RECEPTOR-REGULATED SMADS IN TGF-ß SIGNALING

Fang Liu

Center for Advanced Biotechnology and Medicine, Susan Lehman Cullman Laboratory for Cancer Research, Department of Chemical Biology, Ernest Mario School of Pharmacy, Rutgers, The State University of New Jersey, Cancer Institute of New Jersey, 679 Hoes Lane, Piscataway, NJ 08854

TABLE OF CONTENTS
1. Abstract
2. Introduction
3. Structure-function relationship of Smad2 and Smad3
4. Smad2 and Smad3 activation by TGF-ß receptor
   4.1. Role of SARA in Smad2/3 activation
   4.2. Role of other receptor and Smad-interacting proteins in Smad2/3 activation
5. Smad2 and Smad3 nucleocytoplasm shuttling
6. Transcriptional regulation by Smads
   6.1. DNA binding activities and interacting partners
   6.2. Transcriptional activation
   6.3. Transcriptional repression
7. Ubiquitination mediated proteolysis in TGF-ß/Smad pathway
8. Regulation of Smad2 and Smad3 through non-receptor kinases
9. Physiological roles of Smad2 and Smad3 revealed by knockout studies
10. Smads and cancer
    10.1. Smads are candidate tumor suppressors
    10.2. TGF-ß/Smads and cell cycle control
11. Conclusions and perspectives
12. Acknowledgments
13. References

1. ABSTRACT

Transforming growth factor beta (TGF-ß) and related polypeptides, including activins and bone morphogenetic proteins (BMPs), constitute the largest cytokine family, possessing fascinating features. TGF-ß and related peptides are multifunctional growth factors and they regulate many aspects of cellular processes such as proliferation, differentiation, adhesion and apoptosis. These evolutionarily conserved cytokines play an essential role in the development and homeostasis of virtually every tissue in organisms ranging from fruit flies to humans. Accordingly, inactivating mutations in several components of the TGF-ß signaling pathways have been found to cause a number of human disorders. The TGF-ß family members signal through cell surface serine/threonine kinase receptors. A family of proteins, designated as Smads (mammalian homologues of Drosophila Mad and C. elegans Sma), transduces the TGF-ß signal from cell surface to the nucleus. Upon activation, the TGF-ß type I receptor phosphorylates Smad2 and Smad3, which then form complexes with Smad4 and accumulate in the nucleus to regulate transcription of a variety of genes that encode crucial determinants of cell fate, such as cell cycle components, differentiation factors and cell adhesion molecules. Although Smad2 and Smad3 are highly homologous and share some overlapping activities, they have distinct functions and are regulated differentially. This review is primarily focused on our understanding of the similar as well as distinct function and regulation of Smad2 and Smad3 in TGF-ß signaling, their physiological roles revealed by knockout studies and their tumor suppressive functions.

2. INTRODUCTION

The TGF-ß family plays a fundamental role in cell regulation (1-12). TGF-ß binds and brings together two classes of transmembrane receptors, the type I and type II receptors. The TGF-ß type II receptor is constitutively active. It transphosphorylates the type I receptor, which then transduces the signal to downstream components (Figure 1, refs 3, 4, 6, 8, 9-14).

The Smad family plays a pivotal role in mediating the TGF-ß biological responses (3-14). Smads contain conserved N- and C-terminal regions, also designated as the MH1 (Mad Homology 1) and MH2 (Mad Homology 2) domains, respectively, separated by a divergent proline-rich linker region. The Smad family can be structurally and functionally divided into three groups (Figure 2). One group includes those receptor-regulated Smads (R-Smads, also termed pathway-specific Smads) that are phosphorylated by receptor kinases. Smad2 and Smad3 are phosphorylated by the homologous TGF-ß and activin receptor kinases (Figures 1 and 2, refs 15-18);
Figure 1. TGF-β/Smad signaling pathway overview. TGF-β binding results in the formation of a ligand-receptor complex and activation of the type I receptor. The activated type I receptor then phosphorylates Smad2 and Smad3. Smad4 forms complexes with phosphorylated Smad2 and Smad3, and together they accumulate in the nucleus. Smad3 and Smad4 possess DNA binding activities, whereas Smad2 cannot bind to DNA. Smads are usually recruited to responsive promoters through interaction with transcription partners. Smads can recruit coactivators to stimulate transcription or recruit corepressors to inhibit transcription.

Smads and TGF-β Signaling

Smads and TGF-β Signaling

Smad1, Smad5 and Smad8 are phosphorylated by BMP receptor kinases (Figure 2, refs 19-20). The second group includes common Smads (co-Smads), which are not phosphorylated by receptors but are essential for TGF-β/activin and BMP signaling by associating with a receptor activated Smad (21). The only known member of this group in mammalian cells is Smad4. Smad4 is necessary for the Smad complexes to bind to DNA and to regulate transcription (Figure 1, refs 22, 23). The third group includes inhibitory Smads (I-Smads) that antagonize the function of receptor-activated Smads. For example, Smad7 antagonizes TGF-β signaling by binding to the receptor and thus inhibiting its capacity to phosphorylate Smad2 and Smad3 (Figure 1, refs 24, 25). Similarly, Smad6 binds and inhibits BMP receptor phosphorylation of Smad1 (ref 26). In addition, Smad6 antagonizes BMP signaling by interacting with Smad1, thus preventing Smad1 from forming a complex with Smad4 (ref 27). Interestingly, Smad7 and Smad6 are themselves direct target genes of Smad proteins (28-32). Their transcription is upregulated by the treatment with TGF-β and BMP, respectively, thus providing a negative feedback control of TGF-β family signaling (24, 25, 28-33).

Although Smad3 and Smad4 possess DNA binding activities, they are usually recruited to promoters through interaction with DNA binding cofactors (Figure 1, refs 34, 35). Smad proteins can activate transcription by recruiting transcription coactivators, such as p300/CBP, or repress transcription by recruiting transcriptional corepressors that include TGIF (TG-interacting factor) and the related Ski and SnoN proteins (Figure 1, refs. 34-36). In addition, Smads can also repress transcription by other distinct mechanisms (34).

The TGF-β/Smad signaling pathway is subjected to regulation at multiple steps, such as regulation of receptor activity and activation of Smad2 and Smad3 by TGF-β receptor. These processes are regulated by a variety of TGF-β receptor and/or Smad interacting proteins and by Smad2 and Smad3 homo-oligomerization and hetero-oligomerization with Smad4. Other controls are mediated by subcellular localization, DNA binding properties, interacting partners for transcriptional regulation, phosphorylation by non-receptor kinases, and ubiquitin-mediated degradation. The emphasis of this review is a comparison of similar as well as distinct properties of
Smads and TGF-ß Signaling

**Receptor-regulated Smads:**

- **Smad2, Smad3**
  - (by TGF-ß & activin receptors)
- **Smad1, Smad5, Smad8**
  - (by BMP receptors)

**Co-Smad:**

- **Smad4**
  - (for TGF-ß/activin & BMP signaling)

**Inhibitory Smads:**

- **Smad7**
  - (Anti-TGF-ß/activin signaling)
- **Smad6**
  - (Anti-BMP signaling)

**Figure 2.** The Smad family can be divided into three subgroups. One group includes receptor regulated Smads. Smad2 and Smad3 are phosphorylated by homologous TGF-ß and activin receptors, whereas Smad1, Smad5 and Smad8 are phosphorylated by BMP receptors. Smad4 is a common Smad (Co-Smad) that participates in TGF-ß, activin and BMP signaling by associating with a receptor-phosphorylated Smad. In mammalian cells, Smad4 is the only Co-Smad. The third group includes those inhibitory Smads that antagonize TGF-ß family signaling. Smad7 mostly inhibits TGF-ß/activin signaling. It can also inhibit BMP signaling. In contrast, Smad6 specifically inhibits BMP signaling. The structures of these three types of Smads are shown schematically. The N-terminal and C-terminal domains are also termed as MH1 (Mad Homology 1) and MH2 (Mad Homology 2) domains, respectively, in reference to the *Drosophila* Mad, the founding member of the Smad family. The MH1 and MH2 domains are conserved between receptor regulated Smads and Co-Smads. The MH2 domain is also conserved in the inhibitory Smads, but the MH1 domain is not conserved in the inhibitory Smads.

Smad2 and Smad3 in TGF-ß signaling, which specify their unique physiological roles.

**3. STRUCTURE-FUNCTION RELATIONSHIP OF SMAD2 AND SMAD3**

Smad2 and Smad3 are highly homologous, sharing over 90% homology at the amino acid level (Figure 3). They are relatively divergent in the proline-rich linker region (Figure 3). Most notably, Smad2 contains two stretches of amino acids that are not present in Smad3. The first stretch is 10 amino acids and the second stretch is 30 amino acids, which is encoded by a separate exon (exon 3) (ref 37). The presence of these 30 amino acids in the second stretch interferes with Smad2 binding to DNA (38-41) and also appears to contribute to Smad2 nuclear import by a mechanism different from that used by Smad3 (see section 5). The amino acid sequences that specify other features, such as the nuclear localization signal in Smad3, the DNA binding hairpin in Smad3, the L3 loop that specifies interaction with receptor, the PPXY motif in the linker region that is recognized by the WW domain of Smurf2 (Smad ubiquitin regulatory factor 2), the phosphorylation sites by protein kinase C (PKC) and calcium-calmodulin-dependent-kinase II (CamKII), are also outlined in Figure 3 and are described in detail in the corresponding sections below. A general structure-function relationship of Smad2 and Smad3 is shown in Figure 4 and is also discussed in subsequent sections.

**4. ACTIVATION OF SMAD2 AND SMAD3 BY TGF-ß RECEPTOR**

**4.1. Role of SARA in Smad2/3 activation**

Both Smad2 and Smad3 are phosphorylated by TGF-ß receptor in the C-terminal SSXS motif (17, 18). Smad2/3-receptor interactions are mediated by the L3 loop in the MH2 domain of Smad2/3 (Figure 3) and L45 loop in the TGF-ß type I receptor (42-45). The conserved L3 loop differs by only two amino acids in the TGF-ß receptor regulated Smad2, 3 versus the BMP receptor regulated Smad1,5,8 (Figure 3, ref 44). The difference in these two amino acids is sufficient for discrimination by the TGF-ß receptor versus the BMP receptor (44). The crystal structure of the MH2 domain of Smad2 is similar to that of Smad4, but the MH2 domain of Smad2 contains an extended basic pocket near the L3 loop (46-48). This basic pocket in Smad2 has been proposed to serve as a docking site for the phosphorylated GS domain of the activated TGF-ß type I receptor (47).

SARA (Smad anchor for receptor activation) functions to recruit Smad2 and Smad3 to the TGF-ß receptor (49). SARA contains a FYVE domain for membrane localization, a Smad binding domain (SBD) for binding to Smad2/3, and a carboxyl-terminal domain for interacting with the receptor kinase (49). At basal state, monomeric forms of Smad2 and Smad3 are bound by SARA (47, 50, 51). Although Smad3, and presumably Smad2 as well, have a tendency to trimerize and undergo concentration-dependent trimerization, SARA stabilizes the monomeric forms of Smad2 and Smad3 through the proline-rich structure of SBD, thus inhibiting Smad2 and Smad3 trimerization (50, 51). The structural basis for these effects is that an unphosphorylated, monomeric Smad3 or Smad2 exhibits better surface complementarity and fitness to the SBD of SARA than the oligomeric Smad3 or Smad2 MH2 domains (50, 51). Upon TGF-ß binding and activation of receptor kinase, SARA presents Smad2 and Smad3 for kinase recognition, and precisely positions the phosphorylation sites of Smad2 and Smad3 in the kinase catalytic center (41, 47, 49-51). Phosphorylation of Smad2/3 at the C-terminal SSXS motif by TGF-ß receptor increases the tendency for trimerization (48, 50, 51). This allows Smad2/3 to change from monomeric forms to oligomeric forms, dissociating from SARA and the receptor, and associating with Smad4 (Figure 5, refs 48, 50, 51).

Accordingly, the crystal structure of phosphorylated Smad2 reveals that it is a trimer and that the phosphoserine is recognized by the MH2 domain of Smad2 as well as Smad4 (48). Similar conclusions were also made for Smad1 (52). The crystal structure of pseudophosphorylated Smad3 indicates that it has an increased propensity to homotrimerize and recruits Smad4 to form a
Figure 3. Amino acid sequence comparison between Smad2 and Smad3. The two C-tail serine residues that are phosphorylated by TGF-β receptor are marked with red dots. The L3 loop in Smad2 or Smad3 interacts with TGF-β receptor. The two residues marked with green dots within the L3 loop are different in BMP receptor-regulated Smads. These two residues specify that the L3 loop of Smad2 and Smad3 interact with TGF-β/activin receptors but not BMP receptors. The nuclear localization signal (NLS) and the DNA binding hairpin in Smad3 are indicated. Protein kinase C (PKC) phosphorylation sites in Smad3 and Smad2, and calcium-calmodulin-dependent kinase II (CamKII) phosphorylation sites in Smad2 are indicated by blue dots. The PPXY motif recognized by Smurf (Smad ubiquitin regulatory factor) in Smad2 and Smad3 is indicated. The proline-rich linker regions in the middle of Smad2 and Smad3 are marked with big brackets.
heterotrimer (50, 53). Thus, receptor phosphorylation of pathway specific Smads, such as Smad2 and Smad3, drives their homo-trimerization as well as hetero-oligomerization with the common mediator Smad4 (48, 50, 51, 53, 54). This is also supported by tissue culture studies (55). The Smad3-Smad4 complex is suggested to exist as a heterotrimer containing two Smad3 and one Smad4 molecule (Figure 5, refs 50, 51, 53). The trimeric interaction is mediated through conserved interfaces where tumorigenic mutations map (53). In contrast, Smad2-Smad4 is suggested to exist as a heterodimer (Figure 5, refs 51, 56). The differential mode of oligomerization of Smad2 and Smad3 is also supported by biochemical data obtained using gel chromatography (57, 58). This represents the first point of divergence between the two highly similar receptor regulated Smads of the TGF-ß pathway.

The physiological form of SARA is expected to be dimeric, based on the observation that the membrane-anchoring FYVE domain in early endosome autoantigen EEA1, has a dimeric structure (59). The transmembrane receptor kinase is also dimeric. Thus, it is highly likely that two R-Smad subunits are simultaneously tethered to each receptor complex through dimeric SARA. It will be very interesting to determine whether distinct signaling complexes are formed with Smad2 and Smad3 separately or whether a single complex can form with one Smad2 and one Smad3.

In any case, SARA binding to Smad2 and Smad3 can eliminate aberrant Smad2/3 oligomerization and activation in the absence of TGF-ß. Forced Smad2/3 oligomerization may explain, at least in part, why overexpression of Smad2, and especially Smad3, in mammalian cells often leads to constitutive activation of the TGF-ß pathway. High levels of Smad3 and Smad2 can easily saturate the endogenous levels of SARA, leading to aberrant trimer formation and activation of the signaling pathway.

### 4.2. Role of other receptor and Smad-interacting proteins in Smad2/3 activation

In addition to SARA, several other proteins with anchoring, scaffolding and/or chaperone activity have also been shown to regulate the recruitment of Smads to the TGF-ß receptor complex. Hgs (Hrs) is another FYVE domain protein. There is little homology between Hgs and SARA except for their FYVE domains, which have a 42% identity. Hgs (Hrs) bind Smad2 and Smad3 and cooperate with SARA to stimulate activin and TGF-ß signaling (60).

Disabled-2 (Dab-2) has been identified through a genetic complementation screen as a critical link between TGF-ß receptor and Smad proteins (61). Dab2 contains an N-terminal phosphotyrosine binding domain and a C-terminal proline-rich domain (62), indicating a function as an adaptor molecule (63). Dab2 constitutively interacts with both the type I and type II TGF-ß receptors in vivo, suggesting that Dab2 is part of a multiprotein signaling complex. TGF-ß treatment induces a transient increase in association of Dab2 with the MH2 domains of Smad2 and Smad3. Moreover, expression of Dab2 in a TGF-ß signaling mutant cell restores TGF-ß-mediated Smad2 phosphorylation, Smad translocation to the nucleus and Smad-dependent transcriptional responses. Thus, Dab-2 facilitates the transmission of TGF-ß signal from the receptors to Smads (61).

Axin, a negative regulator in Wnt signaling, may also function as an adapter for Smad3 (64). In the absence of TGF-ß signaling, Axin, and its homologue Axil (also called conductin), interact with Smad3 and colocalize with Smad3 in the cytoplasm. Axin also interacts with Smad2. Upon receptor activation, Smad3 is strongly phosphorylated by the TGF-ß type I receptor in the presence of Axin and dissociates from Axin. In addition, TGF-ß signaling is enhanced by Axin and repressed by an Axin mutant, which is unable to bind to Smad3 (64). Thus, Axin facilitates Smad3 activation by TGF-ß receptors.
Smads and TGF-β Signaling

**Figure 5.** Smad2/3 activation and nuclear import. At basal state, SARA preferentially binds monomeric forms of Smad2 and Smad3. Upon TGF-β binding and activation of receptor kinase, SARA presents Smad2/3 for phosphorylation by the TGF-β receptor. Phosphorylation increases the tendency of Smad2/3 to form homotrimers, dissociation from SARA and the receptor, and association with Smad4. Smad3-Smad4 is suggested to exist as a heterotrimer, whereas Smad2-Smad4 is suggested to exist as a heterodimer. Smad2, and presumably the Smad2-Smad4 complex, are imported into the nucleus by direct interaction of the Smad2 MH2 domain with nucleoporin protein CAN/Nup214, Smad3, and presumably the Smad3-Smad4 complex, are imported into the nucleus by Smad3 MH1 domain direct binding to importin-β.

Caveolin-1, the marker protein for Caveolae, which contains cholesterol-rich membrane microdomains, interacts with the TGF-β type I receptor in a physiologically relevant time frame. Caveolin-1 also cofractionates with TGF-β receptor and Smad2 but not Smad4. Moreover, Caveolin-1 is able to inhibit TGF-β-mediated phosphorylation of Smad2 and subsequent downstream events (65).

Many PDZ (postsynaptic density 95/Discs large/Zona occlusens-1) domain-containing proteins have a role in assembling receptors and signaling molecules in the cell membrane and in the submembranous region (66). ARIPs (activin receptor interacting proteins), identified by the yeast twohybrid screening, contain PDZ domains and associate specifically with the activin type II receptors (67). ARIP1, which contains two WW domains in addition to five PDZ domains, also interacts with Smad3 and this interaction is dependent on the WW domain. ARIP1 is highly expressed only in neuronal cells. It regulates activin-induced and Smad3-mediated signaling in response to activin (68). GIPC (GAIP-interacting protein, C terminus), another PDZ domain-containing protein, can interact with a Class I PDZ binding motif in the cytoplasmic domain of the TGF-β type III receptor, resulting in regulating expression of the type III receptor at the cell surface. Increased expression of the type III receptor mediated by GIPC enhances TGF-β and Smad mediated signaling (69).

STRAP (serine-threonine kinase receptor-associated protein), a WD domain-containing protein, interacts with both TGF-β type I and type II receptors in vivo (70). Overexpression of STRAP inhibits TGF-β transcriptional activation (70). Moreover, STRAP synergizes specifically with Smad7 to inhibit TGF-β-mediated transcriptional responses (71). This synergys is dependent on an intact C-terminal domain of Smad7 that is necessary for the receptor binding. STRAP stably interacts with Smad7 to activate type I receptor to form a complex, and stabilizes the interaction of Smad7 with the activated receptor. Thus, STRAP assists Smad7 in inhibiting Smad2 and Smad3 binding to the receptor (71). Interestingly, STRAP also interacts with Smad2 and Smad3 but does not cooperate with them to transduce the TGF-β signal (71).

Microtubules tether inactive Smads in the cytoplasm (72). Smad2, 3 and 4 bind microtubules in the absence of TGF-β. Treatment with TGF-β triggers Smad2/3 dissociation from microtubules. Presumably, Smad2/3 is then presented to SARA, which in turn presents Smad2/3 for phosphorylation by TGF-β receptor. Pharmacological agents, such as nocadazole, which destabilize the microtubule network, increase Smad2 phosphorylation and transcriptional responses (72).

Filamin, a cytoskeletal actin-binding protein, can interact with several Smad proteins including Smad1-6 (73). TGF-β signaling is defective in filamin-deficient cells. For example, TGF-β receptor-induced serine phosphorylation of Smad2 is impaired in these mutant cells. Re-introduction of filamin into these mutant cells can restore TGF-β responsiveness. These results suggest that filamin plays an important role in Smad-mediated signaling (73).

In addition to these anchoring, scaffolding and chaperone proteins described above for Smad2 and Smad3, a chaperone protein for Smad4 has also been described. TRAP-1 (TGF-β receptor associated protein-1) interacts with inactive TGF-β receptor in the absence of ligand. Upon receptor activation, TRAP-1 dissociates from the receptor complex and associates with Smad4. The
interaction between TRAP-1 and Smad4 is transient and disrupted by activated Smad2 (74). These observations suggest that TRAP1 brings Smad4 into proximity with the receptor complex and assists in the formation of heteromeric complexes between Smad2/3 and Smad4 (74).

5. SMAD2 AND SMAD3 NUCLEOCYTOPLASM SHUTTLING

Nuclear import and export proceed through nuclear pore complexes and occur in a large number of distinct pathways, many of which are mediated by importin-ß-related nuclear transport receptors (75, 76). These receptors shuttle between nucleus and cytoplasm, and they often bind transport substrates via adapter proteins (75, 76). The most studied adapter is importin-ß, which recognizes a lysine-rich and arginine-rich nuclear localization signal (NLS) (75, 76). The typical NLS sequence in SV40 T antigen is KKKRK (76). Alternatively, importin-ß or other members of the importin-ß family bind directly to the transport substrates via a NLS-like motif or via uncharacterized domains in the substrates.

Smad3 contains a NLS at the MH1 domain, 40-KKLKK, which resembles the classic NLS of the SV40 T antigen (Figure 3, ref 77). At basal state, the NLS in Smad3 is largely masked due to the inhibitory intramolecular interaction between the MH1 and MH2 domains (78). TGF-ß receptor phosphorylation of Smad3 induces a conformational change, leading to exposure of the NLS. Smad3, via its MH1 domain, binds directly and specifically to importin-ß (79). This interaction is increased in the presence of TGF-ß and is diminished or abolished by mutations in the NLS (79, 80). In contrast, no interaction between importin-ß and the full length or the MH1 domain of Smad3 can be detected (79). Thus, nuclear import of Smad3 occurs through direct binding to importin-ß (79-81). Ran-GTPase in the nucleus then releases Smad3 from importin-ß in an energy-dependent manner (80, 81). Mutation of the NLS motif has little or no effect on TGF-ß-induced phosphorylation of Smad3, heteromeric complex formation with Smad4 or binding to DNA but abolishes TGF-ß-induced nuclear accumulation (77). In addition, the NLS in Smad3 is responsible for the isolated MH1 to be constitutively localized in the nucleus (77). This classic NLS-like motif in Smad3 is also conserved in Smad2 and all BMP pathway Smads, including Smad1, Smad5, and Smad8, and has been shown indeed to act as a NLS in Smad1 (82).

The MH1 domain of Smad2 cannot bind to importin-ß due to the presence of the extra 30 amino acids encoded by exon 3 (80), even though it contains the exact NLS-like motif KKLKK as in Smad3 (Figure 3). Instead, Smad2 is imported into the nucleus without the participation of importins. Smad2 is anchored in the cytoplasm by SARA (49, 83). TGF-ß-induced phosphorylation enables Smad2 to form a hetero-oligomer, dissociate from SARA and be imported into the nucleus through direct interaction of the MH2 domain with nucleoporin proteins CAN/Nup214 and Nup153 (83, 84). CAN/Nup214 resides on the cytoplasmic side of the nuclear pore complex and participates in nuclear import and export (76). Nup153 is a nucleoporin located at the nuclear side of the nuclear pore complex (76). CAN/Nup214 and Nup153 interaction with Smad2 is constitutive, independent of TGF-ß signaling (84). They compete with SARA in the cytoplasm and DNA binding cofactor(s), such as FAST-1, in the nucleus for binding to Smad2. That the MH2 domain directs Smad2 nuclear accumulation is also supported by the observations that the MH2 domain of Smad2, in a fusion with LacZ, is constitutively present in the nucleus (85), and that Smad2 can be imported into the nucleus without addition of importin in vitro (83, 84). Interestingly, Smad2 does not need intrinsic phosphorylation by receptor for nuclear accumulation. Overexpression of a Smad2 C-tail phosphorylation mutant, which leads to the saturation of SARA, can also result in Smad2 accumulation in the nucleus (83).

At basal state, more Smad3 is present in the nucleus than Smad2. The mechanism underlying this difference is not clear at present. One possibility is that Smad3 possesses DNA binding activity, which may prevent it from being exported to the cytoplasm. Nevertheless, Smad3, and especially Smad2, are nuclear-cytoplasm shuttling proteins. They not only have nuclear import mechanisms but also export systems. How Smad2 and Smad3 are exported is currently unknown. Unlike Smad1 and Smad4, which undergo CRM1-dependent export (82, 86), Smad2 and Smad3 are not exported by a CRM1-dependent mechanism, as overexpression of CRM1 or treatment with leptomycin B, an inhibitor of CRM1-mediated export in mammalian cells and in the fission yeast S. pombe, does not change the subcellular localization of Smad2 and Smad3 (82, 87).

Smad2, and probably to a lesser extent Smad3, are constantly shuttling between the nucleus and the cytoplasm during active signaling, which provides a mechanism for sensing receptor activity (87). When the receptor is active, Smad2 and Smad3 are imported to the nucleus. Part of the activated Smad2 and Smad3 is thought to go through ubiquitin-mediated degradation (see section 7). The majority of the activated Smad2 and Smad3 is thought to be continuously dephosphorylated by a yet to be identified phosphatase, dissociated from Smad4 and exported back to the cytoplasm. If the receptor is still active, Smad2 and Smad3 will be rephosphorylated, form complexes with Smad4 and return to the nucleus. If the receptor is inactive after the TGF-ß signal is shut off, Smad2 and Smad3 will remain in the cytoplasm and interact with SARA again. Thus, for the duration of active signaling, Smad2 and Smad3 constantly monitor the activity of the receptors, which provides a mechanism whereby the levels of active Smad2 and Smad3 in the nucleus directly reflect the levels of active receptors in the cytoplasm (87).

6. TRANSCRIPTIONAL REGULATION BY SMADS

6.1. DNA binding activities and interacting partners

Smad3 and Smad4 possess intrinsic DNA binding activities through the MH1 domains (Figures 3 and 4, and
Smads and TGF-β Signaling

ref 34). In vitro, full length recombinant Smad4 binds DNA, yet full length Smad3 has only weak DNA binding activity even when high doses of recombinant protein are used. The MH1 domain of recombinant Smad3 binds DNA much more efficiently than the intact protein. Using recombinant Smad3 MH1 domain and Smad4 with random oligonucleotides in a PCR-based approach, an 8 base pair palindromic sequence GTCTAGAC was identified as the SMAD binding element (SBE) (88). The crystal structure of the Smad3 MH1 domain binding to the 8 base pair palindromic SBE has been solved. An 11 amino acid ß-hairpin (Figure 3) that is conserved among receptor-regulated Smads and Smad4 is embedded in the major groove of DNA (38, 41). Two molecules of Smad3 MH1 bind to the 8 base pair SBE, with each molecule contacting a single half site, 5′-GTCT-3′ (also called the Smad box) (38, 41). In the GTCT sequence, the G at position 1, the G at position 3 on the complementary strand form hydrogen bonds with amino acids of the ß-hairpin (38). Thus, it is predicted from the crystal structure that the second base in the GTCT motif can tolerate substitutions (38). Indeed, during the selection of the consensus binding site for Smads, it was found that substitution of the second base only modestly reduced DNA binding by Smad3 and Smad4 (88). This is confirmed by observations indicating that Smad3-Smad4 indeed binds a natural promoter with substitution of the second base in the GTCT sequence, such as in the TGF-β inhibitory element (TIE) of the c-myc promoter (see subsection 10.2).

Although Smad2 is highly homologous to Smad3, it cannot bind to DNA due to interference by the exon 3-encoded 30 amino acids present immediately before the DNA binding hairpin (Figure 3, and refs 38-41). If exon 3 is removed, as found in an alternatively spliced Smad2, it can then bind to the SBE sequence (39). The transcript of this alternatively spliced variant of Smad2 is present in certain cells and tissues at a level of about 1/10 of that containing exon 3. In certain cell types that lack or have low levels of expression of Smad3, the variant Smad2 without exon 3 may function as a Smad3-like molecule (39).

A number of TGF-β/Smad responsive promoters, such as the PAI-1, collagenase I, c-Jun, IgA, and Jun B promoters, contain one or multiple copies of the sequence GTCT or AGAC, which can be bound by the Smad3/Smad4 complex (89-101). The GTCT or AGAC sequences have been shown to be critical for TGF-β inducibility of a number of responsive genes. Tandem repeats of GTCT, AGAC, or the 8 bp SBE can confer TGF-β inducibility to heterologous promoters (88, 90, 101, 102). The BMP responsive element of certain BMP regulated genes, such as the Xvent-2 gene (88), also contains the AGAC motif (103). In addition to the GTCT and AGAC elements, Smad3, Smad4, and BMP receptor regulated Smads can also recognize a GC-rich sequence (32, 104-106). For example, the MH1 domains of Smad3 and Smad4 have been shown to bind to a GC-rich sequence of the goosecoid promoter (105), and the MH1 domain of the Drosophila Mad is implicated in binding to a GC-rich sequence of the DPP-responsive vestigial enhancer (104). A GC-rich sequence has also been shown to be essential for BMP-induced activation of the Smad7 promoter (32). Thus, Smad proteins have flexible DNA binding properties.

The palindromic SBE represents an optimal binding site for Smad, which is the basis of its being selected from a pool of random oligonucleotides. A study examined Smad3/Smad4 complex binding to two or three copies of abutting sequences GTCT and AGAC in different combinations (102). Interestingly, Smad3/Smad4 has little or no capacity to bind to two or three copies of the GTCT sequence, or the AGAC sequence followed by one or two copies of the GTCT sequence. These observations further indicate that the SBE is a high affinity binding site for Smad. Smad7 promoter is the only natural TGF-β responsive promoter in vertebrates that has been shown to contain the 8 bp palindromic SBE (28-31). Upon TGF-β induction, endogenous Smad complex binds a Smad7 promoter DNA as short as 14 or 16 base pairs containing the 8 bp palindromic SBE with only 3 or 4 base pairs adjacent sequences on each side, suggesting that Smad complex can bind to the 8 bp palindromic SBE on its own in vivo (30).

However, the vast majority of TGF-β/Smad responsive promoters do not contain high affinity-binding sites for Smad proteins. In fact, Smad proteins are frequently recruited to TGF-β/activin-responsive promoters through interaction with DNA binding cofactors (34). The classic example is from studies of the activin responsive gene Mix.2 (107). The Mix.2 promoter contains a 51 base pair activin responsive element (ARE), which is upregulated by activin and also by TGF-β in the same manner (22, 107). The ARE contains a binding site for FAST-1, a winged helix transcription factor, and a GTCT sequence for binding to Smad proteins (107-109). FAST-1 can bind directly to Smad2, and the interaction is significantly increased in the presence of TGF-β or activin. Moreover, FAST-1, Smad2 and Smad4 form a stable complex, which binds the ARE and activates transcription of the Mix.2 gene (22, 107-109). In addition to FAST-1, a large number of transcription factors (Table 1) have been found to interact with Smad2, Smad3 and/or Smad4 to regulate transcription of diverse genes. Studies using mouse embryonic fibroblasts that are deficient in Smad2 or Smad3 indicate that Smad2 and Smad3 have distinct roles in the regulation of different target genes (143). For example, TGF-β activation of the Smad7 promoter is selectively dependent on Smad3 (143).

6.2. Transcriptional activation

Smad proteins are able to activate transcription. This was first revealed through GAL4 fusion studies in which the C terminal (MH2) domain of receptor regulated Smad or Smad4 along with a small segment of the linker region can activate transcription when fused to the GAL4 DNA binding domain (144). Full length Smads, such as Smad1 or Smad2, have very little activity in the GAL4 fusion assay, but their transcription activities are greatly increased by treatment with the corresponding agonists, BMP or TGF-β (22, 144). Subsequent studies have shown that transcriptional activation by Smad3 and Smad2 occurs,
Table 1. Smad-Interacting Transcription Factors

<table>
<thead>
<tr>
<th>Interacting Proteins</th>
<th>Properties and Functions</th>
<th>Smad</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coactivator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSG-1</td>
<td>Nuclear transcription coactivator</td>
<td>Smad4 (MH2, linker)</td>
<td>154, 155</td>
</tr>
<tr>
<td>p300/CBP</td>
<td>Histone acetyltransferase</td>
<td>Smad1-4 (MH2)</td>
<td>145-150</td>
</tr>
<tr>
<td>P/CAF</td>
<td>Histone acetyltransferase</td>
<td>Smad3 (MH2)</td>
<td>151</td>
</tr>
<tr>
<td>SMIF</td>
<td>EVH1/WH1 protein; Smad4-interacting coactivator</td>
<td>Smad4 (SAD)</td>
<td>156, 157</td>
</tr>
<tr>
<td>Swift</td>
<td>BRCT domain factor; coactivator for Smad2</td>
<td>Smad1,2</td>
<td>125</td>
</tr>
<tr>
<td>For Activation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATF2</td>
<td>ATF/CREB family member; cooperate with Smad3</td>
<td>Smad3,4 (MH1)</td>
<td>110, 111</td>
</tr>
<tr>
<td>c-Jun, JunB, JunD</td>
<td>AP-1 family member; cooperate with Smads to activate c-jun and collagenase promoters</td>
<td>Smad3,4 (MH1, linker)</td>
<td>95-97, 112</td>
</tr>
<tr>
<td>FAST-1</td>
<td>Winged-helix factor; cooperate with Smad2,3</td>
<td>Smad2,3 (MH2)</td>
<td>22, 107-109</td>
</tr>
<tr>
<td>FAST-2</td>
<td>Winged-helix factor; cooperate with Smad2</td>
<td>Smad2 (MH2)</td>
<td>105</td>
</tr>
<tr>
<td>HNF4</td>
<td>Hepatocyte nuclear factor 4; cooperate with Smad3/4</td>
<td>Smad3,4</td>
<td>113</td>
</tr>
<tr>
<td>Lef1/Tcf</td>
<td>HMG box factor; cooperate with Smad2/3/4</td>
<td>Smad2-4 (MH1, MH2)</td>
<td>114, 115</td>
</tr>
<tr>
<td>MEF2</td>
<td>Myocyte enhancer binding factor 2</td>
<td>Smad2</td>
<td>116</td>
</tr>
<tr>
<td>Menin</td>
<td>Nuclear protein; necessary for TGF-beta signaling</td>
<td>Smad2,3 (MH2)</td>
<td>117</td>
</tr>
<tr>
<td>Milk, Mixer</td>
<td>Paired-like homeodomain factor; bind to the distal region of gooscoiid promoter; cooperate with Smad2 to activate p15 and p21 promoters</td>
<td>Smad2 (MH2)</td>
<td>118, 119</td>
</tr>
<tr>
<td>Miz-1</td>
<td>Zinc finger protein; cooperate with Smads and SP1 to activate p15 and p21 promoters</td>
<td>Smad3,4 (MH1)</td>
<td>260-262</td>
</tr>
<tr>
<td>p52 (NFkB)</td>
<td>NFkB/Rel family factor; cooperate with Smad3 to activate JunB promoter</td>
<td>Smad3</td>
<td>120</td>
</tr>
<tr>
<td>Runx/PEBP2/CBFA/AML</td>
<td>Runt-domain protein; cooperate with Smads to activate germline Ig C-alpha promoter</td>
<td>Smad1-4 (MH2)</td>
<td>98-100</td>
</tr>
<tr>
<td>SP1, SP3</td>
<td>Cooperate with Smads and Miz-1 to activate p15 and p21 promoters</td>
<td>Smad2-4 (MH1)</td>
<td>121-124</td>
</tr>
<tr>
<td>TFE3 (µE3)</td>
<td>Helix-loop-helix leucine zipper factor; cooperate with Smads to activate PAI-I promoter</td>
<td>Smad3,4 (MH1)</td>
<td>93, 94</td>
</tr>
<tr>
<td>VDR</td>
<td>Vitamin D receptor; cooperate with Smad3</td>
<td>Smad3 (MH1)</td>
<td>126, 127</td>
</tr>
<tr>
<td>Corepressor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ski</td>
<td>Nuclear proto-oncoprotein; recruit HDAC; bind Smad4 L3 loop to inhibit Smad4 binding to Smad2/3</td>
<td>Smad2-4 (MH2)</td>
<td>36, 164-172</td>
</tr>
<tr>
<td>SnoN</td>
<td>Nuclear proto-oncoprotein and tumor suppressor; recruit HDAC</td>
<td>Smad2-4 (MH2)</td>
<td>36, 169,170</td>
</tr>
<tr>
<td>TGIF</td>
<td>Homeo-domain factor; recruit HDAC, mSin3A, CtBP</td>
<td>Smad2,3 (MH2)</td>
<td>34, 158-161</td>
</tr>
<tr>
<td>For Repression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Androgen receptor</td>
<td>AR inhibits TGF-beta signaling; whether Smad3 activates or inhibits AR activity has conflict results</td>
<td>Smad3 (MH2)</td>
<td>128-131</td>
</tr>
<tr>
<td>BF-1</td>
<td>Winged-helix brain factor-1; protooncogene; bind to DNA binding cofactor, such as FAST-2, thereby inhibiting Smads; also inhibit Smad3 binding to DNA</td>
<td>Smad1-4 (MH2)</td>
<td>132, 133</td>
</tr>
<tr>
<td>CBFA1</td>
<td>Smad3 inhibits osteoblast differentiation through interaction with CBFA1</td>
<td>Smad1-4 (MH1, MH2)</td>
<td>173</td>
</tr>
<tr>
<td>Gli3ΔC</td>
<td>C-terminally truncated Gli3 zinc finger factor; repressor form</td>
<td>Smad1-4</td>
<td>140</td>
</tr>
<tr>
<td>E1a</td>
<td>Adenoviral oncoprotein; compete with p300 to bind to Smad3</td>
<td>Smad1-3 (MH2)</td>
<td>134</td>
</tr>
<tr>
<td>E2F4/5</td>
<td>Interact with Smad3 to downregulate c-myc</td>
<td>Smad3 (MH2)</td>
<td>258, 259</td>
</tr>
<tr>
<td>Estrogen receptor-alpha</td>
<td>ER inhibits Smad3 activity; TGF-beta enhances ER transcription activity</td>
<td>Smad2-4 (MH2)</td>
<td>135</td>
</tr>
<tr>
<td>Evi-1</td>
<td>Zinc finger factor; inhibit Smad3 DNA binding; recruit corepressor CtBP</td>
<td>Smad3 (MH2)</td>
<td>136-138</td>
</tr>
<tr>
<td>Glucocorticoid receptor</td>
<td>Inhibit Smad3 transcriptional activation</td>
<td>Smad3 (MH2)</td>
<td>139</td>
</tr>
<tr>
<td>Myc</td>
<td>Interact with Smad2,3 to inhibit p15 induction</td>
<td>Smad2,3 (MH2)</td>
<td>263</td>
</tr>
<tr>
<td>MyoD</td>
<td>Smad3 inhibits myogenesis by interacting with MyoD</td>
<td>Smad3 (MH2)</td>
<td>174</td>
</tr>
<tr>
<td>SIP1</td>
<td>Zinc finger/homeodomain repressor; repress Xenopus brachyury promoter</td>
<td>Smad1,2,3,5</td>
<td>141</td>
</tr>
<tr>
<td>SNIP1</td>
<td>Forkhead-associated nuclear protein; compete with p300 to bind Smad4</td>
<td>Smad4 (MH1)</td>
<td>142</td>
</tr>
</tbody>
</table>

1288
at least in part, by their ability to recruit general transcriptional coactivator p300/CBP (145-150). This interaction occurs through the MH2 domains of Smad3 or Smad2 and the C-terminal domain of p300/CBP. P300/CBP have intrinsic histone acetyltransferase activity (HAT), which facilitates transcription by altering nucleosome structure through histone acetylation and thereby remodeling the chromatin template (34). In addition, P/CAF, another HAT-containing transcriptional co-activator, has been shown to associate with Smad3 upon TGF-ß receptor activation and to enhance TGF-ß/Smad3 signaling (151).

The transcriptional activities of Smad2, Smad3 and BMP pathway Smad proteins, such as Smad1, are essentially dependent on Smad4. This has been shown by studies using Smad4 deficient cells (22, 23). In such cells, Smad2 and FAST-1 together have minimal ability to stimulate a typical activin/TGF-ß reporter gene. In addition, GAL4-Smad1, GAL4-Smad2, and GAL4-Smad3 have little transcription activities in the Smad4 deficient cells compared to the same cells with transfected Smad4 (22, and our unpublished results). Thus, Smad4 plays a crucial role in Smad-mediated transcriptional activation. This is partly due to the unique Smad activation domain (SAD), a 48 amino acid proline-rich regulatory element in the linker region of Smad4 (152). The crystal structure of a Smad4 fragment containing the SAD and the MH2 domain has been solved (153). The MH2 domain of Smad4 is highly homologous with that of Smad2 and Smad3 (30% identity), except that Smad4 has a unique insert of ~35 amino acids which interacts with the C-terminal tail to form a TOWER-like structural extension from the core. The crystal structure suggests that SAD provides transcriptional capability by reinforcing the structural core and coordinating with the TOWER to present the proline rich surface and a glutamine-rich surface in the TOWER for interaction with transcription partners (153). The SAD domain physically interacts with the N-terminal domain of p300/CBP (152). In addition, the Smad4 interacting protein MSG1, which lacks an intrinsic DNA binding ability, recruits p300/CBP to Smad4 via SAD and functions as a co-activator of Smad4 (154, 155). A recent study shows that at least part of the transcriptional activity of Smad4 requires the association with SMIF, an EVH1/WH1 (enabled VASP (vasodilator-stimulated protein) homology 1)/WASP (Wiskott-Aldrich syndrome protein) homology 1) domain protein that interacts with the SAD domain of Smad4 (156, 157).

6.3. Transcriptional repression

In addition to activating transcription, Smads can also repress transcription through recruitment of corepressors. Three Smad corepressors have been identified: the homeodomain protein TGIF (34) and the two related proteins c-ski and SnoN (36).

TGIF interacts with Smad2 and Smad3, and the interactions are increased by treatment with TGF-ß. In addition to the TGF-ß pathway Smads, TGIF can weakly interact with BMP pathway Smads, such as Smad1. In contrast, TGIF cannot interact with Smad4 (158).

TGIF represses transcription in part by interacting with histone deacetylases (HDACs) (158). In addition, TGIF interacts directly with the paired amphipathic α-helix 2 domain of the mSin3 corepressor, and recruits mSin3 to a TGF-ß activated Smad complex to inhibit transcriptional responses (159, 160). TGIF also interacts with the C-terminal binding protein (CtBP) corepressor, and efficient repression of TGF-ß transcriptional responses by TGIF is dependent on the interaction with CtBP (161). Thus, TGIF uses several modes to act as a corepressor.

TGIF binds to Smad2/3 in the C-terminal domain and is mutually exclusive with p300/CBP association (158). Thus, TGIF levels modulate sensitivity to TGF-ß transcriptional responses. TGIF is a short-lived protein. Small alterations in the physiological levels of TGIF can have profound effects on human development, as shown by the devastating brain and craniofacial developmental defects in heterozygotes carrying a hypomorphic TGIF mutant allele (162). Epidermal growth factor (EGF) signaling can lead to the phosphorylation of TGIF at two Erk MAP kinase sites which results in the stabilization of TGIF and increases the formation of Smad2–TGIF corepressor complexes in response to TGF-ß (163).

Ski was first identified as a viral oncogene (v-ski) from the Sloan-Kettering avian retrovirus that transforms chicken embryonic fibroblasts (36). The related Sno (Ski-related novel gene) is a member of the ski proto-oncogene family. Several isoforms of Sno have been reported in human (36). Sno appears to function both as an oncopogene and as a tumor suppressor (36). Ski/Sno directly binds the N-CoR and mSin3A, which form a complex with HDAC (164, 165). Ski is also required for transcriptional repression by several other proteins, including the Mad repressor in the myc-max-mad network, the thyroid hormone receptor-ß, the Rb protein and the Gli3 repressor (36, 166). Therefore, Ski appears to be an integral part of the cellular transcriptional repression machinery.

Ski and SnoN interact with Smad2 and Smad3 in a TGF-ß dependent manner, whereas Ski and SnoN interaction with Smad4 is constitutive (165, 167-171). The MH2 domains of Smad2, 3, 4 are essential for these interactions (165, 167-171), and Ski has been shown to recognize trimeric Smad3 (50, 51). Ski and SnoN inhibit TGF-ß transcriptional responses in part by interacting with N-CoR and mSin3A, thereby recruiting HDAC to TGF-ß activated Smad complexes (36, 165). Like TGIF, Ski competes with the coactivator p300/CBP for binding to activated Smad3 (168).

The recently solved crystal structure of a Ski fragment bound to the MH2 domain of Smad4 has uncovered a novel mechanism through which Ski can negatively regulate TGF-ß signaling: Ski recognizes the L3 loop of the Smad4 MH2 domain. Because the Ski binding domain on Smad4 overlaps the Smad2 and Smad3 binding domain, high levels of Ski protein may physically interfere with the ability of Smad4 to form a transcriptionally
Smads and TGF-β Signaling

competent complex with phosphorylated Smad2 and Smad3, thus inhibiting TGF-β signaling (172).

The expression of SnoN represents an interesting example of a nuclear negative feedback loop (36, 169, 170). SnoN has been proposed as a nuclear corepressor for Smad4 to maintain TGF-β-responsive genes in a repressed state in the absence of ligand. Treatment with TGF-β leads to nuclear accumulation of activated Smad2/3. Smad3, and to a lesser extent Smad2, target SnoN for degradation (see section 7), thus allowing transcriptional activation of TGF-β-responsive genes including SnoN itself. The increased expression of SnoN then participates in a negative feedback control to turn off TGF-β-signaling (36, 169, 170).

Thus, a balance of expression levels of TGFIF, Ski, and SnoN (co-repressors) and p300/CBP (co-activators) within the cell determines the outcome as well as the intensity of TGF-β/Smad signaling responses (34, 158).

Smad3 can also inhibit transcription by other modes, and in certain cases this has been linked to inhibition of differentiation. For instance, TGF-β inhibits osteoblast differentiation and function. CBFA1, a Runt domain transcription factor, also known as RUNX2, OSF2, PEBP2αA, or AML3, plays a key role in osteoblast differentiation and skeletal formation. It binds and regulates several genes that are activated during osteoblast differentiation, including osteocalcin, alkaline phosphatase, alpha1 and alpha2 (I) collagen, osteopontin and osteoprotegerin ligand. In addition, CBFA1 binds to its own promoter. Smad3 physically interacts with CBFA1 and prevents CBFA1 from activating osteocalcin and its own promoter, thereby inhibiting osteoblast differentiation (173). The inhibition is cell type-dependent. It occurs in mesenchymal but not in epithelial cells.

TGF-β is a potent inhibitor of skeletal muscle differentiation. Smad3, but not Smad2, mediates the inhibitory effect by repressing the activity of the MyoD family of transcription factors. Smad3 physically interacts with the HLH domain of MyoD, thus inhibiting MyoD heterodimerization with an E-box binding protein (such as E12 and E47) and subsequent binding of the heterodimer to the E-box, thereby blocking myogenic differentiation (174).

Smad3 can also compete for DNA binding, thus leading to transcriptional repression. For the goosecoid promoter, FAST-2, Smad2 and Smad4 form a complex that activates expression. In contrast, Smad3 inhibits goosecoid expression (105). This is thought to occur through Smad3 competing with Smad4 binding to a GC-rich sequence. While binding of Smad4 in complex with Smad2 and FAST2 to this GC-rich sequence leads to transcriptional activation, binding of Smad3 to this sequence may alter the conformation of the DNA binding complex, thus leading to inhibition of transcription (105).

Smad3 has also been reported to interact with HDAC through its MH1 domain (175), but whether the interaction is direct remains to be determined. Smad3 downregulation of c-myc expression is described in subsection 10.2. Several other interacting proteins that inhibit Smad-mediated transcription are listed in Table 1.

7. Ubiquitination Mediated Proteolysis in TGF-β/Smad Pathway

Ubiquitin-mediated proteolysis regulates the activities of a variety of proteins (176). TGF-β/Smad signaling is also regulated by this system. Smurfs are C2-WW-HECT domain E3 ubiquitin ligases that interact with Smads and regulate the TGF-β/Smad signaling pathways (177, 178). The WW domains of Smurfs recognize specific PY motifs in the linker region of Smads. All receptor regulated Smad proteins, except Smad8, contain PY motifs (see Figure 3). Smurf1 was initially identified as capable of interacting with the BMP pathway Smads, Smad1 and Smad5, and targeting them for degradation at the basal state (179). Smurf2 interacts with receptor-regulated Smads, including Smad1, Smad2 and Smad3, and mediates the degradation of Smad1 and Smad2 but not Smad3 (180, 181). This may explain, at least in part, why Smad3 is more stable than Smad2 in transient transfection assays.

TGF-β activated Smad2 and Smad3 have also been reported to be subjected to proteasome degradation. Receptor mediated activation of Smad2 leads to mult ubiquitination and subsequent degradation (182). Ubiquitination of activated Smad2, however, does not require the C-terminal tail phosphorylation but is dependent on the accumulation of Smad2 in the nucleus (182). Activated Smad3 can be targeted by the ROC1-SCFFbw1a E3 ubiquitin ligase complex for degradation (183). ROC1 is a RING finger protein and interacts with the MH2 domain of Smad3 in a TGF-β-dependent manner. The SCFFbw1a E3 ubiquitin ligase has previously been shown to be involved in the NF-kB and Wnt/Wingless signaling pathways to target IkB and β-catenin for their degradation (176). Interestingly, recruitment of p300 to nuclear Smad3 facilitates its interaction with the E3 ligase complex. Smad3 bound the ROC1-SCFFbw1a E3 ubiquitin ligase complex is then exported from the nucleus to the cytoplasm for proteasome degradation (183).

Whereas these observations indicate that activated Smad2 and Smad3 can be degraded by ubiquitin-proteasome pathways, the majority of activated Smad2 and Smad3 are thought to be dephosphorylated in the nucleus by a yet to be identified phosphatase and are recycled back into the cytoplasm (see section 5 on nucleocyttoplasmic shuttling of Smad2/3). Thus, ubiquitin-mediated degradation of activated Smad2 and Smad3 may not be an effective way to turn off TGF-β signaling. Recent studies have shown that cells have evolved another mechanism to turn off the TGF-β signal by employing Smurfs. Smurf2, and also Smurfl, bind constitutively to Smad7. Binding to Smad7 induces export of the Smurf-Smad7 complex to the cytoplasm and targets the TGF-β receptor for degradation. Thus, Smurfl/2 uses Smad7 to form a functional E3 ligase to downregulate TGF-β receptor and thereby to terminate TGF-β signaling (177, 178, 184, 185).
8. REGULATION OF SMAD2 AND SMAD3 THROUGH NON-RECEPTOR KINASES

In addition to the TGF-ß receptor, several other kinases have been shown to phosphorylate Smad proteins. The first example is MAP kinase. Thus far, this remains to be further characterized. MAP kinase has been shown to promote Smad2 nuclear translocation (194, 195). MAP kinase has also been shown to phosphorylate Smad1, Smad2 and Smad3 in the linker region, and such phosphorylation inhibits localization to the nucleus (196-197). Two other studies have shown, however, that MAP kinase has no effect on Smad2 and/or Smad3 localization (198, 199). It is unclear at present why different studies have yielded conflicting conclusions. Identification of the exact MAP kinase phosphorylation sites in Smads is necessary to clarify the discrepancy. Nevertheless, it has been consistently shown that activation of the MAP kinase pathways is necessary for transcriptional activation of several TGF-ß responsive genes, such as the CDK inhibitor p21, TGF-ß1 and furin (195, 198, 200, 201).

Calmodulin directly binds to Smad2 and also binds Smad1, Smad3 and Smad4 (202, 203). This binding is calcium-dependent and requires the N-terminal domain of Smad2. Calmodulin downregulates Smad2 activity, whereas it increase Smad1 activity. Binding of calmodulin to Smads inhibits subsequent MAP kinase dependent phosphorylation of Smads and vice versa. These observations suggest cross-talk between calcium/calmodulin, MAP kinase and TGF-ß pathways.

TGF-ß activates p38 via TAK1 (TGF-ß activated kinase 1) (207). The early activation of p38 is independent of Smad proteins (208), whereas the delayed activation of p38 relies on Smad-dependent GADD45ß expression (209). Expression of the angiogenesis inhibitor thrombospondin-1 (TSP-1) is induced by TGF-ß via Smad dependent p38 activation (209), which may play an important role in tumor suppression by TGF-ß. Activated p38 or JNK kinases and activates ATF-2 (110, 111), a basic leucine-zipper transcription factor of the ATF/CREB family, ATF-2 interacts with Smad3 and Smad4, and the interactions are increased by TGF-ß treatment (110, 111). TGF-ß activation of the p38 and JNK pathways plays important roles in apoptosis (12, 208).

CamKII phosphorylates Smad2, Smad4, and to a much less extent, Smad3 (210). The phosphorylation of Smad2 occurs in the linker (amino acid residues 240 and 260) and the N-terminal domain (amino acid 110) (Figure 3), and the phosphorylation results in the inhibition of TGF-ß-induced Smad2 nuclear accumulation and signaling (210). Interestingly, phosphorylation of Ser240 was also observed following treatment of cells with epidermal growth factor (EGF), platelet-derived growth factor (PDGF) and TGF-ß, suggesting that this phosphorylation site can be regulated in multiple ways.

PKC has been shown to phosphorylate Smad3 at Ser37 and Ser70 and Smad2 at the corresponding Ser47 and Ser110 residues both in vivo and in vitro (Figure 3, ref. 211). Phosphorylation of Smad3 by PKC abrogates its DNA-binding activity and thereby inhibits the
transcriptional responses. Loss of DNA binding activity of Smad3 makes cells more sensitive to transformation by PMA and inhibits TGF-ß-mediated cell death (211).

9. PHYSIOLOGICAL ROLES OF SMAD2 AND SMAD3 REVEALED BY KNOCKOUT STUDIES

Both Smad2 and Smad4 homozygous knockout mice are embryonic lethal (212, 213). Smad2 knockout mice show failure in egg cylinder elongation, mesoderm formation, gastrulation and establishment of an anterior-posterior (A-P) axis, and the mice die between E7.5 and E8.5 (214-217). However, in the presence of wild type extraembryonic tissues, Smad2-deficient embryos can develop beyond E7.5 and up to 10.5 days postcoitum (217). Smad4 null mice die between E6.5 and E 8.5 (218, 219). Both Smad2 and Smad3 are ubiquitously expressed. Smad3 is expressed in the brain, thymus, lung and kidney with highest expression in brain (220). Three groups have independently generated Smad3 knockout mice (220-222). Smad3 null mice display impaired mucosal immunity and diminished T cell responsiveness to TGF-ß and die between 1 and 8 months after birth due to a primary defect in immune function (220). One group reported that Smad3 null mice can develop metastatic colon cancer, whereas the other two groups did not or observed much lower incidence of colon cancer (220-222). The differences are difficult to explain. Nevertheless, one important and firm result is that a variety of primary cells examined so far from Smad3 null mice, such as fibroblasts, keratinocytes, astrocytes, T cells and certain splenocytes, are severely resistant to TGF-ß growth-inhibitory effects (143, 220, 222-224). Thus, this firmly establishes the essential role of Smad3 in mediating TGF-ß growth inhibitory responses.

Smad3 also plays an important role in fibrosis. TGF-ß improves wound healing and is critical for fibrosis. Unexpectedly, Smad3 null mice show accelerated cutaneous wound healing characterized by an increased rate of re-epithelialization and a reduced local inflammatory infiltrate (224, 225). During the re-epithelialization process, Smad3 null keratinocytes proliferate faster than wild type keratinocytes, indicating that Smad3 inhibits cell proliferation (224, 225). This is also supported by studies using mouse embryonic fibroblasts (222). The impaired inflammatory response results from a blunted chemotactic response by Smad3 null monocytes to TGF-ß and their diminished ability to induce TGF-ß itself. In addition, studies using Smad3 null mouse embryonic fibroblasts indicate that TGF-ß-induced expression of c-Jun and c-Fos, which is important for the induction of collagen and TGF-ß itself, is mediated by Smad3 (143). Taken together, these findings indicate that Smad3 is a major player in fibrosis.

In addition, Smad3 null mice display skeletal abnormalities shortly after weaning, which become worse with aging. This is due to the inability of Smad3 null chondrocytes to respond to TGF-ß during chondrogenesis, defects in chondrocyte differentiation into hypertrophic chondrocytes and thus resulting in the progressive degenerative cartilage disease, resembling osteoarthritis in human (226).

10. SMADS AND CANCER

10.1. Smads are candidate tumor suppressors

Smad2 has been shown to be a tumor suppressor in colon cancer and lung cancer (227-228). Smad4 is mutated at a high frequency in pancreatic carcinomas and colon carcinomas and to a lesser extent in several other types of cancers (229-234). Both Smad2 and Smad4 are located on human chromosome 18q21, which shows a high frequency of loss of heterozygosity (LOH) in colon and pancreatic cancers (227-229).

Tumor derived missense mutations of Smad2 and Smad4 occur mostly in the C-terminal domains (9). Some of the tumor derived C-terminal point mutations maps to the protein-protein interface of the trimer structure (41, 46, 53). These trimer interface mutations disrupt both homo- and hetero-oligomerization (41, 46, 53). One such trimer interface mutation D537E in Smad4 corresponds to D450E (or D450H) in Smad2. This D450E Smad2 mutant is defective in TGF-ß receptor-mediated phosphorylation. Mutation of the analogous position in Smad3, D407E, inhibits activation of both Smad2 and Smad3 and has a dominant negative effect on TGF-ß (235). Some of the tumor-derived C-terminal missense mutations affect stabilities in Smad4 (190-192). For tumor-derived missense mutations that target the MH1 domain, the R133C in Smad2 and R100T in the analogous position in Smad4 increase autoinhibitory interaction between the MH1 and MH2 domains (78). Moreover, the mutation decreases their stabilities by targeting Smad2 and Smad4 to the ubiquitin-proteasome pathway (190).

The early embryonic lethality of Smad2 null mice prevents the study of its role in tumorigenesis in mouse models. Heterozygotes are fertile and have no apparent abnormality up to at least 1 1/2 years of age. To examine the role of Smad2 inactivation in the process of carcinogenesis, two groups generated compound heterozygous mice, which carry both APC and Smad2 mutations on the same chromosome in the cis-configuration (236, 237). Probably due to the use of genetically different Smad2 heterozygous mice, the two groups observed different results. Although both groups found no difference in the total number of polyps or tumors in the cis-compound APC/Smad2 mice versus the simple APC heterozygous mice, the other observations differed. One group found that APC/Smad2 cis-compound heterozygotes developed multiple invasive cancers not present in APC single heterozygotes (236). In addition, APC/Smad2 cis-compound heterozygotes showed increased sudden death from intestinal obstruction caused by extremely large tumors (236). These observations indicate that loss of Smad2 does not initiate tumorigenesis by itself but accelerates malignant progression of tumors to invasive cancer in the late stages of carcinogenesis (236). In contrast, the other group found that the polyps in the APC/Smad2 cis-compound mice displayed no difference in numbers, size or histopathology compared to the polyps in the APC heterozygous mice, and therefore concluded that Smad2 LOH is insufficient to cause malignant progression of colonic polyps (237). The basis for the discrepancy remains to be determined.
Smads and TGF-β Signaling

The latter group described above had previously constructed APC/Smad4 cis-compound mice, the polyps of which progress to very invasive adenocarcinomas, indicating that Smad4 plays a significant role in the control of malignant progression of colon tumors (238). It also suggests that on human chromosome 18q21, the Smad4 LOH plays a more significant role than Smad2 LOH in cancer progression. Interestingly, Smad4 heterozygous mouse can develop gastric polyposis and tumors in old age. Some of these tumors exhibited loss of LOH of the Smad4 allele, whereas half of those analyzed maintained a normal Smad4 allele. These observations suggest that LOH at the Smad4 locus may not be an obligatory event in Smad4-dependent tumorigenesis (239). Nevertheless, in order to determine whether haploinsufficiency of Smad4 is indeed sufficient for tumor initiation, it is necessary to determine whether these apparently normal Smad4 alleles are indeed wild type, not bearing point mutations.

Human Smad3 is localized near a hot spot mutation area (chromosome 15q21-22) for breast cancer and for a few other types of cancers (228). Thus far, there is only one report suggesting infrequent alterations of the Smad4 locus, whereas half of those analyzed maintained a normal Smad4 allele. These observations suggest that both p27 and Smad3 are important in downregulation of the c-myc proto-oncogene and upregulation of the CDK inhibitors p15 and P21 (see subsection 10.2 below). Accordingly, Smad3 plays an important role in downregulation of the c-myc proto-oncogene and upregulation of the CDK inhibitors p15 and P21 (see subsection 10.2 below).

10.2. TGF-β/Smads and cell cycle control

G1 cell cycle control is intimately associated with cancer (242, 243). Cell cycle progression from G1 to S phase is controlled by G1 cyclin-dependent kinases (CDKs) which include the homologous CDK4 and CDK6 as well as CDK2. CDK4/6 is activated by D type cyclins at early-mid G1 phase, whereas CDK2 is activated by E- and A-type cyclins during late G1 and S phase, respectively (242-244). The G1 CDK activities are negatively regulated by two classes of CDK inhibitors (242-244). The first class includes four members of the INK4 family, such as p16 and p15, which specifically inhibit CDK4/6 activity. The second class includes three members of the p21 family which inhibit CDK2 activity and at high doses also inhibit CDK4/6 function (244). Both CDK4/6 and CDK2 phosphorylate and inactivate the retinoblastoma (Rb) protein (242-244). The Rb pathway, in which p16 inhibits cyclin D-CDK4 and prevents it from phosphorylating Rb, is inactivated in most human cancers (242, 243).

TGF-β potently inhibits cell proliferation by causing cell cycle arrest at the G1 phase through regulation of cell cycle components (9, 245). For example, it has been shown that TGF-β downregulates the expression of c-Myc and the phosphatase CDC25A (246-250) and upregulates the expression of CDK inhibitors p15 and p21 (251-254). The downregulation of c-Myc is the key event in TGF-β mediated growth inhibition (9, 255-259). A recent study shows that treatment with TGF-β induces a preassembled complex containing Smad3, E2F4/5 and DP1, and the Rb-related factor p107 to translocate to the nucleus, associate with Smad3, bind to the TGF-β inhibitory element (TIE) and repress transcription of c-Myc (259). Within the TIE sequence 5′ GGCTTGGCGGGAAA 3′, Smad3 binds to the imperfect site GGCT (258, 259). As described in subsection 6.1, the second position of the TIE element can tolerate substitutions. E2F4/5 and DP1 bind the other part of the TIE sequence, and the associated p107 functions as a corepressor (258, 259).

Downregulation of c-Myc is necessary for subsequent transcriptional induction of p15 and p21 (9, 255, 256, 260, 261). Miz-1, a zinc finger transcription factor, binds to the initiator of p15 promoter and stimulates its transcription (260, 261). In the absence of TGF-β, Myc and Max form a complex with Miz-1 at the p15 initiator, inhibiting transcriptional activation by Miz-1. The binding of Myc to Miz-1 interferes with the recruitment of transcriptional co-activator p300/CBP. Treatment with TGF-β downregulates c-Myc expression, thus relieving the inhibition. Moreover, a Smad complex containing Smad2, Smad3 and Smad4 binds to a Smad binding region containing the AGAC sequence and forms a multi-protein complex with Sp1 and Miz-1, together leading to transcriptional activation of the p15 promoter (260, 261). Smad2, Smad3 and Smad4 are all necessary for maximal induction of p15 promoter (123). In addition, c-Myc can also repress p15 transcription by another mechanism. c-Myc rapidly interacts with Smad2 and Smad3 in response to TGF-β (263). Through its interaction with Smad2/3, c-myc binds the Sp1-Smad complex and inhibits transcriptional activation of the p15 promoter. When the c-Myc expression level is decreased ~2 hours after TGF-β treatment, the interaction between c-Myc and Smad2/3 is decreased, and thereby transcriptional repression is relieved (263).

The p21 promoter is regulated in a similar manner. Although the p21 promoter contains a TATA box and has an overall distinct topology from that of the p15 promoter, Miz-1 can bind the proximal region of the p21 promoter (262). The sequences contacted by Miz-1 on the p21 promoter are different from those on the p15 promoter or the low density lipoprotein receptor promoter, which Miz-1 binds to activate transcription (264). Since Miz-1 contains 13 zinc fingers, it is possible that different zinc fingers recognize different sequence elements. c-Myc is recruited to the p21 promoter by Miz-1 (262) and inhibits p21 induction by TGF-β (9, 256, 262). Smad3 and Smad4 have been shown to cooperate with Sp1 to activate the p21 gene (121, 122). Although Smad2 has not been shown to functionally participate in the activation of p21 promoter, Smad2 has been shown to physically associate with Sp1, depending on an activated TGF-β type I receptor (122). It is also expected that Smad2 plays a role in the TGF-β-
induced transcriptional activation of the p21 gene. The p21 promoter contains several proximal Sp1 binding sites, one of which is critical for TGF-ß induction (122, 254). The p21 promoter contains a distal SBE but appears to be dispensable for TGF-ß induction. Accordingly, the DNA binding activities of Smad3 and Smad4 appear not essential for activation of the p21 promoter, as DNA binding domain mutants of Smad3 and Smad4 are capable of activating the p21 promoter (122). Thus, it still remains to be determined how Smads cooperate with Sp1 and Miz-1 to activate the p21 promoter.

TGF-ß induced p15 binds to CDK4 and CDK6 and prevents their interaction with cyclin D. As a result, p27, which is bound to the cyclin D-CDK4, is displaced and then binds to the cyclin E-CDK2 complex and inhibits its activity. Thus, the coordinated inhibition of CDK4/6 and CDK2 activities by p15 and the p21 family members leads to cell cycle arrest induced by TGF-ß (9, 251, 265, 266).

11. CONCLUSIONS AND PERSPECTIVES

TGF-ß regulates a wide variety of biological activities. Elucidation of the Smad signaling pathway represents a milestone in our understanding of TGF-ß action, yet many interesting questions remain to be addressed. For instance, the early events in the organization of the receptor-SARA-Smad2/3 complex remains to be further investigated. Whether SARA and Smad2/3 bind the receptor complex at the plasma membrane or in the early endosome after receptor endocytosis (267, 268) need to be further addressed. In addition to the C-tail SSXS motif phosphorylation, trypptic peptide analysis indicates that other sites in Smad2, and presumably Smad3 as well, are rapidly phosphorylated in response to TGF-ß by receptor kinase or cytoplasmic kinases (18). It is imperative to identify these phosphorylation sites in Smad2/3 and the kinases as well as to characterize the functional significance of such phosphorylation. The apparently different stoichiometry of the Smad3-Smad4 complex versus the Smad2-Smad4 complex should be further verified, which may be related to distinct activities and regulation of Smad2 and Smad3. Another major issue is how TGF-ß signaling specificity is achieved. In different cells, different combinations of receptors, Smad proteins, and Smad-interacting partners will lead to the activation or repression of distinct target genes, thus eliciting different biological responses. DNA microarray analysis has already revealed a great number of genes that are regulated by TGF-ß (257, 269-271). It is important to determine which Smad proteins regulate which subsets of these TGF-ß-responsive genes. The use of cells from Smad2, Smad3 or Smad4 knockout mice is particularly useful to assign different target genes to distinct Smad proteins (143, 272). This approach is currently limited to the use of mouse embryonic fibroblasts. In the future, conditional knockout of Smad2, Smad3 and Smad4 by Lox-P-mediated recombination or by RNA interference will yield even more valuable information regarding the distinct roles of Smad2, Smad3 and Smad4 in different tissues and diseases. The discovery of Smad has provided the basis for tremendous future work to further understand the multifunctional nature of TGF-ß. New surprises and new findings are fully expected.

12. ACKNOWLEDGEMENT

I thank Y.-G. Chen, A. H. Conney, A. J. Shatkin, Y. Shi and members of my laboratory for critical reading of this review. I also thank the American Association for Cancer Research-National Foundation for Cancer Research, the Burroughs Wellcome Fund, the Sidney Kimmel Foundation for Cancer Research, the Emerald Foundation, the Pharmaceutical Research and Manufacturer of America Foundation, the New Jersey Commission on Cancer Research, and the NIH for support.

13. REFERENCES

BMP receptor kinase. phosphorylated directly and activated functionally by the Massague: The TGF-beta family mediator Smad1 is required for Smad2-Smad4 complex formation and signaling. J Biol Chem 272, 27678-27685 (1997)


77. Xiao Z., X. Liu, Y.I. Henis & H.F. Lodish: A distinct nuclear localization signal in the N terminus of Smad 3
Smads and TGF-β Signaling
Smads and TGF-β Signaling


121. Moustakas A. & D. Kardassis: Regulation of the human p21/WAF1/Cip1 promoter in hepatic cells by functional interaction between Sp1 and Smad family members. Proc Natl Acad Sci USA 95, 6733-6738 (1998)


135. Matsuda T., T. Yamamoto, A. Muraguchi & F. Saatcioglu: Cross-talk between transforming growth factor-


164. Nomura T., M.M. Khan, S.C. Kaul, H.D. Dong, R. Wadhwa, C. Colmenares, I. Kohno & S. Ishii: Ski is a component of the histone deacetylase complex required for...
transcriptional repression by Mad and thyroid hormone receptor. *Genes Dev* 13, 412-423 (1999)
195. Blanchette F., N. Rivard, P. Rudd, F. Grondin, L. Attisano & C.M. Dubois: Cross-talk between the p42/p44
Smads and TGF-ß Signaling


**Abbreviations:** BMP: bone morphogenetic protein, Mad: mothers against Dpp, MH1: Mad Homology 1, MH2: Mad Homology 2, SARA: Smad anchor for receptor activation, Sma: small phenotype in *C. elegans*, Smads: mammalian homologues of *Drosophila* Mad and *C. elegans* Sma, Smurfs: Smad ubiquitin regulatory factors, TGF-β: Transforming growth factor β, TGIF: TG-interacting factor.

**Key Words:** TGF-β, Smad, SARA, Smurf, Cytokine, Signal Transduction, Transcriptional Regulation, Cancer, Tumor Suppressor, Cell Cycle, Review

**Send correspondence to:** Dr. Fang Liu, Center for Advanced Biotechnology and Medicine, Rutgers University, 679 Hoes Lane, Piscataway, NJ 08854. Tel: 732-235-5372, Fax: 732-235-4850, E-Mail: fangliu@cabm.rutgers.edu