Molecular Cell, Vol. 15, 813–823, September 10, 2004, Copyright ©2004 by Cell Press

Structural Basis of Heteromeric Smad Protein Assembly in TGF-β Signaling

Benoy M. Chacko,[1,4] Bin Y. Qin,[1] Ashutosh Tiwari,[2] Genbin Shi,[3] Suvana Lam,[1] Lawrence J. Hayward,[2] Mark de Caestecker,[3] and Kai Lin[1,*]

[1]Department of Biochemistry and Molecular Pharmacology
[2]Department of Neurology
University of Massachusetts Medical School
Worcester, Massachusetts 01655
[3]Department of Nephrology
Vanderbilt University Medical Center
Nashville, Tennessee 37232

Summary

The formation of protein complexes between phosphorylated R-Smads and Smad4 is a central event in the TGF-β signaling pathway. We have determined the crystal structure of two R-Smad/Smad4 complexes, Smad3/Smad4 to 2.5 Å, and Smad2/Smad4 to 2.7 Å. Both complexes are heterotrimers, comprising two phosphorylated R-Smad subunits and one Smad4 subunit, a finding that was corroborated by isothermal titration calorimetry and mutational studies. Preferential formation of the R-Smad/Smad4 heterotrimer over the R-Smad homotrimer is largely enthalpy driven, contributed by the unique presence of strong electrostatic interactions within the heterotrimeric interfaces. The study supports a common mechanism of Smad protein assembly in TGF-β superfamily signaling.

Introduction

The transforming growth factor β (TGF-β) family of secreted ligands, which includes the TGF-βs, bone morphogenetic proteins (BMP), activins, and other factors, has a central role in mediating a host of essential cellular processes, including apoptosis, cell growth and differentiation, and extracellular matrix production (Attisano and Wrana, 2000; Blobe et al., 2000; de Caestecker et al., 2000; Derynck et al., 1998; Heldin et al., 1997; Massague and Wotton, 2000; Roberts, 1999). Binding of ligand to its corresponding transmembrane receptor kinase induces activation of the receptor through oligomerization and subsequent phosphorylation (Wrana et al., 1994). Activated receptor in turn phosphorylates and activates the receptor-specific Smad proteins (R-Smad) (Zhang et al., 1996). R-Smads direct TGF-β signaling in a pathway-specific manner. Smads 1, 5, and 8 mediate BMP signaling, while Smads 2 and 3 mediate TGF-β and activin signaling (Lagna et al., 1996). Following phosphorylation by the receptor, the R-Smads form heteromeric complexes with Smad4, the common mediator Smad. This activated heteromeric complex translocates into the nucleus, binds to specific promoters, and ultimately regulates target gene expression through interactions with other transcriptional coregulators. The R-Smads and Smad4 consist of three distinct domains. The N-terminal MH1 domain is responsible for DNA binding. The C-terminal MH2 domain mediates oligomerization and Smad-dependent transcriptional activation. Phosphorylation of the last two serine residues at the R-Smad C terminus promotes R-Smad/R-Smad and R-Smad/Smad4 complex formation (Abdollah et al., 1997; Souchelnytskyi et al., 1997). A proline-rich linker domain of variable length separates the MH1 and MH2 domains.

The stoichiometry of the active R-Smad/Smad4 heteromeric complex is ambiguous. Despite having been analyzed through various approaches, a direct structural demonstration of a heteromeric Smad complex is lacking. The crystallization of the Smad4 MH2 domain as a homotrimer first led to the suggestion that the active heteromeric Smad complex was a hexamer, composed of homotrimers of an R-Smad and Smad4 (Shi et al., 1997). Subsequent studies of overexpressed, epitope-tagged Smads in mammalian cells led to the proposal that the activated Smad complex was a heterotrimer assembled in a similar manner to the Smad4 homotrimer (Kawabata et al., 1998). This heterotrimer hypothesis was later supported by analytical ultracentrifugation and size exclusion chromatography analysis using purified Smad3 and Smad4, from which the observed stoichiometry was consistently 2:1 in favor of R-Smads (Chacko et al., 2001; Correia et al., 2001). Similar analysis of a BMP Smad, Smad1, gave further support to the heterotrimer model and suggested a general mechanism of Smad assembly in TGF-β signaling (Qin et al., 2001). Consistent with this model, many tumor-derived mutations map to the conserved trimer interfaces, implicating trimer formation as an essential mechanism in Smad signaling (Shi et al., 1997). However, in opposition to the heterotrimer model, other studies have suggested a heterodimeric model for Smad2/Smad4 assembly. Size exclusion chromatography analysis of endogenous Smad2/Smad4 complexes from TGF-β-stimulated cells suggested that these complexes were heterodimers (Jayaraman and Massague, 2000). Subsequent biochemical analysis using purified Smad2 and Smad4 also supported a heterodimeric model (Wu et al., 2001a). These studies may suggest a fundamental difference between the Smad3/Smad4 (heterotrimer) and Smad2/Smad4 (heterodimer) complexes, but subsequent analysis of nuclear extracts of cells transiently cotransfected with differentially tagged Smads suggested the opposite. The DNA-bound Smad2 and Smad4 formed a heterotrimer consisting of two Smad2 and one Smad4 when bound to the cofactors Fast-1/Foxh1 or Fast-3/Foxh1b, while the DNA-bound Smad3/Smad4 complex was a heterodimer (Inman and Hill, 2002).

In addition to the uncertainty surrounding R-Smad/Smad4 complex stoichiometry, the structural basis for the preferential formation of heteromeric Smad complex over the homomeric species is similarly unknown. In the

*Correspondence: kai.lin@umassmed.edu
[4]Present Address: Biotechnology Discovery Research, Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana 46285.

absence of Smad4, R-Smad phosphorylation results in homotrimerization (Chacko et al., 2001; Kawabata et al., 1998; Qin et al., 2001; Wu et al., 2001b). The R-Smad homotrimer is stabilized by the interaction between phosphoserine residues of the C-terminal tail and positively charged residues in the L3 loop/β8 strand pocket of an adjacent subunit, as well as through extensive hydrogen bonds and van der Waals contacts at the conserved trimer interface. However, when Smad4 is present, the R-Smad/Smad4 heteromer forms preferentially over the R-Smad homotrimer. This occurs despite Smad4 not having a phosphorylated C terminus and thus not being able to contribute phosphoserine residues to the L3 loop/β8 strand of its neighboring R-Smad. Therefore Smad4 is presumably able to contribute other favorable interactions that can offset the loss of interaction between the phosphorylated C-terminal tail and the loop-strand pocket.

To resolve the stoichiometric identity of the heteromeric R-Smad/Smad4 complex, and provide insights into the basis for its preferential formation over the R-Smad homotrimer, we determined the crystal structure of both Smad3/Smad4 and Smad2/Smad4 complexes. The structures, in corroboration with functional studies, reveal the presence of unique favorable electrostatic interactions within the heteromeric interfaces, and support a unifying mechanism of heterotrimeric Smad assembly, consisting of two R-Smads and one Smad4, in TGF-β signaling.

Table 1. Summary of Crystal Analysis for Smad2/Smad4 and Smad3/Smad4 Complexes

Parameter	Smad2/Smad4	Smad3/Smad4
Crystal Parameters and Crystallographic Data		
Space group	P2(1)2(1)2(1)	P2(1)2(1)2(1)
Unit cell dimensions	a = 49.24	a = 48.47
	b = 59.95	b = 60.55
	c = 207.44	c = 205.78
Diffraction limit (Å)[a]	100–2.7 (2.76–2.7)	100–2.5 (2.6–2.5)
Total reflections	75,537	111,193
Unique reflections	14,835	21,181
Completeness (%)	85.3 (84.4)	96.5 (87.2)
Intensity/Sigma	16.7 (3.9)	27.5 (3.0)
R_{merge} (%)[b]	7.5	5.2
Refinement Statistics		
Protein atoms	4,777	4,728
Water molecules	0	257
R factor (%)[c]	24.3	23.2
R_{free} factor (%)[d]	27.9	24.7
Rms deviation from ideal		
Bond lengths (Å)	0.011	0.008
Bond angles(°)	1.6	1.7
B factor rms deviation		
Main chain (Å^2)	2.96	1.45
Side chain (Å^2)	4.41	1.74
Overall B factor (Å^2)		
Subunit A	54.3	52.4
Subunit B	84.4	77.9
Subunit C	64.0	61.8

[a] Values in brackets are for the highest resolution shell.
[b] $R_{merge} = \Sigma|I_{hkl} - <I_{hkl}>|/\Sigma\ I_{hkl}$
[c] R factor $= \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$ for all data.
[d] R free $= \Sigma_{hkl}||F_{obs}| - |F_{calc}||/\Sigma_{hkl}|F_{obs}|$ for 10% of the data not used in refinement.

Results

Crystallographic and Calorimetric Evidence for the Stoichiometry of the Heteromeric Smad Complex

The phosphorylated Smad3 and Smad2 constructs used to analyze the R-Smad/Smad4 interaction in these studies include the linker and MH2 domains. The phosphorylated Smad3 construct extends from residues 145–425 and is referred to as S3LC(2P). The phosphorylated Smad2 construct extends from residues 186–467 and is referred to as S2LC(2P). The Smad4 construct used comprises the MH2 domain and part of the linker domain, extending from residues 273–552, and is referred to as S4AF (Qin et al., 1999). The heteromeric Smad complex consisting of S3LC(2P) and S4AF was obtained by combining S4AF in 2-fold molar excess over S3LC(2P) in solution, then separating the complex from free S4AF through size exclusion chromatography. Attempts to crystallize this complex were unsuccessful. The complex was then subjected to a limited chymotrypsin digest and further purified through size exclusion chromatography. This chymotrypsin treatment produced a form of the complex which yielded diffraction-quality crystals. SDS-PAGE, N-terminal sequencing, and mass spectrometry analysis revealed that the final heteromeric Smad complex lost most of the linker sequence as a result of proteolysis. Smad3 extends from residue 228 to the C terminus. Smad4 contains residues 309 to the C terminus in two fragments, as a result of an internal chymotrypsin cut after Leu484. Leu484 is located within a flexible, solvent-accessible region connecting helix 3 and 4 in the Smad4 MH2 domain structure (Qin et al., 1999). This region is presumably also flexible in the heteromeric Smad complex, as indicated by its susceptibility to chymotrypsin digestion. An identical method was used to produce a crystallizable form of the Smad2/Smad4 complex. Both structures were solved through molecular replacement, using the previously solved structure of phosphorylated Smad2 as a search model (Table 1) (Wu et al., 2001b). The Smad3/Smad4 structure was refined to 2.5 Å, and the Smad2/Smad4 structure was refined to 2.7 Å. The crystal structures reveal that both protein complexes are heterotrimeric, consisting of two phosphorylated R-Smad subunits and one Smad4 subunit (Figure 1A). The R-Smad subunits in both crystal structures were designated subunits A and C, and the Smad4 subunit was designated subunit B. These three subunits arrange in a manner similar to the structures of the previously solved homotrimers of Smad4 (Qin et al., 1999; Shi et al., 1997), Smad1 (Qin et al., 2001), and Smad2 (Wu et al., 2001b).

Isothermal titration calorimetry (ITC) was also used to analyze the thermodynamics and stoichiometry of heteromeric Smad complex formation by titrating S2LC(2P) or S3LC(2P) with S4AF (Figure 1B; Table 2). S4AF titrates both S2LC(2P) and S3LC(2P) at a relative molar ratio of 0.5, meaning that a single molecule of Smad4 can interact with two molecules of S2LC(2P) or S3LC(2P). These results are consistent with the stoichi-

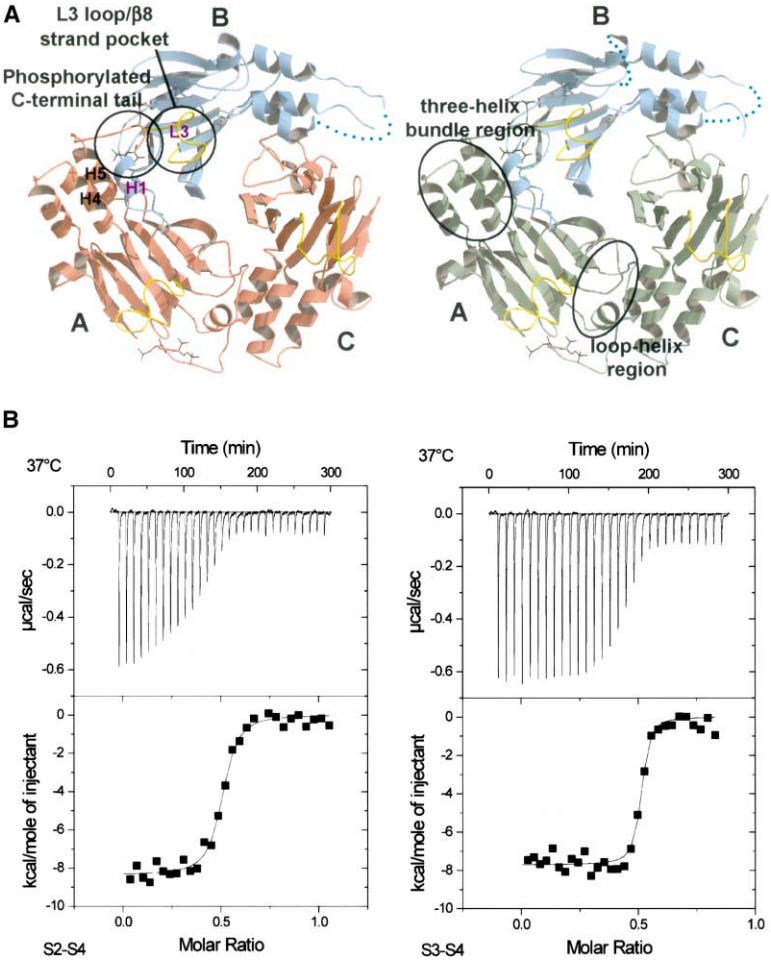

Figure 1. Stoichiometry of the R-Smad/Smad4 Complex

(A) Crystal structures of the Smad2/Smad4 (left) and Smad3/Smad4 (right) complexes. The Smad2 subunits are in red. The Smad3 subunits are in green. The Smad4 subunits in both complexes are in cyan. The L3 loops are in yellow. The phosphoserine side chains are in stick presentation.

(B) ITC analysis of the interaction between Smad2 and Smad4 (left), and Smad3 and Smad4 (right). The thermodynamic parameters derived from the fitting are shown in Table 2.

ometry observed in the crystal structure. The apparent dissociation constant (K_d) values are 60 nM for Smad3/Smad4 and 300 nM for Smad2/Smad4, corresponding to a change in Gibbs' free energy (ΔG) of -10 kcal/mol of Smad3/Smad4 complex formation, and -9 kcal/mol of Smad2/Smad4 complex formation (Table 2). However, it is important to note that both S2LC(2P) and S3LC(2P) exist as homotrimers prior to titration with Smad4, whereupon a single Smad4 subunit displaces a subunit in the R-Smad homotrimer to form the R-Smad/Smad4 heterotrimer. Therefore, the reaction heat obtained from calorimetry represents the difference between the association and dissociation events, as well as the reassociation of displaced Smad3 to form heterotrimers, and the derived K_d values do not represent the true dissociation constants of Smad4 interacting with the R-Smads. The ΔG resulting from heteromeric complex formation is driven primarily by a favorable enthalpy (ΔH), suggesting that polar interactions play significant roles in the preferential interaction between R-Smads and Smad4 (Boulanger et al., 2003; Park et al., 1999; Ye et al., 1999). The favorable entropy term ($T\Delta S$) likely originates in part from the increase in the buried surface area of the reaction. Although the calculated buried surface area for the heterotrimer is smaller than the homotrimer (7100 versus 8000 Å^2) due to Smad4 lacking a tail interaction with the neighboring subunit, there are 1.5-fold more heterotrimers when the reaction is complete than homotrimers at the beginning of the reaction. The small difference in the thermodynamics of interaction

Table 2. Summary of ITC of Smad4 (S4AF) Interaction with Phosphorylated Smad3 or Smad2 and Mutants of Smad3

	Smad3/S4AF	Smad2/S4AF	Smad3/S4AF(R378A)	Smad3/S4AF(D493A)
n	0.50 ± 0.001	0.49 ± 0.01	0.49 ± 0.001	0.64 ± 0.02
ΔH (kcal/mol)	−7.5 ± 0.2	−8.2 ± 0.2	−8.5 ± 0.7	−3.8 ± 0.2
ΔS (cal/mol/K)	8.8 ± 0.6	3.7 ± 0.9	0.01	11.9
ΔG (kcal/mol)	−10.0 ± 0.1	−9.1 ± 0.4	−8.3	−7.3
K_d (nM)	58 ± 2.9	296 ± 165.8	1090	5282
T (K)	310	310	310	310
$T\Delta S$ (kcal/mol)	2.7	1.1	0.003	3.7

between Smad2 and Smad3, which are approximately 97% identical between their respective MH2 domains may be partly explained by the approximately 25% difference in sequence between the linker regions of the two constructs.

Overall Structure of the R-Smad/Smad4 Heterotrimer

The overall structural arrangement of the R-Smad/Smad4 complex is similar to what has been observed in previously solved crystal structures of R-Smad and Smad4 homotrimers. Each of the subunits consists of a central β sandwich flanked on one side by three helices (three helix bundle) and on the other side by three loops and another helix (loop-helix region) (Figure 1A). The previously solved homotrimeric structures of Smad1 (Qin et al., 2001) and Smad2 (Wu et al., 2001b) revealed that subunit-subunit contact is mediated specifically by the interaction of helix H4, helix H5, and the phosphorylated C-terminal tail of one subunit with helix H1, the β4 strand, and the L3 loop/β8 strand pocket, respectively, of a neighboring subunit. This general feature of interface formation is conserved in the R-Smad/Smad4 heterotrimer, with the exception that there are only two phosphorylated C-terminal tail interactions in the heterotrimers, one between subunit A and B and the other between subunit C and A, due to Smad4 lacking a phosphorylated tail (Figure 1A). In addition, Smad4 possesses a unique conformation that results in closer contacts between subunit B and C in the heterotrimer. Superposition of the heterotrimer subunits to those of the previously solved Smad2 homotrimer reveals that the three helix bundle of Smad4 subunit in the heterotrimer undergoes a significant shift toward the neighboring R-Smad subunit when compared to the corresponding subunit in the Smad2 homotrimer (Figure 2A). This unique conformation of Smad4 likely results in improved subunit B to C interactions and contributes to the preferential formation of the R-Smad/Smad4 heterotrimer over the R-Smad homotrimer.

To gain insights into the structural mechanism of preferential heterotrimer formation, the interface contacts of the heterotrimeric crystal structures were compared with those of the previously solved crystal structure of the Smad2 homotrimer. The presence of Smad4 in the heterotrimer results in three interfaces that are distinct from each other, designated AB (between subunit A and B), BC (between subunit B and C), and CA (between subunit C and A). This is in contrast to the Smad2 homotrimer, which contains three identical interfaces. Since the structures of the Smad3/Smad4 and Smad2/Smad4 heterotrimeric complexes are virtually identical, with an overall root-mean-square deviation of only 0.2 Å, most of the structural analysis described below was done on the higher resolution Smad3/Smad4 structure. Overall, there are fewer hydrogen bond interactions in the BC and AB interfaces, although all three heterotrimeric interfaces possess a common subset of hydrogen bond interactions that are also present in the homotrimeric structure (Table 3). Most notably, Smad4 possesses unique amino acid residues not present in the R-Smads that result in new electrostatic interactions in the BC and AB interfaces (see below). These favorable electrostatic interactions likely compensate for the smaller number of hydrogen bonds, thus accounting for the favorable enthalpy change upon heterooligomerization.

Asp493 of Smad4 Coordinates a Buried Electrostatic Interaction in the BC Interface

The smaller number hydrogen bonds in the BC interface is due largely to Smad4 lacking a phosphorylated C-terminal tail; therefore, the interaction between the phosphorylated tail and L3 loop/β8 strand pocket that is characteristic of the R-Smad homotrimeric interfaces does not occur at this interface (Figure 2B, compare heteromeric BC interface with the homomeric interface). The most striking feature of this interface is a network of salt bridges and hydrogen bonds centered on Asp493 of helix H4 of the Smad4 (B) subunit. Asp493 is situated in the center of four arginine residues, Arg321/279 and Arg329/287 of the C subunit (this numbering system will be used to refer to identical residues in Smads 2/3), and Arg496 and Arg497 of the B subunit (Figures 2B, 3A, and 3B). Asp493 is able to form intermolecular salt bridges with the two arginine residues of the C subunit, as well as an intramolecular salt bridge with Arg496. In addition, the backbone oxygen of Asp493 forms an intramolecular hydrogen bond with Arg497. Asp493 of Smad4 was found to be mutated in pancreatic and other cancers, suggesting that this residue has a crucial role in mediating Smad-dependent signaling (Hahn, 1996). The residue in the position equivalent to Asp493 in the R-Smads is a tyrosine. Tyr406/364 forms a single hydrogen bond with Arg329/287 of the neighboring R-Smad subunit in the homotrimer and, therefore, does not appear to provide as much stability to the interface as Asp493 of Smad4 (Figure 3A).

To examine the role of Asp493 in the R-Smad/Smad4 interaction, the ability of a D493A mutant to form heterotrimeric complexes was analyzed through size exclusion chromatography and ITC. As opposed to wild-type S4AF, there is a dramatic decrease in the amount of S4AF(D493A) that coeluted with S3LC(2P) or S2LC(2P) on the size exclusion column, indicating that interaction of the S4AF(D493A) mutant with phosphorylated R-Smads to form heterotrimers is significantly impaired (Figure 3C). The potency of this mutation is underscored by ITC experiments, which showed that S4AF(D493A) has little interaction with S3LC(2P), as evidenced by a much smaller heat of binding detected as compared to the wild-type (Figure 3D; Table 2). These results were further confirmed by immunoprecipitation experiments. COS cells were transfected with full-length versions of Smad2 or Smad3, with full-length WT Smad4 or the Smad4(D493A) mutant, in the presence or absence of a constitutively active TGF-β type I receptor (Wieser et al., 1995). Immunoprecipitation of Smad2 or Smad3 revealed that the Smad4(D493A) mutant exhibits a dramatic decrease in association with either R-Smads as compared to the wild-type Smad4 (Figure 3E). Previous biochemical analysis has suggested a dimeric model for phosphorylation-dependent Smad2/Smad4 interaction, consisting of only subunit A and B (Wu et al., 2001b). Such model would predict that Asp493 of Smad4 not playing a role in heteromeric Smad assembly, as it is located on the surface that is not in contact with the R-Smad. The

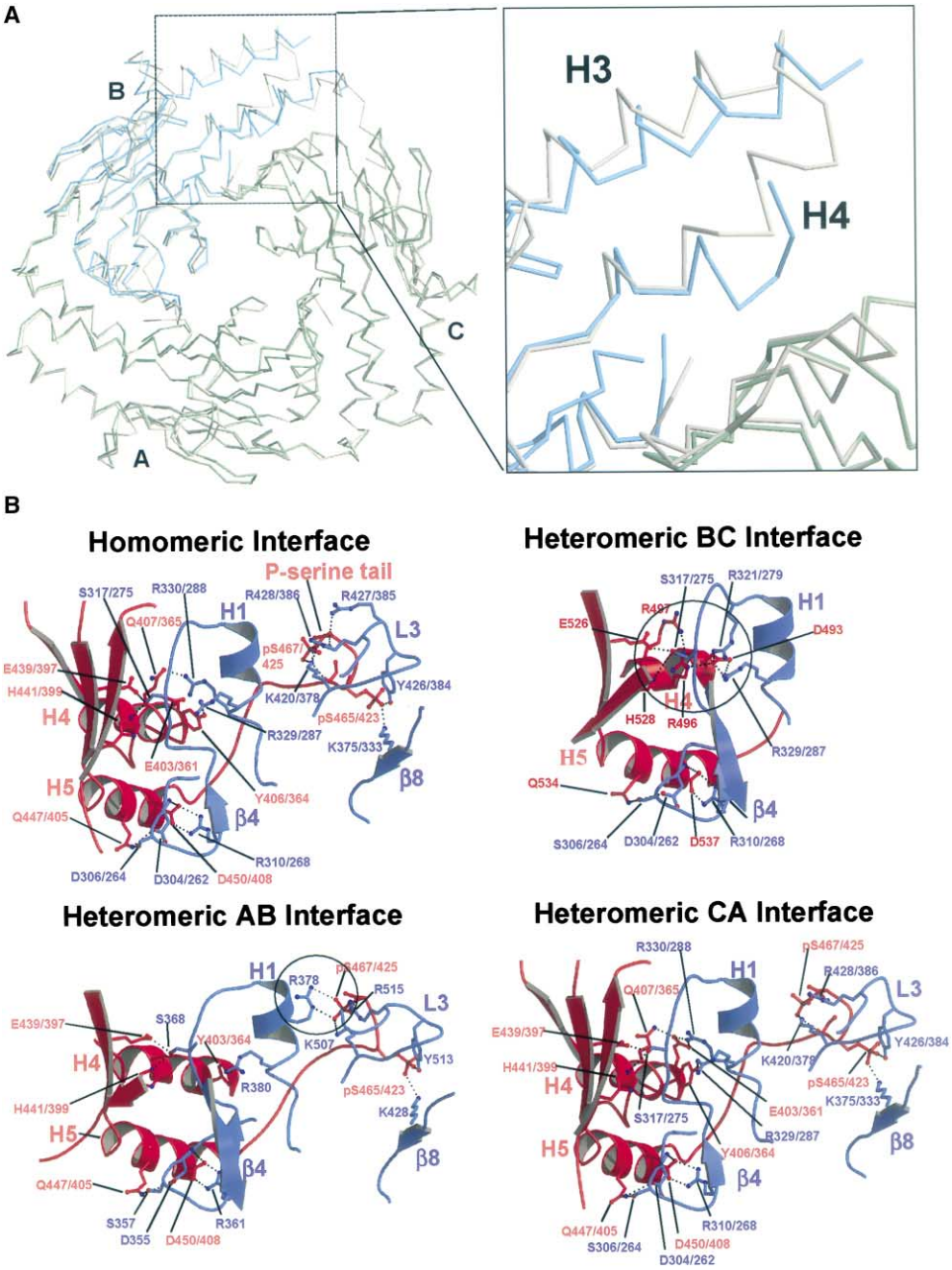

Figure 2. Structural Comparison between the R-Smad/Smad4 Heterotrimer and R-Smad Homotrimer

(A) Left, superposition of the α carbon traces of the Smad3/Smad4 heterotrimer and the Smad2 homotrimer. The homotrimeric structure of Smad2 is in gray. The Smad3 and Smad4 subunits of the heterotrimer are in green and cyan, respectively. The boxed region shows the major conformational difference, involving helix H3 and H4, between the two structures. Right, close-up view of the helix H3/H4 region of Smad4 and the corresponding region in Smad2 is shown. Helices H5, which are in similar position in both structures, are removed in the close-up view for clarity.

(B) Comparison of hydrogen bonding interactions between the R-Smad/Smad4 heterotrimeric interfaces (AB, BC, and CA interfaces) and the Smad2 homotrimeric interface. The three helix bundle regions and the phosphorylated C-terminal tails are in red. The loop-helix regions and the L3 loop/β8 pockets are in blue. The unique features in the BC and AB interfaces, which are likely the key contributors to the preferential formation of the heterotrimer, are circled.

critical role of Smad4(D493) revealed here strongly opposes this dimeric model. The presence of Asp493 at the interface likely coordinates the positive charge contributed by the four surrounding arginine residues, maintaining a charge balance that allows the interface to form. It was reported that mutation of Asp493 in Smad4 to histidine did not affect heteromeric interaction with phosphorylated Smad2 (Wu et al., 2001b). One possibility may be due to His493 able to make some hydrogen bonds with the arginines of the neighboring subunit. Other studies have shown that unphosphorylated Smad2 forms a heterodimer with Smad4 (Wu et al.,

Table 3. Comparison of Intersubunit Hydrogen Bond Interaction between the Homotrimeric Interface and the Heterotrimeric Interfaces

		Smad3/Smad4 Heteromeric Interfaces					
Smad2 Homomeric Interface		AC Interface		BC Interface		AB Interface	
Smad2	*Smad2*	*Smad3*	*Smad3*	*Smad4*	*Smad3*	*Smad3*	*Smad4*
N283 O	N320 OD	N241 O	N278 OD	D332 O	N278 OD		
N283 ND	N320 OD	N241 ND	N278 OD	D332 OD	N278 OD	N241 ND	H371 NE
H291 NE	*E288 OE*[a]	H249 NE	*E246 OE*	K340 NZ	*E246 OE*		
Y406 OH	*R329 NH*	Y364 OH	*R287 NH*			Y364 OH	*R380* NH
Q407 OE	R330 NH						
E439 OE[b]	***S317 OG***	***E397 OE***	***S275 OG***	***E526 OE***	***S275 OG***	***E397 OE***	***S368 OG***
H441 NE	***S317 OG***	***H399 NE***	***S275 OG***	***H528 NE***	***S275 OG***	***H399 NE***	***S368 OG***
L446 O	T303 OG	*L404 O*	T261 OG	*L533 O*	T261 OG		
Q447 NE	***D304 OD***	***Q405 NE***	***D262 OD***	***Q534 NE***	***D262 OD***	***Q405 NE***	***D355 OD***
Q447 NE	***S306 OG***	***Q405 NE***	***S264 OG***	***Q534 NE***	***S264 OG***	***Q405 NE***	***S357 OG***
D450 N	T303 OG	*D408 N*	T261 OG	*D537 N*	T261 OG		
D450 OD	***R310 NH***	***D408 OD***	***R268 NH***	***D537 OD***	***R268 NH***	***D408 OD***	***R361 NH***
D450 OD	T303 N	D408 OD	T261 N	D537 OD	T261 N		
D450 OD	***D304 N***	***D408 OD***	***D262 N***	***D537 OD***	***D262 N***	***D408 OD***	***D355 N***
V461 O	*H331 NE*	I419 O	*H289 NE*			I419 O	*H382 NE*
S464 O	*K420 NZ*	S422 O	*K378 NZ*				
pS465 O	*G421 N*	pS423 O	*G379 N*			pS423 O	*G508 N*
pS465 OP	*K375 NZ*	pS423 OP	*K333 NZ*			pS423 OP	*K428 NZ*
pS465 OP	*Y426 OH*	pS423 OP	*Y384 OH*			pS423 OP	*Y513 OH*
pS467 OP	*K420 NZ*	pS425 OP	*K378 NZ*			pS425 OP	*K507 NZ*
pS467 O	*K420 NZ*	pS425 O	*K378 NZ*				
pS467 O	R427 NH						
pS467 O	*R428 NH*	pS425 O	*R386 NH*			pS425 O	*R515 NH*
		*E361 OE[c]	R288 NH				
				*D493 OD	*R279 NH*		
				*D493 OD	*R287 NH*		
						*pS425 OP	*R378 NE*
						*pS425 OP	*R378 NH*

[a] Residues that are conserved in both the Smad4 and the R-Smads are in italics.
[b] Hydrogen bond interactions conserved in both homomeric and heteromeric interfaces are shown in bold.
[c] Unique hydrogen bond interactions in the heteromeric interfaces.

2001a). The observed heterodimer may be the product of the highly favorable interaction between the B and C subunits in the crystal structure. It is likely that the potent electrostatic interaction in the BC interface allows formation of an intermediate heterodimeric species even in the absence of R-Smad phosphorylation, while phosphorylation drives the complex toward trimerization.

In addition to the favorable electrostatic interactions, the BC interface also displays a better interface packing compared to the R-Smad homotrimeric interfaces. Interface shape complementarity was measured using the gap volume index statistic from the Protein-Protein Interaction Server (Jones and Thornton, 1996). This statistic is a ratio of the interface gap volume to interface buried surface area. More complementary interfaces are indicated by a smaller gap volume index. The gap volume index value for the heterotrimeric BC interface is 2.55, compared to 2.48 for the Smad2 homotrimeric interface. However, the slightly smaller value for the homotrimer is due largely to the highly favorable interactions between the phosphorylated C-terminal tail and the L3 loop/β8 strand pocket of the neighboring subunit, which is lacking in the BC interface. Deletion of the five C-terminal residues of R-Smads results in a homotrimer interface gap volume index of 3.35. This indicates that the core interaction between helices H4 and H5 of Smad4 and helix H1 and the β4 strand of Smad3 in the heterotrimer is significantly tighter than the equivalent interaction in the R-Smad homotrimer. The improvement in surface complementarity of the BC interface is due in part to a shift in the helices H4 and H5 of the three-helix bundle of Smad4 in the heterotrimer as compared to the homotrimer, as described above (Figure 2A).

Arg378 of Smad4 Reinforces Phosphoserine Binding in the AB Interface

The AB interface of the R-Smad/Smad4 heterotrimer also contains fewer hydrogen bonds than the R-Smad homotrimeric interfaces (Table 3). This is primarily due to the lack of hydrogen bond interactions between helix H4 (A subunit) and helix H1 (B subunit), which are present in all Smad interfaces except this one. The unique feature of this interface is the presence of a hydrogen bond interaction network between Arg378 (helix H1) of the Smad4 (B) subunit and the tail phosphoserine residue at position 467/425 of the R-Smad (A) subunit (Figure 2B, compare heteromeric interface AB with the homomeric interface, and Figure 4B). This interaction occurs in the R-Smad/Smad4 heterotrimer, but not in the R-Smad homotrimer because the corresponding arginine residue is not conserved in the R-Smads. Among the R-Smad isoforms, this position is an asparagine, methionine, or leucine, which cannot provide the charge to stabilize the phosphoserine. Among the Smad4 isoforms, this position is almost exclusively arginine, with the exception of *C. elegans* Sma-4, which has a lysine at this position, and can presumably play a similar role to stabilize the phosphoserine. The presence of Arg378 appears

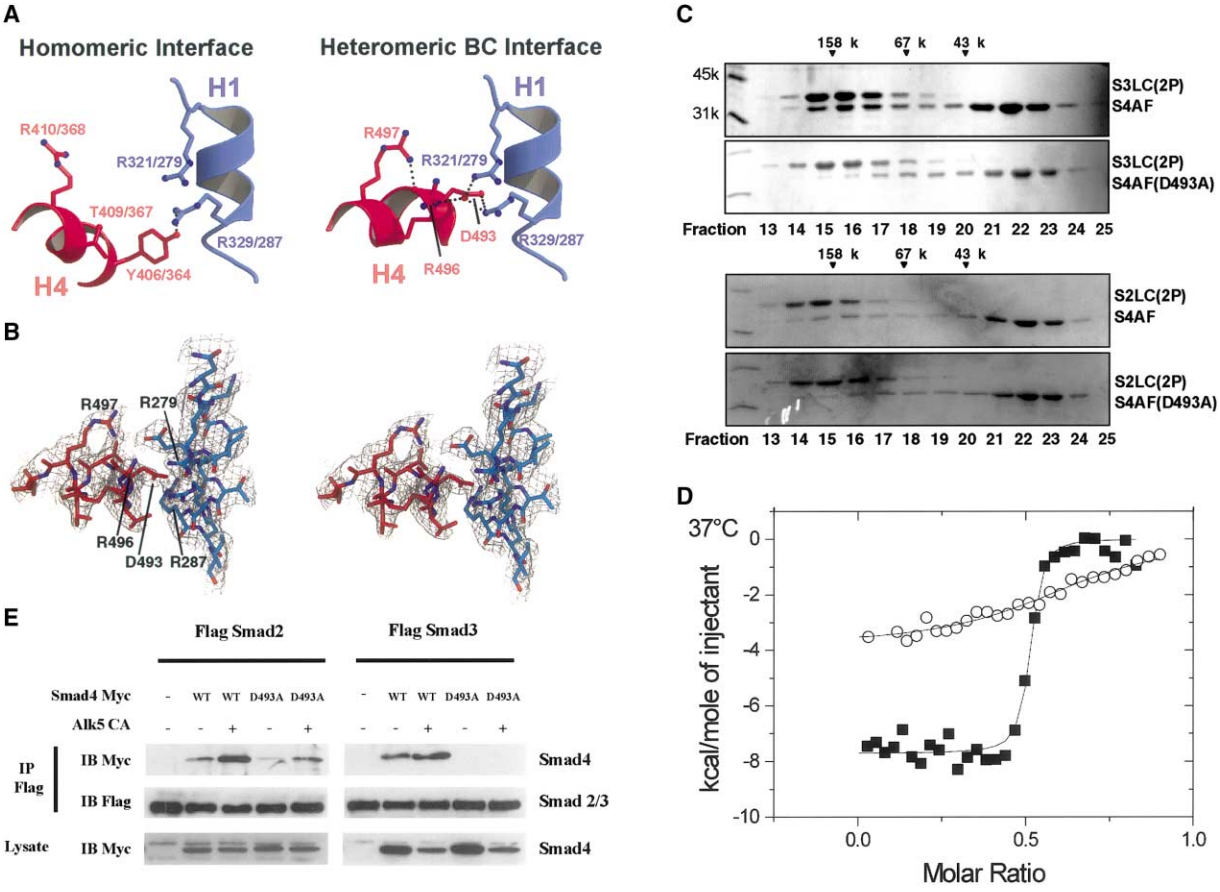

Figure 3. Asp493 of Smad4 Coordinates a Buried Electrostatic Interaction in the BC Interface

(A) Close-up view of the helix 4-helix 1 interaction in the Smad2 homotrimer (left) and the BC interface of the Smad3/Smad4 heterotrimer (right). The presence of aspartic acid at position 493 of Smad4 (as opposed to Tyr406 of Smad2), coupled with the shift in helix H4, results in highly favorable electrostatic interactions between Asp493 and four surrounding arginine residues. This results in improved interface contacts over those in the Smad2 homotrimeric interface.

(B) The stereoview of the $2F_o - F_c$ simulated annealed omit map of the Asp493 region of S4AF.

(C) Size exclusion chromatography analysis of interaction between the wild-type S4AF or S4AF(D493A) mutant with S3LC(2P) or S2LC(2P). S4AF or S4AF(D493A) mutant was mixed equal molar ratio to the S3LC(2P) or S2LC(2P), and the mixture was loaded to the size exclusion column. The eluted fractions were run on a denaturing gel and stained with Coomassie blue. The excess, uncomplexed S4AF or S4AF(D493A) eluted as a monomer with the peak around fraction 22. Interaction between the wild-type S4AF with S2LC(2P) or S3LC(2P) is indicated by their coelution as a heteromeric complex with the peak position around fraction 15. The lack of an interaction in the case of the S4AF(D493A) mutant is indicated by the near absence of the S4AF(D493A) band in the complex fractions.

(D) ITC analysis of the interaction between S4AF(D493A) and S3LC(2P). The S4AF(D493A) data are in open circles. The wild-type S4AF data are in filled squares for comparison.

(E) Coimmunoprecipitation analysis of Smad2/Smad4 and Smad3/Smad4 interaction. COS-1 cells were transfected with FLAG-tagged full-length Smad2 or Smad3 with or without Myc-tagged Smad4 or Smad4(D493A), and with or without the constitutively active TGF-β type I receptor (Wieser et al., 1995). Complexes were immunoprecipitated from cell extracts with an anti-FLAG antibody and immunoblotted with either anti-FLAG or anti-Myc antibody as indicated.

to make Smad4 a better receptor for the phosphorylated C-terminal tail of the adjacent R-Smad subunit (A subunit) by increasing the basic electrostatic potential surface of the phosphoserine binding pocket (Figure 4A). This additional helix H1/phosphorylated tail interaction likely compensates for the lack of hydrogen bonds between helix H4 and helix H1.

The importance of Arg378 was confirmed by mutating Arg378 to alanine and analyzing the effect of the mutant on heteromeric interaction with the phosphorylated R-Smads. On the size exclusion column, S4AF(R378A) was able to interact to some extent with S3LC(2P) or S2LC(2P), as evidenced by its coelution with S3LC(2P) or S2LC(2P) (Figure 4C). However, the interaction mediated by S4AF(R378A) appeared to be weaker, as the coeluted S4AF(R378A) band was consistently of lower intensity as compared to wild-type S4AF (Figure 4C). To more quantitatively analyze the effect of the R378A mutant, ITC analysis was carried out, which showed that the ability of the S4AF(R378A) mutant to interact with phosphorylated Smad3 was substantially reduced from the wild-type protein, as reflected by a near 20-fold increase in the apparent K_d (Figure 4D; Table 2).

The CA Interface Resembles the Homotrimeric Interface

The CA interface formed by the two R-Smad subunits is the only interface that is conserved in both the homo-

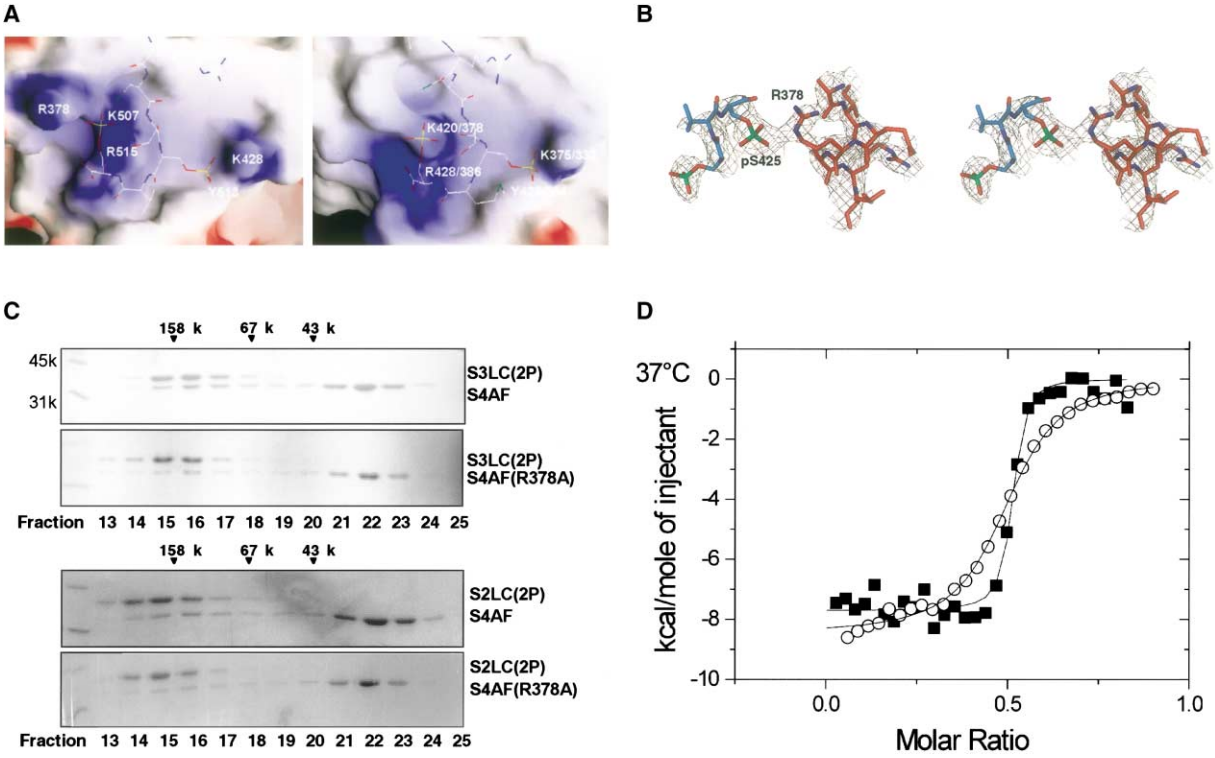

Figure 4. Arg378 of Smad4 Reinforces Phosphoserine Binding in the AB Interface

(A) Surface electrostatic potential presentation of the L3 loop/β8 strand pocket for Smad4 (left) and Smad2 (right). The R-Smad C-terminal phosphoserine structures are shown in stick presentation.

(B) The stereoview of the $2F_o - F_c$ simulated annealed omit electron density map of the Arg378 region of S4AF.

(C) Size exclusion chromatography analysis of interaction between the wild-type S4AF or S4AF(R378A) mutant with S3LC(2P) or S2LC(2P). S4AF or S4AF(R378A) mutant was mixed equal molar ratio to the S3LC(2P) or S2LC(2P), and the mixture was loaded to the size exclusion column. The eluted fractions were run on a denaturing gel and stained with Coomassie blue. The excess, uncomplexed S4AF or S4AF(R378A) eluted as a monomer with the peak around fraction 22. Interaction between the wild-type S4AF with S2LC(2P) or S3LC(2P) is indicated by their coelution as a heteromeric complex with the peak position around fraction 15. The decreased interaction in the case of the S4AF(R378A) mutant is indicated by the lower staining of the S4AF(R378A) band in the complex fractions as compared to the wild-type S4AF.

(D) ITC analysis of the interaction between S4AF(R378A) and S3LC(2P). The S4AF(R378A) data are in open circles. The wild-type S4AF data are in filled squares for comparison.

trimer and heterotrimer, and is essentially identical in both structures (Figure 2B, compare heteromeric interface CA with the homomeric interface, and Table 3). The hydrogen bond networks that mediate the interaction of helix H5 and the phosphorylated tail of the C subunit with the β4 strand and L3 loop/β8 strand pocket of the A subunit are nearly the same between homotrimer and heterotrimer. A subtle difference in the two interfaces is found in the hydrogen bonds present between helix H4 (C subunit) and helix H1 (A subunit). The hydrogen bond that exists between Gln407 (helix H4) of the C subunit and Arg330 (helix H1) of the A subunit of the Smad2 homotrimer is not present in the heterotrimer, but is replaced by a salt bridge between Glu403/361 (Smad2/Smad3) of the C subunit with Arg330/288 of the A subunit. Gln407/365 remains involved in the heterotrimeric interface by forming an intramolecular hydrogen bond with Glu403/361. In the homotrimer, Glu403/361 is positioned toward the solvent and, therefore, appears to have no role in the homotrimeric AC interface. In the heterotrimer, however, Glu403/361 undergoes a conformational change that causes its side chain to flip approximately 6 Å to come in contact with Arg330/288 of the adjacent A subunit. However, mutation of Glu361 to an alanine residue does not result in a noticeable defect in the phosphorylation-induced R-Smad/Smad4 interaction as determined through gel filtration. ITC analysis revealed only an approximately 2-fold increase in K_d compared to wild-type (data not shown). The result suggests that the conformational change of Glu403/361 in the AC interface, and its associated change in H-bonding pattern, does not have a salient effect on the preferential formation of the heterotrimer over the homotrimer.

Discussion

We have shown in the studies presented here that both the Smad2/Smad4 and Smad3/Smad4 complexes are heterotrimers composed of two phosphorylated R-Smad subunits and one Smad4 subunit. This is the first direct structural evidence of the heterotrimeric Smad complex by X-ray crystallography. Structural and functional analysis of the Smad heterotrimer suggests that favorable electrostatic interactions within the BC and AB interfaces play a significant role in the preferential formation of heterotrimer over the R-Smad homotrimer. These

electrostatic interactions exist only in the heterotrimer due to specific charged residues in the Smad4 subunit, Asp493 and Arg378, mediating complementary electrostatic interactions with the neighboring R-Smad subunits. Furthermore, the unique conformation of Smad4, where the three helix bundle structure is tilted more toward the neighboring R-Smad subunit, likely facilitates better subunit-subunit contacts in the heterotrimer.

The stoichiometry of the heteromeric Smad complex and the structural basis for its preferential formation revealed here can likely be generalized to the Smad family of proteins. Although only the TGF-β and activin-specific Smads, Smad2 and Smad3, are analyzed here, previous biochemical analysis of a BMP-specific Smad, Smad1, also revealed a 2:1 ratio between the pseudo-phosphorylated Smad1 and Smad4 (Qin et al., 2001). The generality of the mechanism is supported by the highly conserved subunit interface among the R-Smads and the universal role of Smad4 to serve as a common mediator in Smad signaling. Furthermore, the interactions mediated by Asp493 and Arg378 of Smad4, which play important roles in Smad heterotrimerization, are predicted to be conserved in all R-Smad/Smad4 complexes based on sequence analysis.

Although this study suggests a general model of heterotrimeric Smad assembly in TGF-β signaling, the possibility that other forms of heteromeric Smad assembly can exist in vivo in the presence of other signaling partners cannot be ruled out. For example, Inman and Hill have reported that while the DNA-bound Smad2/Smad4 complex is a heterotrimer when bound to the cofactors Fast-1/Foxh1 or Fast-3/Foxh1b, the Smad3/Smad4 complex is a heterodimer (Inman and Hill, 2002). Also, Jayaraman and Massague reported that the endogenous Smad2/Smad4 complex from TGF-β-stimulated cells is heterodimeric (Jayaraman and Massague, 2000). How these heterodimeric Smad species form in vivo is not clear. One possibility is that Smad cofactors, upon binding the R-Smad/Smad4 complex, can disrupt the core heterotrimer in such a way that one of the R-Smad subunits is replaced by a cofactor, resulting in a Smad heterodimer. For example, one recent report has revealed a functional and direct physical interaction between Smad3 and interferon regulatory factor 7 (IRF-7) (Qing et al., 2004). Based on the crystal structure of IRF-3 and sequence comparison, IRF-7 shares the same structural fold as the Smad family of proteins (Qin et al., 2003). Thus, IRF-7 could potentially replace one of the Smad3 subunits to form a Smad3/Smad4/IRF-7 complex. Further work is necessary to investigate the nature of novel heteromeric Smad complexes and the structural basis for their assembly.

Given the apparent simplicity of the Smad system, it is intriguing that TGF-βs mediate such a diversity of biological activity. The formation of a heterotrimeric Smad complex consisting of two R-Smads and one Smad4 may provide part of the answer, as this assembly mechanism greatly increases the potential diversity of the Smad signaling system. From the structural standpoint, it is feasible for the heterotrimeric Smad complex to include two different R-Smad molecules, as the trimer interface among the R-Smads is highly conserved. Such combinatorial mechanism can greatly increase the number of distinct signaling species, each with a different Smad constitution. These distinct Smad complexes may perform different tasks by recognizing and activating specific sets of gene promoters. This combinatorial mechanism of Smad heterotrimerization is consistent with several recent biological observations (Feng et al., 2000; Goumans et al., 2003; Labbe et al., 1998; Nakao et al., 1997; von Gersdorff et al., 2000).

Finally, Smad trimerization can also increase its specificity for promoter recognition. The promoters of TGF-β-responsive genes generally contain Smad boxes, each with a GTCT sequence and recognized by the N-terminal MH1 domain of an R-Smad or Smad 4 (Shi et al., 1998; Zawel et al., 1998). Smad2 is an exception since its MH1 domain contains an insertion that blocks DNA recognition (Dennler et al., 1999; Yagi et al., 1999). Thus, a heterotrimeric Smad complex consisting of two R-Smads and one Smad4 can contain anywhere from one to three functional DNA binding domains, depending on the number of Smad2 subunits present. This combinatorial mechanism of Smad assembly may explain in part why the TGF-β-responsive promoters contain different number of Smad boxes. The MH1 domains within the heterotrimeric Smad complex can also be at different relative distances and orientations to their MH2 domains due to each Smad having a unique linker domain length. This asymmetric property of the heterotrimeric Smad complex may explain in part the different orientation and spacing of Smad boxes observed in the TGF-β-responsive promoters. Future studies on how different heterotrimeric Smad complexes contact DNAs and transcriptional coregulators will further clarify the role of Smad trimerization in TGF-β signaling.

Experimental Procedures

Protein Expression and Purification

Phosphorylated Smad2, S2LC(2P), and phosphorylated Smad3, S3LC(2P), were prepared using the intein-mediated phosphopeptide ligation method as previously described (Qin et al., 2002; Wu et al., 2001b). The Smad4 fragment, S4AF, was purified as previously described (Qin et al., 1999). To form the heteromeric Smad complex, S4AF was added in 2-fold molar excess to S3LC(2P) or S2LC(2P). Each heteromeric Smad complex was first purified by ion-exchange chromatography on a DEAE column, during which the uncomplexed Smad4 eluted in the flow-through fractions, while the heteromeric Smad complex eluted in a 100–300 mM NaCl gradient. The resulting Smad complex was was further purified by a Superdex 200 size exclusion column (Amersham) on FPLC.

Crystallization and Structure Determination

The S2LC(2P)/S4AF or S3LC(2P)/S4AF complex was subjected to limited chymotrypsin digest (1:100 w/w) for 24 hr on ice, and the complex was purified by a Superdex 200 size exclusion column. Crystals were obtained using the hanging drop vapor diffusion technique. The protein solution (15 mg/ml) was mixed with an equal volume of well solution containing 50 mM Tris-HCl (pH 7.5), 0–15 mM $MgCl_2$, and 5%–15% ethanol. Crystals were transferred to a crysolvent comprised of 24% glycerol and 76% well solution, and flash frozen in liquid nitrogen. Diffraction data for the Smad3/Smad4 complex were collected at the BioCARS beamline at the Advanced Photon Source at Argonne National Laboratory. Diffraction data for the Smad2/Smad4 complex were collected at −170°C using an R-Axis IV image plate system mounted on a Rigaku rotating anode generator, at a detector distance of 200 mm with 1 oscillation per frame. Data were integrated and reduced using DENZO and Scalepack (Otwinowski and Minor, 1997). The structures were determined by molecular replacement using the CNS software package (Brunger

et al., 1998), using the phosphorylated Smad2 MH2 domain as the search model (Wu et al., 2001b). The search located three MH2 domains in the asymmetric unit, designated subunits A, B, and C. Examination of the initial $2F_o - F_c$ electron density map revealed features indicating that subunits A and C corresponded to Smad2/3 while subunit B corresponds to Smad4. For example, the electron density of the phosphorylated C-terminal sequence of Smad2/3 was readily visible in subunits A and C but completely absent in subunit B. Also, extra electron density corresponding to the two-residue insertion in Smad4 (Ser344 and Cys345) between the β2 and β3 strands was present in subunit B but not in subunits A and C. The same observations were made when the Smad4 MH2 domain monomer was used as a search model to locate the three subunits. Rigid body refinement of the three MH2 domain subunits using all possible combinations of Smad2/3 and Smad4 also supported the Smad2/3-Smad4-Smad2/3 arrangement, corresponding to subunit A-B-C. The correct combination, Smad3-Smad4-Smad3, gave an initial R_{free} of 37.5%, while the next best combination, Smad3-Smad3-Smad3, gave an initial R_{free} of 40.7%. Initial solutions were further refined through CNS conjugated gradient, simulated annealing, and B factor refinement protocols. Structures were rebuilt with O (Jones et al., 1991). The final heterotrimeric Smad3/Smad4 complex contains subunit A (residues 228–425 of Smad3), subunit B (residues 314–417, 424–455, and 490–546 of Smad4), and subunit C (residues 230–425 of Smad3). The final heterotrimeric Smad2/Smad4 complex contains subunit A (residues 270–467 of Smad2), subunit B (residues 314–461, 490–549 of Smad4), and subunit C (residues 272–467 of Smad2).

Isothermal Titration Calorimetry

Binding of S4AF to S2LC(2P) or S3LC(2P) was measured by isothermal titration calorimetry (Wiseman et al., 1989) using a VP-ITC calorimeter (MicroCal, Northampton, MA). All samples were dialyzed against an ITC buffer (20 mM HEPES, pH 7.0, 100 mM NaCl, 0.1 mM EDTA, 1 mM TCEP) and degassed prior to titration. Experiments were performed at 37°C. 1.4 ml of 100 μM phosphorylated Smad2 or Smad3 was titrated by 300 μl of 400 μM Smad4 over 30 injections. The heat of dilution of S4AF was measured by titrating S4AF into the ITC buffer and was subtracted for data analysis. Data were analyzed with Origin 7.0 software (MicroCal), using a single-site binding model. ΔH, ΔS, and K_a values were experimentally determined, and ΔG was calculated from these values ($\Delta G = -RT \ln K_a$).

Size Exclusion Chromatography

Size exclusion chromatography was performed on a Superdex 200 HR column using the Äkta Explorer 10 FPLC system (Amersham). All runs were performed at room temperature in an FPLC buffer consisting of 20 mM HEPES (pH 7.4), 0.1 mM EDTA, 100 mM NaCl, and 1 mM DTT. Prior to loading onto the column, protein samples were incubated in 1 mM TCEP for 60 min at room temperature. FPLC operation and data analysis were done with UNICORN software. The column was calibrated with blue dextran (to determine void volume) and molecular weight standards ovalbumin (43 kDa), albumin (67 kDa), aldolase (158 kDa), catalase (232 kDa), and ferritin (440 kDa). Elution fractions (0.5 ml) were collected at room temperature with a flow rate of 0.7 ml/min.

Immunoprecipitation and Western Blots

COS-1 cells were transfected with full-length wild-type versions or FLAG-tagged Smad2 or Smad3 with or without Myc-tagged Smad4 or the Smad4 (D493A) point mutant, and with or without the constitutively active TGF-β type I receptor point mutant (T204D) Alk5 CA (Wieser et al., 1995). After 24 hr, cells were switched to 0.1% serum overnight and lysed in triton X lysis buffer (1% triton X, 150 mM NaCl, 10% glycerol, 5 mM EDTA, and 25 mM HEPES, pH 7.5) in the presence of protease and phosphatase inhibitors. Lysates were either directly separated by SDS-PAGE and transferred onto Immobilon-P membranes (Millipore), or first immunoprecipitated for 2 hr using epitope-specific rabbit anti-FLAG antibodies (Sigma). Immunoblots were performed using mouse 9E10 anti-Myc or anti-FLAG M2 antibodies (Sigma), as indicated, and detected using the appropriate horseradish peroxidase-conjugated secondary antibody, and visualized by chemiluminescence (Pierce).

Acknowledgments

We would like to thank James Knapp and William Royer for X-ray data collection of the Smad3/Smad4 complex and Osman Bilsel, C. Robert Matthews, Celia Schiffer, and Shilpa Shenoy for discussions during the preparation of the manuscript. The work is supported by grant from the National Cancer Institute (K.L.).

Received: March 7, 2004
Revised: June 24, 2004
Accepted: June 29, 2004
Published: September 9, 2004

References

Abdollah, S., Macias-Silva, M., Tsukazaki, T., Hayashi, H., Attisano, L., and Wrana, J.L. (1997). TbetaRI phosphorylation of Smad2 on Ser465 and Ser467 is required for Smad2-Smad4 complex formation and signaling. J. Biol. Chem. *272*, 27678–27685.

Attisano, L., and Wrana, J.L. (2000). Smads as transcriptional comodulators. Curr. Opin. Cell Biol. *12*, 235–243.

Blobe, G.C., Schiemann, W.P., and Lodish, H.F. (2000). Role of transforming growth factor beta in human disease. N. Engl. J. Med. *342*, 1350–1358.

Boulanger, M.J., Bankovich, A.J., Kortemme, T., Baker, D., and Garcia, K.C. (2003). Convergent mechanisms for recognition of divergent cytokines by the shared signaling receptor gp130. Mol. Cell *12*, 577–589.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Chacko, B.M., Qin, B., Correia, J.J., Lam, S.S., de Caestecker, M.P., and Lin, K. (2001). The L3 loop and C-terminal phosphorylation jointly define Smad protein trimerization. Nat. Struct. Biol. *8*, 248–253.

Correia, J.J., Chacko, B.M., Lam, S.S., and Lin, K. (2001). Sedimentation studies reveal a direct role of phosphorylation in Smad3:Smad4 homo- and hetero-trimerization. Biochemistry *40*, 1475–1482.

de Caestecker, M.P., Piek, E., and Roberts, A.B. (2000). Role of transforming growth factor-beta signaling in Cancer. J. Natl. Cancer Inst. *92*, 1388–1402.

Dennler, S., Huet, S., and Gauthier, J.M. (1999). A short amino-acid sequence in MH1 domain is responsible for functional differences between Smad2 and Smad3. Oncogene *18*, 1643–1648.

Derynck, R., Zhang, Y., and Feng, X.H. (1998). Smads: transcriptional activators of TGF-beta responses. Cell *95*, 737–740.

Feng, X.H., Lin, X., and Derynck, R. (2000). Smad2, Smad3 and Smad4 cooperate with Sp1 to induce p15(Ink4B) transcription in response to TGF-beta. EMBO J. *19*, 5178–5193.

Goumans, M.J., Valdimarsdottir, G., Itoh, S., Lebrin, F., Larsson, J., Mummery, C., Karlsson, S., and ten Dijke, P. (2003). Activin receptor-like kinase (ALK)1 is an antagonistic mediator of lateral TGFbeta/ALK5 signaling. Mol. Cell *12*, 817–828.

Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., et al. (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science *271*, 350–353.

Heldin, C.H., Miyazono, K., and ten Dijke, P. (1997). TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature *390*, 465–471.

Inman, G.J., and Hill, C.S. (2002). Stoichiometry of active smad-transcription factor complexes on DNA. J. Biol. Chem. *277*, 51008–51016.

Jayaraman, L., and Massague, J. (2000). Distinct oligomeric states of SMAD proteins in the transforming growth factor-beta pathway. J. Biol. Chem. *275*, 40710–40717.

Jones, S., and Thornton, J.M. (1996). Principles of protein-protein interactions. Proc. Natl. Acad. Sci. USA *93*, 13–20.

Jones, A.T., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building proteins models in electron-density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kawabata, M., Inoue, H., Hanyu, A., Imamura, T., and Miyazono, K. (1998). Smad proteins exist as monomers in vivo and undergo homo- and hetero-oligomerization upon activation by serine/threonine kinase receptors. EMBO J. *17*, 4056–4065.

Labbe, E., Silvestri, C., Hoodless, P.A., Wrana, J.L., and Attisano, L. (1998). Smad2 and Smad3 positively and negatively regulate TGF beta-dependent transcription through the forkhead DNA-binding protein FAST2. Mol. Cell *2*, 109–120.

Lagna, G., Hata, A., Hemmati-Brivanlou, A., and Massague, J. (1996). Partnership between DPC4 and SMAD proteins in TGF-β signalling pathways. Nature *383*, 832–836.

Massague, J., and Wotton, D. (2000). Transcriptional control by the TGF-beta/Smad signaling system. EMBO J. *19*, 1745–1754.

Nakao, A., Imamura, T., Souchelnytskyi, S., Kawabata, M., Ishisaki, A., Oeda, E., Tamaki, K., Hanai, J., Heldin, C.H., Miyazono, K., et al. (1997). TGF-beta receptor-mediated signalling through Smad2, Smad3 and Smad4. EMBO J. *16*, 5353–5362.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation model. Methods Enzymol. *276*, 307–326.

Park, Y.C., Burkitt, V., Villa, A.R., Tong, L., and Wu, H. (1999). Structural basis for self-association and receptor recognition of human TRAF2. Nature *398*, 533–538.

Qin, B., Lam, S.S.W., and Lin, K. (1999). Crystal structure of a transcriptionally active smad4 fragment. Structure *7*, 1493–1503.

Qin, B.Y., Chacko, B.M., Lam, S.S., de Caestecker, M.P., Correia, J.J., and Lin, K. (2001). Structural basis of Smad1 activation by receptor kinase phosphorylation. Mol. Cell *8*, 1303–1312.

Qin, B.Y., Lam, S., Correia, J.J., and Lin, K. (2002). Smad3 allostery links TGF-beta receptor kinase activation to transcriptional control. Genes Dev. *16*, 1950–1963.

Qin, B.Y., Liu, C., Lam, S.S., Srinath, H., Delston, R., Correia, J.J., Derynck, R., and Lin, K. (2003). Crystal structure of IRF-3 reveals mechanism of autoinhibition and virus-induced phosphoactivation. Nat. Struct. Biol. *10*, 913–921.

Qing, J., Liu, C., Choy, L., Wu, R.Y., Pagano, J.S., and Derynck, R. (2004). Transforming growth factor beta/Smad3 signaling regulates IRF-7 function and transcriptional activation of the beta interferon promoter. Mol. Cell. Biol. *24*, 1411–1425.

Roberts, A.B. (1999). TGF-beta signaling from receptors to the nucleus. Microbes Infect. *1*, 1265–1273.

Shi, Y., Hata, A., Lo, R.S., Massague, J., and Pavletich, N.P. (1997). A structural basis for mutational inactivation of the tumour suppressor Smad4. Nature *388*, 87–93.

Shi, Y., Wang, Y.F., Jayaraman, L., Yang, H., Massague, J., and Pavletich, N.P. (1998). Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling. Cell *94*, 589–594.

Souchelnytskyi, S., Tamaki, K., Engstrom, U., Wernstedt, C., ten Dijke, P., and Heldin, C.H. (1997). Phosphorylation of Ser465 and Ser467 in the C terminus of smad2 mediates interaction with smad4 and is required for transforming growth factor-beta signaling. J. Biol. Chem. *272*, 28107–28115.

von Gersdorff, G., Susztak, K., Rezvani, F., Bitzer, M., Liang, D., and Bottinger, E.P. (2000). Smad3 and Smad4 mediate transcriptional activation of the human Smad7 promoter by transforming growth factor beta. J. Biol. Chem. *275*, 11320–11326.

Wieser, R., Wrana, J.L., and Massague, J. (1995). GS domain mutations that constitutively activate TβR-1, the downstream signaling component in the TGF-β receptor complex. EMBO J. *14*, 2199–2208.

Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal. Biochem. *179*, 131–137.

Wrana, J.L., Attisano, L., Wieser, R., Ventura, F., and Massague, J. (1994). Mechanism of activation of the TGF-beta receptor. Nature *370*, 341–347.

Wu, J.W., Fairman, R., Penry, J., and Shi, Y. (2001a). Formation of a stable heterodimer between Smad2 and Smad4. J. Biol. Chem. *276*, 20688–20694.

Wu, J.W., Hu, M., Chai, J., Seoane, J., Huse, M., Li, C., Rigotti, D.J., Kyin, S., Muir, T.W., Fairman, R., et al. (2001b). Crystal structure of a phosphorylated Smad2. Recognition of phosphoserine by the MH2 domain and insights on Smad function in TGF-beta signaling. Mol. Cell *8*, 1277–1289.

Yagi, K., Goto, D., Hamamoto, T., Takenoshita, S., Kato, M., and Miyazono, K. (1999). Alternatively spliced variant of Smad2 lacking exon 3. Comparison with wild-type Smad2 and Smad3. J. Biol. Chem. *274*, 703–709.

Ye, H., Park, Y.C., Kreishman, M., Kieff, E., and Wu, H. (1999). The structural basis for the recognition of diverse receptor sequences by TRAF2. Mol. Cell *4*, 321–330.

Zawel, L., Dai, J.L., Buckhaults, P., Zhou, S., Kinzler, K.W., Vogelstein, B., and Kern, S.E. (1998). Human Smad3 and Smad4 are sequence-specific transcription activators. Mol. Cell *1*, 611–617.

Zhang, Y., Feng, X.H., Wu, R.Y., and Derynck, R. (1996). Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. Nature *383*, 168–172.

Accession Numbers

The Smad3/Smad4 and Smad2/Smad4 structures have been deposited in the Protein Data Bank under ID codes 1U7F and 1U7V, respectively.