Semisynthesis of Phosphovariants of Smad2 Reveals a Substrate Preference of the Activated TβRI Kinase†

Jennifer J. Ottesen, Morgan Huse, Matthew D. Sekedat, and Tom W. Muir*

Laboratory of Synthetic Protein Chemistry, The Rockefeller University, New York, New York 10021

Received January 21, 2004; Revised Manuscript Received March 9, 2004

ABSTRACT: Transforming growth factor-β (TGF-β) signaling regulates a wide range of cellular processes. Aberrant TGF-β signaling has been implicated in various disease states in humans. A key element in this signaling pathway is phosphorylation of R-Smads such as Smad2 at the last two serine residues of the C-terminal sequence CSSXS (residues 463−467 in Smad2) by the TβRI receptor kinase. Phosphorylation results in the release of the R-Smad from the membrane-anchored protein SARA, binding to the co-mediator protein Smad4, translocation into the nucleus, and regulation of target gene expression. Expressed protein ligation was used to probe the contribution of the individual phosphate groups to Smad2 oligomerization and phosphorylation by TβRI. Phosphorylation at both positions was required to generate a stable homotrimer; however, the driving force for Smad2 self-association is provided by pSer465. Additionally, SARA was found to modulate the self-association of partially phosphorylated Smad2, which suggests an added role for this protein in preventing premature release of a monophosphorylated substrate from the receptor complex. In related studies, prephosphorylation of Smad2 at Ser465 was found to significantly increase the rate of phosphorylation at Ser467, suggesting that there may be specific recognition determinants within the kinase for the monophosphorylated intermediate. This information was exploited to design an improved peptide substrate for TβRI, which may prove valuable in the design of inhibitors of the TGF-β pathway.

Transforming growth factor-β (TGF-β)[1] family members, which include TGF-βs, activins, and bone morphogenic proteins (BMPs), are secreted signaling molecules that regulate a wide range of cellular responses including proliferation, differentiation, migration, and apoptosis (*1*). TGF-β family members play a prominent role in the development, homeostasis, and repair of virtually all tissues in organisms from a fruit fly to man. Consequently, deregulated TGF-β signaling has been implicated in multiple developmental disorders and in various human diseases including cancer (*2*).

Tremendous progress has been made over the past decade in elucidating the general mechanism by which TGF-β and related factors activate receptors at the cell surface and transduce signals to the nucleus (3). Specifically, secreted TGF-β family members (which are homodimeric in the active form) engage a tetrameric cell surface receptor complex composed of two "type II" and two "type I" transmembrane serine/threonine kinases. In the case of TGF-β, these receptors are referred to as TβRII and TβRI, respectively. Upon ligand binding, the type II receptor kinase phosphorylates the type I receptor at multiple Ser/Thr residues within a conserved regulatory sequence, called the GS region, situated N terminally to the kinase. This phosphorylation activates the type I receptor which, in turn, phosphorylates specific receptor-activated Smad (R-Smads) proteins (*3*). All R-Smads have the same domain structure, an N-terminal MH1 domain involved in DNA binding attached via a variable linker to a conserved C-terminal MH2 domain involved in various protein−protein interactions. The R-Smad MH2 domain plays a key role in localizing the protein to the activated receptor complex. Specifically, the conserved L3 loop of the MH2 domain interacts with the conserved L45 loop of TβRI (*4*−*6*). In addition, the phosphorylated GS region of the activated TβRI interacts with a positively charged surface patch on the MH2 domain formed by the L3 loop and the β8 strand (the loop/strand pocket) (*7*). The complex is further stabilized by a third protein, the Smad anchor for receptor activation (SARA), which binds to the MH2 domain and the receptor complex (*8*). These interactions are believed to position the R-Smad such that it can be phosphorylated by the activated TβRI on the latter two serines of the conserved C-terminal activation sequence, SSXS. Upon phosphorylation, R-Smads dissociate from TβRI and SARA and form a heterocomplex with a related

† This work was supported by a Rockefeller University Merck Postdoctoral Fellowship (J.J.O) and National Institutes of Health Grants EB001991 (T.W.M.) and GM55843 (T.W.M.).

* To whom correspondence should be addressed: Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Ave, New York, NY 10021. Tel: 212-327-7368. Fax: 212-327-7358. E-mail: muirt@rockefeller.edu.

[1] Abbreviations: TGF-β, transforming growth factor-β; TβRI and II, type I and II TGFβ family receptors; TβRI-0P, nonphosphorylated TβRI; TβRI-4P, semisynthetic, tetraphosphorylated TβRI; R-Smad, receptor-activated Smad; SARA, Smad anchor for receptor activation; SBD, Smad binding domain; Sm2 (also called MH2 domain), residues 241−467 of Smad2; Sm2-0p, nonphosphorylated Sm2; Sm2-p465, Sm2 phosphorylated at Ser465; Sm2-p467, Sm2 phosphorylated at Ser467; Sm2-2p, Sm2 phosphorylated at Ser465 and Ser467; SmadPep-0p, peptide with sequence KMGSPSVRCSSMS; SmadPep-p465, phosphorylated peptide with sequence KMGSPSVRCSpSMS; RP-HPLC, reverse phase HPLC; ESMS, electrospray ionization mass spectrometry; EPL, expressed protein ligation.

10.1021/bi0498407 CCC: $27.50
Published on Web 04/24/2004

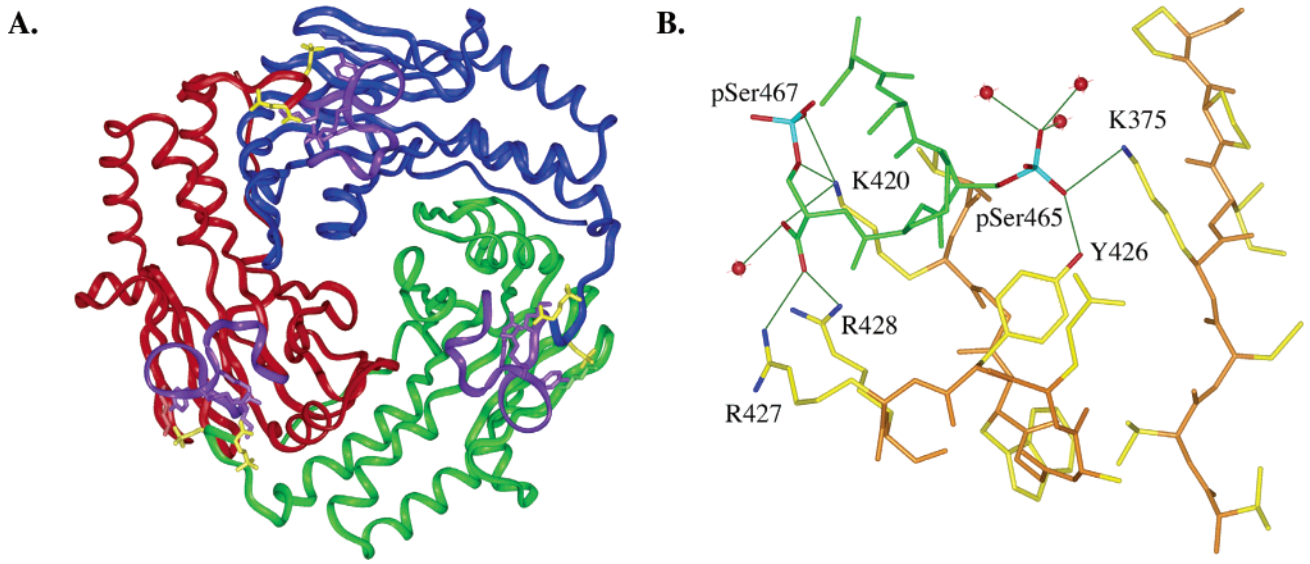


FIGURE 1: Crystal structure of the Sm2-2p trimer (PDB 1KHX). (A) Representation of the Sm2-2p structure. Individual monomers are red, green, and blue. The L3 loop (residues 419−435) in each subunit is purple, and phosphoserines are yellow. Only side chains involved in the hydrogen-bonding networks with the phosphoserines are displayed. (B) Expanded view of the hydrogen-bonding networks at the trimer interface. The phosphorylated C-terminal tail of one subunit is green, and the relevant regions of another subunit are in gold (backbone) and yellow (side chains). Hydrogen bonds (dark green lines) between oxygen (red), nitrogen (blue), and water (red balls) are shown.

protein termed the co-Smad (Smad4) (*8*). Activated R-Smad/Smad4 complexes then translocate into the nucleus where they regulate transcriptional responses, through functional interactions with various transcription factors (*9*).

Several questions remain regarding the mechanism of R-Smad activation and subsequent heterocomplex formation with Smad4. Smad4 shares the same basic domain structure as R-Smads but lacks the C-terminal phosphorylation sequence. Biochemical and structural studies have shown that phosphorylated R-Smads form a stable homotrimer in vitro (*10*, *11*). The crystal structure of the phosphorylated Smad2 MH2 domain shows that the phosphoserine residues form the core of an extensive hydrogen-bonding network involving the loop/strand pocket, thereby stabilizing the trimer interface (Figure 1). It is unclear, however, whether this homotrimer actually forms as a discrete intermediate in vivo, or whether Smad4 directly interacts with the newly phosphorylated R-Smad at the receptor complex. Moreover, the exact composition of the R-Smad/Smad4 heterocomplex is still under debate; extensive biophysical evidence has been presented supporting both heterodimer and heterotrimer models (*12*−*16*).

The details of how R-Smads are doubly phosphorylated by TβRI are not fully understood. Clearly, phosphorylation at the two sites could occur in an ordered fashion, in which one serine must be phosphorylated to prime the kinase−substrate complex for the second step. Alternatively, phosphorylation could take place in a random fashion, in which the C-terminal substrate tail is simply placed near the active site and the two steps occur in no particular order. Indirect evidence for both of these mechanisms has been presented (*17*, *18*). In addition, questions remain regarding the mechanism by which the phosphorylated R-Smad is released from TβRI. R-Smad mutants in which one or both phosphoserine positions are replaced by alanine show a higher affinity for the receptor, based on co-immunoprecipitation experiments, which in turn leads to abrogated signaling. The simplest interpretation of this is that both phosphates are required to effectively compete with the hyperphosphorylated GS region/L45 loop for binding the loop/strand pocket of the MH2 domain. This hypothesis predicts that monophosphorylated R-Smads will form less stable homo- or hetero-oligomers.

In the present paper, we investigate these issues by preparing and characterizing various phosphovariants of Smad2. We find that, indeed, both phosphates are required to give a stable homotrimer but also that the individual phosphoserine moieties make markedly different contributions to the stability of the complex. Additionally, we study phosphorylation of the individual Smad2 phosphovariants from activated TβRI and find that prephosphorylation at the N-terminal position (i.e., residue 465) significantly increases the rate of the second phosphorylation step at position 467. This information is then exploited to design an improved peptide substrate for TβRI that is likely to be useful in inhibitor-screening initiatives.

MATERIALS AND METHODS

General. All amino acid derivatives and resins were purchased from Novabiochem (San Diego, CA) and Peninsula Laboratories (Belmont, CA). All other chemicals were purchased from Sigma−Aldrich Chemical Co. (St. Louis, MO). Restriction endonucleases were purchased from New England Biolabs (Beverly, MA), and an expression plasmid encoding TEV protease was the generous gift of John Kuriyan (University of California, Berkeley, CA). Analytical RP-HPLC was performed on a Hewlett−Packard 1100 series instrument with 214 and 280 nm detection. Semipreparative, preparative, and process RP-HPLC was routinely performed on a Waters DeltaPrep 4000 system fitted with a Waters 486 tunable absorbance detector. All RP-HPLC runs used linear gradients of buffer A [water, 0.1% trifluoroacetic acid (TFA)] versus buffer B (9:1 MeCN/water, 0.1% TFA). Mass spectrometric analysis was routinely applied to all synthetic peptides and components of reaction mixtures. Electrospray ionization mass spectrometry (ESMS) was performed on a

Sciex API-100 single quadrupole electrospray ionization mass spectrometer.

Peptide Synthesis. Peptides were synthesized manually using standard (fluorenylmethoxy)carbonyl (Fmoc)-N^{α} protection strategies and HBTU/HOBT activation protocols with the following modifications: Peptides CSSMS, CSpSMS, and substrate peptides KMGSPSVRCS(S/pS)MS (SmadPep-0p, SmadPep-p465) were synthesized on a preloaded Fmoc-Ser(tBu)-WANG resin. Peptides CSSMpS and CSpSMpS were synthesized on a WANG resin in which the C-terminal phosphoserine was coupled as a symmetric anhydride (*19*); the standard procedure was modified to add an equivalent amount of base to increase the solubility of the protected amino acid. The N-terminal cysteine was coupled as the activated Fmoc-Cys(Trt)-OPfp ester in the absence of a base to minimize racemization (*20*). For SmadPep-0p and SmadPep-p465, cysteine was coupled as Fmoc-Cys(S-tBu)-OPfp; an additional deprotection step was included prior to resin cleavage (*21*). After chain assembly, peptides were cleaved off the resin with standard TFA cleavage cocktails (*22*) and purified by RP-HPLC using Vydac C18 columns. Peptide purity was assessed as ≥ 95% using analytical RP-HPLC, and peptide identity was confirmed with ESMS as follows: CSSMS (14 mg) (found *m*/*z* 514.0 M + H, expected 513.6 M + H); CSpSMS (4 mg) (found *m*/*z* 594.0 M + H, expected 593.6 M + H); CSSMpS (6.5 mg) (found *m*/*z* 594.0 M + H, expected 593.6 M + H); CSpSMpS (6.5 mg) (found *m*/*z* 674.0 M + H, expected 673.5 M + H); SmadPep-0p (1.5 mg) (found *m*/*z* 1356.8 M + H, expected 1356.6 M + H); SmadPep-p465 (1.0 mg) (found *m*/*z* 718.9 M + 2H, expected 1436.6 M + H).

Protein Preparation. Recombinant SARA/SBD (residues 665−721) was expressed and purified as described previously (*10*). Fully recombinant and semisynthetic versions of the TβRI kinase domain were also prepared according to published protocols (*23*−*25*).

Smad2 MH2 (residues 241−462) was cloned as a fusion protein with a modified GyrA intein and a C-terminal chitin binding domain in the pTXB1 vector as previously described (*10*). For protein expression, *Escherichia coli* BL21 (DE3) cells, transformed with this plasmid, were grown to midlog phase at 37 °C in Luria Bertani (LB) media and then induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside for 2.5 h. Cells were harvested by centrifugation, resuspended in lysis buffer (20 mM Tris at pH 7.5, 200 mM NaCl, and 200 mM 1,6-hexanediol), and lysed by passage through a French press. The soluble fraction was equilibrated with chitin beads for 12−24 h at 4 °C, and the beads were then washed extensively with wash buffer (100 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES) at pH 8.0, 200 mM NaCl, and 200 mM 1,6-hexanediol). The desired Smad2 thioester derivative was then released from the beads by treatment with cleavage buffer (wash buffer and 50 mM MESNA) for 24−48 h at 4 °C. Smad2 α-thioester was combined with excess SARA/SBD (1.5−3.0 M equiv), purified over an SP-FF ion-exchange column as the Smad2/SARA complex, and concentrated to 2−10 mg/mL in ligation buffer (100 mM HEPES at pH 8.0, 200 mM NaCl, and 50 mM MESNA). Note, the Smad2 α-thioester derivative was handled exclusively at 4 °C, and MESNA was present in all buffers to minimize hydrolysis.

For expressed protein ligation reactions, the appropriate C-terminal peptide was dissolved in ligation buffer and added in 4−10-fold excess to the Smad2 α-thioester. Ligation reactions were allowed to proceed at 4 °C with gentle agitation and followed by SDS−PAGE and ESMS until completion (12−72 h). TCEP was then added to a final concentration of 25 mM, and the reaction mixture was purified over a Superdex-75 size-exclusion column using a running buffer containing 20 mM Tris at pH 7.5, 100 mM NaCl, 10% (v/v) glycerol, and 2 mM dithiothreitol (DTT). Fractions containing the desired semisynthetic protein were pooled and then concentrated to 2−10 mg/mL. Protein identities were confirmed by ESMS as follows: Sm2-0p (found 25 364 ± 9 Da, expected 25 359 Da); Sm2-p465 (found 25 438 ± 6 Da, expected 25 439 Da); Sm2-p467 (found 25 447 ± 7 Da, expected 25 439 Da); Sm2-2p (found 25 517 ± 3 Da, expected 25 519 Da).

The concentration of an Sm2−2p calibration stock was determined by UV quantitation in 6 M GdmCl. The concentration of Smad2 in all other protein stocks was calculated relative to this calibration stock by integration of the Smad2 peak in RP-HPLC traces.

Gel Filtration. Analytical gel filtration was performed at 25 °C using a Zorbax-GF250 column (Agilent; 4 μm, 4.6 × 250 mm) using a flow rate of 0.5 mL/min and a running buffer containing 20 mM Tris at pH 7.5, 100 mM NaCl, and 0.1 mM DTT. Smad2 protein (2.5 μg) and the appropriate amount of SARA/SBD were brought to a 5-μL total injection volume with running buffer. Standard curves were obtained using the following protein standards: carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), and alcohol dehydrogenase (150 kDa).

Analytical Ultracentrifugation. Sedimentation equilibrium experiments were performed on a Beckman XL-I analytical ultracentrifuge at the State University of New York at Stony Brook (supported by the National Science Foundation Grant DBI-0215690) or at the Biophysical Instrumentation Facility (Massachusetts Institute of Technology) using 6-channel internal- or external-loading cells at 4 °C. Experiments were performed at three rotor speeds (8, 10, and 15k rpm) and typically three protein-loading concentrations (46, 15, and 5 μM) using a sample buffer containing 20 mM HEPES at pH 7.4, 100 mM NaCl, 0.1 mM ethylenediaminetetraacetic acid, 0.1 mM DTT, and 2 mM TCEP. Note, Sm2-p465 was analyzed at a single concentration (9 μM) at three rotor speeds (8, 10, and 15k rpm). Beckman Origin software was used to analyze the data. Partial specific volumes and solution densities were estimated using Sednterp (*26*). Data were globally fit to single ideal species; analysis of the residuals indicated a good fit for Sm2-0p (monomer), Sm2-2p (trimer), and free Sm2-p465 (trimer). A good fit could not be achieved for Sm2-p465 or Sm2-p467 with SARA, most likely because of the complex equilibria between free SARA, Sm2/SARA complex, and Sm2 trimer.

Kinase Assays on Protein Substrates. Assays were performed with 0.25 μg of kinase (TβRI-0P or TβRI-4P) and 2.5 μg of Smad2 protein substrate (10 μM) in 10 μL of kinase buffer (50 mM Tris at pH 7.5, 60 mM NaCl, 5 mM $MgCl_2$, 5 mM $MnCl_2$, and 2 mM DTT) at room temperature. Known amounts of purified SARA/SBD (5−30 μM) were also included into the mixture in certain experiments. All reactions were initiated by the addition of 20 or 100 μM ATP and

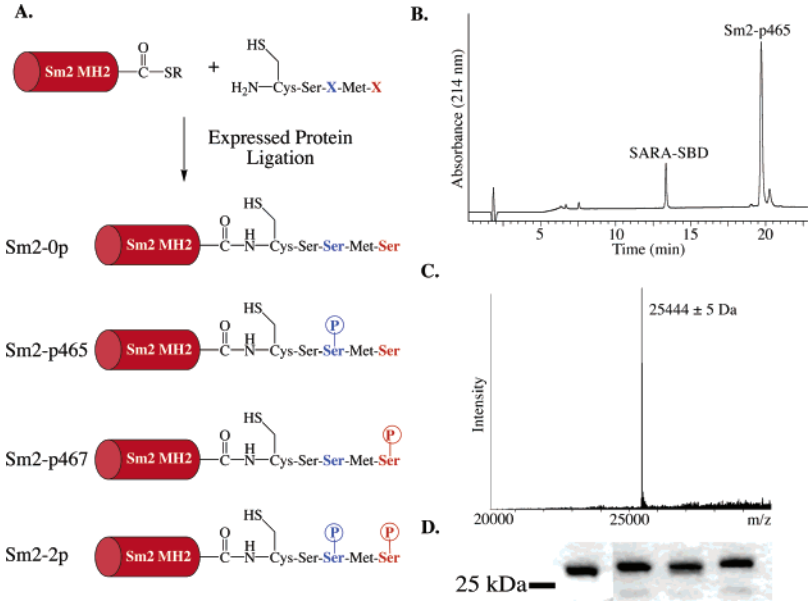


FIGURE 2: Semisynthesis of phosphorylated Smad proteins. (A) Schematic representation of the EPL approach used to prepare semisynthetic Smad proteins for this paper. (B) Typical RP-HPLC chromatogram of the purified Smad protein (Sm2-p465), gradient 0−73% B over 20 min. (C) Representative ESMS of Sm2-p465, isolated by RP-HPLC above (expected molecular weight 25 439). (D) Coomassie stained SDS−PAGE of the purified Sm2 proteins: Sm2-0p (lane 1), Sm2-p465 (lane 2), Sm2-p467 (lane 3), Sm2-2p (lane 4).

trace [γ -^{32}P]ATP and allowed to proceed for 25 min before being quenched by the addition of 9 μL of the reaction mixture to 4 μL of 6× SDS sample buffer. Reaction mixtures were then analyzed by SDS−PAGE (12% gel), stained with Coomassie blue, and dried. Gels were imaged on a Fujifilm BAS-2500 phosphoimager, and phosphorylated bands were quantified using ImageQuant. In addition, bands corresponding to Smad2 were excised and subjected to liquid scintillation counting. All experiments were performed at least two times.

All kinase, Smad2 substrate, and SARA/SBD concentrations were normalized by subjecting known volumes of each protein stock to RP-HPLC and integrating the corresponding protein peaks. These areas were compared and referenced to appropriate calibration stocks of known concentrations.

Filter Binding Assays. Assays were performed using 300 nM kinase (TβRI-0P) and 250 μM peptide (SmadPep-0p, SmadPep-p465) in kinase buffer. Reactions were initiated by adding 100 μM cold ATP and trace [γ -^{32}P]ATP and allowed to proceed at 30 °C. At specific timepoints, aliquots were withdrawn, quenched by adding 0.5% phosphoric acid, and absorbed in duplicates onto P-81 Whatman paper. Filter papers were washed five times in 0.5% phosphoric acid (10 mL per filter paper). Filter papers were then rinsed quickly in acetone, air-dried, and counted in a liquid scintillation counter. ^{32}P incorporation was averaged over the two filter papers per time point; the baseline of kinase autophosphorylation was subtracted, and the amount of phosphate incorporated into peptide in each 10-μL aliquot was calculated from the ratio of cold to hot ATP present in the assay mixture. In vitro kinase assays were performed in duplicates.

HPLC Assays. A total of 0.38 μg of TβRI-0P kinase, 500 μM peptide (SmadPep-0p, SmadPep-p465), and 200 μM ATP were added to 30 μL of kinase buffer at 30 °C. At specified timepoints, aliquots were withdrawn, quenched by the addition of 0.1% TFA, and analyzed by RP-HPLC (0−20% B gradient over 30 min). The product peak was isolated and analyzed by ESMS.

RESULTS

Effect of Individual Phosphate Groups on Oligomerization. Phosphorylation of an R-Smad drives the formation of the active hetero-oligomer with Smad4. This is reflected in the model system consisting of the MH2 domain of Smad2. The unphosphorylated domain acts as a monomer and can heterodimerize with SARA (*27*). However, upon phosphorylation at Ser465 and Ser467, this domain forms a homotrimer mediated by a network of connections involving the phosphoserine groups (Figure 1). Because the MH2 domains of Smad2 and Smad4 are highly conserved, the protein−protein interfaces within the Smad2 homotrimer are likely to be homologous to those in the Smad2−Smad4 heterocomplex (*10*). Thus, the Smad2 MH2 domain can serve as an appropriate model in studying the effect of the individual phosphate modifications on the Smad2 function.

Expressed protein ligation (EPL) (*28*) was used to prepare four variants of the Smad2−MH2 domain (Figure 2): Sm2-0p, with the C-terminal sequence CSSMS; Sm2-p465, with the C-terminal sequence CSpSMS; Sm2-p467, containing the sequence CSSMpS; and doubly phosphorylated Sm2-2p, with the C-terminal sequence CSpSMpS. Each semisynthetic protein was prepared by chemically ligating a purified recombinant protein α-thioester, corresponding to Smad2 residues 241−462, to a molar excess of the appropriate synthetic peptide. The Smad2 thioester fragment was purified in the complex with the SARA/SBD domain, which was found to greatly increase the stability of the protein thioester. Ligation reactions were allowed to proceed for 12−72 h, after which the semisynthetic proteins were purified by gel filtration chromatography. For each protein, purity was assessed at >90% by SDS−PAGE, RP-HPLC, and ESMS (parts B−D of Figure 2).

Analytical gel filtration was used as an initial assessment of the oligomerization state of the four constructs. On the basis of biophysical studies (*10*, *27*, *29*), Sm2-0p should elute as one peak corresponding to the molecular weight of the

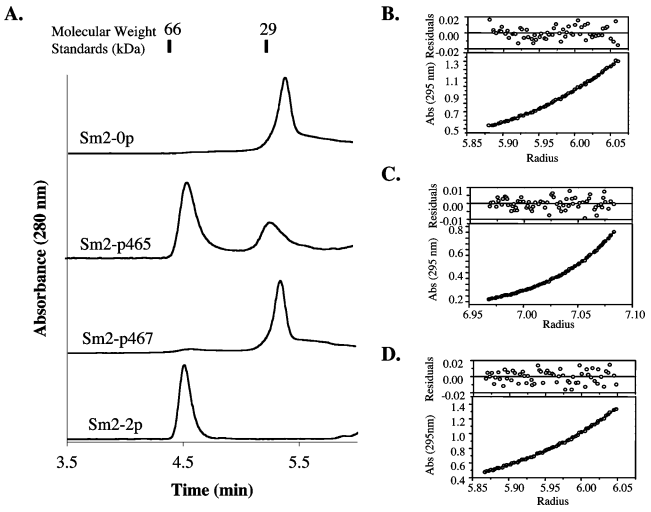


FIGURE 3: Oligomeric state of phospho-Smad constructs. (A) Analytical gel filtration analysis of purified Sm2-0p, Sm2-p465, Sm2-p467, and Sm2-2p. Analytical ultracentrifugation analysis of (B) Sm2-0p (single ideal species of molecular weight = 32 000 ± 2400, fit shown for 46 μM, 15k rpm); (C) Sm2-p465 (single ideal species of molecular weight = 72 000 ± 2700, fit shown for 9μM, 15k rpm); and (D) Sm2−2p (single ideal species of molecular weight = 73 000 ± 3900, fit shown for 46 μM, 10k rpm).

Smad/SARA complex (32 kDa), whereas Sm2-2p should elute as one peak with a retention time corresponding approximately to the trimer state of the domain (75 kDa). While these proteins did elute as single peaks, the retention times observed did not correspond precisely to the expected molecular weights (Figure 3). Therefore, sedimentation equilibrium analytical ultracentrifugation was used to further assess the oligomerization states of the species observed in the gel filtration studies. Accordingly, Sm2-0p behaved as a single ideal species with a molecular weight of 32 000 ± 2400, which corresponds to a 1:1 complex of Sm2-0p and SARA/SBD (henceforth referred to as "monomer"). Sm2-2p also behaved as a single ideal species but with a molecular weight of 73 000 ± 3900, which corresponds to a trimer of the single domain. In contrast to the behavior of Sm2-0p and Sm2-2p, which eluted as single peaks by analytical gel filtration, the two novel monophosphorylated variants eluted as a mixture of the two species (Figure 3). By comparison with the chromatograms of Sm2-0p and Sm2-2p, Sm2-p467 was primarily monomeric with ~5−10% apparent trimer, while Sm2-p465 eluted as a mixture of the two forms.

To better understand the oligomeric states of Sm2-p465 and Sm2-p467, further studies were undertaken. RP-HPLC analysis of each purified Sm2 protein revealed two peaks (Figure 2B), one corresponding to the Smad MH2 domain and one corresponding to SARA/SBD. Because SARA/SBD is known to bind preferentially to the monomeric form of R-Smad over the trimeric form (*29*), the amount of SARA/SBD remaining in the semisynthetic protein preparation after purification is a measure of the oligomerization state of the Smad2 analogue. Indeed, the amount of SARA/SBD varied significantly between the four phosphorylation variants; the ratio of the two peak areas (Smad/SARA) in the RP-HPLC chromatograms was 4.1 ± 0.5 for Sm2-0p[2], 4.6 ± 0.5 for

[2] Equimolar amounts of Smad MH2 and SARA are equivalent to a peak area ratio of 4.0, calculated from an absorbance at 214 nm.

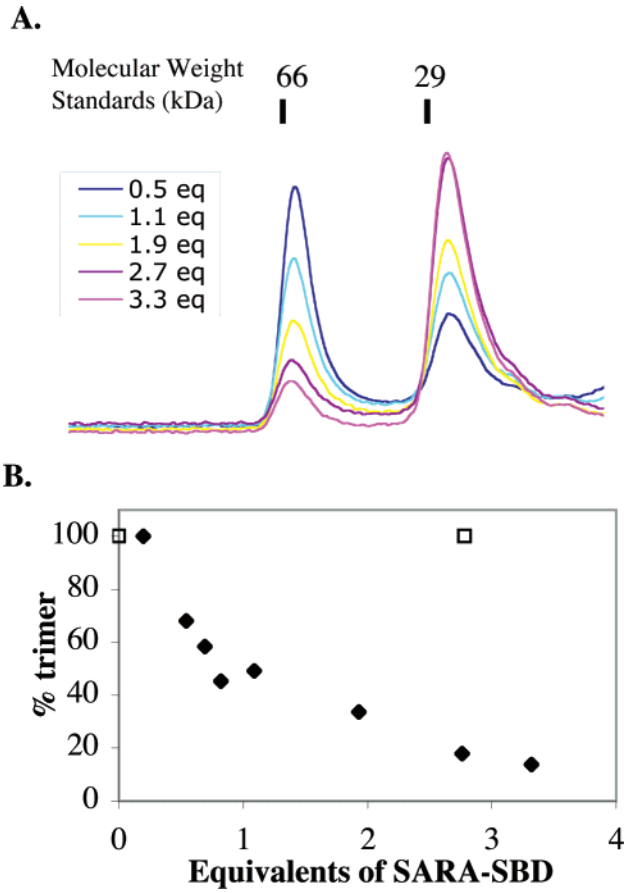


FIGURE 4: SARA/SBD disrupts the Sm2-p465 trimer. (A) Chromatographs of the gel filtration analysis of Sm2-p465 in the presence of increasing amounts of SARA/SBD. Absorbance was monitored at 280 nm. (B) SARA/SBD was added to a solution of Sm2-p465 (♦) or Sm2-2p (□), and the amount of Sm2 homotrimer was assessed by analytical gel filtration. An excess of up to 19-fold of SARA/SBD (data not shown) did not affect the gel filtration profile of Sm2-2p. The percentage of trimer was calculated as (area of trimer peak)/[(area of trimer peak + area of monomer peak) − (area of SARA/SBD peak)].

Sm2-p467 (both primarily monomeric), and 14 ± 9 for Sm2-p465, while Sm2-2p (trimeric) contained only trace amounts of SARA. Thus, there appears to be a strong correlation between the amount of associated SARA/SBD in a given protein and its oligomerization state as observed by analytical gel filtration. Interestingly, the amount of SARA/SBD associated with Sm2-p465 varied significantly between preparations of the protein. Moreover, a direct correlation was observed between the Smad/SARA ratio and the oligomeric state of Sm2-p465 as assessed by gel filtration; preparations with the least SARA appeared purely trimeric both by gel filtration and analytical ultracentrifugation (data could be fit to a single species of 72 000 ± 2700 Da), while preparations with the most SARA contained more of the monomeric form of the protein (up to 55% monomer). When more purified SARA/SBD was added to the Sm2-p465 sample, the protein could be driven further toward the monomer form, up to 86% with an excess of SARA/SBD (Figure 4). Interestingly, while addition of a 3-fold excess of SARA/SBD was sufficient to shift the Sm2-p465 equilibrium nearly completely to monomer, even a 19-fold excess of SARA/SBD did not appear to disrupt the Sm2-2p trimer in gel filtration studies (Figure 4B), indicating that the trimer formed by the doubly

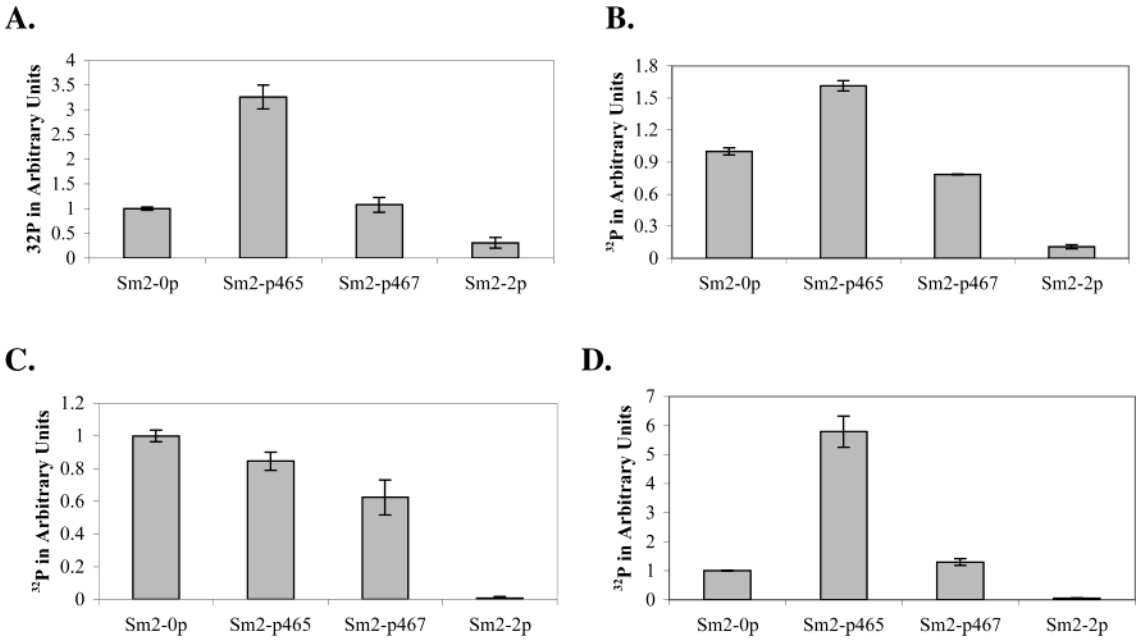


FIGURE 5: In vitro phosphorylation of semisynthetic Smad substrates by TβRI. (A) Phosphorylation of Sm2 substrates by TβRI-0P. Each bar represents the average of three experiments ± standard deviation (SD). Data were normalized to set Sm2-0p = 1.0. (B) Phosphorylation of Sm2 substrates by TβRI-4P. Each bar represents the average of two experiments ± SD. Data were normalized to set Sm2-0p = 1.0. (C) Phosphorylation of Sm2 substrates by TβRI-4P. In this case, the homotrimeric species of Sm2-p465 was used in the assay. Each bar represents the average of three experiments ± SD, and data were normalized to set Sm2-0p = 1.0. (D) Phosphorylation of Sm2 substrates by TβRI-4P in the presence of 3 equiv of SARA/SBD. Under these conditions, Sm2-p465 is ~85% monomer. Each bar represents the average of two experiments ± SD, and data were normalized to set Sm2-0p = 1.0.

phosphorylated protein is more stable than that formed by Sm2-p465.

Our studies show that both Sm2-0p and Sm2-p467 display a preference for binding SARA/SBD over self-association. In contrast, Sm2-2p forms a stable trimer to the exclusion of SARA/SBD. Sm2-p465, however, is balanced between these two extremes. The increased propensity for self-association of Sm2-p465 over Sm2-p467 may be explained by examining the hydrogen-bonding networks observed in the trimer crystal structure (*10*) (Figure 1B). The phosphate group of pSer465 is involved in two intersubunit hydrogen bonds and three more to well-ordered water molecules in the pocket. The phosphate group of pSer467 is involved in only two intersubunit hydrogen bonds, but the C-terminal carboxylate group of this residue contributes three hydrogen bonds across the protein interface, as well as one more to an ordered water. Because the hydrogen bonds mediated by the carboxylate may well exist even in the absence of a phosphate group at Ser467, the intersubunit contacts are thus dominated by pSer465. Thus, structural data supports our observation that pSer465 is sufficient for trimer formation, but pSer467 is necessary to increase the stability of the complex.

Phosphorylation of Smad2 Analogues by Activated TβRI. Phosphorylation of Smad2 at Ser465 and Ser467 must proceed through two successive phosphorylation reactions, although it is not known whether a particular order is preferred. To further probe this process, we studied the in vitro phosphorylation of the semisynthetic Smad2 phosphovariants, Sm2-p465 and Sm2-p467, which are by definition the two potential intermediates in Smad2 phosphorylation. Both the nonphosphorylated TβRI kinase domain (TβRI-0P) and the semisynthetic variant (TβRI-4P), containing four phosphate groups in the GS region, which serve to activate the enzyme (*7*), were used in these studies. This may represent the first time a semisynthetic enzyme has been studied with semisynthetic protein substrates.

Sm2-p465 was a better substrate than Sm2-0p, despite having only one serine site available for phosphorylation as compared to the two sites available for Sm2-0p (Figure 5A). Sm2-p467, which contains a phosphoserine two residues downstream of the available phosphorylation site, does not show a significantly higher rate of phosphorylation than Sm2-0p. As expected, Sm2-2p, in which both of the native phosphorylation sites are prephosphorylated, is a very poor substrate, indicating that phosphorylation of our semisynthetic substrates is specific to Ser465 and Ser467. Thus, the second phosphorylation reaction appears to be primed by the placement of a phosphate group in the substrate P − 2 site but not the P + 2 site. While this basic trend was reproduced for both TβRI-0P (Figure 5A) and the activated TβRI-4P (Figure 5B) enzymes, the size of the difference in phosphorylation of Sm2-p465 compared to Sm2-0p and Sm2-p467 was larger for TβRI-0P. This is consistent with the fact that TβRI-4P is 5−10-fold more active than TβRI-0P toward nonphosphorylated Smad2 (*7*, *24*), an increase in activity that has been attributed to the interaction between the phosphorylated GS region and the loop/strand pocket. However, the fact that Sm2-p465 is still preferentially phosphorylated by the activated TβRI-4P suggests that the increase in binding affinity afforded by phosphorylation of Ser465 is comparable to the increase resulting from the GS region/loop strand pocket interaction.

It is perhaps surprising that Sm2-p465 is such a good substrate for TβRI, considering the observed tendency of this protein to form a homotrimer. In the trimer state, the

C-terminal tail of the protein should be locked into intersubunit interactions, reducing the accessibility of this region to the active site of the kinase. In addition, formation of the homotrimer involves residues in the loop/strand pocket (*10*, *11*), a region of the MH2 domain that is also thought to interact with the L45 loop and phosphorylated GS region of TβRI (*6*, *7*). Thus, homotrimer formation would be expected to reduce the affinity of Sm2-p465 for the activated receptor and therefore make it a poorer substrate, which is exactly the opposite of what we observe (parts A and B of Figure 5). As demonstrated above (Figure 4), the amount of SARA/SBD present can affect the monomer/trimer ratio of Sm2-p465. This provides a tool by which to systematically study the affect of substrate trimerization on phosphorylation. Accordingly, we performed the kinase assay under conditions that altered the monomer/trimer ratio of Sm2-p465. The purified homotrimer form of Sm2-p465 was, if anything, a slightly worse substrate for TβRI-4P than Sm2-0p (Figure 5C) in the absence of additional SARA/SBD. However, when three equivalents of SARA/SBD were included to maximize the amount of the Sm2-p465/SARA/SBD complex present, Sm2-p465 (85% monomer) was a 5−6-fold better substrate than Sm2-0p (Figure 5D). All of the above data are consistent with a mechanism of substrate release from the receptor kinase complex driven by Smad oligomerization.

Development of a Peptide Substrate for TβRI. Previous studies have shown that Smad2 recruitment to TβRI involves protein−protein interactions remote from the kinase active site (*6*, *7*). However, the results above suggest that phosphorylation of Ser465 also contributes to substrate−kinase recognition. To directly test this idea, we synthesized two potential substrate peptides: the nonphosphorylated SmadPep-0p (KMGSPSVRCSSMS) and SmadPep-p465 (KMGSPSVRCSpSMS), containing the 12 C-terminal amino acids [including the last visible residue in the crystal structure (*27*)] and an N-terminal lysine.

SmadPep-0p was a poor substrate for TβRI-0P, showing very little phosphorylation above the background in the in vitro kinase assays. However, the singly phosphorylated SmadPep-p465 was a significantly better substrate for TβRI as shown in the reaction time courses in Figure 6A. To further characterize the phosphorylation product, the reaction was also followed by RP-HPLC. Phosphorylation of SmadPep-0p was so sufficiently poor that no product peaks could be isolated. However, one single product peak appeared over the course of the SmadPep-p465 phosphorylation reaction (Figure 6B). This material was isolated and analyzed by ESMS, which afforded a mass consistent with the addition of a single phosphate group to the peptide (measured *m*/*z* 759.0 M + 2H, expected 1516.5 M + H). Given our studies on the semisynthetic Smad2 proteins, it is reasonable to assume that this phosphate group is being added to the C-terminal serine in the peptide substrate.

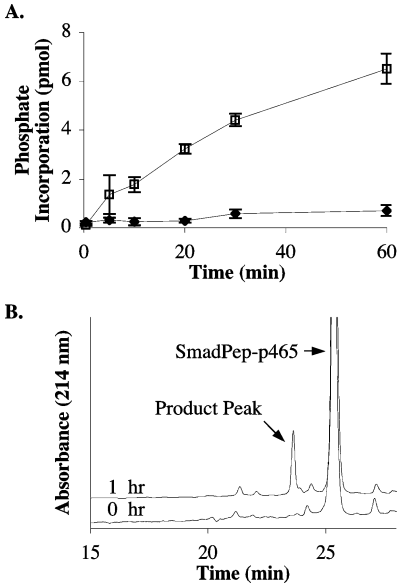


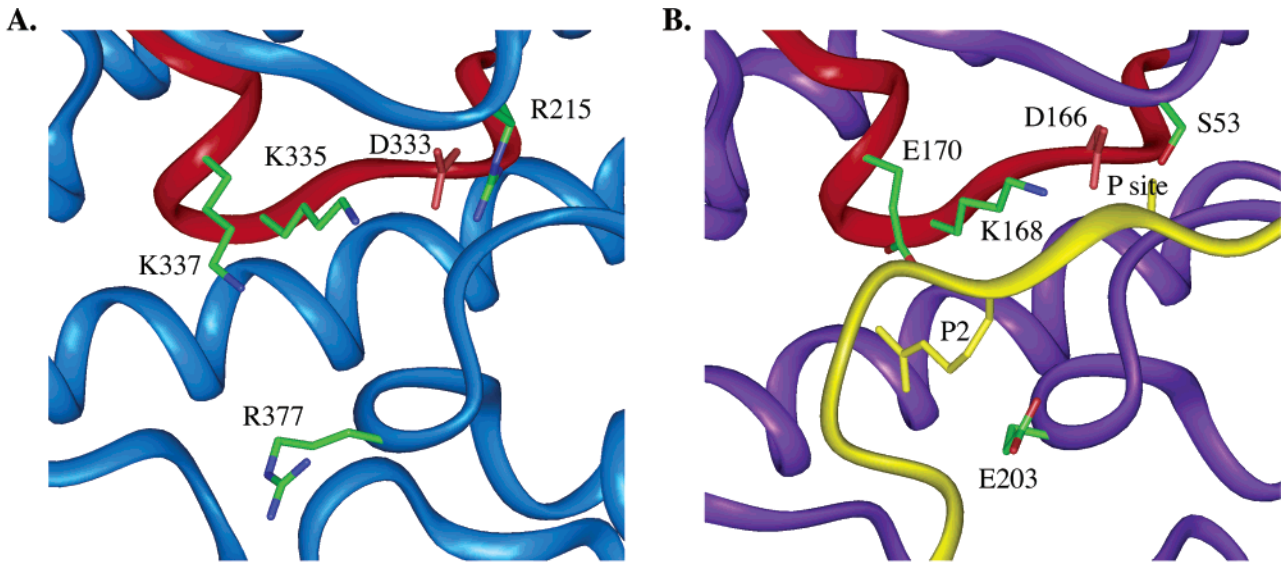
FIGURE 6: In vitro phosphorylation of peptide substrates by TβRI-0P. (A) Phosphorylation of 250 μM SmadPep-0p (♦) and SmadPep-p465 (□) substrates by TβRI-0P. Each data point is the average of two experiments ± SD. (B) RP-HPLC spectrum depicting the accumulation of the phosphorylated product peptide. The peak at 23.6 min appeared over the course of the experiment. ESMS confirmed that the mass of the product at this retention time was consistent with the addition of a single phosphate group to SmadPep-p465.

DISCUSSION

Expressed protein ligation has been used to probe the contribution of the individual phosphate groups to Smad2 oligomerization and phosphorylation by TβRI. Our results indicate that pSer465 provides the predominant driving force for formation of the Smad2 oligomer but that the affinity provided by pSer467 is also essential to the formation of a stable complex. Our data also suggest that SARA plays a role in preventing premature release of monophosphorylated Smad2 from the receptor kinase. Additionally, pSer465 appears to improve the affinity of the C-terminal activation motif for the active site of the receptor kinase.

Phosphorylation at both Ser465 and Ser467 is required for full Smad2-mediated signaling in vivo (*17*, *18*). Thus, it is important to prevent premature release of singly phosphorylated Smad2 from the receptor kinase complex. Our studies suggest a two-pronged solution to this problem. First, release of phosphorylated Smad2 from the receptor complex is thought to be triggered by oligomer formation. We have demonstrated that a single phosphate group at either of the two serine sites is insufficient to form a stable trimer. Although each of the monophosphorylated intermediates displays some propensity for trimerization, Sm2-p465 to a much greater extent than Sm2-p467, both phosphate groups are required to complete Smad oligomerization. Additionally, SARA is presumably present at high local concentrations in the receptor−substrate complex (*8*). If Ser467 is phosphorylated first, the protein remains predominantly monomeric and bound to the receptor until Ser465 can be phosphorylated and the protein can be released. If, however, Ser465 is phosphorylated first, the high local concentrations of SARA could disrupt formation of the trimer as demonstrated above and maintain the receptor-bound form of the monophosphorylated intermediate. Because SARA does not disrupt oligomerization of the doubly phosphorylated Smad2, SARA would not prevent the release of the final product from the kinase. Thus, a complex set of equilibria between the varied binding partners determines the state of the receptor−substrate complex.

The second potential factor precluding the production of singly phosphorylated R-Smad proteins is suggested by our



C.

	215	333 335 337	377
Type I			
TβR-I	K G R F G	A H R D L K S K N I L	G T K R Y M
BMPR-I	K G R Y G	A H R D L K S K N I L	G T K R Y M
ALK1	K G R Y G	A H R D F K S R N V L	G T K R Y M
TKV	K G R Y G	A H R D I K S K N I L	G T R R Y M
SMA-1	Q G R Y G	A H R D I K S K N I I	G T K R Y M
Type II			
TβR-II	K G R F A	V H R D L N S S N I L	G T A R Y M
ACTR-IIB	R G R F G	A H R D F K S K N V L	G T R R Y M
BMPR-II	R G R Y G	S H R D L N S R N V L	G T I R Y M
Other			
CAMK-IIα	K G A F S	V H R D L K P E N L L	G T P G Y L
PKA	T G S F G	I Y R D L K P E N L L	G T P E Y L
PKCα	K G S F G	I Y R D L K L D N V M	G T P D Y I

FIGURE 7: Putative determinants of substrate recognition in the active site of TβRI. (A) Depiction of the active site of TβRI (PDB 1IAS). The catalytic loop is red, and the catalytic aspartate (D333) is illustrated. Side chains that are potentially involved in substrate recognition are depicted in green. (B) For comparison, the corresponding region of PKA in the complex with PKI is shown (PDB 1FMO). The P site alanine and P2 site arginine are illustrated; in Sm2-p465, these residues would be Ser and pSer, respectively. (C) Alignment of the kinase sequences, illustrating that (in TβRI sequence numbering) R215, K335, K337, and R377 are conserved through TGF-β family receptors (to a greater extent in type I than in type II) but not in other Ser/Thr protein kinases.

enzymatic studies of the Smad2 substrates. Simply stated, Sm2-p465 is phosphorylated more rapidly than Sm2-0p or Sm2-p467, the variants that are less likely to undergo oligomerization-driven release from the kinase. Thus, if Ser465 is phosphorylated first, the second phosphorylation step will occur relatively quickly, producing the fully functional form of the protein and eliminating the risk of a single-phosphate release.

Our finding that prephosphorylation of Smad2 at Ser465 significantly increases the rate of phosphorylation at Ser467 (but not vice versa) suggests that there may be specific recognition determinants within the kinase for the monophosphorylated intermediate. This contention is also supported by the peptide substrate studies. SmadPep-p465 differs from SmadPep-0p in the addition of a single phosphate group. The significant enhancement of SmadPep-p465 relative to SmadPep-0p as a TβRI kinase substrate implies that pSer465 participates in local interactions in or around the kinase active site. Analysis of the TβRI crystal structure in the complex with inhibitor NPC-30345 (*7*) reveals a cluster of basic residues, consisting of Lys335, Lys337, and Arg377, close to the active site of the kinase (Figure 7A). On the basis of alignment of the TβRI structure with that of protein kinase A (PKA) in the complex with a substrate-based inhibitor (*30*) (aligned on the catalytic segments), these basic residues appear ideally positioned to engage a phosphate group in the P2 position of the substrate (Figure 7B). Interestingly, this cluster is highly conserved among TGFβ type I receptors, which phosphorylate Smad family substrates, but not as much among TGFβ type II receptors or other Ser/Thr kinases (*31*) (Figure 7C). It is possible that this conserved basic patch forms a phosphate binding pocket, which fits pSer465. Additionally, the side chain of Arg215, which is also conserved among TGFβ receptors, is positioned very close to the carbonyl of the P site residue, which corresponds to the C-terminal carboxylate group of Ser467 in a Smad substrate docked with pSer465 in the P2 position. When taken together, these interactions could provide the enhanced affinity of Sm2-p465 and SmadPep-p465 observed in our studies. However, these intriguing possibilities will

require additional mutagenesis and/or structural studies in order to be confirmed.

TβRI is an attractive target for the development of small-molecule therapeutics because of the link between increased levels of TGFβ signaling and the progression of some human diseases, including certain cancers and hypertension (*1*, *2*). Currently, groups interested in the development of such inhibitors use assays based on kinase autophosphorylation or purified protein substrates to screen compounds (*32*, *33*). A peptide substrate for TβRI with the nonphysiological sequence, KKKKKK(S/T)XXX, has also been identified using a library-based approach (*34*). However, only a modest enhancement of phosphorylation over the background was associated with this peptide. Thus, the identification in this paper of the improved peptide substrate SmadPep-p465 could be of considerable value for future drug-screening initiatives, as well as for basic enzymology studies on the kinase.

ACKNOWLEDGMENT

We thank Elyse Blum, Michael Hahn, Dr. Jean-Philippe Pellois, and Erik Dill for advice and technical help.

REFERENCES

1. Massagué, J. (1998) TGF-β signal transduction, *Annu. Rev. Biochem. 67*, 753−91.
2. Siegel, P. M., and Massagué, J. (2003) Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer, *Nat. Rev. Cancer 3*, 807−820.
3. Shi, Y., and Massagué, J. (2003) Mechanisms of TGF-β signaling from cell membrane to the nucleus, *Cell 113*, 685−700.
4. Lo, R. S., Chen, Y.-G., Shi, Y., Pavletich, N. P., and Massagué, J. (1998) The L3 loop: a structural motif determining specific interactions between SMAD proteins and TGF-β receptors, *EMBO J. 17*, 996−1005.
5. Feng, X.-H., and Derynck, R. (1997) A kinase subdomain of transforming growth factor-β (TGF-β) type I receptor determines the TGF-β intracellular signaling specificity, *EMBO J. 16*, 3912−3923.
6. Chen, Y.-G., Hata, A., Lo, R. S., Wotton, D., Shi, Y., Pavletich, N., and Massagué, J. (1998) Determinants of specificity in TGF-β signal transduction, *Genes Dev. 12*, 2144−2152.
7. Huse, M., Muir, T. W., Xu, L., Chen, Y.-G., Kuriyan, J., and Massagué, J. (2001) The TGFβ receptor activation process: an inhibitor- to substrate-binding switch, *Mol. Cell 8*, 671−682.
8. 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 TGFβ receptor, *Cell 95*, 779−791.
9. Massagué, J. (2000) How cells read TGF-β signals, *Nat. Rev. Mol. Cell Bio. 1*, 169−178.
10. Wu, J.-W., Hu, M., Chai, J., Seoane, J., Huse, M., Li, C., Rigotti, D. J., Kyin, S., Muir, T. W., Fairman, R., Massagué, J., and Shi, Y. (2001) 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.
11. Qin, B. Y., Chacko, B. M., Lam, S. S., de Caestecker, M. P., Correia, J. J., and Lin, K. (2001) Structural basis of Smad1 activation by receptor kinase phosphorylation, *Mol. Cell 8*, 1303−1312.
12. Inman, G. J., and Hill, C. S. (2002) Stoichiometry of active Smad-transcription factor complexes on DNA, *J. Biol. Chem. 277*, 51008−51016.
13. 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.
14. Wu, J.-W., Fairman, R., Penry, J., and Shi, Y. (2001) Formation of a stable heterodimer between Smad2 and Smad4, *J. Biol. Chem. 276*, 20688−20694.
15. 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.
16. 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.
17. Abdollah, S., Macías-Silva, M., Tsukazaki, T., Hayashi, H., Attisano, L., and Wrana, J. L. (1997) TβRI phosphorylation of Smad2 on Ser^{465} and Ser^{467} is required for Smad2−Smad4 complex formation and signaling, *J. Biol. Chem. 44*, 27678−27685.
18. Souchelnystskyi, S., Tamki, K., Engström, U., Wernstedt, C., ten Dijke, P., and Heldin, C.-H. (1997) Phosphorylation of Ser^{465} and Ser^{467} in the C terminus of Smad2 mediates interaction with Smad4 and is required for transforming growth factor-β signaling, *J. Biol. Chem. 272*, 28107−28115.
19. Albericio, F., and Caprino, L. A. (1997) Coupling reagents and activation, *Methods Enzymol. 289*, 104−126.
20. Han, Y., Albericio, F., and Barany, G. (1997) Occurrence and minimization of cysteine racemization during stepwise solid-phase peptide synthesis, *J. Org. Chem. 62*, 4307−4312.
21. Ludolph, B., Eisele, F., and Waldmann, H. (2002) Solid-phase synthesis of lipidated peptides, *J. Am. Chem. Soc. 124*, 5954−5955.
22. Guy, C. A., and Fields, G. B. (1997) Cleavage and deprotection of resin-bound peptides following synthesis by Fmoc chemistry, *Methods Enzymol. 289*, 67−83.
23. Huse, M., Chen, Y.-G., Massagué, J., and Kuriyan, J. (1999) Crystal structure of the cytoplasmic domain of the type I TGFβ receptor in complex with FKBP12, *Cell 96*, 425−436.
24. Huse, M., Holford, M. N., Kuriyan, J., and Muir, T. W. (2000) Semisynthesis of hyperphosphorylated type I TGFβ receptor: addressing the mechanism of kinase activation, *J. Am. Chem. Soc. 122*, 8337−8338.
25. Flavell, R. R., Huse, M., Goger, M., Trester-Zedlitz, M., Kuriyan, J., and Muir, T. W. (2002) Efficient semisynthesis of a tetraphosphorylated analogue of the type I TGFβ receptor, *Org. Lett. 4*, 165−168.
26. Laue, T., Shaw, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-Aided Interpretation of Analytical Sedimentation Data for Proteins, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge, U.K.
27. 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.
28. Muir, T. W. (2003) Semisynthesis of proteins by expressed protein ligation, *Annu. Rev. Biochem. 72*, 249−289.
29. Qin, B. Y., Lam, S. S., Correia, J. J., and Lin, K. (2002) Smad3 allostery links TGF-β receptor kinase activation to transcriptional control, *Genes Dev. 16*, 1950−1963.
30. Narayana, N., Cox, S., Shaltiel, S., Taylor, S. S., and Xuong, N.-h. (1997) Crystal structure of a polyhistidine-tagged recombinant catalytic subunit of cAMP-dependent protein kinase complexed with the peptide inhibitor PKI(5−24) and adenosine, *Biochemistry 36*, 4438−4448.
31. Hanks, S. K., and Hunter, T. (1995) The eukaryotic protein kinase superfamily, in *The Protein Kinase Fact Book: Protein-Serine Kinases* (Hardie, G., and Hanks, S., Eds.) pp 7−47, Academic Press Inc., San Diego, CA.
32. Yakymovych, I., Engström, U., Grimsby, S., Heldin, C.-H., and Souchelnytskyi, S. (2002) Inhibition of transforming growth factor-β signaling by low molecular weight compounds interfering with ATP- or substrate-binding sites of the TGFβ type I receptor kinase, *Biochemistry 41*, 11000−11007.
33. Laping, N. J., Grygielko, E., Mathur, A., Butter, S., Bomberger, J., Tweed, C., Martin, W., Fornwald, J., Lehr, R., Harling, J., Gaster, L., Callahan, J. F., and Olson, B. A. (2002) Inhibition of transforming growth factor (TGF)-β1-induced extracellular matrix with a novel inhibitor of the TGF-β type I receptor kinase activity: SB-431542, *Mol. Pharmacol. 62*, 58−64.
34. Luo, K., Zhou, P., and Lodish, H. F. (1995) The specificity of the transforming growth factor β receptor kinases determined by a spatially addressable peptide library, *Proc. Natl. Acad. Sci. U.S.A. 92*, 11761−11765.

BI0498407