

Tumor Suppressor Smad4 Is a Transforming Growth Factor β -Inducible DNA Binding Protein

JONATHAN M. YINGLING,[†] MICHAEL B. DATTO, CAROLYN WONG, JOSHUA P. FREDERICK, NICOLE T. LIBERATI, AND XIAO-FAN WANG*

Department of Pharmacology and Cancer Biology, Duke University, Durham, North Carolina 27710

Received 8 July 1997/Returned for modification 18 August 1997/Accepted 22 September 1997

Members of the Smad family of proteins are thought to play important roles in transforming growth factor β (TGF- β)-mediated signal transduction. In response to TGF- β , specific Smads become inducibly phosphorylated, form heteromers with Smad4, and undergo nuclear accumulation. In addition, overexpression of specific Smad combinations can mimic the transcriptional effect of TGF- β on both the plasminogen activator inhibitor 1 (PAI-1) promoter and the reporter construct p3TP-Lux. Although these data suggest a role for Smads in regulating transcription, the precise nuclear function of these heteromeric Smad complexes remains largely unknown. Here we show that in Mv1Lu cells Smad3 and Smad4 form a TGF- β -induced, phosphorylation-dependent, DNA binding complex that specifically recognizes a bipartite binding site within p3TP-Lux. Furthermore, we demonstrate that Smad4 itself is a DNA binding protein which recognizes the same sequence. Interestingly, mutations which eliminate the Smad DNA binding site do not interfere with either TGF- β -dependent transcriptional activation or activation by Smad3/Smad4 cooverexpression. In contrast, mutation of adjacent AP1 sites within this context eliminates both TGF- β -dependent transcriptional activation and activation in response to Smad3/Smad4 cooverexpression. Furthermore, concatemered AP1 sites, in isolation, are activated by Smad3/Smad4 cooverexpression and, to a certain extent, by TGF- β . Taken together, these data suggest that the Smad3/Smad4 complex has at least two separable nuclear functions: it forms a rapid, yet transient sequence-specific DNA binding complex, and it potentiates AP1-dependent transcriptional activation.

Transforming growth factor- β (TGF- β) is a multipotent peptide hormone which regulates a diverse array of biological processes (25). The involvement of TGF- β in the pathogenesis of several diseases has resulted in intense investigation of its molecular mechanism of signal transduction (26). Several years ago the signaling receptors for TGF- β were cloned and found to be transmembrane serine/threonine kinases termed “type I” and “type II” receptors (3, 11, 21). Although the molecular nature and mechanism of activation for these TGF- β receptors at the cell surface has been described (33, 34), the intracellular pathways which transduce the TGF- β signal from the membrane to the nucleus have only recently begun to be elucidated.

Genetic studies in *Drosophila* (28) and *Caenorhabditis elegans* (27) identified a conserved family of proteins as playing a critical role in TGF- β superfamily signaling pathways downstream of the receptors. Mammalian homologs of these proteins, now referred to as “Smads” (9), were subsequently cloned and characterized (1, 24). Studies in *Xenopus* embryos have revealed a functional division between the mammalian Smad proteins. Smad1 (12, 22, 32) and Smad5 (30) have been shown to induce ventral mesoderm and thus mediate the BMP signal, while Smad2 transduces TGF- β signals and induces dorsal mesoderm (2, 12). The distantly related Smad4 protein, which was originally identified as a tumor suppressor protein on chromosome 18q (13), induces both ventral and dorsal mesoderm and thus mimics TGF- β and BMP signals (38). Smad4 has been shown to associate with Smad1 in response to BMP and with Smad2 in response to TGF- β and thus is a

common component of these signal transduction pathways (19). The Smads have been found to be inducibly phosphorylated in response to TGF- β and BMP, and the ligand-specific nature of the Smads has been confirmed by these studies. Smad2 and Smad3 are specifically phosphorylated in response to TGF- β (10, 20, 36, 37), while Smad1 is phosphorylated in response to BMP (14, 36). Phosphorylation of the Smads is followed by their heteromerization with Smad4 (19) and the subsequent accumulation of heteromeric Smad complexes in the nucleus (2, 15, 22). Recently, the type I receptor was found to be the kinase responsible for ligand-inducible phosphorylation of C-terminal serine residues of Smad1 in response to BMP (18) and Smad2 in response to TGF- β (23). The C-terminal domain in Smad1 and Smad4 has been shown to possess transcriptional activation activity in the context of Gal4 DNA binding domain fusion proteins (22), thus providing the first indication of a nuclear function for the Smad proteins. Subsequently, overexpression of specific Smad combinations has been found to mimic the transcriptional effect of TGF- β on both the plasminogen activator inhibitor 1 (PAI-1) promoter and the reporter construct p3TP-Lux (6, 19, 22, 23, 37). Smad4 has been shown to be required for this transcriptional activity since Smad4-deficient cell lines are nonresponsive but can be rescued with Smad4 expression (8, 19). Interestingly, the homomeric and heteromeric interactions between Smad3 and Smad4 correlate with their ability to transcriptionally activate the PAI-1 promoter (35). Furthermore, naturally occurring mutations interfere with the ability of Smad4 to associate with Smad3 (37). Although these data suggest a role for Smads in regulating transcription, the precise nuclear function of the heteromeric Smad complexes remains largely unknown.

Here we demonstrate that Smad3 and Smad4 participate in a DNA binding complex on a fragment of the p3TP-Lux reporter and that Smad4 is the DNA binding component of this complex. In the context of this reporter, the Smad binding site

* Corresponding author. Mailing address: Department of Pharmacology and Cancer Biology, Duke University, Durham, NC 27710. Phone: (919) 681-4861. Fax: (919) 681-7152. E-mail: wang@galactose.mc.duke.edu.

[†] Present address: Howard Hughes Medical Institute, Cell Biology Department, Vanderbilt University, Nashville, Tenn.

is not required for transcriptional activation in response to TGF- β or Smad3/Smad4 cooverexpression. However, adjacent AP1 sites are both necessary and sufficient for activation by TGF- β and Smad3/Smad4 cooverexpression. We also show that an endogenous promoter, the PAI-1 promoter, contains a Smad binding site. Thus, the ability of Smad3/Smad4 to directly bind DNA may have physiological relevance in regulating transcription of TGF- β -responsive genes. However, in some contexts, transcriptional activation by Smad cooverexpression may be mediated through the AP1 transcription factor complex, as demonstrated here for the p3TP-Lux reporter.

MATERIALS AND METHODS

Cell culture. Mink lung epithelial cells (Mv1Lu) were obtained from the American Type Culture Collection (ATCC) and maintained in Dulbecco modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS), penicillin, and streptomycin and nonessential amino acids. COS cells were maintained in DMEM with 10% FBS, penicillin, and streptomycin.

Plasmid construction. Flag-tagged human Smad4, human Smad3, and human Smad2 were the generous gift of Rik Derynck. Expression vectors for Smad3 WT and Smad3MT (3S \ddot{A}) were generated by PCR using the following primers: for Smad3 WT, GGATCCGCGATGTCGTCCTTCAC (5' primer) and GGATCCTAAGACACACTGGAACAGC (3' primer); for Smad3MT (3S \ddot{A}), the 5' primer was the same as that for Smad3 WT and the 3' primer was GGATCCTAAGCCACAGCTGCACAGCGGATGCTTGG. The resulting *Bam*HI fragments were cloned in frame with the hemagglutinin (HA) tag in pCGN (31), p3TP-Lux (33) and pGL2-T+I have been previously described (7). The luciferase reporter constructs 4 \times WT, 4 \times SBS Dbl mutant, and 4 \times AP1 Dbl mutant were created by using the following oligonucleotides: for 4 \times WT, GGA TGAGTCAGACACTCTGGGTGTCGGGAAG and TCCCTTCCGGACAG CACAGGTGTCTGACTCA; for 4 \times AP1 Dbl mutant, GGATCAGCAGAG CACTCTGGGTGTCGGGAAG and TCCCTTCCGGACAGCCAGAGGTG TCTGTGTA; for 4 \times SBS Dbl mutant, GGATGAGTCACTGCATTCTGGC TGTCGGGAAG and TCCCTTCCGGACAGCCAGAATGCAGTCACTCA. 2 \times directional constructs were created by first phosphorylating the above oligonucleotide sets, annealing, and ligating in the presence of a 0.5 \times molar ratio of phosphorylated and annealed linker oligonucleotides: short linker, GGCTCGA GAGATCT; long linker 5', TCCAGATCTCTCGAGCC; long linker 3', GGA AGATCTCTCGAGCC. The resulting ligation product was digested with *Bgl*III and cloned into the *Bgl*III site of pGL2-T+I. Constructs which contained two inserts in a backward orientation were then cut with *Xho*I and *Eco*RV. The two site-containing fragments were then cloned into the *Xho*I and *Sma*I sites in reporter constructs which contained two inserts in a forward orientation to produce constructs with four inserts in the same orientation. The 7 \times AP1 reporter construct was made by using the oligonucleotides: GGATGAGTCA GAC and GTCTGACTCATCC, which correspond to the AP1 sites in p3TP-Lux. These oligonucleotides were phosphorylated, annealed, and cloned into the *Sma*I site of PBSK. A resulting construct which contained seven concatemered AP1 sites was digested with *Bam*HI and *Eco*RV, and the AP1 site-containing fragment was cloned into the *Bgl*III and *Sma*I sites of pGL2 T+I. GST-Smad3 and GST-Smad4 were created by PCR from plasmid templates using the following primers: for Smad3, the 5' primer was CCGGATCCCGATGTCGTCCT CCTGCCCTTCAC and the 3' primer was same as that for Smad3 WT; for Smad4, the 5' primer was CCGGATCCCGATGGACAATATGTCTATTACG and the 3' primer was GGATCCTCAGTCTAAAGGTTGTGG. The resulting *Bam*HI fragments were cloned in frame into pGEX3X-HMK (Pharmacia).

Electrophoretic mobility shift assays (EMSA). Extracts were prepared from approximately 2×10^6 COS cells transiently transfected with Smad expression constructs by using a standard DEAE-Dextran transfection protocol (33) or from approximately 2×10^6 Mv1Lu control cells or cells treated with 100 pM TGF- β 1 for 30 min after 2 h of serum starvation. Cells were then lysed, and nuclear extracts were prepared as previously described (7). Whole-cell extracts prepared in a lysis buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5% Nonidet P-40, 50 mM NaF, 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM sodium orthovanadate, and protease inhibitors gave identical gel shift results. For phosphatase treatment, 2 U of CIAP (Boehringer Mannheim) and 0.06 U of PAP (Boehringer Mannheim) were added to 100 μ l of nuclear extract (prepared without phosphatase inhibitors), and the mixtures were incubated at 37 $^{\circ}$ C for 30 min in the presence or absence of the phosphatase inhibitors 1 mM sodium orthovanadate, 10 mM NaF, 10 mM β -glycerophosphate, and 0.2 mM sodium molybdate. [α - 32 P]dTTP Klenow-labeled oligonucleotides used for probes are shown in Fig. 2A. Alternatively, the wild-type probe was created by digesting p3TP-Lux with *Sph*I and *Nde*I and [α - 32 P]dTTP Klenow labeling. The PAI-1 promoter probe was obtained by restriction digestion with *Nco*I and *Eag*I and [α - 32 P]dTTP Klenow labeling. Gel shift conditions were as follows: 1.5 μ l of nuclear extract (or 3 μ l of whole-cell extract) containing approximately 3 μ g of protein, 1 μ g of dI-dC, and 0.5 ng of the probe labeled to an activity of 10,000 to 40,000 cpm/0.5 ng was brought to a final volume of 15 μ l

by using a hypotonic lysis buffer as previously described (7). For supershift analysis a rabbit polyclonal Smad4 antibody was created against full-length GST-Smad4 by standard protocols. Preimmune serum was from the same rabbit. Anti-HA antibody was obtained from Boehringer Mannheim, the Pan Fos antibody was obtained from Santa Cruz (K-25), and anti-Flag antibody (M2) was obtained from Kodak IBI. Two microliters of each antibody was used for supershifts. For gel shifts with eluted GST-Smad3 and GST-Smad4, approximately 100 ng of protein in 1 μ l of a buffer containing 100 mM Tris (pH 8), 120 mM NaCl, and 25 μ M glutathione was used. Complexes were resolved on a 6% acrylamide-0.04% bisacrylamide-0.5 \times Tris-borate-EDTA (TBE) gel as previously described (7), except for the HA supershift panel of Fig. 2B, which was resolved on a 6% acrylamide-0.2% bisacrylamide-0.5 \times TBE gel.

Methylation interference. Methylation interference probes were prepared as above with the following exceptions. Four microliters of DMS was added to 100 μ l of the Klenow labeling reaction mixture which contained 1 μ g of DNA. After a 5-min room temperature incubation, 40 μ l of 1.5 M sodium acetate and 1 M β -mercaptoethanol was added. The probe was then precipitated with the addition of 0.5 ml of 100% ethanol (ETOH). Probe was then gel purified and used in EMSA as described above. After a short -80 $^{\circ}$ C exposure of the unfixed, undried EMSA gel, shifted complexes were cut out and bound probe was electroeluted, precipitated, and resuspended in 100 μ l of 1 M piperidine. Samples were then heated to 90 $^{\circ}$ C for 30 min, and piperidine was subsequently removed by several rounds of lyophilization and resuspension in distilled water. The resulting cleaved products were resolved on a urea-acrylamide sequencing gel.

Luciferase assays. Transfections were performed by using a standard DEAE-dextran transfection protocol (33). Luciferase assays were performed as previously described (7). All transfections were normalized to β -galactosidase activity by cotransfection of 0.5 μ g of a cytomegalovirus β -galactosidase (CMV- β -gal) expression vector. Twelve hours after transfection, 100 pM TGF- β 1 was added and TGF- β -induced luciferase activity was assayed after 20 to 24 h. Quantities of DNA transfected are detailed in the figure legends.

Western blot analysis. Proteins from COS-transfected lysates were resolved by sodium dodecyl sulfate-8% polyacrylamide gel electrophoresis and transferred to Immobilon (Millipore). The blots were blocked in the B/P solution (50 mM Tris [pH 7.5], 150 mM NaCl, 0.1% Tween 20) containing 2% milk. Primary antibody (HA, Flag, or Smad4) was added in the B/P solution at 1:1,000 for 1 h at room temperature. The blots were washed three times with the B/P solution, and the appropriate secondary antibody (Bio-Rad) was added (goat anti-mouse for HA/Flag and goat anti-rabbit for Smad4) for 1 h at room temperature. After three washes with the B/P solution, the blots were developed with ECL (Amersham) and exposed on Kodak XAR5 film.

RESULTS

Smad3/Smad4 cooverexpression regulates transcription.

The p3TP-Lux luciferase reporter is a well-described and widely used artificial promoter construct which was empirically designed to have maximal responsiveness to TGF- β (33). p3TP-Lux has a 31-nucleotide, AP1 site-containing region of the collagenase promoter, concatemered 5' to an \sim 400-nucleotide region of the PAI-1 promoter followed by 70 bp of the adenovirus E4 promoter (Fig. 1A). In agreement with previous findings, we observed a transcriptional activation of p3TP-Lux by both Smad3/Smad4 cooverexpression and TGF- β treatment in Mv1Lu cells (Fig. 1B). In contrast, Smad2 fails to activate transcription of p3TP-Lux when cooverexpressed with Smad4. To define the Smad-responsive region of p3TP-Lux, we created a reporter construct comprised only of the 31-nucleotide AP1 site-containing region concatemered 5' to a minimal promoter. This 4 \times WT reporter (Fig. 1A) not only is TGF- β responsive but is also activated in response to Smad3/Smad4 cooverexpression (Fig. 1C). Thus, this 31-nucleotide repeat contains a DNA sequence which is both TGF- β and Smad responsive.

Smad3/Smad4 participates in a DNA binding complex. To determine if Smad3/Smad4 cooverexpression changes the DNA binding complexes on this 31-nucleotide fragment, we performed gel shifts using a probe consisting of two copies of the 31-nucleotide repeat derived from p3TP-Lux, termed the 2.0 probe (Fig. 3A). When gel shifts were performed with this probe and extracts derived from COS cells cotransfected with epitope-tagged Smad3 and Smad4 (Fig. 2A), we observed not only an AP1-containing complex (complex I) but also a strong additional binding complex (complex II [lane 6]). Overexpres-

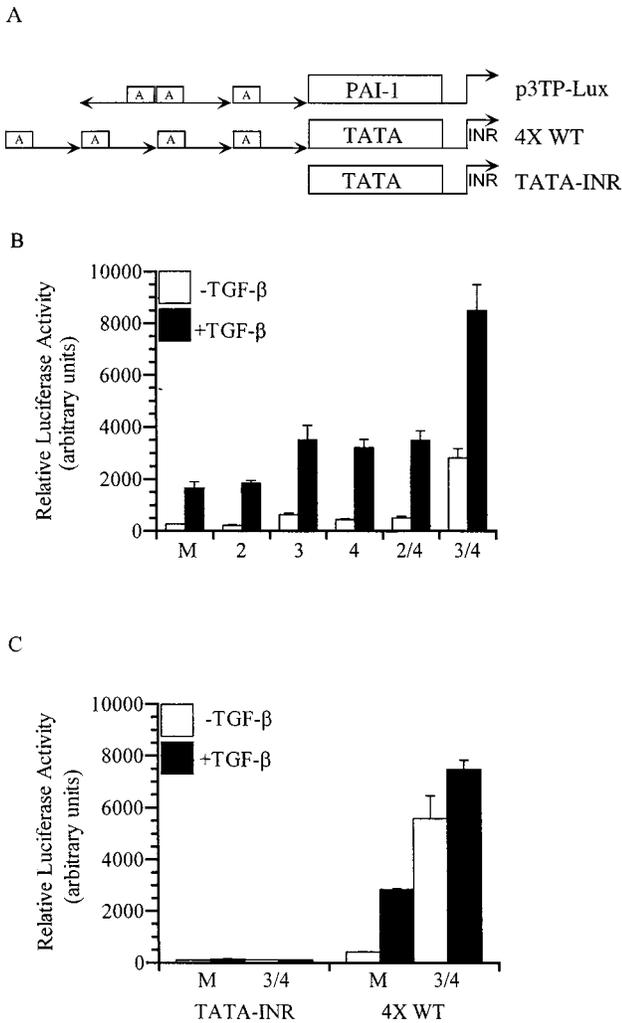


FIG. 1. Smad overexpression regulates transcription. (A) Diagram of the AP1-containing luciferase reporters. (B) p3TP-Lux is activated by both TGF-β and by Smad3/Smad4 cooverexpression. Mv1Lu cells (2×10^6) were cotransfected with 3 μg of 3TP-Lux and 1 μg of the indicated Smad expression construct, either alone (2–4) or in combination with 1 μg of a Smad4 expression construct (2/4 and 3/4). M, transfection with 2 μg of the empty expression vector pCGN. The total amount of DNA transfected was kept constant with the addition of pCGN. Twelve hours after transfection TGF-β was added, and after 20 to 24 h luciferase activity was measured. (C) The 4× wild-type (WT) reporter is activated by TGF-β and Smad3/Smad4 coexpression like p3TP-Lux. Luciferase assays were performed as for panel B with cotransfection of 3 μg of either the minimal TATA-INR reporter construct or the 4× WT reporter construct and either 2 μg of the vector alone (M) or the combination of 1 μg of Smad3 and 1 μg of Smad4 as indicated. After transfection and TGF-β treatment, luciferase activity was measured as above. Error bars represent the standard deviation for triplicate transfections in a single experiment.

sion of Smad3 alone produces a lower level of a complex with similar mobility (Lane 3). Likewise, overexpression of Smad4 produces a complex with similar mobility, as well as a slightly faster-migrating complex (lane 4). In contrast, Smad2/Smad4 coexpression does not produce this complex but appears similar to expression of Smad4 alone (lane 5).

One possible explanation for these observations is that Smad3 and Smad4 form a DNA binding complex. Overexpressed Smad3 alone or Smad4 alone could bind DNA with their endogenous Smad partner, whereas cooverexpression would produce a large amount of Smad3/Smad4 binding complex. To

test this hypothesis, supershift analysis was performed to determine if HA-tagged Smad3 or Flag-tagged Smad4 is present in the additional binding complex (complex II). As shown in Fig. 2B, both HA and Flag antibodies supershift this complex (lanes 7, 12, and 13). As expected, a pan-Fos family member antibody supershifts the faster-migrating AP1 complex (lane 2). This antibody, however, does not shift the Smad3/Smad4 complex (lane 8), suggesting that although the constitutive binding activity contains a Fos family member, the Smad3/Smad4 complex does not. Attempts to supershift the AP1 complex with a number of commercially available anti-Jun family antibodies were not successful due probably to the low affinities of these antibodies (data not shown). Finally, the complexes observed with Smad4 overexpression are all Smad4 containing as demonstrated by Flag supershifts (lane 5). Thus, Smad3 and Smad4, when overexpressed, participate in a DNA binding complex on sequences present in this region of p3TP-Lux.

Recently, the BMP-inducible phosphorylation sites of Smad1 and the TGF-β-inducible phosphorylation sites of Smad2 have been identified (18, 23). Smad3 contains analogous sites of potential phosphorylation at its C terminus. Based on this sequence homology, we created a phosphorylation-deficient mutant of Smad3, Smad3MT, and assayed its ability to participate with Smad4 in a DNA binding complex. Although the expression levels were similar to that of wild-type Smad3, Smad3MT was unable to form a DNA binding complex with Smad4 (Fig. 2A, lane 7). The results with this mutant suggest that an intact carboxyl terminus of Smad3 is essential for formation of the DNA binding complex. Based on the Smad4 crystal structure and interaction studies, this mutation possibly interferes with the ability of Smad3 to form homomeric or heteromeric complex with Smad4 and thus precludes formation of the DNA binding complex (14, 29).

Isolation of the Smad DNA binding element. To more precisely determine the DNA sequences to which the Smad3/Smad4-containing complex binds, we systematically mutated the 2.0 probe (Fig. 3A). As expected, mutation of the AP1 binding sites eliminated the Fos-containing shifted complex. The Smad3/Smad4 complex, however, was still present on the AP1 site mutant probe, although in somewhat decreased amounts (Fig. 3B, lane 3). This further suggests that the Smad3/Smad4 complex is not binding through AP1. We next designed three separate scanning mutants to encompass the entire 2.0 probe in search of the specific sequence which confers Smad3/Smad4 binding (Fig. 3A). As shown in Fig. 3B, scanning mutant 1 eliminates both the AP1 and the Smad3/Smad4 complexes, while scanning mutant 2 specifically eliminates the Smad3/Smad4 complex, leaving the AP1 complex intact. Scanning mutant 3 has no effect on the binding of either complex. Thus, the region necessary for Smad3/Smad4 complex binding lies within the bases mutated in scanning mutants 1 and 2.

Methylation interference was used to more precisely define which guanine residues within the 2.0 probe are contacted by the Smad3/Smad4 complex. The results shown in Fig. 3C confirm the mutagenesis results in that there is a single protected guanine residue that is located within the region predicted by the scanning mutagenesis. Both sites of this two-site probe have almost completely protected guanine residues. This suggests that both sites are being contacted in this single Smad3/Smad4 complex. Mutation of 6 nucleotides surrounding this protected guanine (GACACC) in either the 5' or the 3' site of the 2.0 probe was sufficient to eliminate Smad3/Smad4 binding (Fig. 3D), further indicating the requirement of a bipartite site for Smad3/Smad4 complex formation. In addition, a probe

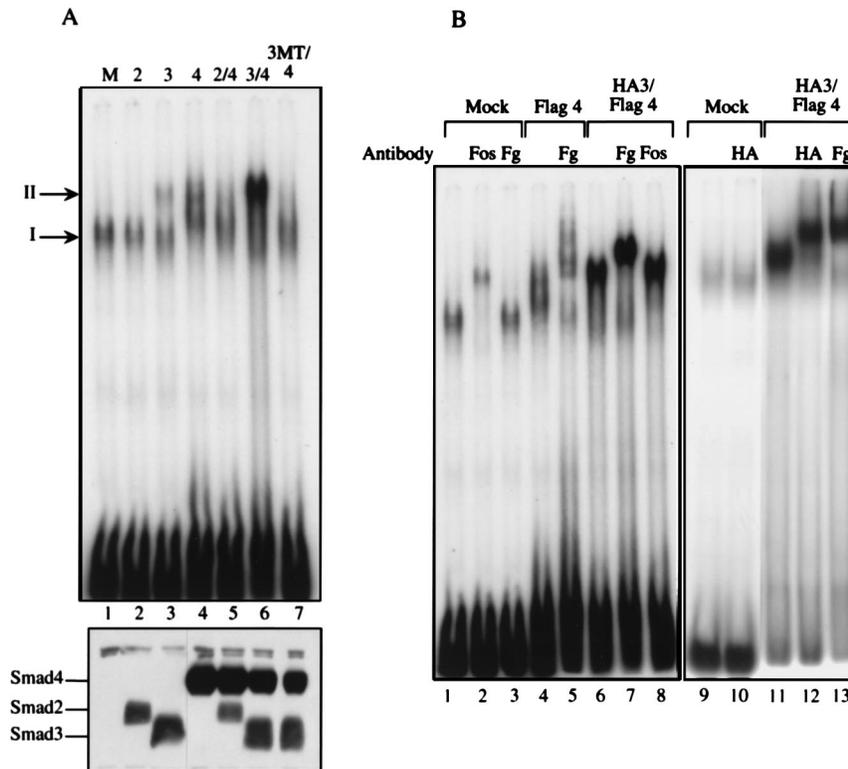


FIG. 2. Cooverexpression of Smad3 and Smad4 in COS cells changes the 2.0 probe binding profile. (A) Gel shifts were performed with the 2.0 probe derived from p3TP-Lux and extracts derived from COS cells transiently transfected with either 7 μ g of pCGN (lane 1), 2 μ g of Smad2 (lane 2), 2 μ g of Smad3 (lane 3), 5 μ g of Smad4 (lane 4), or 2 μ g of Smad2, -3, or -3MT in combination with 5 μ g of Smad4 (lanes 5 to 7). DNA amount was kept constant at 7 μ g with the added pCGN vector. An HA/Flag Western blot was performed to confirm expression of these proteins (lower panel). (B) Smad3 and Smad4 participate in a binding complex. EMSA supershifts using antibodies described in Materials and Methods were performed on COS cells transiently transfected as for panel A with the pCGN vector alone (Mock), Smad4-Flag (Flag 4), HA-Smad3/Smad4-Flag (HA3/Flag 4). Fos, pan-Fos antibody; Fg, Flag epitope antibody; HA, anti-HA epitope antibody.

containing only one of these 31-nucleotide repeats (one half of the probe used in these experiments) was completely unable to bind the Smad3/Smad4 complex in gel shift assays (data not shown).

Smad4 directly binds DNA. Having demonstrated that overexpressed Smad3/Smad4 participates in a DNA binding complex on the 2.0 probe, we next sought to determine if either Smad3 or Smad4 itself was directly binding this DNA sequence. Therefore, we generated GST fusions of both proteins and used these purified reagents in gel shifts with the 2.0 probe (Fig. 4). Although Smad3 is incapable of binding (lanes 1 and 2), GST-Smad4 directly binds the 2.0 wild-type (lanes 3 and 4) and AP1 mutant (lane 6) probes but does not bind the 2.0 Smad binding site mutant probe (lane 5). The DNA binding protein was confirmed to be Smad4 by antibody supershift analysis. The Smad4-specific immune antiserum alone produces a background DNA binding band (lane 7). The Smad4 antibody eliminates the two specific DNA binding complexes (lane 9), while the preimmune serum has no effect (lane 10). Thus, the complex seen with Smad3/Smad4 cooverexpression may be the result of a direct DNA interaction by Smad4. The ability of Smad4 to directly bind DNA may explain the additional shifted complex observed when Smad4 is overexpressed alone in COS cells (Fig. 2, lanes 4); it is Smad4 bound without endogenous Smad3. The ability of the Flag antibody to supershift this complex confirms the presence of Smad4 in this complex (Fig. 2B, lane 5).

TGF- β induces a Smad DNA binding complex in vivo. Mv1Lu cells are highly responsive to TGF- β and have been

used as a model system to define various aspects of TGF- β -mediated signal transduction. Thus, we used Mv1Lu cells as a model system to look in vivo for a Smad-containing DNA binding complex. Since the Smad proteins are cytoplasmic proteins which translocate to the nucleus in response to ligand-induced phosphorylation, an endogenous Smad-containing DNA binding complex would be predicted to be TGF- β inducible and phosphorylation dependent. To examine this question, we performed gel shifts with the 2.0 probe and nuclear extracts prepared from either TGF- β -treated or untreated Mv1Lu cells. In the absence of TGF- β treatment, Mv1Lu cells contain a constitutive binding complex similar to the Fos complex in COS cells. As in COS, this complex is supershifted by the Pan-Fos family member antibody to produce a slower-migrating band. (Fig. 5C, lane 2). Upon TGF- β treatment, a slower-migrating complex appears within 5 min, peaks in 15 min, and disappears after 4 h (Fig. 5A). This time course parallels the TGF- β -dependent phosphorylation kinetics of endogenous Smad proteins (36). In addition, the inducibly bound complex is sensitive to phosphatase treatment, suggesting that its binding is phosphorylation dependent (Fig. 5B). Thus, this inducible complex has the characteristics expected for a Smad-containing DNA binding complex. The presence of Smad4 in this TGF- β -inducible complex was confirmed by the ability of a Smad4-specific antibody to eliminate formation of this complex (Fig. 5C, lane 5). As expected, preimmune serum had no effect on binding (Fig. 5C, lane 6). In addition, the pan-Fos family member antibody did not affect the induced complex binding (Fig. 5C, lane 7) as previously observed with the tran-

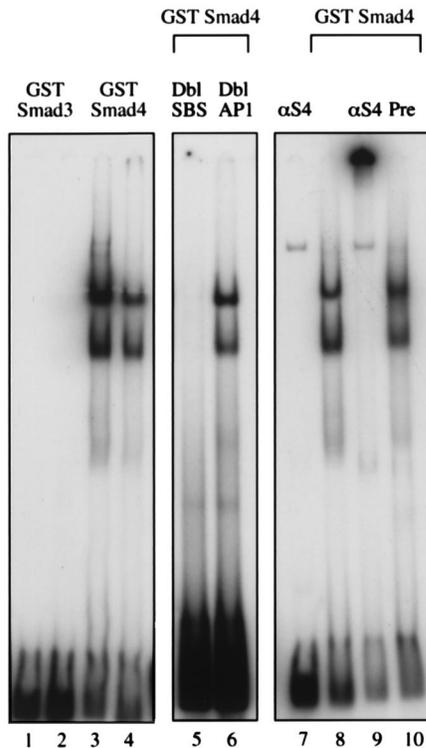


FIG. 4. Purified GST-Smad4, but not GST-Smad3, directly binds DNA. GST fusion protein construction is described in Materials and Methods. Gel shifts were performed with 2 μ l (lanes 1 and 3) or 1 μ l (lanes 2 and 4 to 10) of either GST-Smad3 (lanes 1 and 2) or GST-Smad4 (lanes 3 to 10) at a concentration of \sim 100 ng/ μ l, the indicated probes (diagrammed in Fig. 3A) and the Smad4 antibody (α S4) or preimmune serum (Pre).

Smad4 from COS lysates is able to bind the 2.0 probe, provide strong evidence that the inducible complex in Mv1Lu cells contains Smad3 and Smad4.

Functional analysis of the Smad DNA binding element. Having identified the specific region of the 2.0 probe which was capable of conferring Smad3/Smad4 binding, we examined the functional consequences of Smad binding site and AP1 site mutations in the context of the 4 \times WT reporter in Mv1Lu cells. As shown in Fig. 6A, the AP1 sites are critically important for induction by both TGF- β and Smad3/Smad4 coexpression. Surprisingly, mutation of the Smad binding site had no effect on induction by TGF- β or by Smad3/Smad4 coexpression. These results suggest that the heteromeric Smad3/Smad4 complex has at least two distinct nuclear activities. First, it rapidly forms a transient, sequence-specific DNA binding complex with unknown function, and secondly, it directly or indirectly potentiates AP1-dependent transcriptional regulation in the context of the p3TP-Lux reporter.

Smad overexpression activates transcription from AP1 DNA binding sites. To more firmly establish a role for the Smads in activating transcription from AP1 sites, the ability of Smad3 and Smad4 coexpression to activate a minimal reporter containing only concatemered AP1 sites was assayed. As shown in Fig. 6B, a reporter containing seven concatemered AP1 sites driving a minimal promoter consisting of a TATA box and initiator sequence is activated 10-fold by Smad3/Smad4 coexpression. The control TATA initiator construct is not activated by Smad overexpression. Although the 7 \times AP1 construct is not activated by TGF- β to the same extent as 3TP-Lux, it is consistently activated two- to threefold upon

TGF- β treatment. The ability of overexpressed Smads to transactivate AP1-mediated transcription appears to be specific, as a TGF- β -responsive, Sp1 binding site-driven promoter is not affected by Smad overexpression (20a). Thus, AP1 sites in isolation can support Smad-activated and, to some extent, TGF- β -activated transcription.

The endogenous PAI-1 promoter contains a Smad DNA binding element. The p3TP-Lux reporter is an artificial construct designed empirically for maximum TGF- β responsiveness. Although it has been instructive to biochemically define a novel binding function for Smad complexes, the question remains whether the Smad complex forms on endogenous or native promoter sequences. To address this question, we examined endogenous promoters that are known to be activated by TGF- β and Smad coexpression for their ability to bind a Smad3/Smad4-containing complex. One such DNA sequence is an 800-bp stretch of the PAI-1 promoter. The TGF- β -responsive region of the PAI-1 promoter has been described and surrounds a putative AP1 binding site (Fig. 7A) (16). The similarity of this endogenous sequence to that created in p3TP-Lux made it an ideal candidate for study. Using a probe that encompasses this AP1 site, we discovered that, indeed, Smad3 and Smad4 coexpression in COS cells leads to an additional DNA binding complex (Fig. 7B, complex II). Thus, the ability of Smad3/Smad4 to bind DNA which we originally defined in the context of p3TP-Lux may have physiological relevance in TGF- β 's ability to regulate endogenous promoters such as PAI-1. Studies are under way to investigate the functional role of this Smad binding region within the PAI-1 promoter.

DISCUSSION

An intracellular pathway for mediating TGF- β superfamily signals from the membrane to the nucleus has begun to be elucidated. This highly conserved pathway involves the Smad proteins, which are phosphorylated by the type I receptor, form heteromers with Smad4, and accumulate in the nucleus. In this study, we have investigated the molecular nature of the Smads' ability to transcriptionally activate the p3TP-Lux reporter. As a result of these studies, we have defined a novel function for Smad4 as a DNA binding protein and have demonstrated a functional relationship between the Smads and AP1-dependent transcriptional regulation.

Defining a new DNA binding function for Smad4. As has been previously reported, coexpression of Smad3 and Smad4 was found in this study to activate transcription from p3TP-Lux in a ligand-independent fashion in Mv1Lu cells. Thus, coexpression of Smads may be considered, to some extent, the nonphysiological equivalent to TGF- β -mediated activation through phosphorylation of Smads. It is conceivable that the functional requirement of Smad phosphorylation is lessened when the Smads are overexpressed to high levels. In this context, we show that coexpression of Smad3 and Smad4 in COS cells results in the formation of a specific Smad-containing DNA binding complex in the TGF- β - and Smad-responsive region of 3TP-Lux. Detailed mutagenesis and methylation interference analysis of the binding region identified a bipartite sequence as the Smad binding site.

Although these experiments were informative, Smad overexpression is not equivalent to Smad activation through phosphorylation. Because of this, it was necessary to determine if an endogenous Smad DNA binding complex assembles upon TGF- β treatment. To this end, we defined a TGF- β -induced endogenous complex from Mv1Lu cells, which has the same DNA binding specificity as the overexpressed Smad3/Smad4

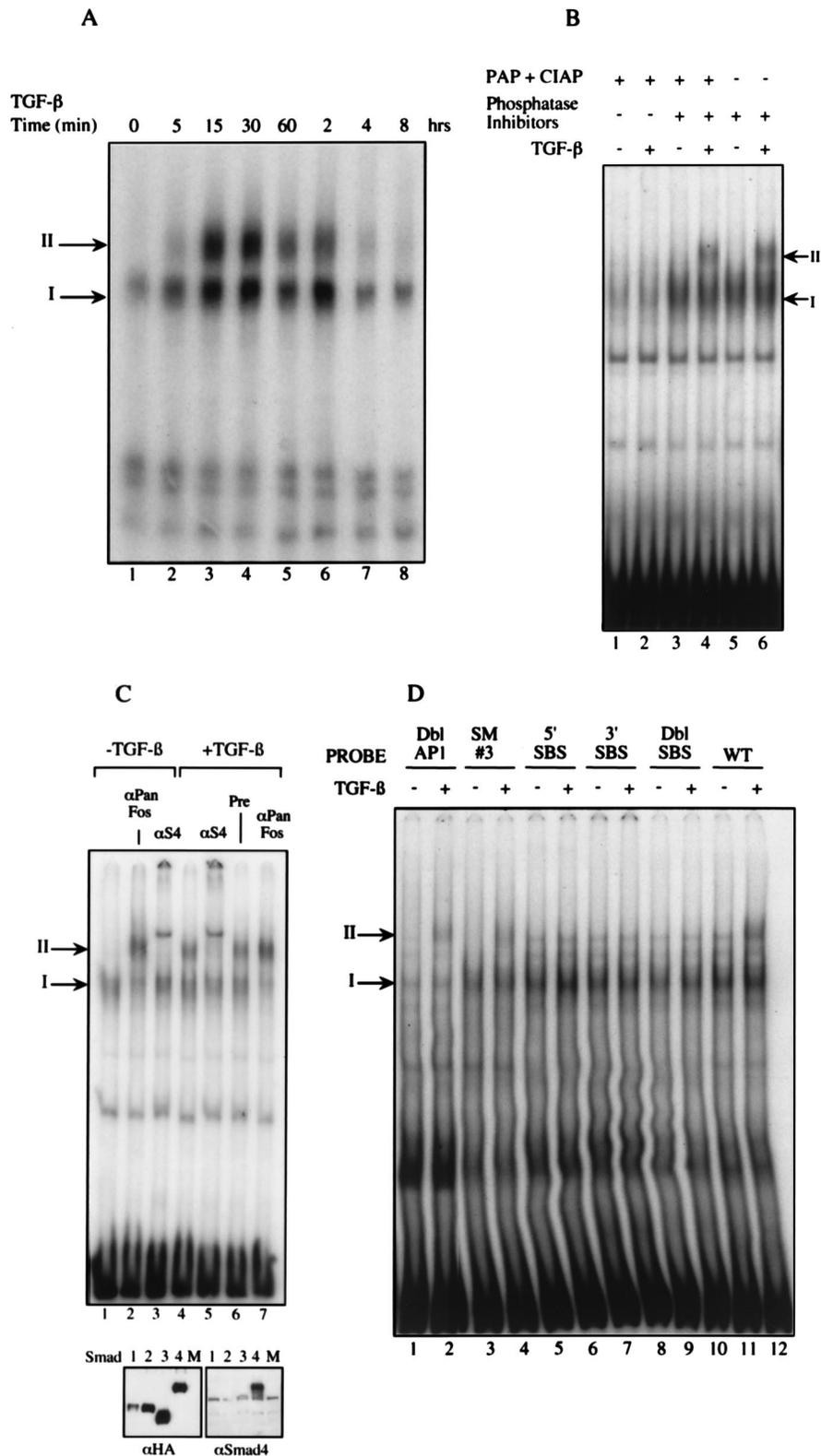


FIG. 5. Induction of a Smad3/Smad4 DNA binding complex in Mv1Lu cells by TGF- β . (A) Gel shift analysis with the 2.0 probe with lysates from a time course of TGF- β treatment of Mv1Lu cells. (B) The inducible complex is sensitive to phosphatase treatment. Mv1Lu lysates from control or TGF- β -treated cells were treated with PAP+CIAP (lanes 1 to 4) or phosphatase inhibitors (lanes 3 to 6) as described in Materials and Methods and used for EMSA studies as for panel A. (C) Supershift analysis on the induced complex in Mv1Lu cells. Gel shifts were performed with the 2.0 probe (diagrammed in Fig. 3A) and the indicated antibodies: α Pan Fos, pan-Fos antibody; α S4, Smad4 antibody; Pre, preimmune serum. Note that the pan-Fos antibody supershifted band migrates at the same place as the Smad-containing complex II. Also, anti-Smad4 antibody produces a nonspecific band above complex II. Lower panels, Western blot showing the specificity of the Smad4 antibody. The left panel is an HA antibody Western blot showing expression of Smad1 to -4; the right panel is a duplicate Western blot with Smad4 antibody. The Smad4 antibody is also specific by immunoprecipitation analysis (data not shown). (D) Analysis of the 2.0 probe mutants with TGF- β -treated Mv1Lu cell lysates. Lysates were prepared from control (odd-numbered lanes) or TGF- β treated (even-numbered lanes) cells, and EMSAs were performed with the indicated probes (diagrammed in Fig. 3A).

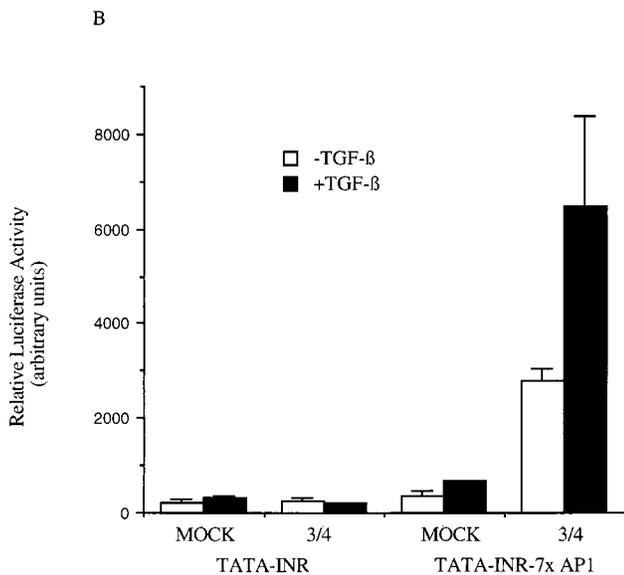
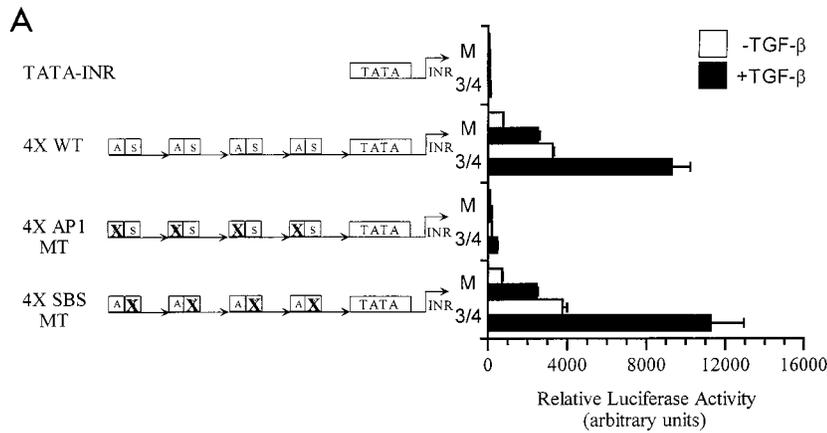


FIG. 6. Smad3/Smad4 coexpression activates transcription through AP1 DNA binding sites. (A) Functional analysis of the 2.0 probe mutants. Mv1Lu cells were cotransfected with 3 μ g of the indicated promoter and either 1 μ g of HA-Smad3 and Smad4-Flag expression vector DNA or 2 μ g of the empty pCGN vector. Luciferase assays were performed as described for Fig. 1. (B) AP1 sites are sufficient for activation by Smad coexpression. Transfections were performed as for panel A with 3 μ g of the 7 \times AP1 reporter and either 2 μ g of the pCGN vector or 1 μ g of HA-Smad3 and Smad4-Flag DNA. Luciferase assays were performed as for Fig. 1.

complex. In addition, we showed that this Mv1Lu complex contains at least Smad4, and its binding is dependent on phosphorylation. Finally, we showed that *in vitro*, recombinant Smad4 binds this DNA sequence with the same binding specificity. With these results, we have defined a novel function for Smad4 as a direct DNA binding factor.

The inability of Smad2 and Smad4 to form a similar DNA binding complex in this Smad3/Smad4 binding region suggests that different Smad complexes could have different DNA binding specificities. Although Smad4 is contained in both complexes, and we have shown that Smad4 alone can bind DNA *in vitro*, this DNA binding ability is likely to be regulated or modified by association with its Smad partners in response to exogenous signals under physiological conditions. This notion is supported by our observation that Smad2/Smad4 coexpression is unable to activate transcription from the Smad3/Smad4-responsive region of the p3TP-Lux promoter.

The surprise result from this study is the apparent dispensability of the Smad binding site within the p3TP-Lux reporter. While the lack of correlation between the DNA binding capability of Smads and their ability to activate transcription makes the interpretation of our findings somewhat difficult, the result could be explained in several ways. Smad complex binding may be having effects which cannot be assayed in these transient

transfection experiments with the use of a large quantity of plasmid DNA as the template of transcription. If, for example, Smad binding plays a role in the recruitment of other transcription factors to adjacent sites (e.g., AP1) or in rearrangement of chromosome structure to provide accessibility of other transcription factors to their binding sites, an effect in the transient transfection assay may be difficult to observe. The transient nature of Smad nuclear accumulation and DNA binding would be consistent with this type of role in transcriptional activation. Alternatively, Smad binding sites may represent enhancer-like regulatory sequences which can function properly only in the context of specific promoters. The proper promoter context may be essential to allow the appropriate interactions between the Smads and the core transcription machinery. Thus, in the context of the artificial p3TP-Lux promoter constructs, Smad binding may not be required, but in the context of wild-type promoters, Smad binding may become indispensable for transcriptional activation. Our demonstration that an endogenous promoter, PAI-1, contains a Smad3/Smad4 binding site provides an opportunity to dissect *in vivo* functions of the Smad3/Smad4 binding site and should provide insight into these important questions.

Smad3 and Smad4 coexpression can activate transcription through AP1 DNA binding sites. Although the functional consequences of Smad binding remain uncertain, we have clearly demonstrated that Smads are involved in this TGF- β -mediated signaling pathway. Other groups have shown that 3TP-Lux is not activated by TGF- β in a Smad4-deficient cell line, but this signaling pathway is restored upon reintroduction of Smad4 (8, 19). In this regard, we have shown that the AP1 binding sites within the TGF- β -responsive region of 3TP-Lux are both necessary and sufficient for activation by either TGF- β or Smad coexpression.

To present possible models explaining this finding, it is informative to review the role of Smad2 in FAST1-mediated transcriptional regulation. Recently, Smad2 has been found in a DNA binding complex with the transcription factor FAST1 (4). In this system, FAST1 is a DNA binding protein which

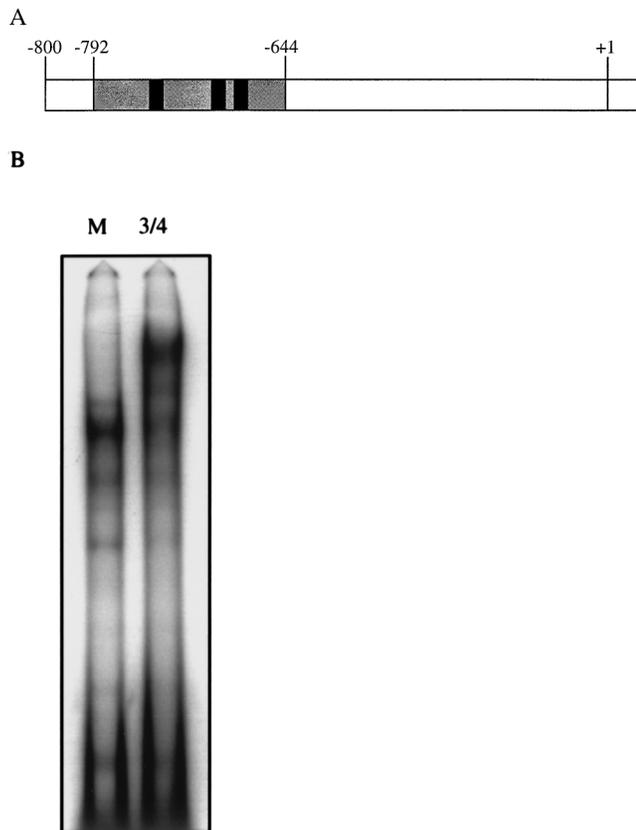


FIG. 7. The endogenous PAI-1 promoter contains a Smad3/Smad4 binding site. (A) Schematic diagram of the PAI-1 promoter. Solid boxes, potential AP1 binding sites; shaded region (-792 to -644), the probe used in gel shift assays. (B) Gel shift analysis using the probe in panel A was performed with COS lysates from cells transfected as for Fig. 2 with the pCGN vector alone (lane M) or HA-Smad3/Smad4-Flag (lane 3/4).

targets Smad2 and Smad4 to specific promoters (4, 5). In this context, the Smads would act as coactivators, which inducibly associate with transcription factors in response to specific signals. An analogy can be drawn to the ability of the Smads to activate transcription through AP1 DNA binding sites. Perhaps the requirement for Smad4 in TGF- β mediated activation of 3TP-Lux results from an interaction between Smad4 and AP1 complexes. In this model, a Smad DNA binding site would be dispensable in transient transfection assays due to the fact that Smad4 would have already been targeted to this promoter through its interaction with AP1.

So why have a DNA binding activity at all? The answer becomes clear in the case of the *Drosophila* Smad homolog, MAD. MAD was recently shown to be a DNA binding protein which regulates the expression of a number of dpp-responsive genes (17). In the Vg promoter, MAD binds to specific DNA sequences to directly mediate activation of the vestigial promoter. Thus, rather than being tethered to a promoter through a second transcription factor, the DNA binding function of MAD is essential. Interestingly, a comparison between the DNA binding site for Smad4 which we have described and the DNA binding site for MAD reveals little sequence similarity. This suggests that different Smads will have different DNA binding specificities and, thus, different target promoters.

These two modes of Smad function, as direct DNA binding proteins and as transcriptional coactivator proteins, broaden

the possible range and diversity of Smad promoter targets. In certain contexts, the interaction of specific Smad heteromers with specific transcription factors may lead to promoter activation. In this context, the DNA binding function of the Smads becomes dispensable, as may be the case with Smad2 and the Mix2 promoter and possibly the 3TP-Lux promoter analyzed in this study. In contrast, other promoter contexts may require the DNA binding function of the Smads to target and activate the promoter, as is the case with MAD and the vestigial promoter.

In conclusion, the finding that Smad4 is a DNA binding protein raises the exciting possibility that this function is at the root of its tumor suppressor activity. In this case, defining additional Smad4 target promoters may be informative in understanding its role as a tumor suppressor and, in a more global sense, the process of tumorigenesis. In addition, the finding that Smad3 and Smad4 can potentiate transcription from AP1 DNA binding sites adds an additional level of complexity to the highly regulated process of AP1-mediated transcription.

ACKNOWLEDGMENTS

J.M.Y., M.B.D., and C.W. contributed equally to this work.

We thank Rik Derynck for his generous gift of human Smad2, Smad3, and Smad4. TGF- β 1 was kindly provided by Amgen, Inc. We appreciate helpful discussions with members of the Wang and Pendergast labs and thank Yong Yu for technical assistance.

This work was supported by grant DK-45746 from the National Institutes of Health to X.-F.W., U.S. Army Breast Cancer Research grant DAMD17-945-4065 to X.-F.W.; a predoctoral U.S. Army Breast Cancer Research grant DAMD17-94-J-4190 to J.M.Y.; and a predoctoral NSF grant to N.T.L. X.-F.W. is a Leukemia Society Scholar.

REFERENCES

- Attisano, L., and J. L. Wrana. 1996. Signal transduction by members of the transforming growth factor- β superfamily. *Cytokine Growth Factor Rev.* 7:327-339.
- Baker, J., and R. Harland. 1996. A novel mesoderm inducer, Madr2, functions in the activin signal transduction pathway. *Genes Dev.* 10:1880-1889.
- Bassing, C. H., J. M. Yingling, D. J. Howe, T. Wang, W. W. He, M. L. Gustafson, P. Shah, P. K. Donahoe, and X. F. Wang. 1994. A transforming growth factor β type I receptor that signals to activate gene expression. *Science* 263:87-89.
- Chen, X., M. J. Rubock, and M. Whitman. 1996. A transcriptional partner for MAD proteins in TGF- β signalling. *Nature* 383:691-696.
- Chen, X., E. Weisberg, V. Fridmacher, M. Watanabe, and M. Whitman. 1997. Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature* 389:85-89.
- Chen, Y., J. J. Lebrun, and W. Vale. 1996. Regulation of transforming growth factor β - and activin-induced transcription by mammalian Mad proteins. *Proc. Natl. Acad. Sci. USA* 93:12992-12997.
- Datto, M. B., Y. Yu, and X. F. Wang. 1995. Functional analysis of the transforming growth factor β responsive elements in the WAF1/Cip1/p21 promoter. *J. Biol. Chem.* 270:28623-28628.
- de Caestecker, M. P., P. Hemmati, S. Larisch-Bloch, R. Ajmera, A. B. Roberts, and R. J. Lechleider. 1997. Characterization of functional domains within Smad4/DPC4. *J. Biol. Chem.* 271:13690-13696.
- Derynck, R., W. M. Gelbart, R. M. Harland, C. H. Heldin, S. E. Kern, J. Massague, D. A. Melton, M. Mlodzik, R. W. Padgett, A. B. Roberts, J. Smith, G. H. Thomsen, B. Vogelstein, and X.-F. Wang. 1996. Nomenclature: vertebrate mediators of TGF- β family signals. *Cell* 87:173.
- Eppeert, K., S. W. Scherer, H. Ozelik, R. Pirone, P. Hoodless, H. Kim, L. C. Tsui, B. Bapat, S. Gallinger, I. L. Andrusis, G. H. Thomsen, J. L. Wrana, and L. Attisano. 1996. MADR2 maps to 18q21 and encodes a TGF β -regulated MAD-related protein that is functionally mutated in colorectal carcinoma. *Cell* 86:543-552.
- Franzen, P., P. ten Dijke, H. Ichijo, H. Yamashita, P. Schulz, C.-H. Heldin, and K. Miyazono. 1993. Cloning of a TGF β type I receptor that forms a heteromeric complex with the TGF β type II receptor. *Cell* 75:681-692.
- Graff, J. M., A. Bansal, and D. A. Melton. 1996. Xenopus Mad proteins transduce distinct subsets of signals for the TGF β superfamily. *Cell* 85:479-487.
- Hahn, S. A., M. Schutte, A. T. Hoque, C. A. Moskaluk, L. T. da Costa, E. Rozenblum, C. L. Weinstein, A. Fischer, C. J. Yeo, R. H. Hruban, and S. E.

- Kern.** 1996. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* **271**:350–353.
14. **Hata, A., R. S. Lo, D. Wotton, G. Lagna, and J. Massague.** 1997. Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4. *Nature* **388**:82–87.
 15. **Hoodless, P. A., T. Haerry, S. Abdollah, M. Stapleton, M. B. O'Connor, L. Attisano, and J. L. Wrana.** 1996. MADR1, a MAD-related protein that functions in BMP2 signaling pathways. *Cell* **85**:489–500.
 16. **Keeton, M. R., S. A. Curriden, A. J. van Zonneveld, and D. J. Loskutoff.** 1991. Identification of regulatory sequences in the type 1 plasminogen activator inhibitor gene responsive to transforming growth factor β . *J. Biol. Chem.* **266**:23048–23052.
 17. **Kim, J., K. Johnson, H. J. Chen, S. Carroll, and A. Laughon.** 1997. Drosophila MAD binds to DNA and directly mediates activation of vestigial by decapentaplegic. *Nature* **388**:304–308.
 18. **Kretschmar, M., F. Liu, A. Hata, J. Doody, and J. Massague.** 1997. The TGF- β family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes Dev.* **11**:984–995.
 19. **Lagna, G., A. Hata, A. Hemmati-Brivanlou, and J. Massague.** 1996. Partnership between DPC4 and SMAD proteins in TGF- β signalling pathways. *Nature* **383**:832–836.
 20. **Lechleider, R. J., M. P. de Caestecker, A. Dehejia, M. H. Polymeropoulos, and A. B. Roberts.** 1996. Serine phosphorylation, chromosomal localization, and transforming growth factor- β signal transduction by human bsp-1. *J. Biol. Chem.* **271**:17617–17620.
 - 20a. **Li, J.-M., and X. F. Wang.** Submitted for publication.
 21. **Lin, H. Y., X. F. Wang, E. Ng-Eaton, R. A. Weinberg, and H. F. Lodish.** 1992. Expression cloning of the TGF- β type II receptor, a functional transmembrane serine/threonine kinase. *Cell* **68**:775–785.
 22. **Liu, F., A. Hata, J. C. Baker, J. Doody, J. Carcamo, R. M. Harland, and J. Massague.** 1996. A human Mad protein acting as a BMP-regulated transcriptional activator. *Nature* **381**:620–623.
 23. **Macias-Silva, M., S. Abdollah, P. Hoodless, R. Pirone, L. Attisano, and J. Wrana.** 1996. MADR2 is a substrate of the TGF- β receptor and its phosphorylation is required for nuclear accumulation and signaling. *Cell* **87**:1215–1224.
 24. **Massague, J.** 1996. TGF β signaling: receptors, transducers, and Mad proteins. *Cell* **85**:947–50.
 25. **Massague, J.** 1990. The transforming growth factor- β family. *Annu. Rev. Cell Biol.* **6**:597–641.
 26. **Roberts, A. B., and M. B. Sporn.** 1993. Physiological actions and clinical applications of transforming growth factor- β (TGF- β). *Growth Factors* **8**: 1–9.
 27. **Savage, C., P. Das, A. L. Finelli, S. R. Townsend, C. Y. Sun, S. E. Baird, and R. W. Padgett.** 1996. *Caenorhabditis elegans* genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor β pathway components. *Proc. Natl. Acad. Sci. USA* **93**:790–794.
 28. **Sekelsky, J. J., S. J. Newfeld, L. A. Raftery, E. H. Chartoff, and W. M. Gelbart.** 1995. Genetic characterization and cloning of mothers against dpp, a gene required for decapentaplegic function in *Drosophila melanogaster*. *Genetics* **139**:1347–1358.
 29. **Shi, Y., A. Hata, R. S. Lo, J. Massague, and N. P. Pavletich.** 1997. A structural basis for mutational inactivation of the tumour suppressor, Smad4. *Nature* **388**:87–93.
 30. **Suzuki, A., C. Chang, J. M. Yingling, X.-F. Wang, and A. Hemmati-Brivanlou.** 1997. Smad5 induces ventral fates in *Xenopus* embryo. *Dev. Biol.* **184**: 402–405.
 31. **Tanaka, M., and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. *Cell* **60**:375–386.
 32. **Thomsen, G. H.** 1996. *Xenopus* mothers against decapentaplegic is an embryonic ventralizing agent that acts downstream of the BMP-2/4 receptor. *Development* **122**:2359–2366.
 33. **Wrana, J. L., L. Attisano, J. Carcamo, A. Zentella, J. Doody, M. Laiho, X. F. Wang, and J. Massague.** 1992. TGF β signals through a heteromeric protein kinase receptor complex. *Cell* **71**:1003–1014.
 34. **Wrana, J. L., L. Attisano, R. Wieser, F. Ventura, and J. Massague.** 1994. Mechanism of activation of the TGF- β receptor. *Nature* **370**:341–347.
 35. **Wu, R. Y., Y. Zhang, X. H. Feng, and R. Derynck.** 1997. Heteromeric and homomeric interactions correlate with signaling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Mol. Cell. Biol.* **17**:2521–2528.
 36. **Yingling, J. M., P. Das, C. Savage, M. Zhang, R. W. Padgett, and X. F. Wang.** 1996. Mammalian dwarfins are phosphorylated in response to TGF- β and are implicated in control of cell growth. *Proc. Natl. Acad. Sci. USA* **93**:8940–8944.
 37. **Zhang, Y., X. Feng, R. We, and R. Derynck.** 1996. Receptor-associated Mad homologues synergize as effectors of the TGF- β response. *Nature* **383**:168–172.
 38. **Zhang, Y., T. Musci, and R. Derynck.** 1997. The tumor suppressor Smad4/DPC 4 as a central mediator of Smad function. *Curr. Biol.* **7**:270–276.