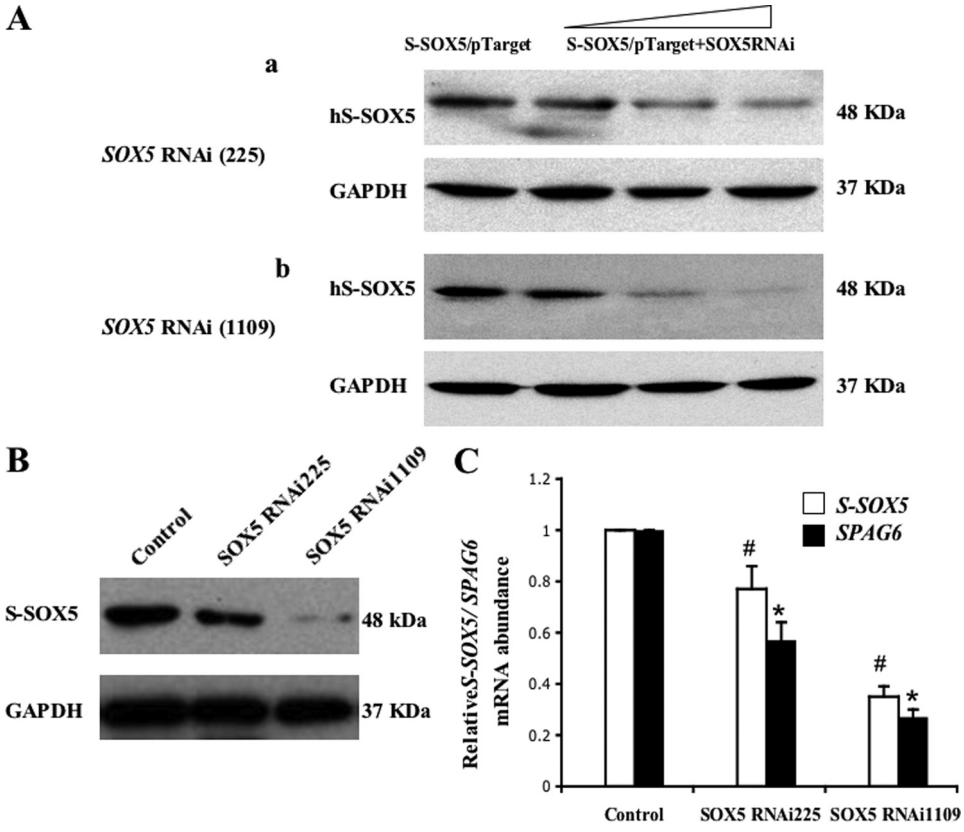

FIGURE 6. **Knockdown of *SOX5* in BEAS-2B cells by RNA interference results in down-regulation of *SPAG6* message.** *A*, Western blot analysis of S-SOX5 protein expression in transfected COS-1 cells after RNAi treatment with Pico system. *a*, *SOX5* RNAi 225 treatment; *b*, *SOX5* RNAi 1109 treatment. *B*, Western blot analysis of endogenous S-SOX5 protein expression in BEAS-2B cells stably expressing the two RNAi constructs with the Femto system. *C*, analysis of *S-SOX5* and *SPAG6* mRNA expression by real-time PCR after *SOX5* RNAi treatment. Values are mean ± S.E. (*error bars*) from three separate experiments. *, $p < 0.05$ compared with controls for S-SOX5 or SPAG6 expression.

bated at room temperature for 30 min and then subjected to an 8% polyacrylamide gel at 250 V for 3 h. The dried gels were then exposed to x-ray films.

RESULTS

Identification of the 5′-UTR of Human SPAG6 Gene—To define the 5′-UTR of human *SPAG6* mRNA and identify transcriptional start sites, 5′-rapid amplification of cDNA ends was performed with a reverse primer located close to the 5′ end of human *SPAG6* mRNA. PCR products were amplified and subcloned into the pCR2.1 TOPO TA vector, and then 15 clones were subjected to DNA sequencing. The 5′-UTR sequence was submitted to GenBank (EF432114). The 5′-UTR is GC-rich, and four potential transcription start sites were identified based on the sequencing results as noted in GenBank.

Both Human and Mouse SPAG6 Promoters Contain Multiple Potential SOX5 Binding Sites—To study transcription regulation of *SPAG6* gene expression further, putative transcription factor binding sites in the proximal promoters of the mouse (AY792595) and human (EF432114) genes were analyzed using ConSite, open source software that determines high probability binding sites for known transcription factors. Multiple putative binding sites for SOX5, an SRY-related transcription factor with a high mobility group DNA binding domain, were identified in both promoter sequences with the default setting, the TF score cutoff being 80% (Fig. 1), these are putative SOX binding sites, and it must be recognized that the program cannot identify those sites which uniquely bind SOX5.

S-Sox5 Transcripts Are Abundant in Mouse Tissues Containing Cells with Motile Cilia—Three mouse *Sox5* transcript variants were identified in GenBank. Transcript variant 1 has 15 exons, and transcript variant 2 is almost identical to transcript variant 1, except that it lacks exon 9 of transcript variant 1. These

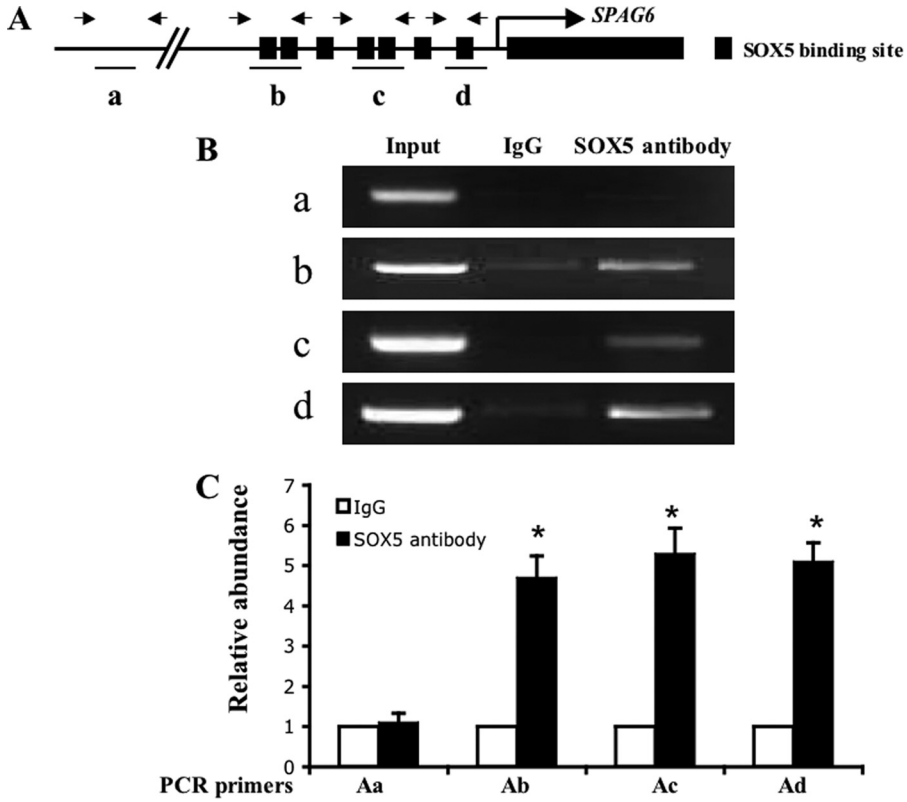

FIGURE 7. **S-SOX5 associates with the human *SPAG6* promoter as revealed by ChIP assay.** *A*, diagram showing the human proximal *SPAG6* promoter and the regions (*a*, *b*, *c*, and *d*) amplified by ChIP primers used in this study. *Arrows* show the location of the primers. *B*, representative ChIP assay results with BEAS-2B cells infected by AdS-SOX5 using a control rabbit IgG or an antibody specifically against SOX5. *a*, with a primer set not flanking any potential SOX5 binding sites; *b–d*, with primer sets flanking potential SOX5 binding sites. *C*, quantitation of ChIP assay results by real-time PCR. Data shown are means ± S.E. (*error bars*) from three separate experiments. *, $p < 0.05$ compared with normal rabbit IgG pulldown group.

two transcripts translate the L-SOX5 protein. The third transcript variant has only eight exons, the first of which is not present in transcript variants 1 and 2, whereas exons 2–8 are identical to exons 8–14 of transcript variant 2 (Fig. 2*A*). This third variant has been reported to be highly expressed in the testis (16) and translates a truncated 48-kDa SOX5 protein (S-SOX5). The unique exon 1 of this variant allowed us to design a specific primer set to amplify this variant. RT-PCR result revealed that this transcript ("testis variant") is only expressed in the testis, lung, and brain (Fig. 2*B*, *top panel*), all tissues that contain motile cilia.

The tissue distribution of *Sox5* encoding L-SOX5 protein was also examined with specific primers (Fig. 2*B*). This transcript is present in heart, brain, kidney, skeletal muscle, uterus, oviduct and ovary, but not in the testis and lung (Fig. 2*B*, *middle panel*). Western blot analysis was conducted with an antibody that recognizes both SOX5 proteins. The 48-kDa S-SOX5 is expressed in the testis, brain, and lung. The 84-kDa L-SOX5 is present in heart, brain, kidney, and skeletal muscle (Fig. 2*C*). Besides the 48-kDa and 84-kDa proteins, the antibody also recognized a 25-kDa protein in some tissues, including liver, spleen, and testis (Fig. 2*C*), which probably represents the translated protein from an additional *Sox5* transcript that was reported in the Ensembl database (ENSMUST00000111746).

S-SOX5 Is Expressed in BEAS-2B Cells—Multiple human *SOX5* transcript variants were identified in databases (Fig. 3*A*). Transcript variants 1 (NM_006940) and 2 (NM_152989) translate into full-length L-SOX5. Transcript variant 3 (NM_178010) translates into a truncated SOX5 protein, and this transcript variant is similar to the mouse *S-Sox5* transcript.

RT-PCR and Western blot analysis were used to investigate endogenous expression of SOX5 in BEAS-2B cells. A single PCR product was amplified with the primer set specific for human *S-SOX5* cDNA (transcript variant 3, Fig. 3*B*). However, no PCR product was amplified with primer sets that are specific for L-SOX5 messages (transcript variants 1 and 2, Fig. 3*B*), indicating that *S-SOX5* message is the predominant isoform in BEAS-2B cells. Western blot analysis was conducted with total lysates from BEAS-2B cells. The 48-kDa S-SOX5 was detected with the Femto Maximum Sensitive substrate system (Fig. 3*C*). The expression of *SOX6*, another member in *SOXD* group, in BEAS-2B cells was also investigated. Even though *SOX6* is in the same group as *SOX5*, and like S-SOX5, it is expressed in the testis, the message was not detected in BEAS-2B cells (Fig. 3*D*).

S-SOX5 Is Localized in the Nuclei in Transfected BEAS-2B Cells—To study the function of S-SOX5 protein, full-length mouse *Sox5* cDNA was subcloned into the mammalian expression vector pTarget. To test the construct, *S-Sox5*/pTarget or empty vector was transfected into COS-1 cells, and Western blot analysis was performed. The 48-kDa S-SOX5 protein was detected only in the cells transfected with *S-Sox5*/pTarget, but not cells transfected with control empty pTarget plasmid when the Pico Chemiluminescent substrate system was used (supplemental Fig. 1*A*). Cellular localization of S-SOX5 following *S-Sox5*/pTarget transfection into BEAS-2B cells was investigated using immunofluorescent staining with a specific antibody against S-SOX5. S-SOX5 was detected only in the nuclei, whereas normal rabbit IgG showed no signal (supplemental Fig. 1*B*).

S-SOX5 Stimulates Both Human and Mouse SPAG6 Promoter Function—To study the effect of S-SOX5 on *SPAG6* gene transcription, the 1.5-kb human *SPAG6* gene promoter, including the transcription start site, was amplified and subcloned into PGL3 basic vector. The S-SOX5 expression plasmids and human or mouse *SPAG6* promoter constructs were co-transfected into bronchial epithelial cell-derived BEAS-2B or renal epithelial cell-derived MDCK cells, and relative luciferase activity was analyzed. In BEAS-2B cells, both the human and mouse *SPAG6* promoters displayed basal activity, which was about

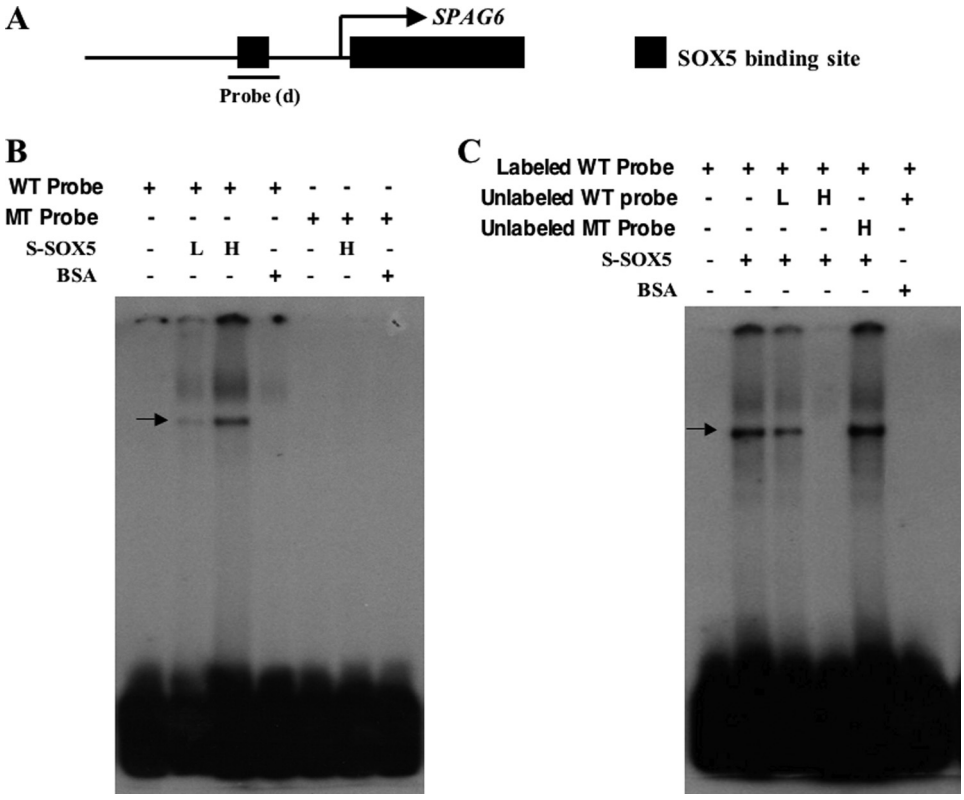

FIGURE 8. **S-SOX5 directly binds to a probe that contains a SOX5 binding site in the human *SPAG6* proximal promoter.** *A*, schematic representation of the human *SPAG6* promoter showing the location of the probe containing a SOX5 binding site used in this study. The sequences of the putative SOX5 binding sites are indicated in supplemental Table 1. *B*, EMSA using purified S-SOX5 protein and a probe containing SOX5 binding site (*WT Probe*) or SOX5 binding site mutated probe (*MT Probe*). *L*, 1 μg of S-SOX5 protein; *H*, 4 μg of S-SOX5 protein. The purified S-SOX5 protein was incubated with the indicated labeled double-stranded oligonucleotide probes, then subsequently applied to 8% polyacrylamide gels. *C*, competition of labeled WT probe with unlabeled WT and unlabeled MT probes. *L*, 50-fold unlabeled probe; *H*, 250-fold unlabeled probe.

4-fold higher than the empty PGL3 basic vector. Co-transfection of S-SOX5, increased relative luciferase activity an additional 3-fold only in BEAS-2B cells (Fig. 4). In MDCK cells, even though both human and mouse *Spag6* promoter activity was detected, expression of S-SOX5 had no effect on the function of both promoters (Fig. 4).

FOXJ1, a transcription factor that is known to be essential for motile ciliogenesis, also activated *SPAG6* promoter function in BEAS-2B cells. The relative luciferase activity was 7-fold greater compared with control (baseline *SPAG6* promoter activity). The combination of S-SOX5 and FOXJ1 resulted in a 13-fold relative luciferase activity in BEAS-2B cells. There was no additional effect of co-transfection of S-SOX5 and FOXJ1 in MDCK cells.

Forced Expression of S-SOX5 Increases SPAG6 mRNA Expression in BEAS-2B Cells—S-SOX5 adenovirus and BEAS-2B cells stably expressing S-SOX5 were generated, and the effect of S-SOX5 on *SPAG6* mRNA expression was investigated. To confirm S-SOX5 protein expression in the BEAS-2B/S-SOX5 stable cell line or BEAS-2B cells infected with ADS-SOX5, Western analysis was performed. The 48-kDa S-SOX5 protein was highly expressed in both cell types, but it was not detected in the control cells with the Pico system (Fig. 5*A*). Real-time PCR was conducted to compare relative *SPAG6* mRNA abundance in the same cells. As expected, in both cell types, relative *SPAG6* mRNA level was significantly higher than in controls (Fig. 5*B*).

Knockdown of S-SOX5 in BEAS-2B Cells Decreases SPAG6 mRNA Expression—To study further the regulation of *SPAG6* by S-SOX5, three human *SOX5* RNAi plasmid constructs were generated; the targeting positions were 167–188, 225–246, and 1109–1130 of *SOX5* transcript variant 3 (NM_178010), respectively. To test the efficiency of these constructs, COS-1 cells were co-transfected with these constructs and *S-Sox5*/pTarget and subjected to Western blot analysis. All three of the RNAi constructs effectively reduced S-SOX5 protein expression, especially the constructs targeting 225–246 and 1109–1130 (Fig. 6*A*). BEAS-2B stable cell lines were generated with these RNAi and control constructs, and real-time PCR was conducted to analyze *S-SOX5* and *SPAG6* mRNA expression. The S-SOX5 protein expression was reduced in the BEAS-2B cells stably expressing the RNAi plasmids (Fig. 6*B*), and the level of reduction was consistent with the reduction of protein level in COS-1 cells (Fig. 6*A*). *S-SOX5* mRNA levels were also reduced in the two specific RNAi cell lines. *SPAG6* mRNA levels were correspondingly reduced in the stable cell lines with the two specific RNAi constructs (Fig. 6*C*).

S-SOX5 Binds Directly to the SPAG6 Promoter—To determine whether S-SOX5 associates with the endogenous *SPAG6* promoter, BEAS-2B cells were infected with AdS-SOX5. 48 h after infection, ChIP assay was performed with an antibody against S-SOX5. All three PCR primer sets that flank potential SOX5 binding sites amplified the DNA fragments only when ChIP was conducted with antibody against S-SOX5 (Fig. 7*A*, *b–d*). The relative abundance of the PCR product was about 5 times higher than in the control (Fig. 7*B*). In contrast, the primer set located 3.5 kb upstream of the transcription start site failed to amplify specific DNA fragments (Fig. 7*Aa*).

To confirm the association of the S-SOX5 with the *SPAG6* promoter region further, EMSAs were performed. S-SOX5 protein was purified from transfected 293T cells (supplemental Fig. 2). One human and two mouse probes that contain putative SOX5 binding sites, and the three probes but with SOX5 binding sites mutated were labeled with [γ-^{32}P]ATP for EMSA experiments, the sequences of these putative SOX5 binding sites are indicated in supplemental Table 1. The locations of the probes are shown in Figs. 8*A* and 9*A*. S-SOX5 bound specifically to all the three probes with SOX5 binding sites,

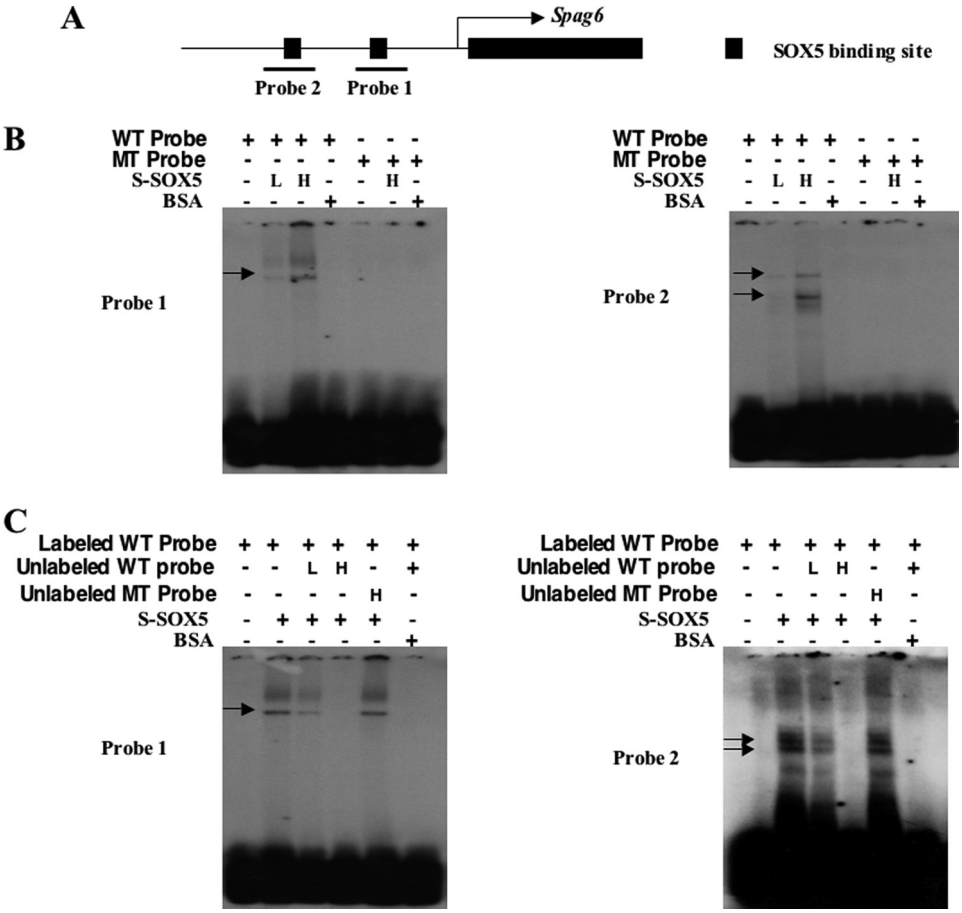

FIGURE 9. **S-SOX5 directly binds to probes that contain SOX5 binding sites in the mouse *Spag6* proximal promoter.** *A*, schematic representation of the mouse *Spag6* promoter showing the location of the two probes containing SOX5 binding sites used in this study. The sequences of the putative SOX5 binding sites are indicated in supplemental Table 1. *B*, EMSA using purified S-SOX5 protein and the two mouse probes containing SOX5 binding sites (*WT Probe*) and SOX5 binding site-mutated probes (*MT Probe*). *L*, 1 μg of S-SOX5 protein; *H*, 4 μg of S-SOX5 protein. *C*, competition of labeled WT probes with unlabeled WT and unlabeled MT probes. *L*, 50-fold unlabeled probes; *H*, 250-fold unlabeled probes.

as revealed by the indicated shifted bands in Figs. 8*B* and 9*B*. The shifted signals represent specific binding of S-SOX5 and the probes because no shifted bands were observed when the probes were incubated with BSA or when the S-SOX5 protein was incubated with the probes with SOX5 binding sites mutated (Figs. 8*B* and 9*B*). The specific signals were eliminated by unlabeled probes containing SOX5 binding sites, but not by probes with mutated SOX5 binding sites (Figs. 8*C* and 9*C*). Besides the three binding sites, S-SOX5 also binds to the two SOX5 binding sites (Fig. 7*A*, *b* and *c* sites) flanked by the ChIP assay primers (supplemental Fig. 3), but S-SOX5 does not bind the probe (Fig. 7*Aa* sites) that does not contain SOX5 binding site (data not shown).

S-SOX5 Fails to Activate SPAG6 Promoter Activity When SOX5 Binding Sites Are Mutated or Deleted—To investigate further whether S-SOX5 activates *SPAG6* promoters through binding to SOX5 binding sites, a 500-bp human *SPAG6* promoter construct containing one SOX5 binding site and a 250-bp *Spag6* promoter construct containing two SOX5 binding sites were generated. Like the 1.5-kb *SPAG6* promoter constructs, S-SOX5 stimulated both promoters. When the SOX5 binding sites were mutated or deleted, the stimulating effect was ablated (Fig. 10 and supplemental Fig. 4).

DISCUSSION

The *SPAG6* gene has been conserved during evolution. The protein plays an indispensable role in the integrity of the axoneme and flagella motility. To study transcriptional regulation of the gene, we characterized the mouse and human *SPAG6* proximal promoters and identified transcription factors that might regulate *SPAG6* gene expression (31). To search for potential transcription factor binding sites, the promoter sequences were analyzed with the ConSite program. It was found that both the human and mouse *SPAG6* promoters have similar potential transcription factor binding sites, including binding sites for CREB/CREM, SOX17, SPZ1, forkhead transcription factors, and SOX5. Thus, both mouse and human *SPAG6* promoter regions contain putative binding sites for similar transcription factors, which strongly suggests that both human and mouse *SPAG6* genes are regulated by similar mechanisms. In this study, we focused on the study of regulation of mammalian *SPAG6* genes by SOX5.

The *SOX5* gene encodes two major proteins, the full-length 84-kDa L-SOX5 (SOX5) and the 48-kDa SOX5 (S-SOX5). The S-SOX5 protein, but not the L-SOX5 protein, is expressed in tissues with motile cilia (Fig. 2), suggesting a role of this transcription factor in motile ciliogenesis. The presence of potential binding sites for SOX5 in both the human and mouse *SPAG6* gene promoters and the fact that the *Spag6* message is also abundant in mouse tissues with motile cilia, paralleling the tissue distribution of S-SOX5, suggested that S-SOX5 regulates *Spag6* gene transcription. To explore this role *in vitro*, we examined the promoter in the context of BEAS-2B cells, a line derived from viral transformed human bronchial epithelial cells. BEAS-2B cells express endogenous *S-SOX5* (Fig. 3), indicating that this cell line is suitable for the study of genes transcriptionally regulated by S-SOX5. Immunofluorescence staining revealed that expressed S-SOX5 protein was located in the nuclei (supplemental Fig. 1), supporting the notion that S-SOX5 functions as a transcription factor.

S-SOX5 stimulated *SPAG6* promoter function in BEAS-2B cells, but not in MDCK cells (Fig. 4), and other cell types, including COS-1 cells and CHO cells (data not shown), that are derived from tissues not bearing motile cilia, implying that S-SOX5 can regulate *SPAG6* gene expression only in the con-

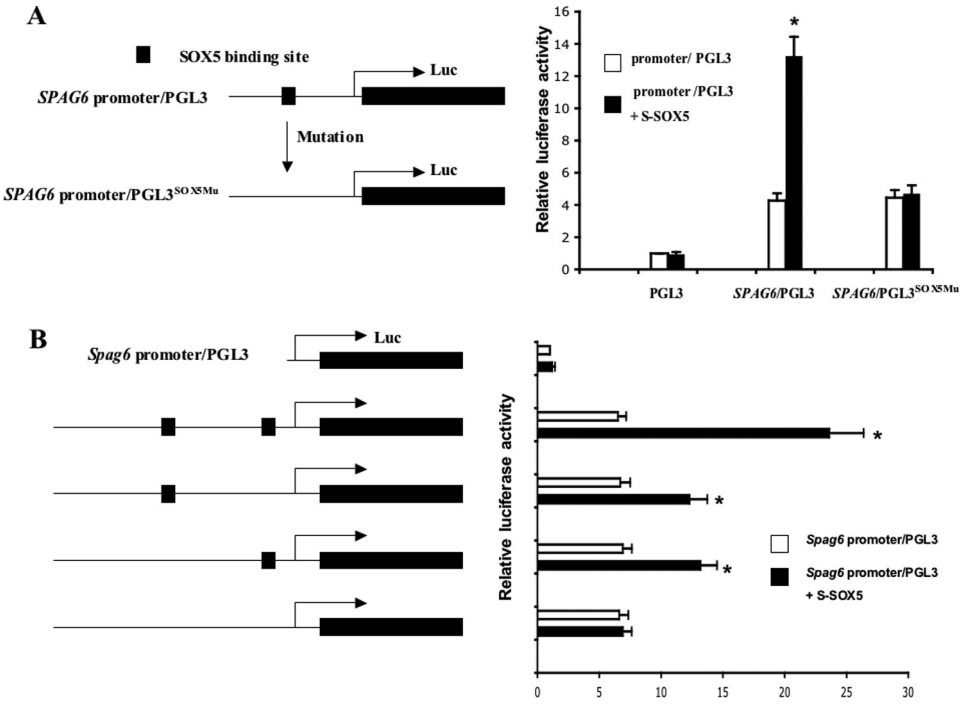

FIGURE 10. **Functional analyses of SOX5 binding sites in human and mouse *SPAG6* promoters.** *A*, effect of S-SOX5 on human *SPAG6* promoter activity. *Left*, maps of the wild type human *SPAG6* promoter construct containing one SOX5 binding site and the construct with mutation of the SOX5 binding site. *Right*, effect of S-SOX5 on the wild type and mutated *SPAG6* promoter function. Values are means ± S.E. (*error bars*) from three separate experiments. *B*, effect of S-SOX5 on mouse *Spag6* promoter activity. *Left*, maps of the mouse *Spag6* promoter constructs that contains two SOX5 binding sites and constructs with series mutation of the SOX5 binding sites. *Right*, effect of S-SOX5 on the wild type and mutated *Spag6* promoter function. Values are means ± S.E. from three separate experiments. *, $p < 0.05$ compared with controls (without S-SOX5 stimulation).

text of the appropriate cells that may contain programs to support production of axonemes with a 9 + 2 structure, such as male germ cells, ependymal cells, and tracheal and bronchial epithelial cells. This cell-specific function is also consistent with the tissue distribution of S-SOX5.

The effect of S-SOX5 on *SPAG6* transcription was confirmed by overexpression of S-SOX5 and silencing *SOX5* in BEAS-2B cells. Overexpression of *S-SOX5* significantly increased *SPAG6* mRNA expression. Conversely, *SPAG6* mRNA expression was dramatically decreased when *SOX5* mRNA was knocked down by RNA interference. Given that S-SOX5 is the predominant SOX5 isoform in BEAS-2B cells, it is unlikely that the down-regulation of *SPAG6* mRNA was the result of knockdown other *SOX5* isoforms.

The fact that *S-SOX5* RNA knockdown reduces *SPAG6* gene expression, whereas forced expression of S-SOX5 increases endogenous *SPAG6* levels, indicates that the abundance of SOX5 in cells is a determinant of the level of *SPAG6* gene transcription. BEAS-2B cells studied under our culture conditions evidently have a steady-state level of S-SOX5 that is sufficient to drive some *SPAG6* transcription, but not at the maximal potential rate because additional S-SOX5 augments *SPAG6* above the basal state.

S-SOX5 associates directly with multiple SOX5 binding sites of the *SPAG6* promoter. This conclusion was supported by the facts that DNA can be amplified by several primer sets that flank SOX5 binding sites (Fig. 7). The EMSA studies indicated that S-SOX5 binds specifically to these *cis*-elements (Figs. 8 and 9 and supplemental Fig. 3). These binding sites have functional significance because S-SOX5 could not stimulate *SPAG6* promoter function when SOX5 binding sites were mutated or deleted (Fig. 10 and supplemental Fig. 4).

Given that there are multiple putative SOX5 binding sites in the promoter region, a comprehensive analysis of all putative sites was not carried out. We cannot quantity the relative contribution of each *cis*-element. Thus, what we can conclude is that S-SOX5 is an important transcription factor controlling the *SPAG6* promoter.

It has been reported that other transcription factors, including FOXJ1 and RFX, regulate ciliogenesis. FOXJ1 is a forkhead/winged-helix transcription factor family. It is expressed in embryonic node, adult lung, choroid plexus, testis, and oviduct (44, 45). Targeted deletion of the *Foxj1* gene resulted in defective ciliogenesis in airway epithelial cells and randomized left-right asymmetry. Half of the mice had situs inversus (46). FOXJ1 is also required for epithelial cell ciliogenesis in the oviduct (47). Microarray analysis indicates that *Xenopus Foxj1* induces the formation of cilia by up-regulating the expression of motile cilia genes, which includes the *Spag6* gene (48). Thus, FOXJ1 plays an essential role in the formation of motile cilia (38). Our studies revealed that FOXJ1 stimulates *SPAG6* promoter function in BEAS-2B and MDCK cells. The stimulatory effect is greater than S-SOX5, suggesting that FOXJ1 might play a dominant role in the regulation of the *SPAG6* gene. Interestingly, in the presence of the combination of S-SOX5 and FOXJ1, significantly higher promoter activity was obtained, suggesting that the two transcription factors function cooperatively to stimulate *SPAG6* promoter function.

Cells with 9 + 2 cilia/flagella presumably express a factor(s) that allows S-SOX5 to transactivate the *SPAG6* gene promoter. Notably, we have previously cloned the promoters of the genes encoding SPAG16L and SPAG17, two other central apparatus proteins. Both of these promoters also contain putative SOX5 binding (supplemental Figs. 5 and 6), and S-SOX5 can also activate these promoters in BEAS-2B cells (data not shown), raising the possibility that the function of S-SOX5 is to regulate formation of the central apparatus through transcriptional control of a suite of genes encoding key components of the central pair of microtubules.

In summary, our study reveals that S-SOX5 regulates *SPAG6* gene expression, acting in concert with FOXJ1. These observations provide the first evidence of S-SOX5 involvement in

motile cilia/flagella through control of expression of an axoneme central apparatus gene.

Acknowledgment—E. A. K. is grateful to Scott Walsh, Ph.D. for assistance in the preparation of her thesis.

REFERENCES

1. Degnan, B. M., Vervoort, M., Larroux, C., and Richards, G. S. (2009) *Curr. Opin. Genet. Dev.* **19,** 591–599
2. Phochanukul, N., and Russell, S. (2010) *Int. J. Biochem. Cell. Biol.* **42,** 453–464
3. Chew, L. J., and Gallo, V. (2009) *J. Neurosci. Res.* **87,** 3277–3287
4. Reiprich, S., Kriesch, J., Schreiner, S., and Wegner, M. (2010) *J. Neurochem.* **112,** 744–754
5. Hernández-Hernández, J. M., Delgado-Olguín, P., Aguillón-Huerta, V., Furlan-Magaril, M., Recillas-Targa, F., and Coral-Vázquez, R. M. (2009) *J. Mol. Biol.* **394,** 1–14
6. Donner, A. L., Episkopou, V., and Maas, R. L. (2007) *Dev. Biol.* **303,** 784–799
7. Savage, J., Conley, A. J., Blais, A., and Skerjanc, I. S. (2009) *Stem Cells* **27,** 1231–1243
8. Cermenati, S., Moleri, S., Cimbro, S., Corti, P., Del Giacco, L., Amodeo, R., Dejana, E., Koopman, P., Cotelli, F., and Beltrame, M. (2008) *Blood* **111,** 2657–2666
9. Kanai-Azuma, M., Kanai, Y., Gad, J. M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. P., and Hayashi, Y. (2002) *Development* **129,** 2367–2379
10. Christiano, A. M. (2008) *Cell Stem Cell* **3,** 3–4
11. Schilham, M. W., Oosterwegel, M. A., Moerer, P., Ya, J., de Boer, P. A., van de Wetering, M., Verbeek, S., Lamers, W. H., Kruisbeek, A. M., Cumano, A., and Clevers, H. (1996) *Nature* **380,** 711–714
12. Haag, J., Gebhard, P. M., and Aigner, T. (2008) *Pathobiology* **75,** 195–199
13. Bonnelye, E., Zirngibl, R. A., Jurdic, P., and Aubin, J. E. (2007) *Endocrinology* **148,** 1195–1205
14. Kormish, J. D., Sinner, D., and Zorn, A. M. (2010) *Dev. Dyn.* **239,** 56–68
15. Lefebvre, V. (2010) *Int. J. Biochem. Cell. Biol.* **42,** 429–432
16. Denny, P., Swift, S., Connor, F., and Ashworth, A. (1992) *EMBO J.* **11,** 3705–3712
17. Hiraoka, Y., Ogawa, M., Sakai, Y., Kido, S., and Aiso, S. (1998) *Biochim. Biophys. Acta* **1399,** 40–46
18. Lefebvre, V., Li, P., and de Crombrugghe, B. (1998) *EMBO J.* **17,** 5718–5733
19. Ikeda, T., Zhang, J., Chano, T., Mabuchi, A., Fukuda, A., Kawaguchi, H., Nakamura, K., and Ikegawa, S. (2002) *Gene* **298,** 59–68
20. Kou, I., and Ikegawa, S. (2004) *J. Biol. Chem.* **279,** 50942–50948
21. Smits, P., Li, P., Mandel, J., Zhang, Z., Deng, J. M., Behringer, R. R., de Crombrugghe, B., and Lefebvre, V. (2001) *Dev. Cell* **1**, 277–290
22. Smits, P., and Lefebvre, V. (2003) *Development* **130,** 1135–1148
23. Dy, P., Han, Y., and Lefebvre, V. (2008) *Genesis* **46,** 294–299
24. Blaise, R., Grober, J., Rouet, P., Tavernier, G., Daegelen, D., and Langin, D. (1999) *J. Biol. Chem.* **274,** 9327–9334
25. Budde, L. M., Wu, C., Tilman, C., Douglas, I., and Ghosh, S. (2002) *Mol. Biol. Cell* **13,** 4179–4194
26. Xu, W., Zhang, S., Qiu, W., He, G., Liu, Y., Sun, Y., Ma, Y., Dong, J., and Zhang, W. (2009) *Mol. Biol. Rep.* **36,** 1187–1193
27. Wunderle, V. M., Critcher, R., Ashworth, A., and Goodfellow, P. N. (1996) *Genomics* **36,** 354–358
28. Neilson, L. I., Schneider, P. A., Van Deerlin, P. G., Kiriakidou, M., Driscoll, D. A., Pellegrini, M. C., Millinder, S., Yamamoto, K. K., French, C. K., and Strauss, J. F., III (1999) *Genomics* **60,** 272–280
29. Sapiro, R., Tarantino, L. M., Velazquez, F., Kiriakidou, M., Hecht, N. B., Bucan, M., and Strauss, J. F., III (2000) *Biol. Reprod.* **62,** 511–518
30. Sapiro, R., Kostetskii, I., Olds-Clarke, P., Gerton, G. L., Radice, G. L., and Strauss, J. F., III (2002) *Mol. Cell. Biol.* **22,** 6298–6305
31. Horowitz, E., Zhang, Z., Jones, B. H., Moss, S. B., Ho, C., Wood, J. R., Wang, X., Sammel, M. D., and Strauss, J. F., III (2005) *Mol. Hum. Reprod.* **11,** 307–317
32. Zhang, Z., Tang, W., Zhou, R., Shen, X., Wei, Z., Patel, A. M., Povlishock, J. T., Bennett, J., and Strauss, J. F., III (2007) *Cell. Motil. Cytoskeleton* **64,** 360–376
33. Lonergan, K. M., Chari, R., Deleeuw, R. J., Shadeo, A., Chi, B., Tsao, M. S., Jones, S., Marra, M., Ling, V., Ng, R., Macaulay, C., Lam, S., and Lam, W. L. (2006) *Am. J. Respir. Cell. Mol. Biol.* **35,** 651–661
34. Smith, E. F., and Lefebvre, P. A. (1996) *J. Cell Biol.* **132,** 359–370
35. Smith, E. F., and Lefebvre, P. A. (2000) *Cell. Motil. Cytoskeleton* **46,** 157–165
36. Ralston, K. S., Lerner, A. G., Diener, D. R., and Hill, K. L. (2006) *Eukaryot. Cell* **5,** 696–711
37. Branche, C., Kohl, L., Toutirais, G., Buisson, J., Cosson, J., and Bastin, P. (2006) *J. Cell Sci.* **119,** 3443–3455
38. Yu, X., Ng, C. P., Habacher, H., and Roy, S. (2008) *Nat. Genet.* **40,** 1445–1453
39. Zhang, Z., Sapiro, R., Kapfhamer, D., Bucan, M., Bray, J., Chennathukuzhi, V., McNamara, P., Curtis, A., Zhang, M., Blanchette-Mackie, E. J., and Strauss, J. F., III (2002) *Mol. Cell. Biol.* **22,** 7993–8004
40. Zhang, Z., Jones, B. H., Tang, W., Moss, S. B., Wei, Z., Ho, C., Pollack, M., Horowitz, E., Bennett, J., Baker, M. E., and Strauss, J. F., III (2005) *Mol. Cell. Proteomics* **4,** 914–923
41. You, Y., Huang, T., Richer, E. J., Schmidt, J. E., Zabner, J., Borok, Z., and Brody, S. L. (2004) *Am. J. Physiol. Lung Cell. Mol. Physiol.* **286,** L650–L657
42. Zhang, Z., Austin, S. C., and Smyth, E. M. (2001) *Mol. Pharmacol.* **60,** 480–487
43. Gonzalez, I. L., and Schmickel, R. D. (1986) *Am. J. Hum. Genet.* **38,** 419–427
44. Brody, S. L., Hackett, B. P., and White, R. A. (1997) *Genomics* **45,** 509–518
45. Blatt, E. N., Yan, X. H., Wuerffel, M. K., Hamilos, D. L., and Brody, S. L. (1999) *Am. J. Respir. Cell Mol. Biol.* **21,** 168–176
46. Brody, S. L., Yan, X. H., Wuerffel, M. K., Song, S. K., and Shapiro, S. D. (2000) *Am J. Respir. Cell Mol. Biol.* **23,** 45–51
47. Okada, A., Ohta, Y., Brody, S. L., Watanabe, H., Krust, A., Chambon, P., and Iguchi, T. (2004) *J. Mol. Endocrinol.* **32,** 615–625
48. Stubbs, J. L., Oishi, I., Izpisúa Belmonte, J. C., and Kintner, C. (2008) *Nat. Genet.* **40,** 1454–1460