Sedimentation Studies Reveal a Direct Role of Phosphorylation in Smad3:Smad4 Homo- and Hetero-Trimerization†

John J. Correia,* Benoy M. Chacko, Suvana S. Lam, and Kai Lin

Department of Biochemistry, University of Mississippi Medical Center, Jackson, Mississippi 39216, and Department of Pharmacology and Molecular Toxicology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

Received August 15, 2000; Revised Manuscript Received November 16, 2000

ABSTRACT: SMAD proteins are known to oligomerize and hetero-associate during their activation and translocation to the nucleus for transcriptional control. Analytical ultracentrifuge studies on Smad3 and Smad4 protein constructs are presented to clarify the model of homo- and hetero-oligomerization and the role of phosphorylation in the activation process. These constructs all exhibit a tendency to form disulfide cross-linked aggregates, primarily dimers, and a strong reducing agent, TCEP, was found to be required to determine the best estimates for reversible association models and equilibrium constants. A Smad4 construct, S4AF, consisting of the middle linker (L) domain and the C-terminal (C) domain, is shown to be a monomer, while a Smad3 construct, S3LC, consisting of the LC domains, is shown to form a trimer with an affinity $K_3 = (1.2-3.1) \times 10^9$ M^{-2}. A Smad3 construct that mimics phosphorylation at the C-terminal target sequence, S3LC(3E), has 17−35-fold enhanced ability to form trimer over that of the wild-type construct, S3LC. S4AF associates with either S3LC or S3LC(3E) to form a hetero-trimer. In each case, the hetero-trimer is favored over the formation of the homo-trimer. Despite high sequence homology between Smad3 and Smad4, a chimeric Smad4 construct with an engineered Smad3 C-terminal pseudo-phosphorylation sequence, S4AF(3E), shows no tendency to form trimer. This suggests a Smad4-specific sequence insert inhibits homo-trimer formation, or other domains or sequences in S3LC are required in addition to the target sequence to mediate the formation of trimer. These results represent a direct molecular measure of the importance of hetero-trimerization and phosphorylation in the TGF-β-activated Smad protein signal transduction process.

The TGF-β superfamily of growth factors activates a family of signal transducers and transcriptional regulators known as Smad proteins. The receptor-mediated Smad proteins (R-Smads)[1] are phosphorylated by the transmembrane receptor serine/threonine kinases at conserved C-terminal serine residues. Smad4 serves as a signaling partner of R-Smads by forming active hetero-oligomers that enter the nucleus to modulate transcription through interaction with co-activators and co-repressors. It has been suggested that R-Smads are trimeric complexes, and upon phosphorylation recruit trimeric Smad4 to form hexameric oligomers (*1, 2*). The hexameric model is based upon the observation that the C-domain of Smad4 crystallized as a trimer. Sequence conservation suggests that R-Smads would also form trimers. Another model suggests that R-Smads and Smad4 are monomers, and interact to form hetero-trimeric complex upon activation (*3*). The hetero-trimer model is based upon co-immunoprecipitation, size exclusion chromatography, and cross-linking studies using lysates from mammalian cells transfected with epitope-tagged Smad proteins (*3*). Based upon these transfection studies, Smad4 and R-Smads are monomers in the basal state, and hetero-oligomerize upon activation. The hetero-oligomer is smaller than the homo-oligomer in size exclusion experiments, an observation that is inconsistent with the hexamer model. There has been no direct molecular investigation of the energetics of Smad protein oligomerization, the role of phosphorylation in stimulating oligomerization, or a determination of the stoichiometry and oligomer size of the R-Smad:Smad4 hetero-complex by a method that measures absolute molecular weight.

Here we present sedimentation studies on Smad3 (a member of the R-Smad family) and Smad4 constructs that verify the size of R-Smad homo- and hetero-oligomers to be trimers, and establish the importance of phosphorylation in stimulating both homo- and hetero-trimer formation. The Smad3 construct, S3LC (residues 145−424 of smad3), consists of the middle linker domain, the C-terminal oligomerization domain, and the $SSVS^{424}$ phosphorylation sequence at the C-terminus. The Smad4 construct, S4AF (residues 273−552 of smad4), consists of the middle Smad4

† The work is supported in part by a Sidney Kimmel Scholar award (K.L.) and the UMMC Analytical Ultracentrifuge Facility (J.J.C.). This is UMMC AUF publication no. 0027.

* Address correspondence to this author at the University of Mississippi Medical Center, Department of Biochemistry, 2500 N. State St., Jackson, MS 39216. Phone (601) 984-1522; fax (601) 984-1501; email jcorreia@biochem.umsmed.edu.

[1] Abbreviations: DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; R-Smad, receptor-mediated Smad proteins; S4AF, Smad4 consisting of the middle L and C domains; S4AF(3E), S4AF with a C-terminal substitution of the CEEVE phosphorylation mimic sequence; S3LC, Smad3 consisting of the LC domains; S3LC(3E), S3LC with a CEEVE pseudo-phosphorylation sequence; TCEP, tris(2-carboxyethyl)phosphine hydrochoride.

10.1021/bi0019343 CCC: $20.00
Published on Web 01/10/2001

activation domain (SAD) and the C-terminal oligomerization domain. These constructs are used because the full-length constructs and the C-terminal oligomerization domains alone are prone to significant irreversible aggregation. The effect of phosphorylation was investigated using the mutant S3LC-(3E) in which the three potential serine phosphorylation sites within the target sequence of S3LC were mutated to glutamic acids. This mutation mimics physiological phosphorylation in a signaling assay (*4*). To further investigate the function of this target sequence, a chimeric Smad4 construct was constructed where the C-terminal sequence was replaced with the Smad3 3E pseudo-phosphorylation sequence, referred to as S4AF(3E). These constructs allow us to investigate the ability of Smad3 and Smad4 to homo-trimerize, when studied alone, and hetero-trimerize, when studied in combination, and the stimulatory role of phosphorylation for each reaction. Our studies reveal a direct function of R-Smad C-terminal phosphorylation in promoting R-Smad homo-trimerization as well as R-Smad:Smad4 hetero-trimerization.

MATERIALS AND METHODS

Reagents. TCEP was purchased from Pierce; Hepes was from Fisher; NaCl was from EM Science; EDTA was from Sigma; DTT was from BioRad.

Smad Constructs. The S3LC, S3LC(3E), and S4AF constructs were generated as described in ref *4*. The S4AF-(3E) construct was generated using cassette mutagenesis, in which the Smad4 C-terminal sequence PIADPQPLD^{525} was replaced by the pseudo-phosphorylated Smad3 C-terminal sequence GSPSIRCEEVE. To facilitate cassette mutagenesis, a unique *Nco*I site was created between codons 542 and 544 of Smad4 cDNA by silencing mutation. A double-stranded DNA cassette encoding the Smad3 C-terminal pseudo-phosphorylation sequence and the stop codon were created by annealing of two oligonucleotides, which produces flanking *Nco*I and *Eco*RI sites. The cassette was used to replace the *Nco*I/*Eco*RI fragments of the wild-type construct.

Analytical Ultracentrifugation. Experiments were conducted at appropriate speeds in a Beckman Optima XLA analytical ultracentrifuge equipped with absorbance optics and an An60Ti rotor. Samples were dialyzed into the sedimentation buffer conditions, aliquoted, and frozen. Buffer conditions were 20 mM Hepes, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, pH 7.4 at 24.7 °C. For each construct, experiments were repeated in the presence of 2 mM TCEP to suppress nonspecific disulfide cross-linking and aggregation. Neutralized 500 mM TCEP was diluted directly into each sample and dialysate reference buffer to achieve the desired concentration. Temperature was calibrated by the method of Liu and Stafford (*5*). Sedimentation velocity experiments were performed at 42K rpm in charcoal-filled Epon double-sector centerpieces. Velocity data were collected at 280 nm, and occasionally 250 nm, at a spacing of 0.002 cm with no averaging in a continuous scan mode. Velocity data were analyzed[2] using DCDT+, version 1.12 (*6*, *7*), and, where appropriate, SVEDBERG, version 6.37 (*8*, *9*). For presentation purposes in each figure panel, the range of scans analyzed with DCDT+ were adjusted to correspond to the same extent of centrifugation, as determined by a time range test for broadening available in DCDT+. This ensures the sedimentation coefficient distribution or $g(s)$ curves have similar width/height characteristics and best convey visually the concentration-dependent trend of the data. The reported weight average sedimentation coefficient values ($\bar{s}_{20,w}$) obtained from DCDT+ are calculated by a weighted integration over the entire range of sedimentation coefficients covered by the $g(s)$ distribution (*10*) and corrected for the solution density and viscosity (*11*). The weight average sedimentation coefficient values were fit to both a monomer−trimer and a monomer−dimer−trimer model with Fitall (MTR Software, Toronto, Canada) as described (*10*). In the monomer−trimer model, the overall K_3 corresponds to an isodesmic polymer growth where each subunit addition is of equal affinity K, and thus $K_3 = K^2$. In this model, only the monomer and the n-mer species are included in the $\bar{s}_{20,w}$ calculation. In the monomer−dimer−trimer model, the overall affinity K_3 equals $K^2\gamma$, where K corresponds to isodesmic polymer growth and closure of the trimer structure incorporates cooperativity, γ (*12*, *13*). Based upon their similar size and sequence, we assumed the S3LC and the S3LC(3E) monomers have similar s values to that of S4AF, 2.46 $S_{20,w}$ (2.65 S_{app}). The trimer s value was determined to be 4.9 $S_{20,w}$ (5.3 S_{app}) by extrapolation of a $1/c$ vs $1/s$ plot utilizing peak positions from the $g(s)$ patterns.[3] Note, this is slightly smaller than the largest s value for trimer expected by an $n^{2/3}$ or constant axial ratio assumption, i.e., 5.12 $S_{20,w}$. In the 1-2-3 model, the s value for dimer was assumed to obey the $n^{2/3}$ rule. For comparative purposes, the S3LC − TCEP data were also fit with SEDFIT (*16*, *17*) to a monomer−trimer model using a constrained value of the $s_{1,app}$ (2.65 S) and $s_{3,app}$ (5.3 S) values for each experiment.

Sedimentation equilibrium experiments were performed at 28K rpm for S4AF, 20K and 22K rpm for S3LC, 20K rpm for S3LC(3E), and 24K for S4AF(3E), in charcoal-filled Epon 6 channel centerpieces. Equilibrium data were collected at a spacing of 0.001 cm with 9 averages in a step scan mode. Data sets were edited with REEDIT to extract each channel of data and fit jointly with NONLIN (*18*) to an appropriate association scheme as described in detail elsewhere (*19*). Nonlin fits to an effective reduced molecular weight [$\sigma = M(1 - \upsilon\rho)\omega^2/RT$] were for an associating system where σ_1 is held at a value corresponding to the monomer. The monomer molecular weight of these constructs is based upon the amino acid composition and calculated in Sednterp (*11*) to be 30 884 for S4AF, 31 124 for S4AF(3E), 31 565 for

[2] Note, the software used in this analysis (NONLIN, REEDIT, SEDFIT, and Sednterp) is available through the RASMB site located at http://www.bbri.org/RASMB/rasmb.html. or (DCDT+, SVEDBERG) at http://www.jphilo.mailway.com/download.htm.

[3] An elegant, state-of-the-art method in use today is to estimate the hydrodynamic behavior of a macromolecular complex from its crystal structure using bead modeling with HYDRO (*14*). Unfortunately, only Smad4 constructs have been crystallized as trimers (*1*, *15*), and they do not form homo-trimers in solution. Using the pdb file for the trimeric crystal stucture of S4AF and a Sednterp calculated value of hydration and vbar, we calculate a value for S_3 of 5.47 $S_{20,w}$. This appears to be too large of an estimate for the Smad3 trimer, being in excess of the extrapolated value, 4.9 $S_{20,w}$, and the value predicted by the $n^{2/3}$ rule, 5.12 $S_{20,w}$. This suggests that crystal packing forces compact the S4AF structure relative to the solution structure, and either S3LC homo-trimers and S3LC:S4AF hetero-trimers are more expanded than the S4AF trimeric structure predicts, or some floppy domain is significantly increasing the frictional drag on the solution complex. Using the HYDRO estimated value of S_3 in the weight average fitting (Table 1) decreases the average overall K_3 by 42.4%.

S3LC, and 31 691 for S3LC(3E). The molecular weight and association constants determined by NONLIN fitting are present as 95% confidence intervals and correspond to two standard deviations. There is evidence of heterogeneity in many of these data sets, especially in the absence of TCEP. Thus, the best fit often involves separate *K* values for each data set, an option in NONLIN, rather than a single global *K* that describes all data simultaneously. In some instances, i.e., S3LC + TCEP, the separate *K*'s are narrowly spread about the global *K*, possibly reflecting trace amounts of aggregation or stable conformational and energetic microheterogeneity (*20*), while in other instances, S4AF(3E) − TCEP, the *K* value systematically decreases with increasing loading concentration, reflecting the presence of significant amounts of cross-linked dimers. Sedimentation equilibrium experiments were also performed on mixtures of S4AF plus the Smad3 constructs at various ratios. In these instances, short column methods at low speeds (14K rpm) were employed to reduce fractionation of the components and thus, by single species fits with NONLIN, obtain the best whole cell estimates of the hetero-complex molecular weight.

The density of the buffer was measured in an Anton-Paar DMA 5000 density meter to be 1.00228 ± 0.00057 at 24.7 °C. Buffer viscosity was measured in a Cannon-Manning semimicro viscometer to be 0.9185 ± 0.0016 at 24.7 °C. These are the average values for different dialysate buffer solutions used in these experiments. The actual values for each buffer used in an experiment were applied to correct data to $s_{20,w}$ and MW (*11*). The partial specific volume of each construct was calculated in Sednterp (*11*; based upon the method of Cohen and Edsall as justified in ref *21*) to be 0.7320 for S4AF, 0.7307 for S4AF(3E), 0.7231 for S3LC, and 0.7231 for S3LC(3E). The extinction coefficient at 280 nm, used to convert absorbance to molar and weight concentration scales, was calculated in Sednterp [*11*; based upon the method of Pace et al. (*22*)] to be 4.24×10^4 M^{-1} cm^{-1} for each construct. This implies that for the analysis of mixed sedimentation velocity samples, i.e., S4AF + S3LC, integration of the *g*(*s*) patterns properly converts them to a weight average sedimentation coefficient (see Figure 3), rather than an absorbance average which is expected if the species have different extinctions.

RESULTS

S4AF Sedimentation Studies. Figure 1A presents *g*(*s*) patterns for S4AF. The peak shows no shift with concentration and is consistent with a noninteracting, primarily monomeric species in solution. The mean weight average sedimentation coefficient ($\bar{s}_{20,w}$) determined by DCDT+ analysis of these data is 2.46 (±0.03) $S_{20,w}$ (see Figure 3). The highest concentration patterns display a small amount of centrifugal skewing, consistent with aggregated material, and thus single and double Gaussian fitting of these patterns was conducted with DCDT+ to estimate the molecular weight of the species present. The average molecular weight from single-species fits was 26 905 ± 2197, while the average of the best fits, single-species fits at low concentration and two-species fits at higher concentrations, was 27 802 ± 1527 Da with $s_{20,w} = 2.43 \pm 0.01$ for the major species. To verify these results, the data were also fit with SVEDBERG in a similar one- and two-species manner. The averages from single-species fits (2.56 ± 0.03 $S_{20,w}$ and 28 740 ± 1390 Da) and two-species fits (2.55 ± 0.02 $S_{20,w}$ and 29 780 ± 809 Da) are consistent with S4AF behaving as a monomer with the presence of a small amount (2.6%) of 4.7−4.9 S_{app} aggregate in these solution conditions. [The amount and size of the aggregate were constant, within error, and thus not consistent with reversible self-association. In size exclusion chromatography experiments, these aggregates elute between the monomer and reversible trimer zones (*4*).] These results were further verified by two sedimentation equilibrium runs (Figure 2) that gave global best fits of the data consistent with a monomeric species in solution (expt 1 = 30 133 ± 690; expt 2 = 29 902 ± 503; expected MW = 30 844). For reasons expanded on below, sedimentation velocity experiments were also preformed in the presence of 2 mM TCEP (the +++ curve in Figure 1A), a strong reductant that conveniently has very little absorbance at 280 nm. The results were identical except the 4.7−4.8 S aggregate was undetectable (<1%), suggesting they represented nonspecific disulfide-cross-linked species. Thus, we can conclude that, up to approximately 33 μM in these solution conditions, S4AF behaves as a monomer.

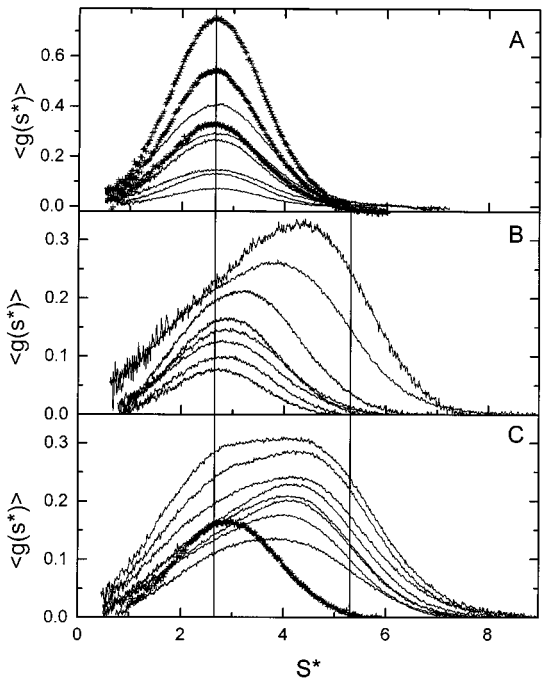

FIGURE 1: Sedimentation velocity analysis of S4AF, S3LC, and mixtures of the two Smad constructs. Panel A: *g*(*s*) profiles from sedimentation velocity runs with S4AF; (- - -) minus TCEP; (+++) plus TCEP. The apparent weight average *s* value of these data is 2.67 (±0.03) S, corresponding to the left vertical line, and consistent with a noninteracting monomer. Panel B: *g*(*s*) profiles from sedimentation velocity runs with S3LC + 2 mM TCEP. The shift with increasing concentration is consistent with a reversible monomer−trimer self-association with pure trimer corresponding to the right vertical line (see text). The highest concentration pattern was collected at 250 nm and rescaled to 280 nm, accounting for the noisey appearance of the curve. Panel C: *g*(*s*) profiles from sedimentation velocity runs with mixtures of a fixed amount of S3LC (7 μM) and increasing amounts of S4AF. The (+++) line identifies the S3LC sample in the absence of S4AF. The addition of S4AF immediately causes a shift in the distribution to higher *s* values. With further addition of S4AF, a centripetal, or slower, zone develops near 2.7 S, consistent with excess S4AF running as monomer. The $\bar{s}_{20,w}$ values of the data in all three panels are plotted in Figure 3. Similar results were obtained in the absence of TCEP.

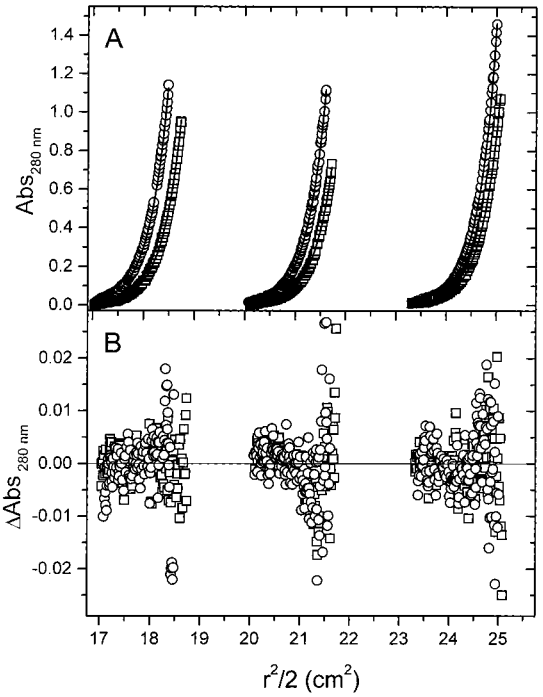

FIGURE 2: Sedimentation equilibrium data for S4AF performed at 28 K, 24.7 °C, in the absence of TCEP. The loading concentrations were 2, 4, 6, 8, 10, and 12 μM. Panel A: Equilibrium data and the best fits. Panel B: The residuals. The best global fit of the data is consistent with a monomeric species in solution giving a measured molecular weight of 29 902 ± 503 (expected MW = 30 844) with an rms deviation of 0.00557. A separate experiment at 28 K on a different preparation gave a similar result, MW = 30 133 ± 690, with an rms deviation of 0.00271, while experiments in the presence of TCEP at lower speeds are also consistent with S4AF behaving as a monomer in solution (*4*).

S3LC Sedimentation Studies. Figure 1B presents the sedimentation velocity $g(s)$ data for S3LC in the presence of 2 mM TCEP. The $g(s)$ patterns clearly shift to higher values with increasing concentration, consistent with a reversible self-association process. The $\bar{s}_{20,w}$ values (±TCEP) are plotted in Figure 3, and while both data sets are consistent with a concentration-dependent process, the data in the presence of TCEP are shifted to slightly lower values. To investigate this, further sedimentation equilibrium data were collected (±TCEP) to verify reversible self-association and to determine the best model that describes these data (Figure 4). In both cases (±TCEP), the data are best fit by a monomer−trimer model with a global fit giving an overall trimerization constant K_3 of (3.09−2.99) × 10^9 M^{-2} (Table 1 and Figure 4). Fitting to both N and K returns a value of N = 2.96 ⟨2.81, 3.09⟩ (+TCEP) and 2.98 ⟨2.84, 3.12⟩ (−TCEP), verifying that trimer is the major species in both fits. There is no improvement in the fits by including a dimer term, K_2, thus implying the formation of trimer is cooperatively favored over dimer formation. [Note, for example, in the presence of TCEP the value of K_2 returned is 2.67 × 10^3 M^{-1}, and thus while not statistically significant nor required to describe the data, this implies the value is simply too small to produce a significant fraction of dimer (see Discussion).] The rms of each fit is significantly improved by using separate K_3 values for each channel, consistent with some form of heterogeneity. For example, for the +TCEP data, a significantly better fit is obtained, rms = 0.0051, where the average K_3 value is 3.213 (±0.899) × 10^9 M^{-2},

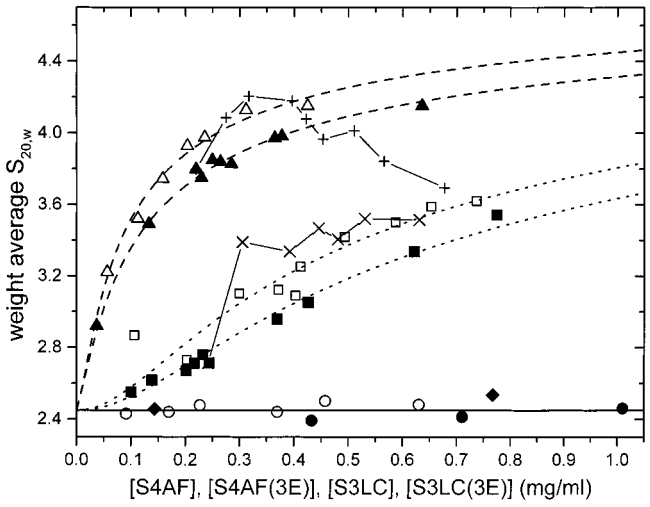

FIGURE 3: Weight average $\bar{s}_{20,w}$ values derived from $g(s)$ analysis of sedimentation velocity runs with S4AF (open and closed circles), S4AF(3E) (closed diamonds), S3LC (open and closed squares), and S3LC(3E) (open and closed triangles). The closed symbols represent data collected in the presence of 2 mM TCEP. The trend of the S4AF data is consistent with a monomer of 2.45 (±0.04) $S_{20,w}$. The S3LC and S3LC(3E) data were best fit to either a monomer−trimer (dotted line) or a cooperative monomer−dimer−trimer model (dashed lined), although the differences are slight (see Table 1). Experiments presented on mixtures of S4AF + S3LC (-x-x-x) or + S3LC(3E) (-+-+-+) were performed +TCEP. The initial addition of S4AF to S3LC or S3LC(3E) causes a shift to more trimer and a larger $\bar{s}_{20,w}$ value, consistent with a tighter hetero-association than homo-association mechanism.

a value reasonably consistent with the global value (Table 1). In the absence of TCEP, the heterogeneity is at least in part due to the presence of larger aggregates. These aggregates are larger than hexamers because attempts to fit to a 1−3-nmer model return a value of N > 14 with a modest improvement in the rms. In the presence of TCEP, a 1−3-nmer model does not improve the fit. Thus, S3LC is best described by a monomer−trimer equilibrium, but in the absence of TCEP larger aggregates are evident, while under both ±TCEP conditions there is some residual micro-heterogeneity possibly due to the presence of small aggregates or static conformational heterogeneity. It is worth noting that in size exclusion chromatography experiments these aggregates coelute near the reversible trimer zone (*4*).

Knowing the data represent a monomer−trimer equilibrium, we next converted the $g(s)$ patterns (±TCEP) to $\bar{s}_{20,w}$ values (Figure 3) and fit them to a monomer−trimer model and a monomer−dimer−trimer model as described under Materials and Methods (Table 1 and Figure 3). Both models give similar fits with an overall equilibrium constant of (2.10−1.15) × 10^9 M^{-2}. This is in reasonable agreement with the sedimentation equilibrium results, (3.10−2.99) × 10^9 M^{-2}, and, as is evident in the plot of $\bar{s}_{20,w}$ vs protein concentration for S3LC (Figure 3), reveals that wild-type trimer formation is 50% complete at 18.0−29.5 μM. (This range of values is estimated by taking the $1/\sqrt{}$ of the K_3 values.)

Sedimentation Studies of S4AF/S3LC Mixtures. Figure 1C shows the effect of titrating S4AF into a fixed amount of S3LC. Rather than producing a sum of the S4AF monomer zone and the S3LC interacting zone, the boundary shifts to a higher degree of sedimentation, consistent with the formation of a hetero-complex. As more S4AF is added, this

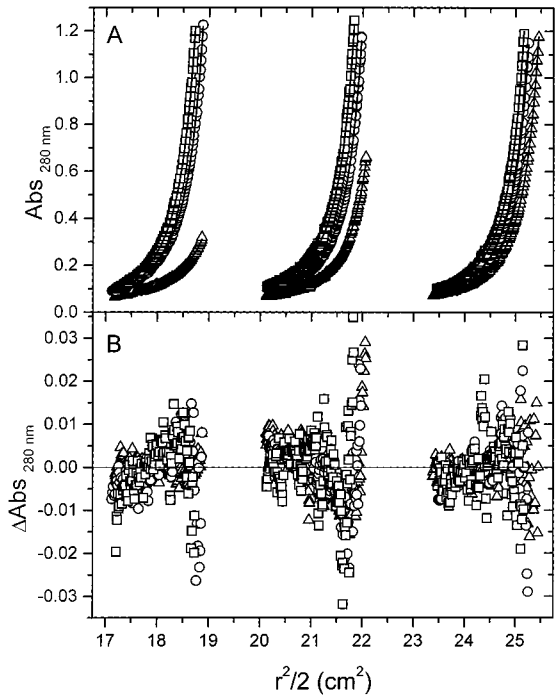

FIGURE 4: Sedimentation equilibrium experiment of S3LC at loading concentrations of 2−16 μM + 2 mM TCEP at 20 K, 24.7 °C, 280 nm. Panel A: Equilibrium data and the best fits. Panel B: The residuals. The best global fit of these data, shown below, is a monomer−trimer model where $K_3 = 2.991\ \langle 2.588, 3.482\rangle \times 10^9$ M^{-2} with an rms deviation of 0.0067. A better fit is obtained (rms deviation = 0.0051) if each channel of data is fit with a separate K_3 value, where the average K_3 value is 3.213 ($\pm$0.899) $\times 10^9$ M^{-2}. This implies that even in the presence of TCEP there is some residual aggregation or micro-heterogeneity. This result compares favorably with the −TCEP equilibrium data, and within a factor of 3 with the weight average fitting of the data $\pm$TCEP [K_3 = (1.15−2.10) $\times 10^9$ M^{-2}; see Table 1].

Table 1

construct	TCEP	K_3 ($\times 10^9$ M^{-2})	rms[a]	method
S3LC	−	3.090 ⟨2.597, 3.691⟩	0.0056	equilibrium
	+	2.991 ⟨2.588, 3.482⟩	0.0067	equilibrium
	−	2.097	0.1240	$\hat{s}_{20,w}$ (1-3)
	−	1.478 ± 0.326	0.0117	SEDFIT (1-3)[b]
	−	2.109	0.1265	$\hat{s}_{20,w}$ (1-2-3)
	+	1.150	0.0437	$\hat{s}_{20,w}$ (1-3)
	+	1.150	0.0465	$\hat{s}_{20,w}$ (1-2-3)
S3LC(3E)	−	54.76 ⟨42.99, 70.90⟩	0.0062	equilibrium
	+	51.94 ⟨40.71, 67.60⟩	0.0058	equilibrium
	−	74.33	0.0686	$\hat{s}_{20,w}$ (1-3)
	−	92.04	0.0519	$\hat{s}_{20,w}$ (1-2-3)
	+	36.33	0.0908	$\hat{s}_{20,w}$ (1-3)
	+	52.29	0.0356	$\hat{s}_{20,w}$ (1-2-3)

[a] rms deviation of the fit; in units of ΔOD for equilibrium and SEDFIT analysis; in units of ΔS for weight average fitting. [b] Average values (10 data sets with S3LC − TCEP) from single-experiment fits to a monomer−trimer model.

associating zone continues to grow in area, or concentration, with a peak position that is consistent with trimer formation. Eventually the addition of S4AF also gives rise to a slowly sedimenting zone near 2.65 S_{app}, corresponding to excess monomeric S4AF. These data are converted to a weight average $s_{20,w}$ value and plotted in Figure 3 with the S3LC results. It is apparent that at the same total protein concentration, i.e., S3LC vs S3LC + S4AF, the mixed sample has a larger average size, consistent with a tighter hetero-trimer complex than homo-trimer complex. [We cannot exclude some contribution from the shape of the hetero-complex to this increase, but given the similar size and sequence of S3LC and S4AF, the hetero-trimer should be very similar hydrodynamically to the homo-trimer (*15*).] At higher S4AF concentrations, the $\bar{s}_{20,w}$ plateaus and begins to decrease, due to the contribution of excess S4AF monomer on the weight average molecular weight (*23*, *24*). Short-column sedimentation equilibrium experiments on mixtures of S4AF and S3LC approach but do not exceed the molecular weight of the trimer (data not shown), confirming the formation of a hetero-trimer complex and excluding the possibility that this hetero-zone (Figure 1C) uniquely corresponds to dimeric or hexameric species. By comparison with the $\pm$TCEP S3LC weight average data (Table 1), the data on mixtures of S4AF and S3LC are consistent with a slight, 2−3-fold, increase in the equilibrium constant for hetero-trimer formation over homo-trimer formation. Analysis of similar zones purified off a sizing column and run on an SDS gel suggests a 1:2 stoichiometry for this complex, containing 1 molecule of S4AF and 2 molecules of S3LC (*4*). Note, in the $g(s)$ patterns in Figure 1C, even at the highest stoichiometry the interacting zone is still shifting toward trimer, and thus we cannot easily interpret these data in terms of the stoichiometry.

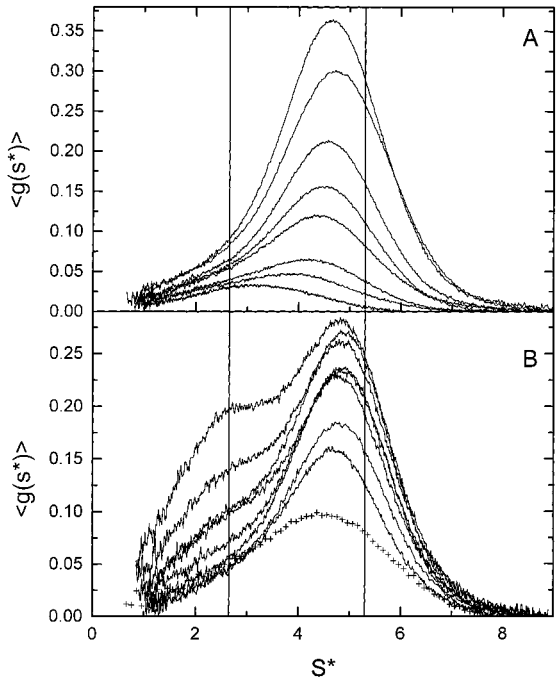

FIGURE 5: Sedimentation velocity analysis of S3LC(3E) and mixtures of S4AF and S3LC(3E). Panel A: $g(s)$ patterns from S3LC(3E) in the presence of TCEP. Panel B: $g(s)$ patterns from a fixed amount of S3LC(3E) (+++, 7 μM) and increasing amounts of S4AF. The $\bar{s}_{20,w}$ values of the data in these panels are plotted in Figure 3. Similar results were obtained in the absence of TCEP (*4*).

S3LC(3E) Sedimentation Studies. Figure 5A presents the sedimentation velocity $g(s)$ data for S3LC(3E) in the presence of TCEP. The $g(s)$ patterns clearly shift to higher values with increasing concentration, consistent with a reversible self-association process, and the curves are shifted dramatically more to trimer formation than the S3LC data (Figure 1B). The data ($\pm$TCEP) are converted to $\bar{s}_{20,w}$ values and fit to a monomer−trimer and a cooperative monomer−dimer−trimer model (Table 1), and the best fits are presented in Figure 3.

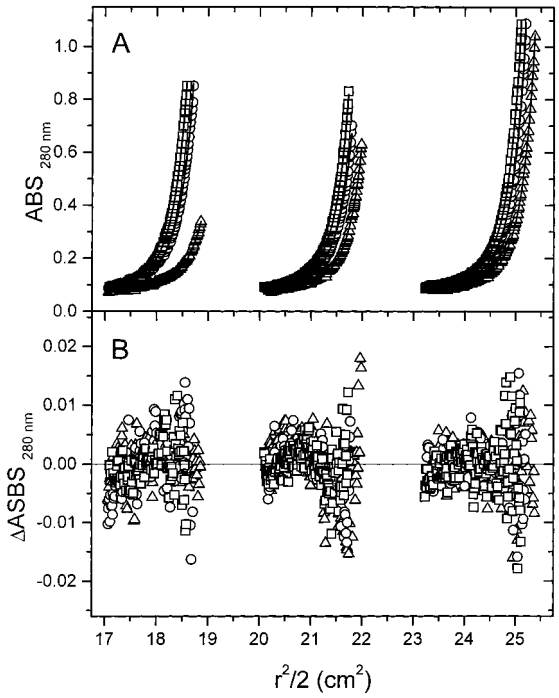

FIGURE 6: Sedimentation equilibrium experiment of S3-25(3E) at loading concentrations of 2−16 μM + 2 mM TCEP at 20 K, 24.7 °C, 280 nm. Panel A: Equilibrium data and the best fits. Panel B: The residuals. The best global fit of these data is a monomer−trimer model where $K_3 = 5.194\ \langle 4.071, 6.760\rangle \times 10^{10}\ M^{-2}$ with an rms deviation of 0.00580. Better fits are obtained if each channel of data is fit with a separate K_3 value, again implying that even in the presence of TCEP there is some residual aggregation or micro-heterogeneity occurring. This result compares favorably with the −TCEP equilibrium data (Table 1), and within a factor of 2 with the weight average fitting of the data ±TCEP [$K_3 = (3.63-9.20) \times 10^{10}\ M^{-2}$, Table 1]. These data demonstrate that pseudo-phosphorylation enhances trimerization 17−35-fold over the wild type.

It is readily apparent that S3LC(3E) associates with much tighter affinity due to the presence of the pseudo-phosphorylation sequence. The best monomer−trimer fits (±TCEP) are consistent with an overall equilibrium constant of $(3.63-7.43) \times 10^{10}\ M^{-2}$, approximately 32−35-fold larger than the corresponding wild-type or S3LC data (Table 1). To verify these results, sedimentation equilibrium experiments (±TCEP) were also performed (Figure 6). The best fit is to a monomer−trimer model with K_3 varying from 5.19×10^{10} to $5.48 \times 10^{10}\ M^{-2}$ (Table 1). This is excellent agreement with the best monomer−trimer fits of weight average data and corresponds to 17.4−17.7-fold enhancement of trimerization relative to the wild-type S3LC sedimentation equilibrium data. As is evident in the plot of $\bar{s}_{20,w}$ vs protein concentration for S3LC(3E) (Figure 3), these affinity constants imply that trimer formation is 50% complete at 3.3−5.2 μM for this pseudo-phosphorylation construct. Thus, the substitution of a pseudo-phosphorylation sequence enhances homo-trimer formation 17−35-fold over the wild type, and contributes 1.68−2.10 kcal of trimer stabilization.

S3LC(3E)/S4AF Mixtures. Figure 5B presents the $g(s)$ patterns for mixtures of fixed amounts of S3LC(3E) plus increasing amounts of S4AF. As with the S3LC experiment, the addition of S4AF does not give the sum of the $g(s)$ patterns, but rather causes a shift to a larger extent of association. The trimer zone grows in area or concentration first, consistent with the formation of hetero-trimer, and then with increasing amounts of S4AF, the excess S4AF begins to resolve into a 2.65 S_{app} zone. As with the wild-type data (Figure 1C), the mixed sample initially has a larger average size at the same total protein concentration as S3LC(3E) alone, consistent with a tighter hetero-trimer complex than homo-trimer complex. This places the overall affinity of the hetero-trimer complex at a value $>5.5 \times 10^{10}\ M^{-2}$, although by comparison with the ±TCEP data, it may very likely be only a factor of 2−3-fold larger than S3LC(3E) homo-trimerization. As excess monomeric S4AF accumulates, the weight average value decreases as expected (*23, 24*). Note, in this tighter mixed-association reaction, the excess S4AF becomes evident above a 1:1 stoichiometry, consistent with the sizing column data that the hetero-complex is 1 S4AF molecules and 2 S3LC(3E) molecules (*4*). [We must caution that this apparent qualitative agreement requires more quantitative analysis of the heterogeneous boundary shapes (Stafford and Correia, in development).]

S4AF(3E) Sedimentation Studies. Taken together, our data suggest that phosphorylation of the Smad3 C-terminal target sequence directly promotes Smad3 homo-trimerization and Smad3:Smad4 hetero-trimerization. Despite high sequence homology between Smad3 and Smad4 over the C-terminal oligomerization domain (50% identity), Smad4 lacks the C-terminal phosphorylation sequence and does not undergo reversible oligomerization. To investigate whether the C-terminal phosphorylated sequence can induce trimerization in Smad4, a S4AF(3E) construct was generated where the Smad4 C-terminal sequence was replaced by the pseudo-phosphorylated Smad3 sequence. Surprisingly, in the absence of TCEP, this construct sedimented as a broad, concentration-independent zone of 3.04 S, a value larger than a monomer, but smaller than a trimer (Figure 7). Two-species fits with DCDT+ (where the sedimentation coefficient of the monomer is constrained to the S4AF value, 2.65 S) describe the data, with the second species, presumably dimer, running at $S_2 = 4.11 \pm 0.04$ and making up 34.4 ± 4.4% of the material. (However, molecular mass estimates for the two species from this analysis are generally low, 22.3 ± 1.9 and 33.9 ± 5.5 kDa, suggesting some additional boundary spreading, possibly due to multiple cross-linked forms.) Sedimentation equilibrium data were best fit (data not shown) by a heterogeneous monomer−dimer model; i.e., each data set was best fit by a separate K_2 value that systematically decreased with increasing loading concentration, suggesting the presence of a significant concentration of cross-linked dimers. In the presence of 2 mM TCEP, sedimentation velocity and equilibrium data were consistent with a homogeneous monomer, giving a weight average s value of 2.7 S (the closed triangles in Figure 3 and the left vertical line in Figure 7) and a MW = 30 818 ⟨30 083, 31 536⟩ (Figure 8; expected MW = 31 124). [Note, these data clearly highlight the need for a strong reducing agent with in vitro studies of Smad proteins, because otherwise we might mistakenly conclude S4AF(3E) specifically forms dimers.] Even at 25 μM, S4AF(3E) + TCEP runs as a monomer of 2.74 S and 29.11 kDa, determined by fitting with DCDT+ (see Figures 3 and 7). Thus, the pseudo-phosphorylated target sequence alone cannot induce homo-trimerization of S4AF, possibly due to an inhibitory effect of the Smad4 specific sequence, referred to as the TOWER, and known to interact with the C-terminal sequence of Smad4 (*15, 25*). It is also possible

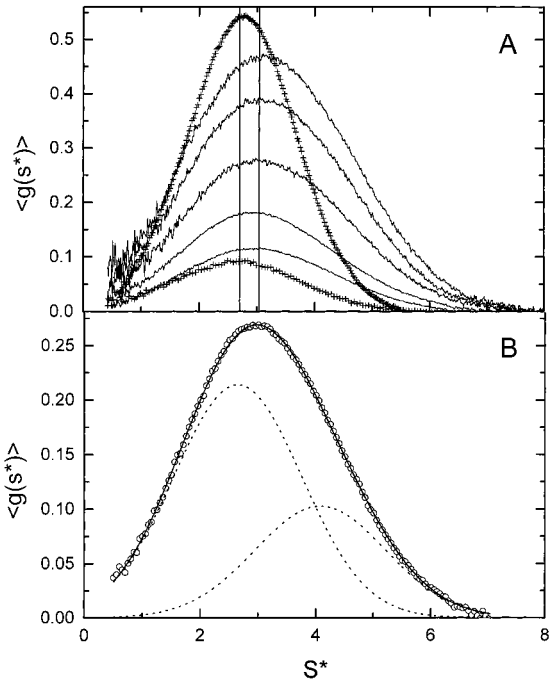

FIGURE 7: Panel A: Sedimentation velocity $g(s)$ data for S4AF-(3E) at 42 K, 24.7 °C, in the absence (solid lines) and presence (+++) of 2 mM TCEP. The average $\bar{s}_{app}$ values are 3.04 and 2.70 S, respectively, as indicated by the two vertical lines. These data are consistent with a mixture of monomers and cross-linked dimers in the absence of TCEP, and homogeneous monomers in the presence of TCEP. The $\bar{s}_{20,w}$ values of the +TCEP data are plotted in Figure 3. Panel B: An example of a two-species Gaussian fit (solid line) performed with DCDT+ demonstrating the quality of the fit with a monomer and a cross-linked dimer. The two species are represented by the dotted lines, while the actual data (every fifth point) are plotted as open circles.

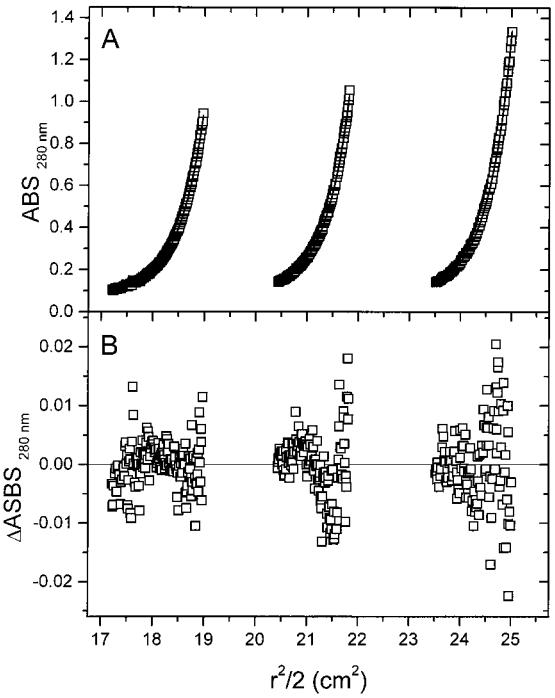

FIGURE 8: Sedimentation equilibrium data for 5, 10, and 15 μM S4AF(3E) run at 24 K, 24.7 °C, in the presence of 2 mM TCEP. Panel A: Equilibrium data and the best fits. Panel B: The residuals. The best global fit is a single species with MW = 30 818 ⟨30 083, 31 536⟩ (expected MW = 31 124) and an rms deviation of 0.00583.

that the MH2 domain of S3LC forms unique structures absent in S4AF that interact with the target sequence and facilitate trimer formation. Despite this inhibitory effect on homo-trimerization, S4AF is able to hetero-trimerize, suggesting the TOWER sequence has favorable interactions in the hetero-complex (*15*).

DISCUSSION

The Complementary Use of Sedimentation Equilibrium and Velocity Methods. The sedimentation results presented involve a combination of equilibrium and velocity techniques. The results complement one another and in general are in excellent quantitative agreement. As reviewed elsewhere (*10*, *26*), this is not a surprising result and further demonstrates that sedimentation velocity analysis reflects the equilibrium properties of the system. For example, the concentration dependence of the S3LC and S3LC(3E) $g(s)$ patterns (Figures 1 and 5) suggests a monomer−n-mer model, and the extrapolated maximum s value is consistent with trimer formation. [It is worth noting that Gilbert theory predicts monomer−n-mer models where $n > 2$ should resolve into a bimodel pattern (*12*, *13*, *27*). However, this is based upon simulations conducted in the absence of diffusion (*27*). The small size of these Smad proteins is consistent with rapid diffusion and the spreading of the reacting zone into a centrifugally skewed boundary (*28*).] Sedimentation equilibrium results robustly verify a monomer−trimer model, and fitting of the equilibrium data and the weight average $s_{20,w}$ data agree within a factor of 2−3 for both the wild-type and the pseudo-phosphorylated forms of the protein. With this as a basis, interpretation of the $g(s)$ patterns of mixtures of S4AF with S3LC or S3LC(3E) (Figures 1C and 4B) can thus be undertaken with great confidence to suggest the formation of hetero-trimer. This was also verified for mixtures of S4AF and S3LC by short-column sedimentation equilibrium methods. These results also allow us to definitively interpret the chromatography patterns obtained during size exclusion experiments as being diagnostic of the formation of homo- and hetero-trimers (*4*).

The excellent quantitative agreement between the weight average and the sedimentation equilibrium fitting (Table 1) is subject to certain caveats (*10*). First, it is extremely difficult to span a wide enough concentration range to absolutely determine the sedimentation coefficient of the endpoints, the monomer and trimer species. The lowest concentration point for S3LC (Figures 1B and 3) seems to be in reasonable agreement with the assignment that the monomeric form sediments quantitatively like S4AF. However, at high concentration, Stafford (*29*) has suggested, based upon simulations, that hydrodynamic behavior consistent with the complex, in this case trimer, requires experiments at concentrations at least 2 orders of magnitude above the K_d. This would correspond to 2.37 mM S3LC and 425 μM S3LC-(3E). Determination of the s value for the trimer is problematic because it is subject to the influence of hydrodynamic nonideality at high concentration, ignored in the fitting presented in Table 1. Consistent with this, the peak positions in Figure 5A,B for the highest protein concentrations of S3LC(3E) seem to be slowing down, exhibiting effects due to nonideality and in fact suggesting the correct trimer value could only be determined by an extrapolation procedure (see extrapolation in refs *12*, *13*, and *30* as

examples). Hydrodynamic nonideality is included in the fitting function in the form of $(1 - gc)S_i$ where c is the total weight concentration (mg/mL) and g has units of mL/mg (*10*, *12*, *13*). A reasonable estimate of g for Smad proteins is 0.006 mL/mg (*31*). If we use this value of g and refit the data in Figure 3, the K_3 values are raised by less than 3%, although the fits are only slightly improved in three of the eight cases. This marginally improves the agreement between the weight average and the equilibrium data for the S3LC construct, and has no significant effect on the agreement of the S3LC(3E) estimates. Thus, nonideality is not a significant contributor to the fitting of these data, as was already evident in the S4AF weight average data (Figure 3). It is also likely that small errors in the estimates of S_1 and S_3 partially compensate for the omission of hydrodynamic nonideality in the weight average fitting, and thus improve the agreement with the sedimentation equilibrium fits (see Table 1).

The best fit of the S3LC(3E) weight average data by the criteria of rms is to a 1-2-3 model which gives an overall estimate for K_3 of $(5.23-9.20) \times 10^{10}$ M^{-2}, a range also in reasonable agreement with the sedimentation equilibrium values. Consistent with this, we observed that a 1-2-3 model also fits the sedimentation equilibrium data, although it is not statistically better than a 1-3 model and thus not the best model. In the 1-2-3 fits of the ±TCEP data, the value for K_2 is in the range of $(5.11-5.90) \times 10^4$ M^{-1} (with slightly larger K_3 values than the 1-3 fits), suggesting very low concentrations of dimer in solution.[4] Thus, we can ask the following question: Why are the weight average data for S3LC(3E) so much better described by the 1-2-3 model? The simplest answer is those data include few points and have a larger rms than the equilibrium data, and are thus more subject to random bias. However, the heterogeneity we described in the 1-3 fits of the equilibrium data suggests some additional complexity in the system. We have referred to this as small amounts of aggregation or stable conformational and energetic micro-heterogeneity, and emphasize these separate K_3 values are consistent with the global K_3 value (Table 1). Static (and not rapidly interconverting) conformational heterogeneity will have more impact on the shape and thus s values of dimers and trimers than on their molecular weight, and thus it is possible the S3LC(3E) weight average data are simply more sensitive to the presence of small amounts of aggregate and/or reversible dimer. This further emphasizes the point that sedimentation velocity analysis of self-association processes can have advantages over equilibrium measurements for systems that are undergoing shape as well as molecular weight changes (*10*, *32*).

The excellent agreement between the equilibrium and velocity data ±TCEP seems to contradict the shift in the $g(s)$ patterns. The slightly faster weight average s values in the absence of TCEP are consistent with the presence of nonreversible aggregation. Why do these aggregates not have large effects on the equilibrium fitting? One factor is the equilibrium data are edited to be less than about 1.0−1.25 ODU to avoid nonlinear optical effects, while the velocity data include the contribution from the entire boundary. The data at the base of the equilibrium run will be enriched for aggregates, and thus those aggregates will selectively spin out of view during the run. It is also possible that the aggregates further aggregate during the time of the equilibrium experiment, effectively pelleting them from view. However, the residual heterogeneity observed in the system suggests the small aggregates ±TCEP are of a similar size and contribute equally to both data sets. These results further justify the complementary use of sedimentation velocity in these studies.

Finally, a new analysis method (SEDFIT)[2] that directly fits a sedimentation velocity boundary to a monomer−trimer model has been developed by Schuck (*16*). This technique has the advantage that it can remove time-invariant noise from the data set and thus in principle provide a less biased estimate of K_3. This method has the disadvantage that it only fits one or two data sets at a time. Inspection of the $g(s)$ patterns for reversible association (Figures 1B and 5A) clearly shows that the end points of the reaction, $S_{1,app}$ and $S_{3,app}$, are not well-determined from a single data set, and thus cannot be estimated as fitted parameters. Initial attempts to float these parameters gave negative values of $S_{1,app}$ and values for $S_{3,app}$ above 7 S. A more robust approach, analogous to the discriminatory power of fitting multiple equilibrium data sets with NONLIN, would appear to require a global analysis of many sedimentation velocity data sets simultaneously (*33*). However, constraining the apparent sedimentation coefficients to the monomer and trimer values, as described above for the weight average fitting method, does allow excellent SEDFIT analysis of the data that quantitatively agree with the equilibrium and weight average fits reported in Table 1. For example, the average value of K_3 determined by analysis with SEDFIT for 10 data sets of S3LC − TCEP was 1.478 (±0.326) × 10^9 M^{-2}.

Implications for Smad Protein Trimerization. The Smad3 and Smad4 constructs used in this study are prone to aggregation and disulfide cross-linking, although the extent of cross-linking is construct-specific. Even though DTT is present during protein purification and analysis, we suspect that weak, nonspecific aggregation occurs during the concentration, typically to 300−400 μM, and dialysis steps. From the comparative sedimentation analysis with and without TCEP, it was evident that TCEP was required to reduce the extent of cross-linking and establish the reversible behavior and accurate energetics of Smad protein trimerization. Mutations that alter global features or disrupt normal association interfaces may especially be anticipated to induce these nonspecific effects. The relatively strong reducing environment of the cytoplasm and the low endogenous level of Smad proteins suggest this phenonenon is an in vitro artifact that is controllable by the presence of TCEP. However, very recently, Jayaraman and Massague (*34*) reported the presence of very large aggregates >650 kDa of Smad2, -3, and -4 in extracts from HaCaT and COS-1 cells. These aggregates may be partially be due to the presence of other protein factors, but our results with TCEP suggest the possibility of covalent cross-linking in these experiments. Thus, we recommend all in vitro and extraction experiments with Smad proteins be performed in a strong reducing environment if biologically relevant data are to be collected. The strongest evidence that the reversible phenomena

[4] The cooperativity or γ values in these fits are typically 4−21, but one value is 6×10^5, suggesting they are not reliable estimates. The cooperativity parameter can also be inferred from 1-2-3 fits of the sedimentation equilibrium data. Those values are typically 3×10^3, a magnitude more consistent with the expected structure of a closed trimer and the energy gained from formation of an additional interface.

observed in our in vitro study have biologically relevance are transfection and activation studies where the full-length pseudo-phosphorylated constructs, including the MH1 DNA binding domain, activate transcription (*4*). These data clearly imply that the energetic enhancement caused by pseudo-phosphorylation of S3LC has biological significance.

Both sedimentation velocity and equilibrium results are consistent with S4AF behaving as a monomer in solution. Qin et al. (*15*) crystallized this same construct as an asymmetric trimer. Is it a contradiction that the solution behavior of S4AF be monomeric while S4AF crystallizes as an asymmetric trimer? First, it is an assumption that the crystallized form of a protein is identical to the solution structure. In this instance, S4AF was crystallized in 200 mM $(Li)_2SO_4$, 10% PEG 4000. It is well established that PEG favors the oligomerization of many protein complexes, primarily by an excluded volume mechanism and effects on hydration. In addition, there were numerous sulfate binding sites in the refined structure, thus implying that sulfate binding favors trimerization, in part overcoming inhibitory interactions in the structure or mimicking phosphorylation sites in the target sequence. In addition, crystal packing forces often contribute to the stability of the unit cell. Thus, while in solution, S4AF at $<33\ \mu$M is a monomer; at much higher concentrations and in favorable crystallization conditions, S4AF readily forms trimers. As discussed previously, the Smad4-specific C-domain sequence, referred to as the TOWER, interacts with the C-terminus of Smad4 and clearly plays a role in inhibiting homo-trimerization. This is further supported by the observation that Smad4 in the basal state in cells is a monomer (*3*). Since the hetero-trimers of S4AF and S3LC or S3LC(3E) form so readily at micromolar concentrations, we anticipate, as suggested by Qin et al. (*15*), the hetero-trimer crystal structure will reveal unique favorable interactions between the Smad4-specific C-domain sequence and the Smad3 C-terminus.

The surprising tendency of S4AF(3E) to form disulfide-cross-linked dimers reflects a gain of an undesirable function caused by an unusual conformation or tertiary fold of the chimeric protein. This may reflect unfavorable intramolecular interactions with the target sequence, mediated by the Smad4-specific C-domain sequences, which then allows for intermolecular interactions leading to disulfide cross-linking. (Note, we do not assume this cross-linked dimer is similar to a reversible dimeric species, although its apparent sedimentation coefficient, 4.11S, is very similar to the expected value of 4.21 S.) While it is reasonable to suggest the cysteine in the Smad3 C-terminus added to the Smad4 sequence is the site of disulfide cross-linking, we cannot exclude the involvement of other Cys residues, mediated by local or global effects on the tertiary fold. Nonetheless, the simplest interpretation is that the Smad3 target sequence added to S4AF(3E) is the site of both nonspecific aggregation and disulfide cross-linking. An obvious future test of this interpretation is to mutate the target sequence Cys residue to Ala.

The tendency of a mutation to induce a gain of nonspecific association or aggregation in part reflects the complexity of protein folding and the linkage between various domains at the level of secondary, tertiary, or quaternary structure. The classic example of this is sickle cell hemoglobin, where a single mutation at the β6 position introduces a hydrophobic patch on the tetramer surface. This gives rise to an alternate polymer form, a heterologous rodlike polymer of tetramers, which then induces sickling of red blood cells and the pathophysiology of the disease. In the case of Smad proteins, mutations at the trimer interface are associated with disruption of homo-oligomerization. The majority of tumor-associated mutations in the R-Smad family map to this critical region of the protein and suggest a connection between the loss of homo-oligomerization and the loss of tumor supressor function (*1*). Sedimentation experiments on an interface mutant D407E (data not shown) did in fact reveal the presence of monomer, and the disruption of trimer formation. But surprisingly, the samples also have a 12S aggregate that was not sensitive to the presence of TCEP. Aggregates have also been observed in chromatography studies with a (3E) series of interface mutants (*4*). Thus, the disruption of trimer formation appears to introduce nonspecific aggregation that is strong enough to be observed at micromolar concentrations even in the presence of strong reducing agent. It is not known if this aggregation event has any direct in vivo consequences. However, it is very clear that loss of normal homo- and hetero-trimer formation is a critical aspect of the loss of receptor-mediated signaling and biological function, and the presence of aggregation does not change this interpretation.

Using a combination of sedimentation and equilibrium techniques, we have demonstrated that S4AF is a monomer, S3LC is a weakly associating trimer, and that pseudo-phosphorylation of S3LC enhances trimer formation by 17−35-fold. This directly verfies the importance of Smad3-target sequence phosphorylation in the activation process by TGF-β. In addition, we have demonstrated the ability of S4AF to hetero-trimerize with both wild-type and pseudo-phosphorylated S3LC, with phosphorylation enhancing complex formation. These results are consistent with the conclusion of Kawabata et al. (*3*) where the full-length Smad proteins were analyzed in mammalian cell extracts and appear to rule out the presence of hexamer formation in the signal transduction process. The pseudo-phosphorylation sequence employed in this work involves a conversion of three serines in the target sequence, $SSVS^{467}$, to glutamic acid, $EEVE^{467}$. The active in vivo phosphorylated form is believed to involved modifications at two target site serines, 465 and 467, with a critical role for serine 464 (*35*−*37*). Phosphorylation at serines 465 and 467 appears to be critical for release from the receptor. It is conceivable that homo- and hetero-trimer formation is stimulated by some combination of single, double, or triple phosphorylation at the target sequence serines. To investigate this question, we are currently constructing all seven pseudo-phosphorylated target sequence forms of S3LC. This includes the one zero site

[5] Funaba and Mathews (*38*) recently reported similar results on full-length Smad2 pseudo-phosphorylation constructs. The 2E construct formed a hetero-oligomer with Smad4 that was 3−4-fold tighter than the wild-type hetero-complex, a result consistent with the importance of phosphorylation in activating signal transduction, but significantly smaller than the 17−35-fold effect measured with our Smad3 constructs. These authors estimated the K_d for Smad4 dissociation from the hetero-oligomer (79 and 270 nM, respectively) by fitting fluorescence data to a one-site binding model. It appears they were measuring the overall formation of trimer from subunits, and obtained values that are qualitatively consistent with ours, although quantitatively weaker (Table 1). The difference may be intrinsic to full-length Smad2, or due to the influence of the dansyl aziridine used in the assay.

(wild type or S3LC), three one site, three two site, and one three site [S3LC(3E)] S to E mutations. These will be studied by sedimentation and chromatography techniques (*4*) to decipher the energetic role of each site in trimer activation. For example, the 3E mutation induces a 1.7−2.1 kcal enhancement of trimerization. Does this derive from a single site with the other two sites having no effect, or is there a cooperativity between sites, with each mutation increasing the effect in an additive manner? Preliminary sedimentation velocity experiments with the SEVE construct (+TCEP) reveal it trimerizes 7.7-fold tighter than the 3E construct,[5] suggesting that an E or phosphorylation at position 464 inhibits trimerization.

In addition to the events outlined above that describe the hetero-oligomerization of Smad4 and R-Smads, there are numerous other factors involved in Smad protein regulation (*39*, *40*). A recent report demonstrates that Smad2, -3, and -4 localize to cytoplasmic microtubules and suggests a novel mechanism of negative TGF-β regulation (*41*). Also within the cytoplasm, a Smad binding protein, SARA, binds to Smad2 and Smad3, controlling localization and transport to type I receptors for phosphorylation (*39*, *40*). There is also a set of inhibitory Smads, Smad6 and Smad7, that antagonize signaling (*42*), possibly by disrupting trimer formation. Once activated, the nuclear hetero-trimers interact with numerous activator and/or repressor molecules for selective gene regulation (*39*, *40*). This stage of the Smad regulation process appears to be remarkably complex and involves numerous nuclear factors. Future molecular studies of multi-complex formation and specific DNA binding should further contribute to an understanding of the important events in site-specific gene regulation.

ACKNOWLEDGMENT

We thank Walter Stafford for assistance with the HYDRO calculation and insightful discussions during this work. We also thank Sharon Lobert for critical proofing of the manuscript. All sedimentation experiments were performed in the University of Mississippi Medical Center Analytical Ultracentrifuge Facility.

REFERENCES

1. Shi, Y., Hata, A., Lo, R. S., Massague, J., and Pavletich, N. P. (1997) *Nature 388*, 87−93.
2. Hata, A., Shi, Y., and Massague, J. (1998) *Mol. Med. Today 4*, 257−262.
3. Kawabata, M., Inoue, H., Hanyu, A., Imanura, T., and Miyazono, K. (1998) *EMBO J.* 4056−4065.
4. Chacko, B., Qin, B., Lam, S. S., Correia, J. J., and Lin, K. (2000) *Nat. Struct. Biol.* (in press).
5. Liu, S., and Stafford, W. F., III (1995) *Anal. Biochem. 224*, 199−202.
6. Philo, J. S. (2000) *Anal. Biochem. 279*, 151−163.
7. Stafford, W. F., III (1992) *Anal. Biochem. 203*, 295−301.
8. Philo, J. (1994) in *Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems* (Schuster, T. M., and Laue, T. M., Eds.) pp 156−170, Birkhauser Boston, Inc., Cambridge, MA.
9. Philo, J. S. (1997) *Biophys. J. 72*, 435−444.
10. Correia, J. J. (2000) *Methods Enzymol. 321*, 81−100.
11. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge, U.K.
12. Frigon, R. P., and Timasheff, S. N. (1975) *Biochemistry 14*, 4559−4566.
13. Frigon, R. P., and Timasheff, S. N. (1975) *Biochemistry 14*, 4567−4573.
14. Garcia de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) *Biophys. J. 78*, 719−730.
15. Qin, B., Lam, S. S. W., and Lin, K. (1999) *Structure 7*, 1493−1503.
16. Schuck, P. (1999) *Anal. Biochem. 272*, 199−208.
17. Schuck, P., and Demeler, B. (1999) *Biophys. J. 76*, 2288−2296.
18. Johnson, M. L., Correia, J. J., Halvorson, H., and Yphantis, D. A. (1981) *Biophys. J. 36*, 575−588.
19. Correia, J. J., Gilbert, S. P., Moyer, M. L., and Johnson, K. A. (1995) *Biochemistry 34*, 4898−4907.
20. Yphantis, D. A., Correia, J. J., Johnson, M. L., and Wu, G.-M. (1978) in *Physical Aspects of Protein Interactions* (Catsimpoolas, N., Ed.) pp 275−303, Elsevier/North-Holland, Amsterdam, The Netherlands.
21. Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169−180.
22. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
23. Philo, J. S. (1997) *Proc. SPIE—Int. Soc. Opt. Eng. 2985*, 170−177.
24. Philo, J. S. (2000) *Methods Enzymol. 321*, 100−120.
25. Tada, K., Inoue, H., Ebisawa, T., Makuuchi, M., Kawabata, M., Imamura, T., and Miyazono, K. (1999) *Genes Cells 4*, 731−741.
26. Stafford, W. F., III (2000) *Methods Enzymol. 323*, 302−325.
27. Gilbert, L. M., and Gilbert, G. A. (1973) *Methods Enzymol. 27*, 273−196.
28. Cann, J. R. (1970) *Interacting Macromolecules: The Theory and Practice of Their Electrophoresis, Ultracentrifugation, and Chromatography*, Academic Press, New York.
29. Stafford, W. F., III (1994) in *Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems* (Schuster, T. M., and Laue, T. M., Eds.) pp 119−137, Birkhauser Boston, Inc., Cambridge, MA.
30. Trujillo, J. L., and Deal, W. C., Jr. (1977) *Biochemistry 16*, 3098−3104.
31. Rowe, R. J. (1977) *Biopolymers 16*, 2595−2611.
32. Hansen, J. C., Krieder, I. K., Demeler, B., and Fletcher, T. M. (1997) *Methods: Compan. Methods Enzymol. 12*, 62−72.
33. Stafford, W. F., III (1992) *Biophys. J. 74*, A301.
34. Jayaraman, L., and Massague, J. (2000) *J. Biol. Chem. 275*, 40710−40717.
35. Abdollah, S., Macias-Silva, M., Tsukazaki, T., Hayashi, H., Attisano, L., and Wrana, J. L. (1997) *J. Biol. Chem. 272*, 27678−27685.
36. Liu, X., Sun, Y., Constantinescu, S. N., Karam, E., Weinberg, R. A., and Lodish, H. F. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 10669−10674.
37. Souchelnytskyi, S., Tamaki, K., Engstrom, U., Wernstedt, C., ten Dijke, P., and Heldin, C.-H. (1997) *J. Biol. Chem. 272*, 28107−28115.
38. Funaba, M., and Mathews, L. S. (2000) *Mol. Endocrinol. 14*, 1583−1591.
39. Attisano, L., and Wrana, J. L. (2000) *Curr. Opin. Cell Biol. 12*, 235−243.
40. Massague, J., Blain, S. W., and Lo, R. S. (2000) *Cell 103*, 295−309.
41. Dong, C., Li, Z., Alvarez, R., Feng, X.-H., and Goldschmidt-Clermony, P. J. (2000) *Mol. Cell 5*, 27−34.
42. Hata, A., Lo, R. S., Wotton, D. Lagna, G., and Massague, J. (1997) *Nature 388*, 82−87.

BI0019343