A Novel CpG-free Vertebrate Insulator Silences the Testis-specific SP-10 Gene in Somatic Tissues

ROLE FOR TDP-43 IN INSULATOR FUNCTION

Mayuresh M. Abhyankar1, Craig Urekar1, and Prabhakara P. Reddi

From the Department of Pathology, University of Virginia School of Medicine, Charlottesville, Virginia 22908

Regulation of cell type-specific gene transcription is central to cellular differentiation and development. During spermatogenesis, a number of testis-specific genes are expressed in a precise spatiotemporal order. How these genes remain silent in the somatic tissues is not well understood. Our previous studies using the round spermatid-specific mouse SP-10 gene, which codes for an acrosomal protein, revealed that its proximal promoter acts as an insulator and prevents expression in the somatic tissues. Here we report that the insulator tethers the SP-10 gene to the nuclear matrix in somatic tissues, sequestering the core promoter in the process, thus preventing transcription. In round spermatids where the SP-10 gene is expressed, this tethering is released. TAR DNA-binding protein of 43 kDa (TDP-43), previously shown to interact with the SP-10 insulator, was found to be in the 2 M NaCl-insoluble nuclear matrix fraction. TDP-43 prevented enhancer-promoter interactions when artificially recruited between the two by Gal4 strategy. Knockdown of TDP-43 using small interfering RNA released the enhancer-blocking effect of the SP-10 insulator in a stable cell culture model. Mutation of TDP-43 binding sites abolished this effect. Finally, a 50-bp subfragment of the SP-10 insulator, which includes TDP-43 binding sites, functioned as a minimal insulator in transgenic mice and silenced an otherwise ectopically expressed transgene in somatic tissues. The SP-10 insulator lacks CpG dinucleotides or CTCF binding sites. Thus, the present study characterized a novel vertebrate insulator in a physiological context and showed for the first time how a testis-specific gene is silenced in the somatic tissues by an insulator.

Regulation of transcription of a tissue-specific gene is central to cellular differentiation and development. Mechanisms must be in place not only to activate the gene in the correct cell type but also to keep the gene silenced in all other tissues. We study this problem using the mouse testis-specific SP-10 gene as a model. The SP-10 gene codes for an acrosomal protein conserved in mammals and is exclusively expressed in round spermatids (1). In this report, we address the mechanism of transcriptional silencing of the SP-10 gene in the somatic tissues. Our previous work in transgenic mice showed that the SP-10 proximal promoter performs dual functions; it activates testis-specific transcription and also acts as an insulator preventing the possibility of transcription in the somatic tissues (2).

Insulators are DNA sequences located at gene boundaries (3). Their action prevents ectopic expression of genes they flank. Insulators are operationally defined as enhancer blockers and barrier elements. Enhancer-blocking insulators prevent a foreign enhancer from inappropriately encroaching into the promoter of the neighboring gene. The barrier insulators prevent the spread of heterochromatin and thus prevent untimely shut down of gene expression. Insulators may possess only one or both of the above properties (4).

The chicken β-globin HS4 insulator (cHS4),3 a prototypic vertebrate insulator, possesses both of the above activities (5, 6). Binding sites for the transcription factors CTCF and USF1 have been shown to be responsible for the enhancer-blocking and barrier functions of the cHS4 insulator, respectively (7, 8). The imprinting control region located between the Igf2 and H19 genes and the Tsix imprinting/choice center on the mammalian X chromosome are two other well characterized vertebrate insulators with enhancer-blocking properties. The imprinting control region governs the parent of origin-specific expression of Igf2 and H19, depending on whether or not it is bound by the 11-zinc finger protein, CTCF (9, 10). Similarly, the Tsix imprinting/choice center acts as an enhancer blocker when bound by CTCF and prevents Xist gene expression, thus marking the active X chromosome (11). Both the imprinting control region and the Tsix imprinting/choice center insulators contain CpG dinucleotides, the methylation status of which dictates occupancy by CTCF. A number of other vertebrate insulators also contain CpG dinucleotides and binding sites for CTCF. These include the insulators at the ribosomal RNA genes of Xenopus (12), the chicken 3′-HS insulator of the β-globin gene (13), the BEAD-1 element at the human TCRα/δ locus (5), the human β-globin HS5 insulator (14), the apoB insulator of the human apolipoprotein B chromatin domain (15), and the DM1 insulator located between the human DMPK and S5X genes (16).

In contrast, the 319-bp SP-10 insulator, which functions as an enhancer blocker in transgenic mice, does not contain CpG dinucleotides or CTCF binding sites (2). We reported that the

---

1 These authors contributed equally to this work.
2 To whom correspondence should be addressed: Dept. of Pathology, University of Virginia School of Medicine, P.O. Box 800904, Charlottesville, VA 22908. Tel.: 434-982-0007; Fax: 434-924-2151; E-mail: ppr5s@virginia.edu.

3 The abbreviations used are: cHS4, chicken β-globin HS4 insulator; DBD, DNA binding domain; CMV, cytomegalovirus; TK, thymidine kinase; siRNA, small interfering RNA; TDP-43, TAR DNA-binding protein of 43 kDa.

Received for publication, July 16, 2007, and in revised form, September 24, 2007. Published, JBC Papers in Press, October 11, 2007, DOI 10.1074/jbc.M705811200
A Novel Vertebrate Insulator

SP-10 insulator binds to TDP-43 (TAR DNA-binding protein of 43 kDa) and that mutation of TDP-43 binding sites compromised enhancer-blocking function (17). Therefore, we anticipate that the study of the SP-10 insulator will uncover novel mechanisms of insulator function.

Here, we have addressed the mechanism of SP-10 insulator function in the endogenous context; investigated the requirement of TDP-43 for SP-10 insulator function; and, finally, determined whether a 50-bp subfragment can function as a minimal insulator capable of silencing a transgene in the somatic tissues.

EXPERIMENTAL PROCEDURES

Cell Lines and Culture Conditions—Mouse spermatocyte cells (GC-2; ATCC catalog number CRL-2196) and monkey kidney cells (CV-1) were cultured in Dulbecco’s modified Eagle’s medium with 10% fetal calf serum, 1% nonessential amino acids. COS-7 cells were maintained in Dulbecco’s modified Eagle’s medium with 10% fetal calf serum. Flp-In CV-1 stable cell lines were purchased from Invitrogen. Dulaoco’s modified Eagle’s medium with 10% fetal calf serum, 1% nonessential amino acids. COS-7 cells were maintained in Dulbecco’s modified Eagle’s medium with 10% fetal calf serum. Flp-In CV-1 stable cell lines were purchased from Invitrogen.

In Vivo Matrix Assay—Low ionic strength matrices from cultured cells, mouse liver cells, and round spermatids were prepared essentially as described (18, 19). Following lithium dioosaliculate-mediated extraction, the nuclei were digested with 100 μg/ml RNase-free DNase I (Roche Applied Science) for 3 h at room temperature. Matrices were collected by centrifugation and resuspended in Tris-EDTA (TE), 0.1% SDS, and 1 mg/ml proteinase K. Extracts were incubated at 55 °C overnight. The DNA associated with these preparations was precipitated after phenol/chloroform extraction (matrix-bound DNA). Total genomic DNA was also extracted in parallel. PCR was performed using matrix bound as well as genomic DNA as template. The SP-10 −1.9/+1.6 kb genomic region was analyzed using primer sets, each yielding an ~270-bp ampiclon at 100-bp intervals as shown in Fig. 1A. A typical PCR contained 10 pmol of each primer, 1 μl of 10 mm dNTP mix, 0.5 μl of Taq, and 5 ng of template in a 50-μl volume. The cycling conditions were 94 °C for 3 min, followed by 27 cycles of 94 °C for 30 s/55 °C for 30 s/72 °C for 30 s with a final extension at 72 °C for 5 min. The amplification was in a linear range under these conditions. Fold enrichment values were obtained by taking ratios of band intensities of ampiclons generated from matrix DNA/genomic DNA. Band intensities were quantitated using ImageQuant software.

Isolation of Mouse Spermatids—Pure population (>95%) of spermatids was obtained using StaPut gradient as described previously (17, 20). Typically, testes from 10–12-week-old Swiss Webster mice were subjected to StaPut to obtain ~20–40 × 10⁶ round spermatids. Alternatively, testicular germ cells from 25-day-old male mice were used as a source of round spermatids (85–90% pure). To obtain day 25 germ cells, testes from 10 mice were decapsulated and subjected to enzymatic treatment (collagenase, hyaluronidase, tryspin, and DNase) to release the cells. These were collected and washed twice with phosphate-buffered saline. Round spermatids obtained from either source were used to prepare nuclear matrix.

High Salt Matrix Preparation—High salt matrices from cultured cells were essentially prepared as described (19). In short, cells were incubated in RSB buffer for 10 min on ice, homogenized 10 times with a Dounce homogenizer by using a loose pestle, and centrifuged at 1000 × g for 5 min at 4 °C. Pelleted nuclei were washed twice in RSB and 0.25 M sucrose, resuspended in RSB and 2 M sucrose, and centrifuged at 34,000 × g for 10 min. Pelleted nuclei were washed once in RSB and 0.25 M sucrose; resuspended in RSB, 0.25 M sucrose, 1 mM CaCl₂ and phenylmethylsulfonyl fluoride; and digested for 3 h with 100 μg/ml RNase free DNase I (Roche Applied Science) at room temperature on a nutator mixer. After digestion, the nuclei were pelleted at 1500 × g for 5 min, and supernant was removed. These nuclei were then extracted with extraction buffer (20 mM Tris, pH 7.4, 10 mM EDTA) containing an increasing amount of NaCl ranging from 0.1 to 2.0 M. For each concentration of salt, the digested nuclei were washed three times with 3 pellet volumes of extraction buffer, followed by centrifugation at 1500 × g for 5 min. The three washes were combined at each step and mixed with loading buffer. An equal volume of each was used for Western blotting. The final 2 M salt-insoluble nuclear matrix was solubilized directly in loading buffer in an equivalent volume.

Western Blotting and Antibodies—Polyclonal antibodies to TDP-43 were raised in guinea pigs as mentioned (17). Anti-Gal4DBD (catalog number sc510) and anti-lamin A/C (catalog number sc6215) antibodies were purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA). Antibody to histone H3 (catalog number sc10809) was from Upstate Biotechnology, Inc. All blots were developed using an Immobilon Western kit from Millipore.

DNA Constructs—Full-length murine TDP-43 (NM_145556) was amplified from mouse testis Marathon-ready cDNA (Clontech) and cloned in pCR II-TOPO vector. This clone was used as a template to generate NH2-terminal deletion (ΔRRM1; amino acids 169–414) or COOH-terminal deletion (ΔGly; amino acids 1–274) mutants using PCR.

Gal4 Assay Constructs—To make DNA binding domain (DBD) fusion proteins, full-length or mutant forms of TDP-43 were cloned in pFA-CMV vector (Strategene) as BglI-XbaI fragments in frame with the Gal4 BD. Expression of the fusion proteins in GC-2 cells was confirmed by Western blotting. The control SP-10 open reading frame was cloned similarly. A construct expressing the p53-BD fusion was a gift from Dr. Rong Li (Dept. of Biochemistry, University of Virginia). The reporter plasmid was constructed as follows. The 5× Gal element was PCR-amplified from pFR-Luc plasmid (Strategene) and was ligated into BglII site of pGL3 −91/+28 Luc (as described in Ref. 2). Thus, the 5× Gal site separates CMV enhancer and mouse SP-10 core promoter, allowing recruitment of the Gal4DBD fusion proteins between the enhancer and the core promoter.

Luciferase Reporter Gene Constructs—Luciferase reporter constructs were built using the pGL3basic vector (Promega). Refer to Fig. 5A for a schematic illustration of all constructs. The coordinates for mouse SP-10 promoter sequence are as per Fig. 7 in Ref. 2 (GenBank™ accession number AF 133710). The SP-10 core promoter (−91/+28) was PCR-amplified with BglII (5’ end) and HindIII (3’ end) sites using the −408SP10−gfp
able upon request). siRNA was introduced into CV-1 cells by hybridization (data not shown).

mediated site-specific integration was confirmed by Southern instructions. For all clones, flippase recombinase target (FRT)-were then cut out and subcloned into pCDNA5FRT vector partially cloned in pGL3basic vector (Promega). These cassettes mutated), and (iv) CMV-stuffer-SP10 core-luciferase) were ini-

orientation) was selected, giving rise to construct 3 of Fig. 5A. The 50-bp insulator bearing KpnI sites on both ends was cloned into KpnI-cleaved construct 1, which placed the insulator upstream of the CMV enhancer in the correct as well as reverse orientations, giving rise to constructs 6 and 7, respectively. To generate reporter constructs with the thymidine kinase (TK) core promoter, a 92-base pair fragment corresponding to the –61/+31 (wherein +1 corresponds to the transcriptional start site) region of the HSV TK gene promoter was PCR-amplified from pRL-TK (Promega, WI) with BglII and HindIII sites and cloned into CMV-pGL3 (2). This base construct (construct 8; Fig. 5A) was then used to clone in the 50-bp SP-10 insulator in both orientations as described above, creating constructs 9 and 10. All constructs were confirmed by sequencing.

Transient Transfection Assays—Transient transfections were performed in COS-7 or GC-2 cells using Mirus TransIT®-LT1 (Mirus Corp.) transfection reagent. 1.5 × 105 Cells were plated in 6-well plates, and typically, the total amount of DNA transfected per well was 1 μg. Cells were harvested 48 h post-transfection. The plasmid pRL-TK (Promega) was cotransfec-
ted at a 1:10 ratio with the reporter to normalize for transfection efficiency. Luciferase activities were measured by the Dual-Luciferase reporter assay system (Promega) according to the instructions provided with the kit. Normalized luciferase values were plotted in histograms. For Gal4DBD-based assays, 2.0 × 105 cells were seeded and transfected with 1 μg each of reporter and effector construct. The luciferase assay was done 48 h post-transfection.

Generation of Stable CV-1 Cell Lines—CV-1 stable cell lines were generated using the Flp-In system (Invitrogen) as per the manufacturer’s instructions. In short, four cassettes ((i) CMV-SP10 core-luciferase, (ii) CMV-(−408/−92) insulator-SP10 core-luciferase, (iii) CMV-mutant −408/−92 insulator-SP10 core-luciferase (the two GTGTTGT motifs at −172 and −160 mutated), and (iv) CMV-stuffer-SP10 core-luciferase) were ini-

ally cloned in pGL3basic vector (Promega). These cassettes were then cut out and subcloned into pCDNA5FRT vector (Invitrogen). Flp-In CV-1 cells (Invitrogen) were used to co-
transfect these expression vectors with pOG44, and hygromycin-resistant clones were selected as per the manufacturer’s instructions. For all clones, flippase recombinase target (FRT)-mediated site-specific integration was confirmed by Southern hybridization (data not shown).

RNA Interference—TDP-43-specific or control siRNA reagents were purchased from Dharmacon (sequences available upon request). siRNA was introduced into CV-1 cells by reverse transfection using Lipofectamine RNAiMAX (Invitro-
gen) as per the manufacturer’s instructions. Cells were harvested 48 h post-transfection for luciferase assay.

Generation of Transgenic Mice—Transgenic mouse lines were generated using constructs 1–4 (Fig. 5A). Each transgene was purified as a BamHI fragment. Pronuclear injections were performed by the University of Virginia’s Gene Targeting and Transgenic Facility, using standard procedures (21). Tail DNA was isolated from founder mice using the Qiagen DNeasy® tissue kit (Qiagen), and transgenic founders were identified by PCR using primers designed to amplify the entire open reading frame of the luciferase cDNA. Primers LucF (5’-GGTAAAGCCACCATGGAAGACGCCA) and Lucr (5’-TTACACGGCGATCTTTCGCCCTTCT) were used as forward and reverse primers, respectively. Amplification of a 1665-bp product iden-
tified putative positive founders. The transgenic lines were sub-
sequently confirmed by Southern hybridization using the Luciferase cDNA (Fig. 5B) and the CMV enhancer as probes.

Copy Number Estimation for Transgenic Mice—Southern blot hybridization was used to determine the copy number of the transgene in all of the mouse lines that were used for the analysis of luciferase expression. Ten micrograms of tail DNA was digested to completion with BglIII, electrophoresed through a 1% agarose gel, and transferred to Duralon-UV™ membrane (Stratagene). Full-length luciferase coding sequence (bp 54 –1746 from vector pGL3basic; Promega, WI) was radio-
labeled using [α-32P]dCTP and the Prime-a-Gene® labeling system (Promega). DNA hybridization was performed using ExpressHyb (Clontech) per the manufacturer’s recommended protocol. The probed blot was exposed to an Amersham Bio-
sciences storage phosphor screen and scanned in on an Amer-
sham Biosciences Storm 860. For copy number analysis, the blot image was imported into Amersham Biosciences ImageQuant version 5.0. Within ImageQuant, a box was placed over each band on the blot, and volume quantification was per-
formed on all bands after the subtraction of background. The Southern blot image (Fig. 5B) was carefully analyzed to identify bands corresponding to single copy integration versus multiple copy integration of the transgene. Because of the location of a BgIII site at the 5′ end of the core promoter in all constructs used for making transgenic mice, single copy integration would result in a hybridization signal above 2 kb. Because of the pres-
ence of a second BgIII site at the 3′ end of the CMV enhancer, head to tail ligated multiple copy transgene integration would result in a strong hybridization signal at 2.4 kb for constructs 1–4. Head to tail ligated multiple copy integration would result in a hybridization signal at 4 kb for all of the constructs. The transgenic mouse line (line 540), which generated a single hybridization signal that is at neither 2.4 nor 4 kb (Fig. 5B, lane 5), represented single copy integration, and the volume of that band was taken as a standard single integration volume to determine the copy number of the rest of the lines. The volume of each hybridization band from each line was divided by that single-integration volume to determine the number of copies each band represented on the blot. The total value for each of the bands was then added up to give the total copy number of that particular mouse line. Copy number was independently estimated and confirmed using the CMV enhancer region
A Novel Vertebrate Insulator

(Xhol-released fragment from CMV-TOPO) (2) as probe on the same blot as described above (data not shown).

Luciferase Assays on Tissues from Transgenic Mice—Tissues were harvested from 8–12-week-old transgenic male mice as well as nontransgenic littermates and immediately placed in Promega passive lysis buffer, manually homogenized, and left to sit at room temperature for 15 min. The homogenized tissues were then frozen until analysis. Upon thawing, lysed tissue samples were centrifuged at 16,000 × g for 2 min to remove debris; the supernatant was transferred to a new microcentrifuge tube. Soluble protein was quantified using the Bio-Rad protein quantification reagent (catalog number 500-0006) using the manufacturer’s recommended protocol. The luciferase analysis was performed on a Turner Designs TD20/20 luminometer with Promega luciferase assay reagent. Luciferase assays were standardized by protein amount; 50 μg of protein was used for each tissue analyzed. The highest luciferase activity obtained from any tissue of nontransgenic littermates was 0.15. Therefore, we considered 0.15 as the background cut-off and subtracted this from the luciferase activities of transgenic mice.

RESULTS

The SP-10 Insulator Is Associated with the Nuclear Matrix—We previously demonstrated using transgenic mice as an assay system that the proximal promoter of the spermid-specific mouse SP-10 gene functions as an enhancer-blocking insulator in the somatic tissues. In round spermatids, where the SP-10 gene is expressed, the insulator function is relieved. Here we investigated the mechanism of the SP-10 insulator function in its native context. Previous studies reported a close association between insulator elements and components of nuclear matrix (22–24). To determine whether the SP-10 insulator operates by tethering to the nuclear matrix, we employed an in vivo matrix attachment assay. In this assay, a combination of mild detergent and lithium diiodosalicylate is used to detach proteins from nuclei without disturbing the native attachment sites (18, 19). Nuclei obtained from mouse liver cells (SP-10 gene silent) or round spermatids (SP-10 gene expressed) were washed with mild detergent, extracted with lithium diiodosalicylate, and then digested with DNsase I. After digestion, the nuclear matrices were prepared as described, and the associated DNA was purified. To determine whether the SP-10 gene insulator is attached to the matrix, we PCR-amplified a series of 270-bp segments at 100-bp intervals spanning the −2.0 to +1.6 kb region of the mouse SP-10 gene and plotted the -fold enrichment (Fig. 1A). In liver cells where the SP-10 gene is not expressed, the region corresponding to −266/+28 was found enriched up to 20-fold in the nuclear matrix fraction as compared with the adjacent sequences on either side (Fig. 1B). PCR amplification using primers corresponding to the larger −408/+28 region also showed similar enrichment as the −266/+28 region (data not presented). This specific enrichment, however, was completely lost in the case of round spermatids where the SP-10 gene is expressed (Fig. 1B). These results suggest that the proximal promoter region of the SP-10 gene is sequestered within the nuclear matrix in cells that do not express SP-10 and that it is released within the differentiating spermatogenic cells, thus allowing SP-10 gene transcription. We confirmed this differential tethering by varying the source of the cell type. Similar to liver cells, enrichment of the SP-10 proximal promoter was observed within the nuclear matrix in GC2 tissue culture cells as well (data not shown). We used day 25 postnatal mouse testicular germ cells as an alternate source of SP-10-expressing cells. At postnatal day 25, a majority of the cells in mouse testis will be round spermatids because of the synchronous nature of the first wave of seminiferous epithelium. Similar to the round spermatids isolated from adult mice, the day 25 spermatids also showed no enrichment of SP-10 promoter region in the nuclear matrix (data not shown). To our knowledge, this is the first direct comparison of attachment properties of an insulator in two different scenarios of gene expression.

TDP-43, the SP-10 Insulator-binding Protein, Is Co-localized in the Matrix—We previously reported identification of TDP-43 (TAR DNA-binding protein of 43 kDa) using a portion of the SP-10 insulator to screen a cDNA expression library (17). TDP-43 was originally cloned using the HIV TAR region and shown to be a transcriptional repressor (25). Buratti and Baralle (26) showed a role for TDP-43 in mRNA splicing and determined that it binds to RNA/DNA via UG/TG repeats. We previously showed that TDP-43 binds to the −186/−148 SP-10 promoter via two 5′-GGTGTG motifs located on the opposite strand. Disruption of these motifs resulted in loss of TDP-43 binding in vitro and compromised the stage- and cell-specific expression of a reporter gene in transgenic mice (17). These results prompted us to investigate the potential of TDP-43 as a candidate insulator protein. To begin, we first asked if TDP-43 was in the nuclear matrix. Nuclear matrix proteins are typically resistant to extraction with 2 M NaCl. We used the serial extraction method as described by Yusufzai and Felsenfeld (19). This procedure avoids miscellaneous behavior of proteins as a result of direct treatment with 2 M NaCl (27). Purified nuclei from

![FIGURE 1. The SP-10 insulator tethers the proximal promoter of the gene to the nuclear matrix. A, schematic representation of the SP-10 region analyzed for matrix attachment. Each vertical bar numbered 1–11 represents an individual PCR amplicon spanning the −1.9 kb/+1.6 kb region of the SP-10 gene. The shaded box represents the −408/+28 region of the SP-10 promoter, +1 being the transcriptional start point. B, in vivo nuclear matrix attachment assay. Low ionic strength matrices were prepared from liver cells, and round spermatids and matrix-bound DNA fragments were extracted after extensive DNsase I digestion. Enrichment of the SP-10 insulator fragment was analyzed by PCR using the primer sets depicted in A. Liver cell nuclear matrix showed selective enrichment of the region corresponding to the SP-10 insulator. Round spermatids showed no such enrichment. Data shown represent the average of at least three independent experiments represented as mean ± S.E. Primer sequences are available upon request.](image-url)
GC-2 cells were extensively digested with DNase I and subjected to extraction with buffers of increasing ionic strength. The release of TDP-43 was monitored by Western blotting. (Fig. 2). The majority of TDP-43 was retained in the 2 M salt-resistant nuclear matrix fraction, although small amounts of the protein were extractable at lower salt concentrations. Known soluble and insoluble nuclear proteins were assayed in parallel as controls. A majority of nuclear lamins were detected in the matrix fraction as expected. In contrast, histone H3 was readily extractable with a lower amount of salt, being a nonmatrix protein. These results indicated that TDP-43 is a component of the nuclear matrix.

TDP-43 Blocks Enhancer-Promoter Interaction—In order to test the candidacy of TDP-43 as an insulator protein, we performed an enhancer-blocking assay wherein TDP-43 was artificially recruited to occupy a position between the CMV enhancer and the SP-10 core promoter. A reporter gene construct was generated in which five tandem repeats of Gal4 DNA binding sites were inserted between the CMV enhancer and the −91/+28 SP-10 core promoter (Fig. 3A). TDP-43 contains two RNA recognition motifs and a carboxyl-terminal glycine-rich (Gly) domain (26). We fused the Gal4 DNA binding domain to the wild type TDP-43 or its truncated versions lacking either the NH2-terminal RRM1 domain or the COOH-terminal Gly domain (Fig. 3A). The Gal4 DBD fusion protein constructs were cotransfected with the above reporter plasmid into GC2 cells, and the ability of the DBD fusion proteins to block the CMV enhancer activity was measured. Luciferase activities resulting from cotransfecting DBD alone were used as 100%. Wild type full-length TDP-43 could block 60% of the enhancer activity (Fig. 3B). This effect was due to site-specific recruitment, because the untargeted TDP-43 did not show a similar blocking effect. Deletion of the COOH-terminal Gly domain restored the enhancer activity to near control level, indicating a role for this domain in blocking function. In contrast, the RRM1 domain appeared to play a lesser role in blocking. The SP10-DBD negative control showed no enhancer blocking, whereas the p53-DBD fusion protein, a potent transcriptional activator, elevated reporter gene activity (Fig. 3B). These results showed that TDP-43 could prevent the enhancer-promoter interaction and hence has a potential to act as an insulator protein in vivo.

Functional Link between TDP-43 and SP-10 Insulator; Knockdown of TDP-43 Released Repression of an Insulated Transgene—In order to directly test whether TDP-43 is involved in mediating SP-10 insulator function, we used the siRNA-mediated knockdown approach. To this end, first we generated a series of stable cell lines carrying in the same genomic location single copy integration of the enhancer-promoter interactions. The constructs used to generate these lines and their luciferase activities are shown in Fig. 4A. The CMV enhancer plus SP-10 core promoter-driven transgene expressed a high amount of luciferase. The presence of the −408/−92 SP-10 insulator but not a stuffer fragment of similar length blocked the CMV enhancer activity. Mutation of the TDP-43 binding sites (two GTGTTG motifs at −172 and −160 mutated) compromised the enhancer-blocking ability of the SP-10 insulator. Next, to test the requirement of TDP-43 for the above enhancer-blocking function, we performed siRNA-mediated knockdown of TDP-43 in all of the above stable cell lines. To account for any off-target effects, we used a commercially available control siRNA reagent. Western blots indicated...
that TDP-43-specific siRNA completely knocked down the TDP-43 protein, whereas the control siRNA had little effect (Fig. 4B). For each stable cell line, luciferase values resulting from the off target siRNA were plotted as 100%. Knockdown of TDP-43 significantly elevated \( (p < 0.005) \) the reporter gene activity in the cells carrying the SP-10 insulator as compared with the cells containing the stuffer DNA. TDP-43 knockdown had no effect when TDP-43 binding sites in the SP-10 insulator were mutated (Fig. 4C). Taken together, these results indicated that TDP-43 plays a role, at least in part, in the enhancer-blocking activity of the \(-408/-92\) SP-10 insulator.

A 50-bp Fragment of the SP-10 Insulator Functions as a Minimal Insulator—Next, we asked whether a smaller segment within the \(-408/-92\) SP-10 insulator can function as a minimal insulator. Our previous work using progressive 5’ deletions of the \(-408/-92\) insulator established that the \(-186/-92\) but not the \(-135/-92\) region acted as an enhancer blocker (2). To determine if the \(-186/-135\) region by itself, which includes TDP-43 recognition sites, could act as a minimal insulator, we inserted the \(-186/-135\) fragment between the potent CMV enhancer (28) and the \(-91/+28\) SP-10 core promoter and performed enhancer-blocking assays in COS cells. The construct containing the 319-bp full-length SP-10 insulator \((-408/-92\) was also used for comparing the enhancer-blocking activity. Circular plasmids (see constructs 1–7 in Fig. 5A) were transfected into COS cells, and luciferase values were measured 48 h post-transfection. The construct containing the CMV enhancer and SP-10 core promoter showed robust luciferase expression, as expected. The SP-10 insulator (INS) when present between the enhancer and the core promoter repressed luciferase expression by more than 5-fold, as compared with the uninsulated clone containing a stuffer fragment (stuffer) of identical length. Mutation of TDP-43 binding sites within the insulator (mutant) compromised the enhancer-blocking ability by 3-fold. The values represent the average of at least three independent assays. The above CV-1 stable cell lines were transfected with 10 nm control or TDP-43-specific siRNAs, as described under “Experimental Procedures.” Protein extracts were prepared 48 h post-transfection. Western blots using anti-TDP-43 (top) or anti-lamin A/C (bottom) antibodies showed specific knockdown of TDP-43 protein. C, knockdown of TDP-43 released the enhancer-blocking effect of the SP-10 insulator. Luciferase levels in cells treated with nontargeting control siRNAs were considered as 100 for normalization. Values are shown as the means of three independent experiments ± S.E. The increase in luciferase activity of cells carrying insulated transgene is statistically significant compared with the cells lacking the insulator by Student’s t-test (\( p < 0.005 \)). However, the luciferase activities of mutant insulator (\( p > 0.1 \)) or stuffer (\( p > 0.49 \)) are not statistically significant.
A Novel Vertebrate Insulator

**FIGURE 5.** Schematic representation of the reporter gene constructs (not drawn to scale) used in enhancer-blocking assays in tissue culture cells and in transgenic mice. A, the base vector for all of the above constructs was pGL3basic (Promega, WI). The CMV enhancer (406 bp) corresponds to coordinates 59–465 bp of pEGFPN1 (Clontech). The SP-10 core promoter (-91/+28) fragments used in the above constructs originated from the −408SP10-gfp plasmid (16), which contains the −408/+28 SP-10 promoter (+1 corresponds to the transcriptional start site). The −60/+31 HSV-thymidine kinase core promoter (TK Core; 92 bp in size) corresponds to coordinates 668–759 bp of pRLTK vector (Promega, WI). The size of the luciferase gene is 1916 bp. Transgenic mice were generated using constructs 1–4. Although depicted in only some, BglII sites flank all of the constructs. CMV enh, CMV enhancer; SP-10 core, SP-10 core promoter, INS, SP-10 insulator of indicated size; TK core, HSV-thymidine kinase core promoter; Luc, luciferase coding region and poly(A) signals. B, evidence for reporter gene integration in transgenic mice and determination of copy number. BglII-digested genomic DNA from all of the transgenic mouse lines were probed with radiolabeled luciferase cDNA. Construct numbers correspond to A. Line numbers denote independent mouse lines for each construct. Estimated copy number is indicated for each mouse line (see “Experimental Procedures” for an explanation of the size of hybridization signals and how the copy number was determined). The strong hybridization bands at the 2.4- and 2-kb region in many of the lines correspond to multiple copies of the transgene. Lack of hybridization in nontransgenic controls (NTG) indicates the specificity of the luciferase probe. Molecular size markers are indicated in base pairs to the left.

**FIGURE 6.** The 50-bp (−186/−135) SP-10 insulator acts as a minimal enhancer blocker in transiently transfected cells. Constructs 1–7 from Fig. 5A were used in enhancer-blocking assays. COS cells were transiently transfected with circular plasmids, and cells were harvested 48 h later. The resulting luciferase activities were normalized and plotted. Note the high CMV enhancer activity in the absence of an insulator (construct 1). The presence of one (construct 2) or three copies (construct 3) of the 50-bp SP-10 insulator between the enhancer and the core promoter reduced luciferase activity by 2.9- and 13-fold, respectively. Surprisingly, when inserted in the opposite orientation (construct 4), the 50-bp SP-10 insulator lost the ability to block the CMV enhancer. The construct containing the full-length 319 bp (−408/−92) SP-10 insulator (construct 5) was used as a positive control for enhancer blocking. When placed upstream of the CMV enhancer, the 50-bp SP-10 insulator failed to block enhancer activity, thus proving that the 50-bp fragment is a true insulator and not a transcriptional repressor (constructs 6 and 7). Normalized luciferase values obtained from six independent experiments are plotted. The mean values with S.E. are indicated.

A stuffer fragment containing unrelated DNA failed to inhibit CMV enhancer activity (data not shown). Enhancer blocking by the 50-bp minimal insulator was less efficient compared with that of the 319-bp full-length SP-10 insulator, suggesting that the −408/−186 region contains additional cis-elements for enhancer blocking. Three tandem copies of the 50-bp fragment, however, showed an additive effect on enhancer blocking and reduced luciferase values 13-fold. In contrast, when placed in the opposite orientation (−135/−186), the minimal insulator altogether lost its ability to block the upstream enhancer (Fig. 6). Finally, for a definitive proof of insulator function, the 50-bp SP-10 fragment was placed upstream of the CMV enhancer. A true insulator will not interfere with enhancer activity when placed upstream of the enhancer, whereas a transcriptional repressor would be expected to repress the enhancer from this location. Enhancer-blocking assays showed that the 50-bp SP-10 fragment did not block enhancer activity when placed upstream of the CMV enhancer in either orientation (Fig. 6). These data provide proof of the insulator function of the 50-bp SP-10 fragment. Taken together, the above results showed that the 50-bp (−186/−135) SP-10 fragment functions as a minimal insulator.

The 50-bp Minimal Insulator Is both Necessary and Sufficient for Enhancer Blocking—Since the 50-bp insulator originates from a testis-specific gene, we investigated if the enhancer-blocking activity of the SP-10 insulator was restricted to the
context of the SP-10 core promoter elements only. To test this, we replaced the −91/+28 SP-10 core promoter in constructs 1, 2, and 4 of Fig. 5A with the heterologous −60/+31 TK core promoter (constructs 8−10; Fig. 5A). Nearly identical enhancer-blocking activities were observed in the context of either core promoter, suggesting that the mechanism of enhancer blocking by the SP-10 minimal insulator is not limited to the context of the native SP-10 core promoter (Fig. 7A). Consistent with the notion of a generic mechanism, the reverse orientation of the 50-bp SP-10 insulator failed to block enhancer activity in the context of the TK promoter as well (construct 10; Figs. 5 and 7A). Thus, the above experiment using the TK core promoter proved that the 50-bp SP-10 insulator is both necessary and sufficient for enhancer blocking.

A comparison of the 50-bp SP-10 minimal insulator and the chicken β-globin cHS4 insulator 42-bp FII portion (7) showed no sequence similarities. FII contains the canonical consensus site for CTCF, as do other well characterized vertebrate insulators, but the nucleotide sequence of the 50-bp SP-10 insulator does not bear any resemblance to the consensus CTCF site (Fig. 7B). The FII insulator is GC-rich and contains two CpG dinucleotides, whereas the 50-bp SP-10 insulator is devoid of CpG dinucleotides. In fact, the 50-bp SP-10 insulator is A-rich (36% of nucleotides are A; FII contains 19% A) (Fig. 7B).

Test for Insulator Function in a Physiological Context—The above enhancer-blocking assays were performed using circular plasmids and transient transfection assays in COS cells. Although there is evidence that plasmids do chromatinize following transfection, the higher order structure of chromatin may be considerably different in vivo. To test for enhancer-blocking function in native chromatin and to evaluate the potential of the 50-bp insulator in preventing ectopic expression of the testis-specific SP-10 gene, we generated transgenic mice. The experimental paradigm consisted of first generating transgenic mice in which a luciferase reporter transgene was placed under the control of CMV enhancer and the SP-10 core promoter (construct 1; Fig. 5A). This was meant to mimic a situation wherein a foreign enhancer gains direct access to the SP-10 core promoter in vivo. Ubiquitous expression of the transgene was expected in these uninsulated mice because of the panactive nature of the CMV enhancer (29). A second set of transgenic mice harboring the
The present study is aimed at understanding how an insulator might regulate testis-specific transcription of the mouse SP-10 gene. We show that an insulator located in the proximal promoter of the SP-10 gene tethers the gene to the nuclear matrix in non-germ cells. This tethering sequesters the SP-10 gene expression away from other somatic tissues, resulting in a more localized expression pattern. The insulator function in vivo is not orientation-dependent, as demonstrated by the results presented in this study.
A Novel Vertebrate Insulator

The SP-10 insulator is facultative; therefore, its function must be modulated in the testis. What might be the nature of the modulation that untethers the SP-10 insulator from the nuclear matrix in the spermatogenic cells? Lack of CpG dinucleotides rules out DNA methylation as a possible mechanism. Instead, our results implicate TDP-43 in SP-10 insulator function. TDP-43 has been shown to be a transcriptional repressor and a splicing regulator (25, 26). Insulator function may be an additional role for TDP-43. It is not uncommon for transcription factors to have multiple roles in gene regulation.

Importantly, the present study shows that TDP-43 is part of the nuclear matrix where a putative insulator protein would be expected to be present (Fig. 2). Our previous study showed that mutation of TDP-43 binding sites within the SP-10 insulator resulted in erratic expression (17). Consistent with this, knockdown of TDP-43 resulted in the release of SP-10-insulated transgene in a stable cell line model (Fig. 4). Taken together, this predicts that TDP-43, which is present in testis, must undergo either post-translational changes or splice variation in the male germ line to alter its function and permit SP-10 gene expression. In fact, the present study shows that TDP-43 lacking the Gly domain can no longer be effective as an enhancer blocker (Fig. 3B). Our previous study showed that round spermatids express only the short form of TDP-43 mRNA compared with other cell types (Figs. 3 and 4 in Ref. 17). In addition, recent studies suggest that phosphorylation and ubiquitylation modify TDP-43 function in neuronal cells (34). Our future studies will focus on the biology of TDP-43 in spermatogenic cells. It must be noted here that knockdown of TDP-43 in GC2 tissue culture cells did not result in expression of SP-10 as determined by Western blots (data not shown). Our explanation for this is that in addition to release of repression, spermatid-specific activators will be required for the activation of the SP-10 gene.

The significance of this study is that it characterized a 50-bp subfragment of the SP-10 insulator and showed its enhancer-blocking properties in a true physiological context. To our knowledge, this is the first demonstration in a mouse model that such a small 50-bp vertebrate insulator functions as a tissue-selective enhancer blocker. The evidence that the 50-bp minimal insulator can silence an otherwise ubiquitously expressed transgene in somatic tissues (Fig. 9) highlights the transcriptional regulatory strategies used by testis-specific genes. In light of its ability to act as an enhancer blocker in the heterologous context of the TK core promoter (Fig. 7), we propose a role for the 50-bp SP-10 insulator, which is conserved between mice and humans, in gene therapy vectors and for targeted gene delivery to testis, in particular.

The enhancer-blocking activity of the SP-10 insulator observed in the somatic tissues of transgenic mice (Fig. 9) begs the question as to whether the SP-10 gene requires protection from the enhancer(s) of other gene(s) located nearby. In fact, on mouse chromosome 9 (NCBI, Map Viewer), the testis-specific SP-10 gene is flanked by the ubiquitously expressed check point kinase 1 (Chek1) gene (35) on one side (less than 10 kb away) and a seminal vesicle-specific gene Gmi191 (symbol A6330095E13Rik), which codes for SSLP-1 (secreted seminal vesicle Ly-6 protein 1) (36) on the other side (50 kb away). Thus, contiguous location of three genes with divergent tissue-spe-
cific expression patterns (1, 35, 36) warrants the presence of boundary elements capable of preventing inappropriate enhancer-promoter interactions. Therefore, it is conceivable that the SP-10 insulator characterized in the present study may play a role in protecting the SP-10 gene from the influence of the enhancers of neighboring genes. The loci for mouse and human SP-10 genes on chromosome 9 and chromosome 11, respectively, are syntenic, and the proximal promoters of mouse and human SP-10 genes exhibit sequence homology (1), including the 50-bp insulator region. Evolutionary conservation of the sequence highlights the significance of SP-10 insulator function.

The SP-10 insulator, however, differs from other vertebrate insulators in terms of its location in the gene. Typically, insulators are located at gene boundaries upstream of the promoter and enhancer region. The HS4 chicken globin insulator is located several kilobase pairs upstream of the globin gene cluster and the LCR region (5). At the H19 locus, the imprinting control region containing the insulator is located far from the promoters of the imprinted H19 and Igf2 genes (10). Similarly, in the human apoB gene locus, the boundary element is located 55 kb upstream of the promoter (15). In contrast, the SP-10 insulator is located in close proximity to the core promoter (2). We hypothesize that the proximal promoters of some testis-specific genes have adapted insulator function during evolution. Testis-specific genes code for unique proteins that make up the acrosome, the neck, and the flagellum of spermatozoa and as such are not required by any other cell type. Sequestration of the core promoter by tethering to the nuclear matrix must have offered a fail-proof mechanism for these genes to remain silent in somatic tissues.

In conclusion, the present study provides a working model for testis-specific gene transcription. Although previous promoter analyses in transgenic mice have established that short proximal promoters of many testis-specific genes are sufficient for recapitulation of tight regulation shown by the endogenous genes (2, 37–41), the actual mechanism by which testis-specific promoters operate is poorly understood. Previous studies established a correlation between the methylation status of the promoter and gene expression for some testis-specific genes (42). A recent study implicated CTCF and Sp3 in mediating somatic repression of a testis-specific gene (43). Our study shows how an insulator located in a CpG-free proximal promoter can regulate testis-specific gene transcription. Specific cis-regulatory information on proximal promoters is expected to contribute toward restructuring gene regulatory networks (44–46). The compact nature of testis-specific promoters in general and the apparent ability of insulators to silence gene expression in somatic tissues (present study) prompt the prediction that insulators may be more prevalent in testis-specific gene promoters than has been realized.

Acknowledgments—We thank the anonymous reviewers for helpful comments and members of the Reddi laboratory for useful discussions and advice.

REFERENCES

1. Reddi, P. P., Flickinger, C. J., and Herr, J. C. (1999) Biol. Reprod. 61, 1256–1266.
2. Reddi, P. P., Shore, A. N., Shapiro, J. A., Anderson, A., Stoler, M. H., and Acharya, K. K. (2003) Dev. Biol. 262, 173–182.
3. Kellum, R., and Schedl, P. (1991) Cell 64, 941–950.
4. West, A. G., Gaszner, M., and Felsenfeld, G. (2002) Genes Dev. 16, 271–288.
5. Chung, I. H., Whitney, M., and Felsenfeld, G. (1993) Cell 74, 505–514.
6. Pikaart, M. J., Recillas-Targa, F., and Felsenfeld, G. (1998) Genes Dev. 12, 2852–2862.
7. Bell, A. C., West, A. G., and Felsenfeld, G. (1999) Cell 98, 387–396.
8. West, A. G., Huang, S., Gaszner, M., Litt, M. D., and Felsenfeld, G. (2004) Mol. Cell 16, 453–463.
9. Bell, A. C., and Felsenfeld, G. (2000) Nature 405, 482–485.
10. Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000) Nature 405, 486–489.
11. Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S., and Lee, J. T. (2002) Science 295, 345–347.
12. Robinett, C. C., O’Connor, A., and Dunaway, M. (1997) Mol. Cell Biol. 17, 2866–2875.
13. Saitoh, N., Bell, A. C., Recillas-Targa, F., West, A. G., Simpson, M., Pikaart, M., and Felsenfeld, G. (2000) EMBO J. 19, 2315–2322.
14. Li, Q., and Stamatoyannopoulos, G. (1994) Blood 84, 1399–1401.
15. Antes, T., Nacmici, S. J., Fournier, R. E. K., and Levy-Wilson, B. (2001) Biochemistry 40, 6731–6742.
16. Filippova, G. N., Thienes, C. P., Penn, B. H., Cho, D. H., Hu, Y. J., Moore, J. M., Klessert, T. R., Lobanenkov, V. V., and Tapscott, S. J. (2001) Nat. Genet. 28, 335–343.
17. Acharya, K. K., Govind, C. K., Shore, A. N., Stoler, M., and Reddi, P. P. (2006) Dev. Biol. 295, 781–790.
18. Mirkovitch, J., Mirault, M. E., and Laermans, U. K. (1984) Cell 39, 223–232.
19. Yusufzai, T. M., and Felsenfeld, G. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 8620–8624.
20. Bellvé, A. R., Cavicchia, J. C., Millette, C. F., O’Brien, D. A., Bhatnagar, Y. M., and Dym, M. (1977) J. Cell Biol. 74, 68–85.
21. Hogan, B. (1994) Manipulating the Mouse Embryo: A Laboratory Manual, pp. 226–244, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
22. Gerasimova, T. I., Byrd, K., and Corces, V. G. (2000) Mol. Cell 6, 1025–1035.
23. Gilbert, N., Gilchrist, S., and Bickmore, W. A. (2005) Int. Rev. Cytol. 242, 283–336.
24. Gaszner, M., and Felsenfeld, G. (2006) Nat. Rev. Genet. 7, 703–713.
25. Ou, S. H., Wu, F., Harrich, D., Garcia-Martinez, L. F., and Gaynor, R. B. (1995) J. Virol. 69, 3584–3596.
26. Buratti, E., and Baralle, F. E. (2001) J. Biol. Chem. 276, 36337–36343.
27. Hancock, R. (2000) Chromosoma 109, 219–225.
28. Boshart, M., Weber, F., Jahn, G., Dorsch-Hasler, K., Fleckenstein, B., and Schaffner, W. (1985) Cell 41, 521–530.
29. Schmidt, E. V., Christop, G., Zeller, R., and Leder, P. (1990) Mol. Cell Biol. 10, 4406–4411.
30. Geus, M. E., Fletcher, C., Block, G. D., Straume, M., Copeland, N. G., Jenkins, N. A., Kay, S. A., and Day, R. N. (1997)Curr. Biol. 7, 758–766.
31. Ikawa, M., Kominnami, K., Yoshimura, Y., Tanaka, K., Nishimune, Y., and Okabe, M. (1995) Dev. Growth Differ. 37, 455–459.
32. Marber, M. S., Mestrel, R., Shen-Hua Chi., Sayen, M. R., Yellon, D. M., and Dillman, W. H. (1995) J. Clin. Invest. 95, 1446–1456.
33. Yusufzai, T. M., Tagami, H., Nakatani, Y., and Felsenfeld, G. (2004) Mol. Cell 13, 291–298.
34. Neumann, M., Sampathu, D. M., Kwong, L. K., Truax, A. C., Micsenyi, M. C., Chou, T. T., Bruce, J., Schuck, T., Grossman, M., Clark, C. M., McCluskey, I. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q., and Lee, V. M. (2006) Science 314, 130–133.
35. Sanchez, Y., Wong, C., Thoma, R. S., Richman, R., Wu, Z., Pinnica-Worms, H., and Elledge, S. J. (1997) Science 277, 1497–1501.
A Novel Vertebrate Insulator

36. Li, S. H., Lee, R. K., Lin, M. H., Hwu, Y. M., Lu, C. H., Chen, Y. J., Chen, H. C., Chang, W. H., and Chang, W. C. (2006) Reproduction 132, 493–500
37. Robinson, M. O., McCarrey, J. R., and Simon, M. I. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 8437–8441
38. Langford, K. G., Shai, S. Y., Howard, T. E., Kovac, M. J., Overbeek, P. A., and Bernstein, K. E. (1991) J. Biol. Chem. 266, 15559–15562
39. Zambrowicz, B. P., Harendza, C. J., Zimmermann, J. W., Brinster, R. L., and Palmiter, R. D. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5071–5075
40. Dix, D. J., Rosario-Herrle, M., Gotoh, H., Mori, C., Goulding, E. H., Barrett, C. V., and Eddy, E. M. (1996) Dev. Biol. 174, 310–321
41. Li, S., Zhou, W., Doglio, L., and Goldberg, E. (1998) J. Biol. Chem. 273, 31191–31194
42. Geyer, C. B., Kiefer, C. M., Yang, T. P., and McCarrey, J. R. (2004) Biol. Reprod. 71, 837–844
43. Kim, M., Li, D., Cui, Y., Mueller, K., Chears, W. C., and DeJong, J. (2006) J. Biol. Chem. 281, 34288–34298
44. Levine, M., and Davidson, E. H. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 4936–4942
45. Istrail, S., and Davidson, E. H. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 4954–4959
46. Smith, A. D., Sumazin, P., Xuan, Z., and Zhang, M. Q. (2006) Proc. Natl. Acad. Sci. 103, 6275–6280