THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 23, Issue of June 8, pp. 20688–20694, 2001

Printed in U.S.A.

Formation of a Stable Heterodimer between Smad2 and Smad4*

Received for publication, January 9, 2001, and in revised form, March 23, 2001
Published, JBC Papers in Press, March 27, 2001, DOI 10.1074/jbc.M100174200

Jia-Wei Wu‡, Robert Fairman§, Jack Penry‡, and Yigong Shi‡¶‖

From the ‡Department of Molecular Biology, Princeton University, Lewis Thomas Laboratory, Princeton, New Jersey 08544, and §Department of Molecular, Cell, and Developmental Biology, Haverford College, Haverford, Pennsylvania 19041

Smad proteins mediate transforming growth factor β signaling from the cell membrane to the nucleus. Upon phosphorylation by the activated receptor kinases, the receptor-regulated Smad, such as Smad2, forms a heterocomplex with the co-mediator Smad, Smad4. This heterocomplex is then translocated into the nucleus, where it associates with other transcription factors and regulates expression of ligand-responsive genes. The stoichiometry between receptor-regulated Smad and co-mediator Smad is important for understanding the molecular mechanisms of the signaling process. Using purified recombinant proteins, we demonstrate that Smad2 and Smad4 form a stable heterodimer and that the Smad4 activation domain is important for the formation of this complex. Many tumor-derived missense mutations disrupt the formation of this heterocomplex in *in vitro* interaction assays. Mapping these mutations onto the structures of Smad4 and Smad2 identifies a symmetric interface between these two Smad proteins. Importantly, two previous models on the formation of a heterocomplex are incompatible with our observations and other reported evidence.

TGF-β[1] signaling regulates a broad range of cellular responses, including growth, differentiation, and cell fate specification, in all animals (1, 2). TGF-β signaling from the cell membrane to the nucleus is mediated by the Smad family of proteins, which contains at least nine distinct members in vertebrates and two of which, Smad2 and Smad4, have been identified as tumor suppressors in humans (3–6).

The Smad proteins are functionally divided into three distinct classes: (i) co-mediator Smads (Co-Smads), namely, Smad4 in mammals and Smad10 (also known as Smad4β) in *Xenopus*, which participate in signaling by diverse TGF-β family members, (ii) receptor-regulated Smads (R-Smads), including Smad1, Smad2, Smad3, Smad5, and Smad8, each of which is involved in a specific signaling pathway, and (iii) inhibitory Smads, which include Smad6 and Smad7 and negatively regulate these pathways (3–5, 7).

A TGF-β response is initiated by the binding of a specific TGF-β ligand to a pair of specific transmembrane receptors, the type I and II receptors, leading to the activation of the Ser/Thr kinase in the cytoplasmic domain of the type I receptor (8). The signal is then propagated by the type I receptor-mediated phosphorylation of specific R-Smads. For example, Smad1, Smad5, and Smad8 are phosphorylated by the bone morphogenetic protein receptors, whereas Smad2 and Smad3 are phosphorylated by the activin and TGF-β receptors. The phosphorylated R-Smad hetero-oligomerizes with Co-Smad Smad4, translocates into the nucleus, and associates with sequence-specific DNA-binding protein(s), resulting in the positive or negative regulation of ligand-responsive genes.

The Smad proteins are conserved across species, particularly in the N-terminal MH1 domain and the C-terminal MH2 domain. The MH2 domain, to which most of the tumor-derived mutations map, is responsible for receptor recognition, transactivation, interaction with transcription factors, and homo- and hetero-oligomerization among Smads. The MH1 domain, on the other hand, exhibits sequence-specific DNA binding activity and negatively regulates the functions of the MH2 domain.

Formation of a heterocomplex between Co-Smad and R-Smad is indispensable for the signaling process (2). The only known Co-Smad in mammals, Smad4, forms a homotrimer in a concentration-dependent manner both *in vivo* and *in vitro* (9–11). The R-Smads, however, exhibit several distinct oligomeric states at the basal state (11, 12). The complex between Co-Smad and R-Smad has been suggested to be a heterohexamer (9), a heterotrimer (10, 12), or, more recently, a heterodimer (11). Understanding this stoichiometry has important implications for understanding the molecular mechanisms of transcriptional regulation by Smad proteins. To address this controversy, we have undertaken a biochemical and biophysical approach, using purified homogeneous Smad proteins. Results from both gel filtration and ultracentrifugation analyses demonstrate that Smad2 and Smad4 form a stable heterodimer. In addition, 15 tumor-derived missense mutations were introduced into these two Smad proteins to assess several prevailing models of heterocomplex formation. Our results suggest a novel arrangement between Smad2 and Smad4.

MATERIALS AND METHODS

Site-directed Mutagenesis and Protein Preparation—Point mutations were generated using a standard polymerase chain reaction-based cloning strategy, and the identities of individual clones were verified through double-strand plasmid sequencing. All Smad4 proteins were overexpressed in *Escherichia coli* strain BL21(DE3) at room temperature as a GST-fusion protein using a pGEX-2T vector (Amersham Pharmacia Biotech). The soluble fraction of the GST-Smad4 fusion in the *E. coli* lysate was purified over a glutathione-Sepharose column, cleaved by thrombin, and further purified by anion-exchange chromatography (Source-15Q; Amersham Pharmacia Biotech) and gel filtration chromatography (Superdex-75, Amersham Pharmacia Biotech). All recombinant Smad2 proteins were overexpressed in *E. coli* strain

* This research was supported by National Institutes of Health Grant R01-CA82171. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ A Searle Scholar and a Rita Allen Scholar.

‖ To whom correspondence should be addressed: Dept. of Molecular Biology, Princeton University, Washington Road, Princeton, NJ 08544 Tel.: 609-258-6071; Fax: 609-258-6730; E-mail: yshi@molbio.princeton.edu.

[1] The abbreviations used are: TGF-β, transforming growth factor β; R-Smad, receptor-regulated Smad; Co-Smad, co-mediator Smad; SAD, Smad4 activation domain; GST, glutathione *S*-transferase.

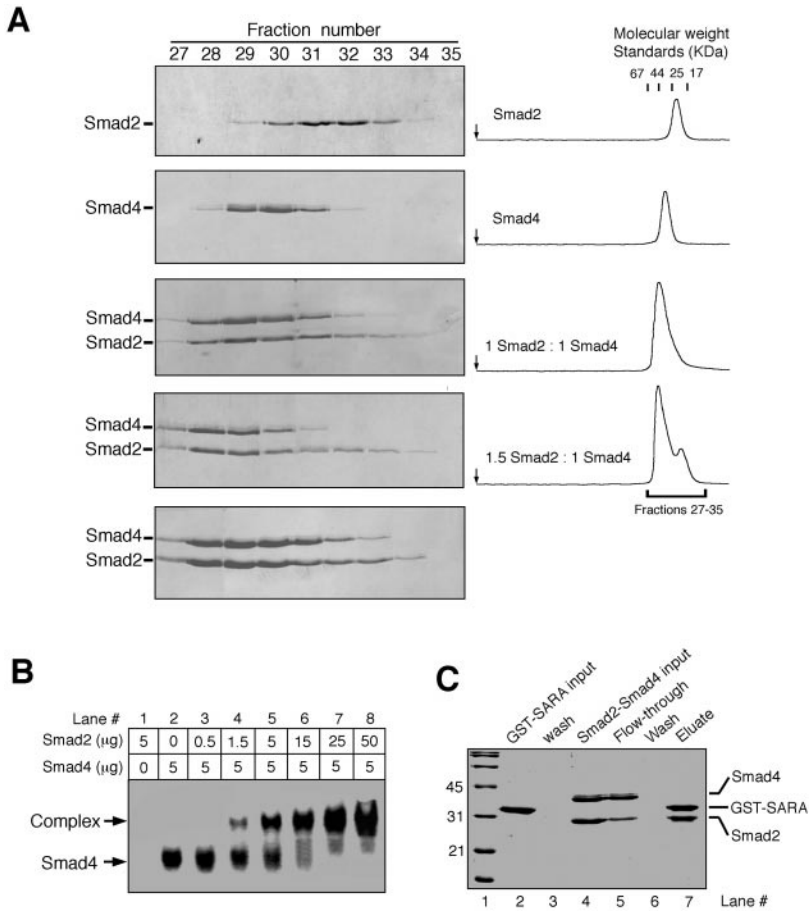

FIG. 1. **Formation of a stable heterodimer between Smad2 and Smad4.** The Smad2 and Smad4 fragments contained residues 241–467 and 251–552, respectively. *A,* size exclusion chromatography of Smad2 (250 μg; *panel 1*; *panels* in *A* are numbered from *top* to *bottom*), Smad4 (330 μg; *panel 2*), a 1:1 complex of Smad2-Smad4 (250 + 330 μg; *panel 3*), a 1.5:1 complex of Smad2-Smad4 (370 + 330 μg; *panel 4*), and a concentrated 1:1 complex of Smad2-Smad4 (total input 2.5 mg; *panel 5*). Relevant fractions were visualized on SDS-polyacrylamide gel electrophoresis. The chromatographs and the calibration of the Superdex-200 column are shown on the *right*. The *arrows* indicate the starting points for all four chromatographic runs. *B,* electrophoretic mobility shift assays under nondenaturing conditions. *C*, mutual exclusion of a Smad2-Smad4 complex and a Smad2-SARA complex. GST-SARA was first bound to glutathione resin, and a stoichiometric amount of a 1:1 Smad2-Smad4 complex was allowed to flow through the resin. The resin was washed four times with assay buffer, and aliquots of the last wash were visualized on SDS-polyacrylamide gel electrophoresis.

BL21(DE3) at room temperature using a pET-3d vector (Novagen). The soluble fraction of the cell lysate was fractionated over a SP-Sepharose column and further purified by anion-exchange chromatography (Source-15Q) and gel filtration chromatography (Superdex-75). The identities of all proteins were confirmed by mass spectroscopy. The concentrations of Smad4 and Smad2 were determined by spectroscopic measurement at 280 nm. All recombinant proteins were characterized by gel filtration and dynamic light scattering.

Interaction Assay by Size Exclusion Chromatography—Size exclusion chromatography, using a Superdex-200 column (10/30; Amersham Pharmacia Biotech), was employed to examine the interaction between Smad4 and Smad2. In all cases, Smad2 is incubated with Smad4 at 4 °C for at least 45 min to allow equilibrium to be reached. The flow rate was 0.5 ml/min, and the buffer contained 25 mM Tris, pH 8.0, 150 mM NaCl, and 2 mM dithiothreitol. All fractions were collected at 0.5 ml each. Aliquots of relevant fractions were mixed with SDS sample buffer and subjected to SDS-polyacrylamide gel electrophoresis. The proteins were visualized by Coomassie Blue staining. The column was calibrated with molecular mass standards.

Analytical Ultracentrifugation—Protein samples were prepared in 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 1 mm dithiothreitol. All sedimentation equilibrium experiments were carried out at 4 °C using a Beckman Optima XL-A analytical ultracentrifuge equipped with an An60 Ti rotor and using six-channel, 12-mm path length, charcoal-filled Epon centerpieces and quartz windows. Loading concentrations included 2.5, 5, 10, and 20 μM complex. Data were collected at four rotor speeds (8,000, 11,000, 14,000, and 17,000 rpm) and represent the average of 20 scans using a scan step size of 0.001 cm. Partial specific volumes and solution density were calculated using the Sednterp program (13). Data were analyzed using the HID program from the Analytical Ultracentrifugation Facility at the University of Connecticut (Storrs, CT). The ultracentrifuge figure was composed using IGOR Pro version 3.16 (WaveMetrics Inc., Lake Oswego, OR).

Electrophoretic Mobility Shift Assays—All assays were performed at 4 °C to offset the heat generated by electrophoresis. The 6% polyacrylamide (37.5:1 acrylamide:bisacrylamide) gels were used under the buffer condition of 65 mM Tris, pH 8.5, and 65 mM boric acid. All protein samples were prepared in a 40 μl volume containing 25 mM Tris, pH 8.0, 150 mM NaCl, 2 mM dithiothreitol, and 5% glycerol. After prerunning the gels for 15 min, half of each sample (20 μl) was loaded into each lane and subjected to electrophoresis with a constant electric field of 15 V/cm. The gels were stained with Coomassie Blue, destained, and dried for photography. After scanning the gels with a densitometer, the unbound fractions of Smad4 were used to fit a binary interaction equation, from which the estimate for the binding affinity was obtained.

GST-mediated Pull-down Assay—Approximately 0.4 mg of a recombinant SARA fragment (residues 665–721) was bound to 200 μl of glutathione resin as a GST-fusion protein. The resin was washed with 400 μl of buffer four times to remove excess unbound SARA (Smad anchor for receptor activation) or other contaminants, and then 600 μg of a Smad2-Smad4 complex was allowed to flow through the resin. After extensive washing with an assay buffer containing 25 mM Tris, pH 8.0, 150 mM NaCl, and 2 mM dithiothreitol, GST-SARA was eluted with 5 mM reduced glutathione, and all fractions were visualized by SDS-polyacrylamide gel electrophoresis with Coomassie Blue staining.

RESULTS AND DISCUSSION

Smad2 Forms a Stable Heterodimer with Smad4—The full-length Smad2 is unable to form a complex with Smad4 in the absence of phosphorylation in its C-terminal SS*MS* sequence. This is likely due to an autoinhibitory interaction between the MH1 and MH2 domains that can be relieved by phosphorylation (14). Indeed, removal of the MH1 domain in Smads results in constitutively active transcriptional activity (15) and allows the formation of a stable heterocomplex in the absence of phosphorylation (9, 14, 16). Hence, we chose to focus on the MH1-deleted proteins of Smad2 (residues 241–467) and Smad4 (residues 251–552). These recombinant proteins were overexpressed in bacteria and purified to homogeneity. After confirmation of protein identity by mass spectroscopy, both proteins were subjected to gel filtration and dynamic light scattering analysis to ensure that they are well folded and exhibit good solution properties.

To examine the stoichiometry between Smad2 and Smad4, we devised an *in vitro* interaction assay employing size exclu-

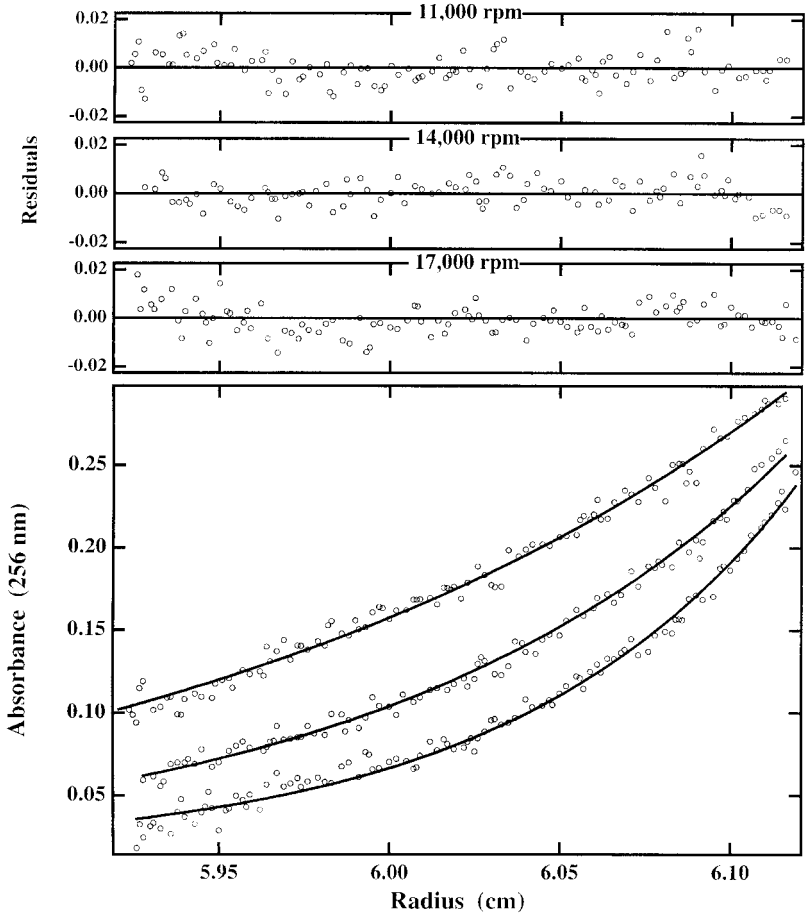

FIG. 2. **Smad2 and Smad4 form heterodimers by ultracentrifugation analysis.** Sedimentation equilibrium data and fit derived from a one-state model where the molecular mass is fixed to 58,915 daltons, that of a heterodimer.

sion chromatography. We also used equilibrium sedimentation analytical ultracentrifugation to characterize the molecular mass of the hetero-oligomer. Because Smad proteins exhibit concentration-dependent homo-oligomerization, we selected the concentration ranges in which both proteins behave as monomers on gel filtration analysis (Fig. 1*A*). The elution volume for Smad2 corresponds to a molecular mass of ~22 kDa (Fig. 1*A, panel 1* (panels in Fig. 1*A* are numbered from *top* to *bottom*)), consistent with its calculated molecular mass of 25 kDa. The elution volume for Smad4, which contains a 30-residue flexible loop (residues 460–490) between helices H3 and H4 (9), corresponds to a molecular mass of about 39 kDa (Fig. 1*A, panel 2*), in reasonable agreement with its calculated molecular mass of 33 kDa. When equimolar amounts of Smad2 and Smad4 were used, the vast majority of Smad2 was shifted to earlier fractions, indicating a 1:1 stoichiometry (Fig. 1*A, panel 3*). In addition, the elution volume for the complex corresponds to an apparent molecular mass of ~54 kDa, consistent with the expected mass of a complex composed of one Smad2 and one Smad4 (Fig. 1*A, panel 3*). To further demonstrate the formation of a heterodimer, 1.5 molar equivalents of Smad2 were mixed with one molar equivalent of Smad4. In this case, the excess amount of Smad2 was eluted from the size exclusion column as a monomer (Fig. 1*A, panel 4*). During the course of size exclusion chromatography, there is little dissociation between Smad2 and Smad4, suggesting stable complex formation. Nevertheless, excess Smad2 (Fig. 1*A, panel 4*) or higher concentrations of both proteins (Fig. 1*A, panel 5*) led to more complete formation of a heterodimer.

To estimate the binding affinity between Smad2 and Smad4 and to further confirm the formation of a heterodimer, electrophoretic mobility shift assays under nondenaturing conditions were employed (Fig. 1*B*). Under these conditions, Smad2 does not enter the gel (Fig. 1*B, lane 1*) whereas Smad4 migrates as a single band (Fig. 1*B, lane 2*). With increasing concentrations of Smad2, a distinct heterocomplex is formed (Fig. 1*B, lanes 3–8*). Neither excess Smad4 (*lanes 3* and *4*) nor excess Smad2 (*lanes 6–8*) resulted in more than one heterocomplex. Quantitation of the binding experiment revealed a dissociation constant of ~1 μM.

Formation of a functional Smad2-Smad4 complex may be antagonized by the formation of a complex between SARA and Smad2 (17, 18). To further demonstrate the functional relevance of the observed heterodimer, the Smad2-Smad4 complex was applied to glutathione resin preimmobilized with GST-SARA (Fig. 1*C*). As expected, the flow-through fraction contained less Smad2 than the input (Fig. 1*C, lane 5*), suggesting disruption of a Smad2-Smad4 complex by SARA. Indeed, the eluted fraction contained a SARA-Smad2 complex (Fig. 1*C, lane 7*), demonstrating that SARA does compete with Smad4 for binding to Smad2.

To complement the gel filtration analysis, the molecular mass of the hetero-oligomer was analyzed by analytical ultracentrifugation. The complex was prepared in 1:1 ratio, based on the conclusion from the gel filtration results that this complex forms a heterodimer. The complex was analyzed at four loading concentrations and four rotor speeds. At 5 and 10 μM, the protein complex is fully consistent with that of a heterodimer, with molecular masses of 58,900 and 59,800 daltons, respectively (Table I). If Smad2 and Smad4 form heterotrimers instead, then the apparent molecular mass would be significantly reduced because one of the two proteins would be in significant excess and would have contributed to a significant reduction in the reported molecular mass. In Fig. 2, we show the fit of the 10 μM data to a heterodimer model (Fig. 2). At 2.5 μM, we see evidence for a dynamic equilibrium between monomers and

TABLE I

Analytical ultracentrifugation analysis of a Smad2-Smad4 complex

The Smad2-Smad4 complex exhibits concentration-dependent molecular masses, indicating higher order oligomerization. Two-state models were analyzed for these complexes to assess which model would best fit the data. Models tested included heterodimer⟨-⟩heterotrimer (where both 2:1 or 1:2 ratios of Smad2:Smad4 were considered) and both heterodimer and heterotrimer self-association (1⟨-⟩N where 1 equals either the heterodimer or heterotrimer unit and N = 2, 3, 4, 5, and 6 units of each heteromer). Least squares analysis disfavors all models involving heterotrimer formation. The best two-state model is a self-associating model for the heterodimer with N = 4.

Concentration of complex	Molecular mass
μM	*daltons*
2.5	37,800 ± 3,900
5.0	58,900 ± 4,200
10	59,800 ± 7,900
20	72,200 ± 5,800

heterodimers, whereas at 20 μM, there is some evidence for minor aggregation (Table I). Indeed, Smad2 by itself starts to homo-oligomerize and/or aggregate at the 20 μM concentration. The best-fit two-state model to account for the aggregation is a heterodimer-hetero-octamer model. In summary, ultracentrifugation analysis demonstrates that Smad2 and Smad4 form a heterodimer.

The Smad4 Activation Domain Is Required for the Formation of a Stable Heterocomplex—To investigate the sequence requirement for the formation of a stable Smad2-Smad4 complex, we created deletion mutants in Smad2 and Smad4 and examined their interaction with each other. The results indicate that neither MH2 domain is sufficient for heterocomplex formation (Fig. 3*A*). In the case of Smad2, a 28-residue extension N-terminal to the MH2 domain (residues 269–467) is required for stable interaction with Smad4 (Fig. 3*A*), presumably due to the structural requirement that this fragment contribute an additional β-strand that packs against a hydrophobic surface (19). On the other hand, neither the MH2 domain of Smad4 (residues 319–552) (Fig. 3*B, upper panel*) nor a longer fragment with 19 additional amino acids (300–552) was able to form a stable heterocomplex with Smad2, as judged by their progressive propensity toward dissociation upon size exclusion chromatography. Inclusion of the full Smad4 activation domain (SAD) (20) restored its ability to form a stable heterodimer with Smad2 (Fig. 3*B, bottom panel*). The coupling of efficient formation of a heterocomplex with the requirement for SAD may have functional implications for TGF-β signaling.

Tumorigenic Mutations Disrupt the Formation of the Smad2-Smad4 Heterodimer—Most of the tumor-derived missense mutations map to the MH2 domain in Smad proteins (6, 9, 21). Because the formation of a functional Smad2-Smad4 complex is important for signaling, some of these missense mutations may act by disrupting this complex. Although previous work shows that this is indeed the case *in vivo* (14), only four mutations were examined by immunoprecipitation, which precluded a conclusion on whether or not these mutations directly prevented formation of a heterocomplex.

To address this issue systematically, we generated a total of 18 missense mutations into either Smad2 (residues 241–467) or Smad4 (residues 251–552), purified the mutant proteins to homogeneity, and examined their interactions with their wild-type counterparts. To rule out the possibility of misfolding or aggregation, each mutant protein was carefully compared with the wild-type Smad2 or Smad4 by gel filtration, dynamic light scattering, and thermodenaturation analyses. With the exception of three insoluble mutants (L440R and P445H in Smad2 and I527R in Smad4), these analyses demonstrated that each of the mutant proteins was well folded and exhibited very

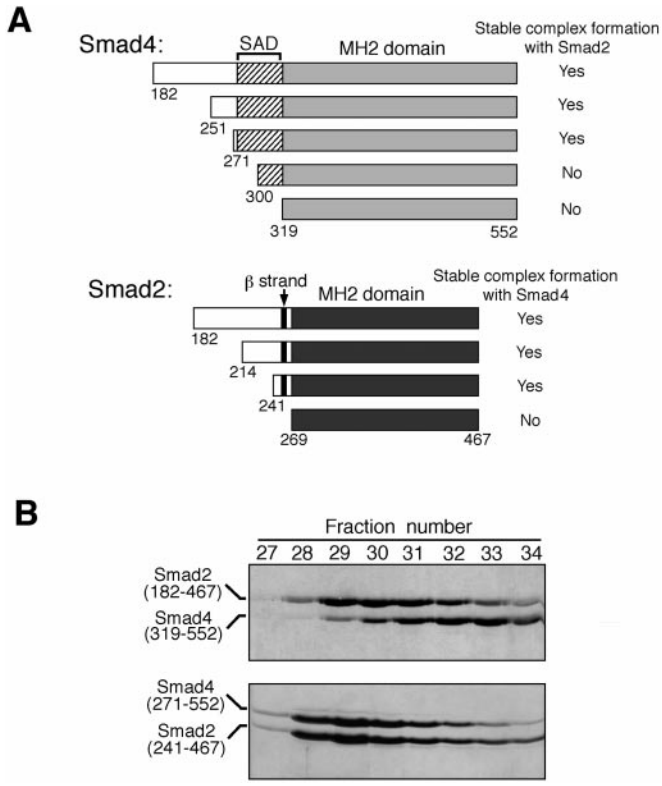

FIG. 3. **The SAD (residues 275–322) in Smad4 is required for the formation of a stable heterodimer with Smad2.** *A,* a summary of the results obtained with Smad2 and Smad4 deletion mutants. The Smad4 mutants were assayed for their interaction with Smad2 (residues 241–467), whereas the Smad2 mutants were examined using Smad4 (residues 251–552). Gel filtration and GST-mediated pull-down assays yielded consistent results. *B,* two representative results showing the requirement for SAD in Smad4 for complex formation with Smad2. In the *top panel,* equimolar amounts of Smad4 (319–552) and Smad2 (182–467) were analyzed on gel filtration after a 45-min preincubation. They dissociate during the course of gel filtration. In the *bottom panel,* equimolar amounts of the SAD-containing Smad4 (271–552) and Smad2 (241–467) were co-eluted as a stable heterodimer.

similar solution properties. Each of eight tumor-derived mutations was introduced in both Smad2 and Smad4. For example, D450H in Smad2 has been reported in colon cancers (22). Both D450H in Smad2 and the corresponding D537H in Smad4 were generated. Each of the three tumor-derived mutations in Smad4 (D351H, R361H, and V370D), as well as their corresponding mutations in Smad2 (D300H, R310H, and V319D), disrupted the formation of a heterocomplex (Table II; Fig. 4*A*). On the other hand, D450H in Smad2 and D537H in Smad4 failed to disrupt the formation of a heterodimer (Table II; Fig. 4*A*). Two additional tumorigenic mutations (F346V in Smad2 and R420H in Smad4) and their corresponding mutations (W398V in Smad4 and W368H in Smad2) exhibited no effects on heterocomplex formation (Table II; Fig. 4). Interestingly, each of the tumorigenic mutations in the original Smad and the corresponding mutation in the other Smad has identical effects on heterocomplex formation (Fig. 4*B*). This result suggests that Smad2 and Smad4 may form a pseudosymmetric heterodimer, in which each of the two Smad proteins uses similar surface motifs to interact with the other. Because three of the four deleterious mutations affect residues that are located in the L1 and L2 loop region (9) (Fig. 4*B*), this loop-helix region must be directly involved in mediating the formation of a heterodimer between Smad2 and Smad4. In support of this observation, this loop-helix region contains the overwhelming majority of the most highly conserved and solvent-exposed residues in Smad proteins (9).

TABLE II
Mutational analysis of a Smad2-Smad4 complex

The identities of all mutant proteins were confirmed by double-stranded plasmid sequencing and mass spectroscopic analysis. Three mutant proteins were insoluble, as predicted by previous structural studies (9). The original tumorigenic mutations are shown in bold.

Mutation in Smad4	Interaction with Smad2	Mutation in Smad2	Interaction with Smad4
WT	Yes (~1 μM)	SE*ME*	Yes (~0.4 μM)
D351H	No	D300H	No
R361H	No	R310H	No
V370D	No	V319D	No
R420H	Yes	W368H	Yes
R441P	Yes	N387P	n/a
D493H	No	Y406H	n/a
W398V	Yes	**F346V**	Yes
I527R	n/a (insoluble)	**L440R**	n/a (insoluble)
A532H	No	**P445H**	n/a (insoluble)
D537H	Yes	**D450H**	Yes

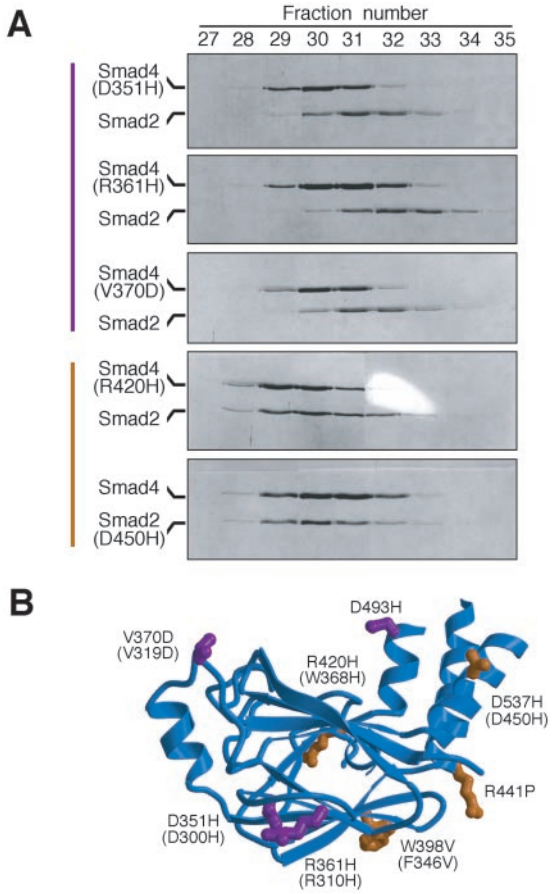

FIG. 4. **Tumorigenic mutations inactivate the heterodimer between Smad2 and Smad4.** Ten tumor-derived missense mutations were introduced in Smad4 (251–552) and Smad2 (241–467). These mutant proteins were purified to homogeneity and assayed for their ability to interact with their wild-type counterparts by gel filtration. *A,* representative results showing that three Smad4 mutants (indicated by a *purple line*) are unable to form a heterocomplex with Smad2, whereas one Smad4 mutant and one Smad2 mutant (highlighted by an *orange line*) retain their ability to form a heterodimer. *B,* schematic representation of the residues affected by the missense mutations. The MH2 domain of Smad4 is shown in *blue*. Mutation of the *purple* residues results in disruption of heterodimer formation. Mutation of the *orange* residues does not affect the formation of a heterodimer. Mutations in *parentheses* are introduced in Smad2. The two corresponding mutations in Smad2 and Smad4 exhibit an identical effect on the formation of a heterocomplex.

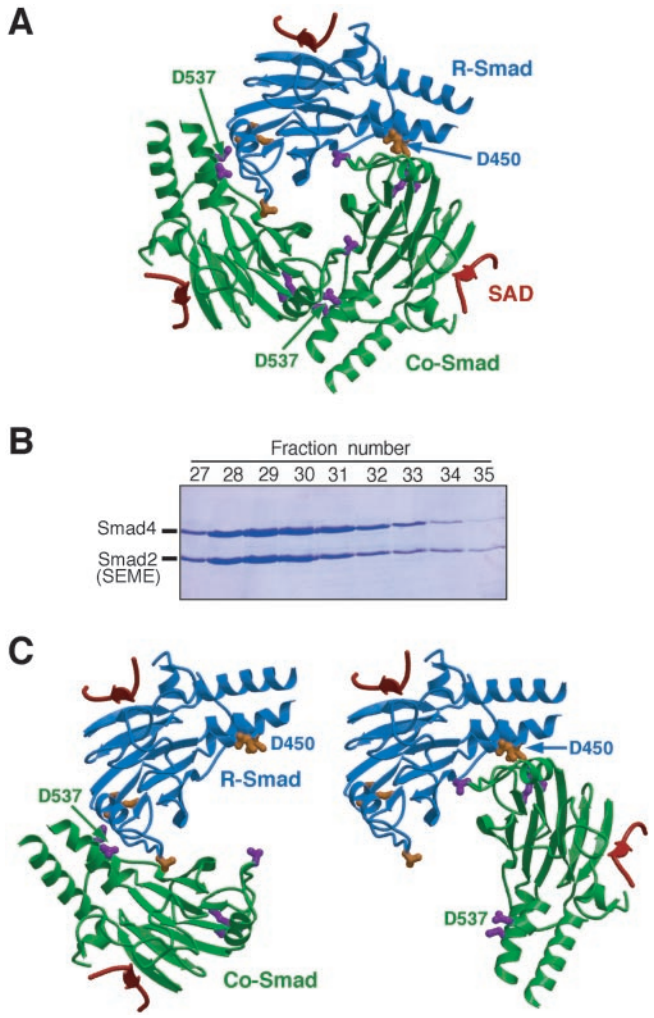

FIG. 5. **Comparison of proposed models of heterocomplex between Co-Smad and R-Smad.** In *A* and *C*, Co-Smad and R-Smad are colored *green* and *blue,* respectively. The critical interface residues from Co-Smad and R-Smad are shown in *purple* and *orange,* respectively. The SAD fragment is shown in *red*. *A*, proposed heterotrimer model. This model is inconsistent with several experimental observations (see "Results and Discussion"). This model shows a 2:1 complex between Co-Smad and R-Smad. The discussion in the text also applies to the other scenario, in which Co-Smad and R-Smad form a 1:2 complex. *B*, formation of a stable heterodimer between Smad4 (251–552) and a mutant Smad2 (241–467) in which the two C-terminal Ser residues (465 and 467) are replaced by Glu (*SEME*). Equimolar amounts of these two proteins were incubated together for 45 min before assay by size exclusion chromatography. *C*, two possible models of a heterodimer. Neither is consistent with our mutational analysis.

Comparison of Several Distinct Models for Heterocomplex Formation—Assessment of the effects of these tumorigenic mutations and other reported evidence allows us to evaluate several existing models on the formation of a heterocomplex between Smad2 and Smad4. The initial heterohexamer model was proposed on the basis of structural features of a homotrimeric Smad4 (9) but lacks experimental support. Subsequently, using gel filtration to assay overexpressed and epitope-tagged Smad proteins in mammalian cells, it was shown that the size of the heterocomplex between Smad2 and Smad4 is smaller than that of a Smad2 homotrimer (12); however, for reasons that are unclear, a heterotrimer model was proposed (12). More recently, using gel filtration of endogenous Smad proteins in mammalian cells, it was suggested that Smad2 forms a heterodimer with Smad4 (11).

In the heterotrimer model, the organization of the three subunits is believed to be very similar to that of the Smad4 homotrimer, except that one or two copies of Smad4 are now replaced by Smad2 (10, 12) (Fig. 5*A*). This model is inconsistent with several lines of experimental data. First, our mutational

data contradict this model. Asp-450 in Smad2 and the corresponding Asp-537 in Smad4 play a central role in this model, each making three intersubunit hydrogen bonds (9). Mutation of this residue is expected to completely disrupt a network of hydrogen bonds, leading to the disruption of the heterotrimeric packing (Fig. 5*A*). However, neither D450H in Smad2 nor D537H in Smad4 disrupted the formation of a heterocomplex (Table II; Fig. 5). Second, according to the heterotrimer model, SARA binding to Smad2 should not interfere with Smad2-Smad4 interactions because the SARA-binding surface of Smad2 is far away from the proposed heterotrimeric interface (19); however, SARA-Smad2 and Smad2-Smad4 complexes appear to antagonize each other (17, 18) (Fig. 1). Third, the heterotrimer model is inconsistent with the observation that the SAD domain is important for the formation of a heterocomplex (Fig. 3) because SAD is located in the periphery of the proposed heterotrimer (Fig. 5*A*). Fourth, a Smad2-Smad4 heterotrimer would have a larger molecular mass than that of a Smad2 homotrimer, which is inconsistent with the observation that the apparent molecular mass of a functional complex between Smad2 and Smad4 is appreciably smaller than that of a Smad2 homotrimer or a Smad4 homotrimer (11, 12). Fifth, according to the crystal structure of a transcriptionally active Smad4 fragment, the C-terminal phosphorylated SS*MS* motif in Smad2 may bind a specific site where two sulfate ions were found within 4 Å of each other (10). The linear distance between Cα of residue 545 in Smad4 and the closest sulfate ion of the two is over 32 Å, more than the head-to-tail distance of 8 residues, even if they are in their most extended conformations. Thus, in order to reach this site, a minimum of 8 residues is required between the Smad2 residue that corresponds to residue 545 in Smad4 and the first phosphorylated Ser residue in Smad2; however, there are only 6 residues in between them (9, 10). Finally, perhaps most importantly, if there were a heterotrimer between Smad2 and Smad4, then we should have observed this heterotrimer by either gel filtration or analytical ultracentrifugation because the heterotrimer formation should be heavily favored over an unstable heterodimer due to the cooperativity involved. However, we did not obtain any evidence supporting the existence of a heterotrimer.

In our studies, we used the unphosphorylated Smad2 and Smad4 proteins. Could this affect our final conclusion? We think that the answer is likely to be no. The major role of phosphorylation is to relieve the inhibitory effect of the MH1 domain and to release R-Smads from SARA and other proteins (14, 17). With the removal of the MH1 domain, the resulting Smad proteins are fully able to form heterocomplex between R-Smad and Co-Smad and are constitutively active in transcriptional assays (14–16). In fact, the MH2 domain of Smad2 in *Xenopus* is constitutively localized to the nucleus, and its developmental phenotype closely resembles that of the full-length Smad2 upon activin signaling (23). In addition, phosphorylation occurs in the C-terminal end of R-Smads, which is flexible in solution and disordered in all crystal structures (9, 10, 19). Thus, phosphorylation is very unlikely to alter the stoichiometry of a stable heterocomplex. It has been reported that substitution of the C-terminal Ser residues in R-Smad by Glu resulted in increased binding affinities between R-Smad and Co-Smad (24), presumably because the carboxylate side chain in Glu residues mimics the phosphate group. To assess the effect of such substitutions, we created a mutant Smad2 with the two C-terminal Ser residues replaced by Glu and purified this protein to homogeneity. Using gel filtration (Fig. 5*B*) and electrophoretic mobility shift assays (data not shown), this mutant Smad2 (SEME) is shown to form a stable heterodimer with Smad4 with a modest increase in binding affinity (Table II). Using this mutant Smad2, we also repeated all of the experiments reported in this study. In all cases, the results are identical to those for the wild-type Smad2. Despite this agreement, we caution against interpretation of results by using this mutant to substitute for the fully phosphorylated Smad2 protein. The reason is clear: a phosphorylated Ser is stereochemically different from the Glu residue, and the phosphorylated SS*MS* motif in Smad may show stringent specificity for binding.

On the basis of our biochemical and biophysical analysis, Smad2 clearly forms a heterodimer with Smad4. There are three scenarios for a heterodimer. Two scenarios involve packing interactions similar to that of the homotrimer, except that the relative positions of Smad2 and Smad4 could be switched (Fig. 5*C*); a third scenario involves a novel interaction interface, possibly psuedosymmetric. Our current data support the third scenario because the results with mutations D450H in Smad2 (D537H in Smad4) and D351H (D300H in Smad2) and R361H (R310H in Smad2) in Smad4 are not compatible with either of the first two scenarios.

Smad2 versus Smad3—Given the strong sequence similarity among R-Smads, it appears likely that Smad4 forms a heterodimer with other R-Smads. For example, Smad2 shares 92% sequence identity with Smad3, and both proteins are involved in signaling by TGF-β and activin. Nevertheless, the basal states of Smad2 and Smad3 differ considerably (11). Thus, the states of their heterocomplexes with Smad4 may also be different, as is the case for their biological functions. For example, Smad3 exhibits the highest sequence-specific DNA binding affinity; but Smad2 does not bind DNA because of an obstructing insertion immediately before the DNA-binding β-hairpin (25). Ectopic expression of the ubiquitin E3 ligase, Smurf2, selectively reduces the steady-state levels of Smad2 but not Smad3 (26). More importantly, Smad3-null mice are viable, but Smad2-null mice are not. Thus, it remains to be seen how Smad3 or Smad1 forms a heterocomplex with Smad4.

In summary, we conclude that the R-Smad Smad2 forms a stable heterodimer with Co-Smad Smad4. The formation of this complex requires the SAD domain in Smad4 and can be disrupted by a number of tumor-derived missense mutations in both Smad2 and Smad4. This finding should have broad implications in the interpretations of a range of biological experiments.

Acknowledgments—We thank F. Hughson for critical reading of the manuscript and S. Kyin for peptide sequencing, DNA synthesis, and mass spectroscopy.

REFERENCES

1. Roberts, A. B., and Sporn, M. B. (1990) in *Peptide Growth Factors and Their Receptors* (Sporn, M. B., and Roberts, A. B., eds), pp. 419–472, Springer-Verlag, Heidelberg
2. Massague, J. (1998) *Annu. Rev. Biochem.* **67,** 753–791
3. Heldin, C.-H., Miyazomo, K., and ten Dijke, P. (1997) *Nature* **390,** 465–471
4. Miyazono, K., ten Dijke, P., and Heldin, C. H. (2000) *Adv. Immunol.* **75,** 115–157
5. Wrana, J. L., and Attisano, L. (2000) *Cytokine Growth Factor Rev.* **11,** 5–13
6. Massague, J., Blain, S. W., and Lo, R. S. (2000) *Cell* **103,** 295–309
7. Massague, J., and Wotton, D. (2000) *EMBO J.* **19,** 1745–1754
8. Wrana, J. L., Attisano, L., Wieser, R., Ventura, F., and Massague, J. (1994) *Nature* **370,** 341–347
9. Shi, Y., Hata, A., Lo, R. S., Massague, J., and Pavletich, N. P. (1997) *Nature* **388,** 87–93
10. Qin, B., Lam, S. S. W., and Lin, K. (1999) *Structure* **7,** 1493–1503
11. Jayaraman, L., and Massague, J. (2000) *J. Biol. Chem.* **275,** 40710–40717
12. Kawabata, M., Inoue, H., Hanyu, A., Imamura, T., and Miyazono, K. (1998) *EMBO J.* **17,** 4056–4065
13. Laue, T., Shaw, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., eds), pp. 90–125, The Royal Society of Chemistry, Cambridge, United Kingdom
14. Hata, A., Lo, R. A., Wotton, D., and Massague, J. (1997) *Nature* **388,** 82–87
15. Liu, F., Hata, A., Baker, J. C., Doody, J., Carcamo, J., Harland, R. M., and Massague, J. (1996) *Nature* **381,** 620–623

16. Wu, R.-Y., Zhang, Y., Feng, X.-H., and Derynck, R. (1997) *Mol. Cell. Biol.* **17,** 2521–2528
17. Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L., and Wrana, J. L. (1998) *Cell* **95,** 779–791
18. Xu, L., Chen, Y.-G., and Massague, J. (2000) *Nat. Cell Biol.* **2,** 559–562
19. Wu, G., Chen, Y.-G., Ozdamar, B., Gyuricza, C. A., Chong, P. A., Wrana, J. L., Massague, J., and Shi, Y. (2000) *Science* **287,** 92–97
20. de Caestecker, M. P., Hemmati, P., Larisch-Bloch, S., Ajmera, R., Roberts, A. B., and Lechleider, R. J. (1997) *J. Biol. Chem.* **272,** 13690–13696
21. Schutte, M., Hruban, R. H., Hedrick, L., Cho, K. R., Nadasdy, G. M., Weinstein, C. L., Bova, G. S., Isaacs, W. B., Cairns, P., Nawroz, H., Sidransky, D., Casero, R. A., Meltzer, P. S., Hahn, S. A., and Kern, S. E. (1996) *Cancer Res.* **56,** 2527–2530
22. Uchida, K., Nagatake, M., Osaka, H., Yatabe, Y., Kondo, M., Mitsudomi, T., Masuda, A., and Takahashi, T. (1996) *Cancer Res.* **56,** 5583–5585
23. Baker, J. C., and Harland, R. M. (1996) *Genes Dev.* **10,** 1880–1889
24. Funaba, M., and Mathews, L. S. (2000) *Mol. Endocrinol.* **14,** 1583–1591
25. Shi, Y., Wang, Y.-F., Jayaraman, L., Yang, H., Massague, J., and Pavletich, N. P. (1998) *Cell* **94,** 585–594
26. Lin, X., Liang, M., and Feng, X.-H. (2000) *J. Biol. Chem.* **275,** 36818–36822