PERSPECTIVE

From mono- to oligo-Smads: The heart of the matter in TGF- β signal transduction

Aristidis Moustakas¹ and Carl-Henrik Heldin

Ludwig Institute for Cancer Research, SE-751 24 Uppsala, Sweden

Understanding the mechanisms that govern specificity in signal transduction networks is of paramount importance as the ever increasing power of molecular and cellular techniques brings to the surface more and more complex associations between pathways of cell communication. A primary determinant of specificity in all signaling protein networks is the powerful but flexible biochemical "glue" of interacting protein interfaces. When X-ray crystallographers explore this principle at the atomic level, they often reveal beautiful examples of molecular design. One such novel revelation is presented in this issue of *Genes & Development* by Qin et al. (2002), who describe novel protein interfaces and their dynamic role in controlling specificity during signal transduction by members of the transforming growth factor- β (TGF- β) family.

TGF- β and its related growth and morphogenetic protein factors regulate cell proliferation, differentiation, migration, and programmed cell death as well as embryonic development and tissue interactions. Over the past 10 years, the three-dimensional structures of TGF- β s, themselves, as well as of several components in the signaling pathway, have been solved by X-ray crystallography (for review, see Souchelnytskyi et al. 2002). The pathway initiates with the activation of receptor serine/threonine kinases (RS/TK) on the cell surface by extracellular ligands, and propagates inside the cell primarily via the Smad family of signal transducers that rapidly translocate to the nucleus (for review, see Moustakas et al. 2001; Attisano and Wrana 2002). Alternative signaling cascades may also be activated by TGF- β family members; however, the Smads represent the major and most direct route of communication between the extracellular milieu and the cell nucleus, the ultimate target of signal transduction pathways (Massagué 2000; Wakefield and Roberts 2002).

The findings by Qin et al. (2002) suggest a model that integrates molecular events proximal to the cell surface receptors with those occurring in the nucleus during

pathway-instructed regulation of gene expression. A major transition that Smad proteins undergo as they move from the receptors to the nucleus is to change from inactive monomers to active trimers. This stoichiometric change regulates molecular specificity in an interesting manner. Pivotal to the understanding of the new model is the modular structure of Smads, which consists of the conserved, globular Mad-homology (MH) domains, MH1 and MH2, in the N and C termini, respectively, bridged by a less conserved and flexible, proline-rich linker (L) domain (Shi 2001).

A signaling module made by RS/TKs, SARA, and Smads

Ligand binding induces the assembly of two type II RS/TKs into heterotetramers with two type I RS/TKs. The constitutively active type II receptor kinase phosphorylates and activates the type I kinase, which sequentially phosphorylates and activates different receptor-regulated (R)-Smads (Wrana 2002). This initial step in TGF- β signal transduction is organized architecturally by the contribution of at least one major adaptor protein, the Smad anchor for receptor activation (SARA) (Tsukazaki et al. 1998). The ligand-receptor complex forms in the plasma membrane, it is then endocytosed and the functional signaling module assembles on the lipid bilayer of early endosomes, where the receptor complex meets SARA, which tethers R-Smads to the bilayer via its phosphoinositide-binding FYVE domain (Fig. 1; Itoh et al. 2002; Panopoulou et al. 2002). These early steps in the organization of the signaling complex have only recently been experimentally addressed, and therefore whether SARA plus R-Smads meet the receptor complex at the plasma membrane or in the early endosome remains controversial (Lu et al. 2002; Penheiter et al. 2002). Interesting questions relating to the model include whether the proximity of Smads to the membrane is important and why SARA prefers monomeric Smads. Qin et al. (2002) provide satisfactory data to answer both questions.

Based on the analysis of the new co-crystal between the SARA Smad-binding domain (SBD) and the MH2 domain of Smad3, one of the R-Smads activated by TGF- β ,

¹Corresponding author.

E-MAIL Aris.Moustakas@licr.uu.se; FAX (46) 18 160420.

Article and publication are at <http://www.genesdev.org/cgi/doi/10.1101/gad.1016802>.

Figure 1. TGF- β signal transduction from the cell surface to the nucleus. Five snapshots of the signaling pathway are highlighted with an emphasis on the new findings by Qin et al. (2002). Binding of ligand to the heterotetrameric receptor complex on the plasma membrane leads to trans-phosphorylation (flat, pink arrows) of the GS-box of each type I receptor (I) by the constitutive type II receptor kinase (II). The ligand-receptor complex is endocytosed to early endosomes where dimeric SARA, tethered to the lipid bilayer via its FYVE domain (diamond), anchors two subunits of an R-Smad (Smad3) in a "closed" conformation via its SBD domain (triangle). The MH1 domain is shown as an oval, the linker domain is shown as a thin line, and the MH2 domain is shown as a rectangle with a pocket that represents the combined L3 loop-basic patch surface that recognizes phospho-serines either on the GS-box of the type I receptor or on the C-terminal tail of another Smad. Assembly of the ligand-receptor-SARA-Smad machinery on the endosomal bilayer leads to a conformational change of Smad3 and its proper positioning towards the type I receptor kinase, so that the C-terminal tail becomes phosphorylated (flat arrows). The linker and MH1 domains are not known to interact with the receptors and in this cartoon they can be visualized as orienting away from the receptors in the third dimension. The receptor-proximal steps of Smad oligomeric assembly are not fully resolved and are indicated by broken arrows. Smad3 homotrimer and Smad3-Smad4 heterotrimer are shown. Ultimately, the heterotrimer is imported into the nucleus, where it primes the assembly of a nuclear complex in association with chromatin. The Smad heterotrimer is depicted as an adaptor platform that recruits the corepressor (c-Ski) with histone deacetylases (HDAC), coactivators (CBP/p300), DNA-binding transcription factors (TF) linked through the MH1 or MH2 domains, and E3 ubiquitin ligases (Smurf) through the linker domain. Arrows point to processes of acetylation or deacetylation, and a positive or negative outcome on promoter activity is indicated.

the authors determined that Smad3 was in monomeric configuration in the complex, much like the isolated Smad3 MH2 domain that the authors also crystallized. Although the crystal data of the Smad3–SARA complex faithfully reproduce the previously published structure of the Smad2–SARA complex (Wu et al. 2000), the new study goes one step further. We now learn that an unphosphorylated, monomeric Smad exhibits better surface complementarity and fitness to the SBD of SARA than the oligomeric Smad3 MH2 domains since the receptor-induced activating phosphorylation of two C-terminal serine residues induces a significant shift in the core MH2 fold. In agreement with this, Qin et al. (2002) show that the SARA SBD can serve as an inhibitor of Smad3 trimerization based on its ability to lock Smad3 in its monomeric conformation. Thus, SARA can also serve as guardian of R-Smads by eliminating the possibility of aberrant R-Smad oligomerization and activation in the absence of a proper signal from the receptors. This would explain why experiments of R-Smad overexpression in mammalian cells so often lead to constitutive activation of the pathway; high levels of R-Smad possibly overcome the limiting levels of endogenous SARA allowing aberrant trimer formation despite the lack of proper activation by the ligand–receptor system.

Proximity to the lipid bilayer is also critical, as this most probably guarantees proper assembly of the receptor kinase–SARA–Smad complex that will result in Smad C-terminal phosphorylation. Based on molecular modeling of the structures of the type I receptor ectodomain bound to BMP-2 (Kirsch et al. 2000), the type I RS/TK in its active and inactive form (Huse et al. 2001), the FYVE domain of early endosome autoantigen 1 (EEA1) (Dumas et al. 2001), which exhibits a high degree of similarity with the FYVE domain of SARA, and finally the current Smad3–SARA co-crystal, the authors provide the first atomic view of the signaling machinery that transmits the TGF- β signal from the extracellular environment to the R-Smad (Fig. 1). The modeling data reveal a high degree of surface complementarity that makes the model very likely to reflect the *in vivo* protein machinery. The model also provides the appealing prediction that two R-Smad subunits are simultaneously tethered to each receptor complex through dimeric SARA, and it will be important to define whether distinct signaling complexes are formed with each R-Smad type (Smad2 versus Smad3) or whether a single complex can form with one Smad2 and one Smad3. In addition, the model emphasizes the need for proper docking and orientation of the various subunits of the machinery. This is consistent with experimental data showing that the receptor kinases must be oriented properly, presumably through rotational adjustment induced by the ligand binding to the receptor ectodomains and transmission through the α -helical rod of the transmembrane domain (Zhu and Sizeland 1999). In addition, SARA may assist in the proper orientation of the kinases, since it interacts with the receptors directly (Tsukazaki et al. 1998), but this possibility remains to be proven. In the complex between receptors and the SARA-bound R-Smad, the un-

structured C-terminal tail of Smad must be properly oriented to fit into the active center of the kinase. That SARA mediates proper presentation of the R-Smad C-terminal tail to the catalytic site is strongly suggested by the current modeling data; however, it requires additional biochemical verification.

A sensitive transition step and a point of controversy

Upon R-Smad phosphorylation, the signaling cascade continues with the incorporation of Smad4 into the trimeric complex (for review, see Moustakas et al. 2001). Smad4 is the common mediator Smad (Co-Smad) of all TGF- β superfamily signaling pathways, and it lacks the C-terminal serines that are phosphorylated by the type I receptors. Lin and colleagues (Chacko et al. 2001) have previously reported on the formation of the Smad3–Smad4 heterotrimer with a stoichiometry of two Smad3 subunits to one Smad4 subunit. How is Smad4 incorporated into the functional oligomer? Although this question is not addressed by the current report, it is appropriate to discuss it briefly here. We can envision two mechanisms that are not necessarily mutually exclusive (Fig. 1). In both scenarios Smad4 may be presented to the receptor complex by an adaptor such as TGF- β receptor interacting protein (TRIP)-1 (Wurthner et al. 2001). One possibility is that a heterotrimer is built from the two phospho-Smad3 subunits that are presented to the receptor complex by the dimeric SARA and are released after phosphorylation, and one Smad4 subunit. Alternatively, a homotrimer of phospho-Smad3 is first formed, then a Smad4 subunit is incorporated into the trimer by displacing one phospho-Smad3 subunit. The latter event is possible based on thermodynamic criteria established by previous ultracentrifugation studies and the current crystallographic and gel chromatographic analyses using purified Smad3 MH2 domains (Correia et al. 2001; Qin et al. 2002). Additional work is needed to clarify these possibilities especially under *in vivo* conditions. Moreover, evidence exists suggesting a distinct role of activated R-Smads in the nucleus in the absence of Smad4 (for review, see Moustakas et al. 2001). Thus, the incorporation of a Smad4 subunit in a heterotrimeric Smad complex affects Smad function and is likely to be a carefully regulated event. Future structural and cell biological experiments are needed to elucidate this central regulatory mechanism in the Smad pathways.

Interestingly, data coming from another group of crystallographers suggests that other R-Smads may work via different mechanisms. For example, studies of Smad2, another R-Smad of the TGF- β pathway, point to a different model of oligomerization. Smad2 can form a heterodimer with Smad4 despite the fact that phospho-Smad2 crystallizes as a trimer much like Smad3 and Smad4 (Wu et al. 2001a,b). The differential mode of oligomerization of the two R-Smads, Smad2 and Smad3, is also supported by biochemical data using gel chromatography (Jayaraman and Massagué 2000). This represents the first point of divergence between the two highly similar R-Smads of the TGF- β pathway. In addition, Smad2 and Smad3 use

different mechanisms of nuclear import and exhibit distinct abilities to interact with transcription factors and to bind to DNA elements, termed Smad binding elements (SBEs), suggesting that they may play distinct roles in the regulation of various gene targets of the TGF- β pathway (for review, see Moustakas et al. 2001). A comparative analysis of the mode of oligomerization of Smad2 and Smad3 by both structural and biochemical means is highly warranted.

The nuclear Smad trimer as a transcriptional adaptor

Qin et al. (2002) go even further in their analysis of the significance of their allosteric model of Smad3 that depends on C-terminal phosphorylation and subsequent oligomerization. They proceed by studying a major nuclear partner of the Smad complex, namely the proto-oncogene product c-Ski, which is known to repress the Smad signal by recruitment of corepressors and histone deacetylases (for review, see Liu et al. 2001). Ski contacts Smads at their MH2 domains and Qin et al. have mapped the minimal sequence in c-Ski that defines this interaction. Using biochemical approaches they then demonstrate that the c-Ski epitope recognizes the same amino acid interface of Smad3 MH2 as SARA. As a consequence, the two interactions are mutually exclusive. This is reminiscent of a recent study of Smad2 that defined the interface between the MH2 domain of Smad2 and transcription factors of the FoxH1 and Mix families that regulate critical gene targets of the TGF- β pathway in early embryonic development (Randall et al. 2002). The Smad interaction motif (SIM) of FoxH1 and Mix proteins resembles the SARA SBD and all groups of proteins interact with overlapping epitopes on the Smad2 MH2 domain, much like c-Ski and SARA. The current study goes once again one step further and proves that the Smad-binding sequence of c-Ski preferentially binds to Smad3 present in a trimeric complex. This observation is beautifully confirmed by mutagenesis of those key amino acids that contribute to the Smad3-oligomer interface (first revealed by the earlier crystallographic work of Chacko et al. 2001), leading to complete loss of interaction with the c-Ski peptide. Thus, a peptide derived from the Smad-binding sequence in c-Ski can serve as a molecular tool to probe cellular levels of Smads present within the oligomeric complex. Finally, molecular modeling of the c-Ski peptide, using the structure of the SARA SBD as a template, revealed conserved elements in the secondary structure of the two proteins. The c-Ski peptide is shown to fit best if associated with Smad3 incorporated into trimeric complexes, whose core MH2 fold has been shifted, while the fit of the c-Ski peptide to monomeric Smad3 MH2 is poor. In contrast, the SARA SBD recognizes the conformation of the MH2 domain in monomeric Smad3. An oligo-proline motif in SARA, that is replaced by glycines in c-Ski, seems to explain this critical differential affinity of the two partners of Smads.

Thus, the structural analysis gives a molecular explanation for the exquisite specificity of interaction be-

tween two important cofactors of the Smad pathway. SARA will only see the Smads in the cytoplasm where it resides tethered to the lipid bilayers of early endosomes and this is dictated by its high affinity for monomeric Smads. Once the activated Smad trimer is formed, the interaction with SARA is lost due to steric hindrance. However, the trimers provide new surfaces for interactions with transcription factors such as c-Ski (Fig. 1). One important prediction that the new model makes regarding the nuclear function of the Smad complex is that it can serve as a multi-adaptor for several nuclear factors. Thus, Smad4 may preferentially bind to coactivators of the p300 family through its Smad activation domain (SAD) (Qin et al. 1999), one Smad3 subunit may contact corepressors such as c-Ski, and another Smad3 subunit may bind to another transcription factor such as a Runx family member (Fig. 1). All current structural data in support of this hypothesis are derived from studies of the C-terminal MH2 domain of Smads; however, one has to remember that Smads also contain the N-terminal MH1 and intermediate linker domains. The MH1 domain not only binds to DNA sequences (the SBEs) but also interacts with additional transcription factors such as c-Jun or Sp1. Finally, the linker can recruit E3 ubiquitin ligases of the Smurf family (Fig. 1). Thus, under physiological conditions a nuclear Smad heterotrimer may provide a multitude of protein interfaces that are able to recruit to a single genomic locus DNA-binding transcription factors, coactivators, corepressors, acetyltransferases, deacetylases, and ubiquitin ligases in order to execute the transient transcriptional program of the signaling pathway in an efficient and highly regulated way. If this prediction stands true, we may hypothesize that many of the known nuclear components that interact with Smad proteins may exhibit a considerable degree of cross talk or coregulation. A major task for the future of the Smad signaling field will be to decipher the precise and sequential mode of Smad interactions. The basis for this work will be the Smad heterotrimer as the current study of Qin et al. (2002) firmly establishes. Much excitement is thus anticipated for the near future of this young signal transduction field.

References

- Attisano, L. and Wrana, J.L. 2002. Signal transduction by the TGF- β superfamily. *Science* **296**: 1646–1647.
- 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**: 1473–1482.
- Dumas, J.J., Merithew, E., Sudharshan, E., Rajamani, D., Hayes, S., Lawe, D., Corvera, S., and Lambright, D.G. 2001. Multivalent endosome targeting by homodimeric EEA1. *Mol. Cell* **8**: 947–958.
- Huse, M., Muir, T.W., Xu, L., Chen, Y.G., Kuriyan, J., and Massagué, J. 2001. The TGF β receptor activation process: an

- inhibitor- to substrate- binding switch. *Mol. Cell* **8**: 671–682.
- Itoh, F., Divecha, N., Brocks, L., Oomen, L., Janssen, H., Calafat, J., Itoh, S., and ten Dijke, P. 2002. The FYVE domain in Smad anchor for receptor activation (SARA) is sufficient for localization of SARA in early endosomes and regulates TGF- β /Smad signalling. *Genes Cells* **7**: 321–331.
- Jayaraman, L. and Massagué, J. 2000. Distinct oligomeric states of SMAD proteins in the transforming growth factor- β pathway. *J. Biol. Chem.* **275**: 40710–40717.
- Kirsch, T., Sebald, W., and Dreyer, M.K. 2000. Crystal structure of the BMP-2-BRIA ectodomain complex. *Nat. Struct. Biol.* **7**: 492–496.
- Liu, X., Sun, Y., Weinberg, R.A., and Lodish, H.F. 2001. Ski/Sno and TGF- β signaling. *Cytokine Growth Factor Rev.* **12**: 1–8.
- Lu, Z., Murray, J.T., Luo, W., Li, H., Wu, X., Xu, H., Backer, J.M., and Chen, Y.G. 2002. TGF β activates Smad2 in the absence of receptor endocytosis. *J. Biol. Chem.* (In press).
- Massagué, J. 2000. How cells read TGF- β signals. *Nat. Rev. Mol. Cell Biol.* **1**: 169–178.
- Moustakas, A., Souchelnytskyi, S., and Heldin, C.H. 2001. Smad regulation in TGF- β signal transduction. *J. Cell Sci.* **114**: 4359–4369.
- Panopoulou, E., Gillooly, D.J., Wrana, J.L., Zerial, M.H., Stenmark, H., Murphy, C., and Fotsis, T. 2002. Early endosomal regulation of Smad-dependent signaling in endothelial cells. *J. Biol. Chem.* **277**: 18046–18052.
- Penheiter, S.G., Mitchell, H., Garamszegi, N., Edens, M., Doré Jr., J.J., and Leof, E.B. 2002. Internalization-dependent and -independent requirements for transforming growth factor β receptor signaling via the Smad pathway. *Mol. Cell Biol.* **22**: 4750–4759.
- Qin, B., Lam, S.S., and Lin, K. 1999. Crystal structure of a transcriptionally active Smad4 fragment. *Structure Fold Des.* **7**: 1493–1503.
- Qin, B.Y., Lam, S.S., Correia, J.J., and Lin, K. 2002. Smad3 allosterically links TGF- β receptor kinase activation to transcriptional control. *Genes & Dev.* (this issue).
- Randall, R.A., Germain, S., Inman, G.J., Bates, P.A., and Hill, C.S. 2002. Different Smad2 partners bind a common hydrophobic pocket in Smad2 via a defined proline-rich motif. *EMBO J.* **21**: 145–156.
- Shi, Y. 2001. Structural insights on Smad function in TGF β signaling. *Bioessays* **23**: 223–232.
- Souchelnytskyi, S., Moustakas, A., and Heldin, C.-H. 2002. TGF- β signaling from a three-dimensional perspective: insight into selection of partners. *Trends Cell Biol.* **12**: 304–307.
- Tsukazaki, T., Chiang, T.A., Davison, A.F., Attisano, L., and Wrana, J.L. 1998. SARA, a FYVE domain protein that recruits Smad2 to the TGF β receptor. *Cell* **95**: 779–791.
- Wakefield, L.M. and Roberts, A.B. 2002. TGF- β signaling: positive and negative effects on tumorigenesis. *Curr. Opin. Genet. Dev.* **12**: 22–29.
- Wrana, J.L. 2002. Phosphoserine-dependent regulation of protein-protein interactions in the Smad pathway. *Structure (Camb)* **10**: 5–7.
- Wu, G., Chen, Y.G., Ozdamar, B., Gyuricza, C.A., Chong, P.A., Wrana, J.L., Massagué, J., and Shi, Y. 2000. Structural basis of Smad2 recognition by the Smad anchor for receptor activation. *Science* **287**: 92–97.
- 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- β signaling. *Mol. Cell* **8**: 1277–1289.
- Wurthner, J.U., Frank, D.B., Felici, A., Green, H.M., Cao, Z., Schneider, M.D., McNally, J.G., Lechleider, R.J., and Roberts, A.B. 2001. Transforming growth factor- β receptor-associated protein 1 is a Smad4 chaperone. *J. Biol. Chem.* **276**: 19495–19502.
- Zhu, H.J. and Sizeland, A.M. 1999. A pivotal role for the transmembrane domain in transforming growth factor- β receptor activation. *J. Biol. Chem.* **274**: 11773–11781.

From mono- to oligo-Smads: The heart of the matter in TGF- β signal transduction

Aristidis Moustakas and Carl-Henrik Heldin

Genes Dev. 2002, **16**:

Access the most recent version at doi:[10.1101/gad.1016802](https://doi.org/10.1101/gad.1016802)

Related Content **Swapping Smad Partners**
[Sci. STKE August , 2002 2002: tw294-TW294](#)

References This article cites 25 articles, 10 of which can be accessed free at:
<http://genesdev.cshlp.org/content/16/15/1867.full.html#ref-list-1>
Articles cited in:
<http://genesdev.cshlp.org/content/16/15/1867.full.html#related-urls>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).