Characterization of the Structure and Function of W → F WW Domain Variants: Identification of a Natively Unfolded Protein That Folds upon Ligand Binding†

Edward K. Koepf,‡ H. Michael Petrassi,‡ Gayathri Ratnaswamy,‡ Mary E. Huff,‡ Marius Sudol,§ and Jeffery W. Kelly*,‡

Department of Chemistry and The Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, MB 12, La Jolla, California 92037, and Department of Biochemistry, Mount Sinai School of Medicine, New York, New York 10029

Received May 13, 1999; Revised Manuscript Received August 25, 1999

ABSTRACT: The WW domain adopts a compact, three-stranded, antiparallel β-sheet structure that mediates protein−protein interactions by binding to xPPxY-based protein ligands, such as the PY-ligand (EYPPYPPPPYPSG) derived from p53 binding protein-2. The conserved Trp residues, after which this domain was named, were replaced with Phe so their importance in structural integrity and for ligand binding could be evaluated. A biophysical approach was employed to compare the W17F, W39F, and W17F/W39F WW domains to the wild-type protein. The data demonstrate that replacement of Trp39 with Phe (W39F) does not disrupt the structure of the WW domain variant, but does abolish ligand binding. In contrast, the W17F WW domain variant is largely if not completely unfolded; however, this variant undergoes a PY-ligand induced disorder to order (folding) transition. The dissociation constant for the W17F WW domain−PY-ligand interaction is 15.1 ± 1.2 μM, only slightly higher than that observed for the wild-type WW domain interaction (5.9 ± 0.33 μM). The W17F WW domain is a natively unfolded protein which adopts a native conformation upon PY-ligand binding.

The WW domain of the human Yes-associated protein (hYAP)[1] is a small, 38-residue, independently folded protein module which adopts a three-stranded, antiparallel β-sheet conformation (Figure 1). Initially discovered by computer-aided analysis of a short imperfectly repeated sequence in the mouse isoform of YAP, the WW domain has since been identified in nearly 170 unrelated proteins which exhibit a broad spectrum of biological functions (*1*−*3*). Named for the two highly conserved tryptophan residues that are spaced 20−22 residues apart in the sequence, the WW domain is implicated in mediating protein−protein interactions by binding polyproline-rich ligands. These ligands have a minimal consensus sequence of xPPxY such as the PY-ligand used in this study (EYPPYPPPPYPSG), where P is proline, Y is tyrosine, and x is any other amino acid (*4*). The YAP WW domain is a transcriptional coactivator that functions by activating target genes (*5*). The discovery of the xPPxY ligand motif or the WW domain in the activation domains of many transcription factors suggests that this interaction is important in transcription.

Although they are functionally similar to the SH3 domain which binds polyproline-rich ligands with the sequence PxxP, recent high-resolution structural studies demonstrate that the SH3 and WW domains have unique structures. The WW domain fold is a bent and slightly twisted three-stranded, antiparallel β-sheet (Figure 1), while the SH3 domain is a compact β-barrel composed of two orthogonal β-sheets made up of five antiparallel β-strands (*6*−*10*). The NMR solution structure of the hYAP WW domain in complex with a proline-rich peptide ligand (GTPPPPYTVG) (*6*) and preliminary binding studies involving potential interactions between mutant GST−WW fusion proteins and GST−xPPxY-rich ligands identified a number of WW domain residues that appear to be essential for ligand binding (*11*). The Trp39 residue, which is located on the concave face of the WW β-sheet (Figure 1), defines the xPPxY ligand binding surface together with Tyr28, Leu30, His32, and the rest of the residues of β-strand 3, making contacts with the central prolines (Pro4 and Pro5) of the peptide ligand (*6*). Introduction of Phe at position 39 appears to abolish xPPxY ligand binding, presumably as a consequence of disrupting the critical intermolecular contacts identified in the structure of the WW domain−xPPxY complex (*6*, *11*). In contrast, Trp17, which is located on the convex face of the β-sheet opposite the ligand binding surface, does not make any intermolecular contacts with the peptide ligand. Substitution of Trp17 with Phe in the context of the GST fusions did not

† We gratefully acknowledge primary financial support from the National Institutes of Health (Grant R01GM51105) and secondary support from The Skaggs Institute of Chemical Biology and the Lita Annenberg Hazen Foundation.

* To whom correspondence should be addressed. Telephone: (858) 784-9605. Fax: (858) 784-9610. E-mail: jkelly@scripps.edu.

‡ The Scripps Research Institute.

§ Mount Sinai School of Medicine.

[1] Abbreviations: buffer A, 10 mM potassium phosphate (pH 6.0), 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, and 0.02% NaN_3; DabcylSE, 4-{[4-(dimethylamino)phenyl]azo}benzoic acid succinimidyl ester; GST, glutathione *S*-transferase; HPLC, high-performance liquid chromatography; hYAP, human Yes-associated protein; ITC, isothermal titration calorimetry; IPTG, isopropyl β-D-thiogalactopyranoside; NMP, *N*-methylpyrrolidone; PCR, polymerase chain reaction; PMSF, phenylmethanesulfonyl fluoride; PY, synthetic 13-residue proline-rich internal peptide from p53 binding protein-2 having the sequence EYPPYPPPPYPSG; WW, wild-type Trp-Trp domain; W17F, WW domain with a Trp17 to Phe mutation; W39F, WW domain with a Trp39 to Phe mutation; W17F/W39F, WW domain having both Trp17 and Trp39 mutated to Phe.

10.1021/bi991105l CCC: $18.00
Published on Web 10/08/1999

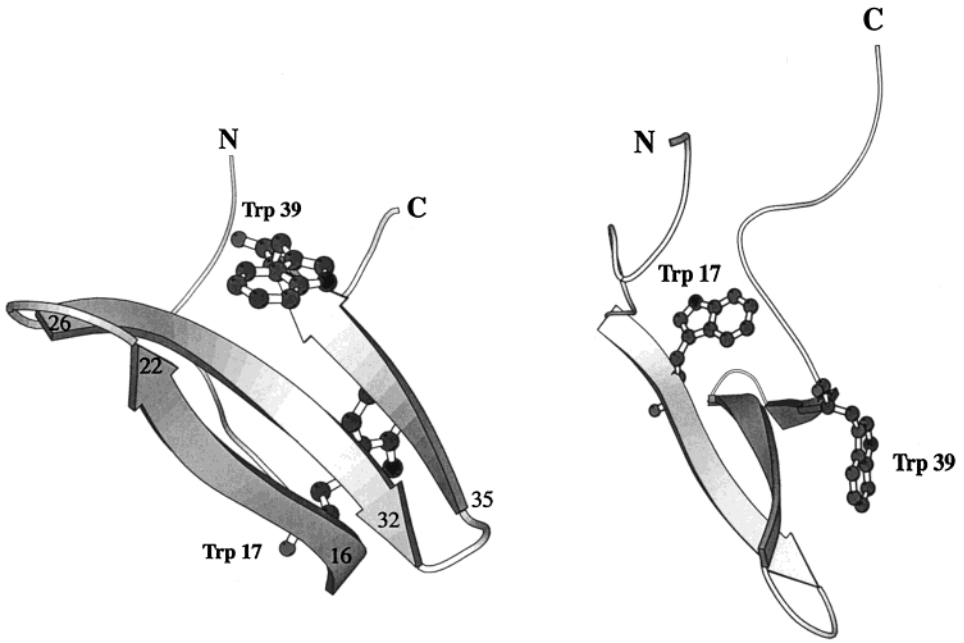

FIGURE 1: Ribbon diagram representations of two views of the isolated WW domain (57 residues) from the human YAP protein illustrating the side chains of Trp17 and Trp39 in a ball-and-stick representation. Molscript was used to prepare the figure (*40*) using coordinates provided by Macias et al. (*6*).

Table 1: Sequence Composition and Molecular Masses (Acquired by Electrospray Ionization Mass Spectrometry) of the Wild-Type, W39F, W17F, and W17F/W39F WW Domains Used in This Study

	Sequence	Mass (Da)
Wild Type	GSMSFEIPDDVPLPAG**W**EMAKTSSGQRYFLNHIDQTTT**W**QDPRKAMLSQMNVTAPTS	6372.2
W39F	GSHMSFEIPDDVPLPAG**W**EMAKTSSGQRYFLNHIDQTTTFQDPRKAMLSQMNVTAPTS	6470.7
W17F	GSQSSFEIPDDVPLPAGFEMAKTSSGQRYFLNHIDQTTT**W**QDPRKAMLSQMNVTAPTSRIHRD	7095.5
W17FW39F	GSHMSFEIPDDVPLPAGFEMAKTSSGQRYFLNHIDQTTTFQDPRKAMLSQMNVTAPTS	6431.3

noticeably attenuate binding of the WW domain to the xPPxY ligand (*11*). On the other hand, Trp17 does participate in the formation of a hydrophobic buckle through intramolecular interactions with Pro14 and Pro42, which may be essential for maintaining the structural integrity of the WW domain (*6*).

To understand the relationship between the conserved Trp residues and the structural and functional (xPPxY ligand binding) properties of the WW domain, we used a biophysical approach to compare the W17F, W39F, and W17F/W39F WW domains to the wild-type WW protein. The results herein demonstrate that the W39F variant adopts a native-like conformation but has lost its ligand binding activity, unlike the W17F variant which appears to be largely unstructured but retains its ability to bind to a xPPxY-based ligand. Upon ligand binding, the W17F WW domain folds into a well-defined WW domain structure.

A generally held view concerning protein−ligand recognition is that well-defined structural elements are required for ligand binding. However, recent examples of ligands binding to largely, if not completely unordered, proteins suggest that folding and binding can be linked in some cases (*12*−*16*). Characterization of unfolded proteins and theoretical folding studies support the existence of residual structure in the denatured state(s), which may retain the ability to recognize a ligand (*17*, *18*). Alternatively, ligands could bind to the transition state(s) associated with folding and to the ground state, inducing folding primarily by lowering the activation barrier (*19*). Further investigations into linked binding and folding transitions are necessary for understanding this potentially important mode of binding.

EXPERIMENTAL PROCEDURES

Protein Expression Systems and Purification. The recombinant wild-type WW domain used in this study was derived from the human Yes-associated protein (hYAP), GenBank accession number X80507, and was purified according to previously published protocols (*6*, *20*). The wild-type WW domain is comprised of 57 residues with a mass of 6372.2 Da (calculated mass of 6372.2 Da) and an N-terminal sequence of GSMSFE (Table 1). The concentration of this protein was determined using its extinction coefficient of 12 660 M^{-1} cm^{-1} at 280 nm (*21*, *22*).

The W39F WW domain expression system was generated by polymerase chain reaction (PCR) using W39F cDNA in pGEX-2TK, a construct used by the Sudol group (Amersham

Pharmacia Biotech, Piscataway, NJ) as the template source (see the Supporting Information for details) (*11*). The DNA generated in the PCR was cloned into the *Nde*I and *Bam*HI sites of pET-15b (Novagen, Madison, WI) and expressed in *Escherichia coli* BL21(DE3) Epicurian Gold cells (Stratagene, La Jolla, CA). The presence of the W39F mutation in this construct as well as the mutations in all the other constructs that were made was confirmed by DNA sequencing. The cell cultures used to produce the various recombinant WW domains were grown for 5−6 h at 37 °C after induction with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). The target W39F WW domain, expressed as a N-terminal His-tagged fusion protein, was separated from the *E. coli* proteins by passing the total bacterial cell lysate through Ni−NTA agarose resin (Qiagen, Inc., Valencia, CA). Elution of the protein from the resin was carried out using imidazole-based buffers as recommended by the manufacturer. Following overnight dialysis against 25 mM Tris-HCl (pH 8.0, 100 mM NaCl), the His tag was cleaved from the target W39F WW domain by thrombin proteolysis. Bovine thrombin (Sigma T4648, St. Louis, MO) was added to the protein at a ratio of 2 units of enzyme per milligram of His-tagged fusion protein and allowed to react for 4 h at room temperature. The target protein was purified to homogeneity by gel permeation chromatography on a HiLoad 16/60 Superdex 30 column (Amersham Pharmacia Biotech). The first six amino acids at the N-terminus of this 58-residue W39F WW variant are GSHMSF according to sequencing (Table 1). The protein has a mass of 6470.7 Da (calculated mass of 6470.3 Da), and an extinction coefficient of 7020 M^{-1} cm^{-1} at 280 nm was used to determine its concentration.

The expression system for the W17F WW variant was constructed by subcloning the W17F DNA in pGEX-2TK into pGEX-2T (Amersham Pharmacia Biotech) (*11*). The pGEX-2TK vector produces a GST−W17F fusion protein with five additional residues added to the N-terminus of the target peptide following thrombin proteolysis. By subcloning the *Bam*HI−*Eco*RI DNA fragment from the pGEXT-2TK construct into pGEX-2T, we generated a 63-residue protein with a mass of 7095.5 Da (calculated mass of 7095.0 Da) having an N-terminal sequence of GSQSSF (Table 1). Purification of the GST−W17F fusion protein was achieved by passing the total *E. coli* BL21(DE3) cell lysate through glutathione−Sepharose 4B affinity resin (Amersham Pharmacia Biotech) as recommended by the manufacturer. Following extensive column washing, the fusion protein was eluted from the resin with a solution of 10 mM reduced glutathione in 50 mM Tris-HCl (pH 8.0). The peptide was cleaved from the GST binding domain of the fusion protein using the thrombin proteolysis procedure described in detail above. Following proteolysis, phenylmethanesulfonyl fluoride (PMSF) was added to the reaction mixture to a final concentration of 1 mM. Final peptide purification was achieved by gel permeation chromatography as described previously. To minimize endogenous proteolysis, cell lysis and glutathione chromatography were carried out in the presence of 1 mM PMSF (4 °C). The W17F protein concentration was determined by absorbance measurements taken at 280 nm using its extinction coefficient of 7020 M^{-1} cm^{-1}.

The W17F/W39F WW double mutant DNA was synthesized using the QuickChange Site-Directed Mutagenesis protocol from Stratagene employing W17F DNA in pET 15b as the template (see the Supporting Information for details). Protein expression, enzymatic cleavage, and purification were carried out using the protocol described in detail for the W39F WW domain. The double mutant WW variant consisted of 58 residues having a molecular mass of 6431.3 Da (calculated mass of 6431.3 Da) and an N-terminal sequence of GSHMSF (Table 1). The peptide concentration was determined with a Coomassie dye binding assay (Pierce, Rockford, IL) using the wild-type WW domain to prepare the standard curve.

PY-Ligand Synthesis (EYPPYPPPPYPSG) and Peptide Purification. All reagents used for peptide synthesis were purchased from either Fisher, Aldrich, or Applied Biosystems Inc. and were used without further purification. Amino acids and Rink amide resin were acquired from Nova Biochem (Calbiochem-Novabiochem Corp., La Jolla, CA), while the PY-acylating agent 4-{[4-(dimethylamino)phenyl]azo}-benzoic acid succinimidyl ester (DabcylSE) was procured from Molecular Probes (Eugene, OR). Automated peptide synthesis using the FMOC strategy was performed on an ABI 433A Peptide Synthesizer (Applied Biosystems, Inc., Foster City, CA) using Fmoc-Glu(*t*-Butyl), Fmoc-Ser(*t*-Butyl), and Fmoc-Tyr(*t*-Butyl) side chain-protected amino acids. Upon completion of the automated synthesis, the N-terminal FMOC protecting group was removed. In one of the two automated syntheses, the resin was transferred to a manual peptide synthesis vessel and swelled for 30 min in approximately 5 mL of *N*-methylpyrrolidone (NMP). Acylation at the N-terminus of the PY-peptide was carried out by the addition of 147 mg of DabcylSE (0.4 mM, 4 equiv) and 70 μL of diisoproplyethylamine (0.44 mM, 1.1 equiv per DabcylSE) to the peptide synthesis vessel. The dabcylation reaction proceeded for 16 h at room temperature with vigorous agitation using a mechanical shaker. The resin was washed seven times each with 5 mL of NMP and subjected to the cleavage protocol outlined below. Both labeled and unlabeled peptides were cleaved from the resin and deprotected with 10 mL of 95% trifluoroacetic acid, 2% triisopropyl silane, 2% water, and 1% ethanedithiol. After 3 h, the cleavage−deprotection reaction mixture was passed through a sintered glass frit into a 50 mL round-bottom flask, and the filtrate was placed under reduced pressure to remove the volatile components of the reaction cocktail. After the contents of the round-bottom flask had been cooled to 0 °C for 10 min, peptide precipitation was carried out by the addition of 20 mL of methyl *tert*-butyl ether and subsequent storage at −20 °C for approximately 12 h. The precipitated peptide suspension was transferred to 15 mL polyethylene falcon tubes and pelleted to a white solid using a benchtop centrifuge. The solid was resuspended in 10 mL of diethyl ether and centrifuged again, repeating the cycle five times to remove all cleavage and deprotection reagents from the precipitate. The white solid was taken up in 10% acetic acid, lyophilized, and purified via high-performance liquid chromatography (HPLC).

Preparative HPLC was carried out on a Waters 600E Multisolvent Delivery System (Waters Corp., Milford, MA) using a VYDAC (Hesperia, CA) C18 protein and peptide reverse phase column (218TP1022, 300 Å, 5 μm pore size, 22 mm inside diameter × 250 mm). Detection was accomplished by absorbance at both 230 and 280 nm using a

Waters 486 tunable absorbance detector. Purification of the unlabeled PY-peptide was achieved using a linear (15 to 30% B) gradient over the course of 25 min, while purification of the dabcylated peptide was accomplished with a linear (20 to 80% B) gradient over the course of 45 min, both purifications being performed at a flow rate of 12 mL/min. Solvent A consisted of 95% H_2O, 4.8% acetonitrile, and 0.2% trifluoroacetic acid, while solvent B was composed of 95% acetonitrile, 4.8% H_2O, and 0.2% trifluoroacetic acid. Fractions containing pure peptide were pooled, placed under reduced pressure to remove the volatile solvent, and lyophilized to yield a white powder which was analyzed by analytical HPLC, ^{1}H NMR, and electrospray mass spectrometry. The molecular mass of the PY-ligand peptide was determined to be 1459.63 Da (calculated mass of 1459.60 Da), and the mass of the dabcylated PY-peptide was 1710.73 Da (calculated mass of 1710.88 Da).

Fluorescence Spectroscopy. The ability of the wild-type WW domain as well as the W17F, W39F, and W17F/W39F WW variants to bind the PY-ligand was characterized by monitoring the changes in the intrinsic fluorescence properties of the WW domains upon the addition of the 13-residue PY-ligand. Emission spectra of 5 μM samples of each of the WW domains in buffer A (pH 6.0) were recorded with an Aviv model ATF105 automated titrating differential/ratio spectrofluorometer (Lakewood, NJ), both in the presence and in the absence of a 50 μM PY-ligand solution. All Trp-containing proteins were analyzed at 25 °C using an excitation wavelength of 295 nm with the excitation and emission bandwidths set to 2 nm. A set of emission spectra with the apo and PY-ligand-saturated forms of the four WW domain constructs were also acquired in the presence of 6.0 M GdnHCl.

Circular Dichroism Spectroscopy. All CD measurements were performed on an Aviv model 202SF stopped flow circular dichroism spectrometer (Lakewood, NJ) equipped with a Peltier temperature-controlled cell holder. Far-UV CD spectra were recorded using 50 μM solutions of the WW domains in buffer A (pH 6.0) employing a 0.1 cm path length Suprasil quartz cell (Hellma, Forest Hills, NY), while near-UV CD measurements were taken with 100 μM WW domain solutions in a 1.0 cm path length quartz cell. The spectra were recorded in 0.5 nm increments using an integration time of 1 s. Each spectrum presented in Figures 3−6 represents an average of three consecutive scans measured at 25 °C which have been corrected by subtracting a buffer blank. The far-UV CD titrations (buffer A, pH 6.0) of the WW domains (50 μM) with added PY-ligand (100 μM) are presented in Figure 4. Analogous near-UV CD spectra of WW domains (100 μM) with added PY-ligand (200 μM) are shown in Figure 6.

^{1}H NMR Spectroscopy. The ^{1}H NMR experiments were recorded on either a Bruker DMX 750 MHz or an AMX 600 MHz spectrometer. The ID spectra were acquired with a spectral width of 12 000 Hz at an operating frequency of 750 MHz or 10 000 Hz at a frequency of 600 MHz using 8192 points at 5 °C. Water suppression was achieved using the Watergate pulse sequence (*23*). All spectra were referenced relative to the methyl proton resonances of the internal standard 3-(trimethylsilyl)propionate-2,2,3,3-d_4 (Aldrich, Milwaukee, WI) at 0.0 ppm. The data were processed using XWIN NMR software version 6.0 (Bruker), using a line broadening parameter of 3 Hz. All the samples were evaluated at a concentration of 600 μM in buffer A (pH 6.0). PY-ligand was added directly to the wild-type, W17F, and W39F WW domains to yield samples containing 0.5 (data not shown) and 2 molar equiv of ligand relative to WW domain. After the addition of each aliquot of PY-peptide ligand, the sample was mixed manually and returned to the spectrometer for spectral acquisition within 5 min.

Analytical Ultracentrifugation. Sedimentation equilibrium measurements were taken on a temperature-controlled Beckman XL-I analytical ultracentrifuge equipped with an An60Ti rotor and photoelectric scanner (Beckman Instruments, Palo Alto, CA). The sample compartment of a double-sector cell fitted with a 12 mm Epon centerpiece and sapphire windows was charged with 140 μL of sample using a blunt-end Hamilton syringe, while 160 μL of protein dialysis buffer was loaded into the reference compartment of the cell. Data collection was initially carried out with a rotor speed of 3000 rpm at 20 °C, and then at speeds between 30 000 and 40 000 rpm using a radial step size of 0.001 cm. Absorbance measurements were made over the λ range of 470−520 nm until equilibrium had been established, determined by overlaying absorbance scans taken at 3 h intervals during the course of the sedimentation equilibrium run (typically 18−24 h). The data were analyzed by a nonlinear least-squares approach using the Origin software module (Microcal Software, Inc., Northhampton, MA) provided by Beckman. The data were fit to either a single-ideal species model or a two-species nonassociating model (eqs 1 and 2, respectively)

$$A_r = \exp\{\ln(A_0) + [M\omega^2(1 - \bar{v}\rho]/2RT)(x^2 - {x_0}^2)\} + E \quad (1)$$

$$A_r = \exp\{\ln(A_0) + [M\omega^2(1 - \bar{v}\rho)/2RT](x^2 - {x_0}^2)\} + \exp\{\ln(A_1) + [M\omega^2(1 - \bar{v}\rho)/2RT)(x^2 - {x_0}^2)\} + E \quad (2)$$

where A_r is the sample absorbance at radial position x, A_0 and A_1 are the absorbances at a reference radius, x_0 (generally the meniscus), $\bar{v}$ is the partial specific volume of the protein (milliliters per gram), ρ is the density of the solvent (grams per milliliter), ω is the angular velocity of the rotor (radian per second), E is a baseline error correction factor, M is the molecular weight, R is the universal gas constant (8.314×10^7 erg/mol), and T is the temperature (kelvin). The partial specific volumes of the dabcyl-PY−WW and dabcyl-PY−W17F binary complexes (0.722 and 0.721 mL/g, respectively) were calculated on the basis of amino acid composition according to the method of Cohn and Edsall as implemented in the XL-I software (*24*), while 1.0 g/mL was used for ρ. The goodness of the fit was evaluated on the basis of the randomness and magnitude of the residuals, expressed as the difference between the theoretical curve and the experimental data, and by checking the fit parameters for physical reasonability.

Determination of Binding Constants by Isothermal Titration Calorimetry. The binding constant for the PY-ligand interacting with the wild-type, W17F, and W39F WW domains was measured using a MicroCal MCS isothermal titration calorimeter (ITC) (MicroCal Inc.). All samples were dialyzed against buffer A (pH 6.0), filtered, and degassed by stirring under vacuum prior to use. The PY-ligand at 1.0 mM was titrated into the ITC cell (volume of 1.3407 mL)

containing a solution of either 50 μM wild-type WW domain, 44.5 μM W17F WW domain, or 50 μM W39F WW domain at 25 °C. A total of 32 consecutive 5 μL injections were performed following an initial 2 μL PY-peptide injection which was discarded during data analysis. A total of five independent titration experiments were conducted to determine the binding constant for the PY-ligand to the wild-type WW domain, while four runs were carried out with the W17F and W39F WW domain variants. Titration data liberating heat as a result of ligand binding were corrected for dilution artifacts by subtracting the heat response acquired for the PY-ligand titrated into buffer (heat of the dilution of the ligand) and by subtracting the heat generated from buffer titrated into the protein solution (heat of dilution of the protein sample). Equilibrium association constants (reported as dissociation constants) and binding stoichiometries were determined by fitting the calorimetric data to a single-binding site model described by eq 3 using the ITC Data Analysis Module in Origin version 5.0 provided by MicroCal. The total heat $\Delta Q(i)$ released by the solution at the end of the ith injection is given by

$$\Delta Q(i) = Q(i) + \frac{dV_i}{V_0}\left[\frac{Q(i) + Q(i-1)}{2}\right] - Q(i-1) \quad (3)$$

where Q (see equation below) is related to n, the number of ligand binding sites, V_0, the cell volume, and M_t and X_t, the bulk concentrations of the protein and the ligand in the calorimeter cell, respectively (ITC Data Analysis in Origin, 1993, MicroCal Inc.).

$$Q = \frac{nM_t\Delta HV_0}{2}\left[1 + \frac{X_t}{nM_t} + \frac{1}{nKM_t} - \sqrt{\left(1 + \frac{X_t}{nM_t} + \frac{1}{nKM_t}\right)^2 - \frac{4X_t}{nM_t}}\right]$$

No heat in addition to the heat of dilution of the PY-ligand was observed for the W39F titration.

RESULTS

Fluorescence Properties of the WW Domains. The fluorescence spectrum of a 5 μM solution of the wild-type WW domain in buffer A (pH 6.0) is typical for a tryptophan-containing protein in an aqueous solvent, exhibiting an emission maximum (λ_{max}) at 342 nm upon excitation at 295 nm (Figure 2a) (*25*, *26*). When this measurement is repeated in the presence of 6.0 M GdnHCl, the λ_{max} of the WW domain red shifts 13 nm to 355 nm, and the quantum yield at that wavelength diminishes by nearly 57% relative to the maximum emission intensity observed at 342 nm under nondenaturing conditions. The observed differences indicate a structural unraveling of the WW domain β-sheet architecture in 6.0 M GdnHCl, facilitating exposure of the Trp chromophores to solvent-induced collisional quenching (*20*). Addition of a 10-fold molar excess of PY-ligand to the folded WW domain is accompanied by a nearly 2-fold increase in the Trp emission intensity and a 6 nm blue shift in λ_{max} (Figure 2a), observations consistent with the WW domain binding the PY-ligand and placement of Trp39 into a more hydrophobic environment.

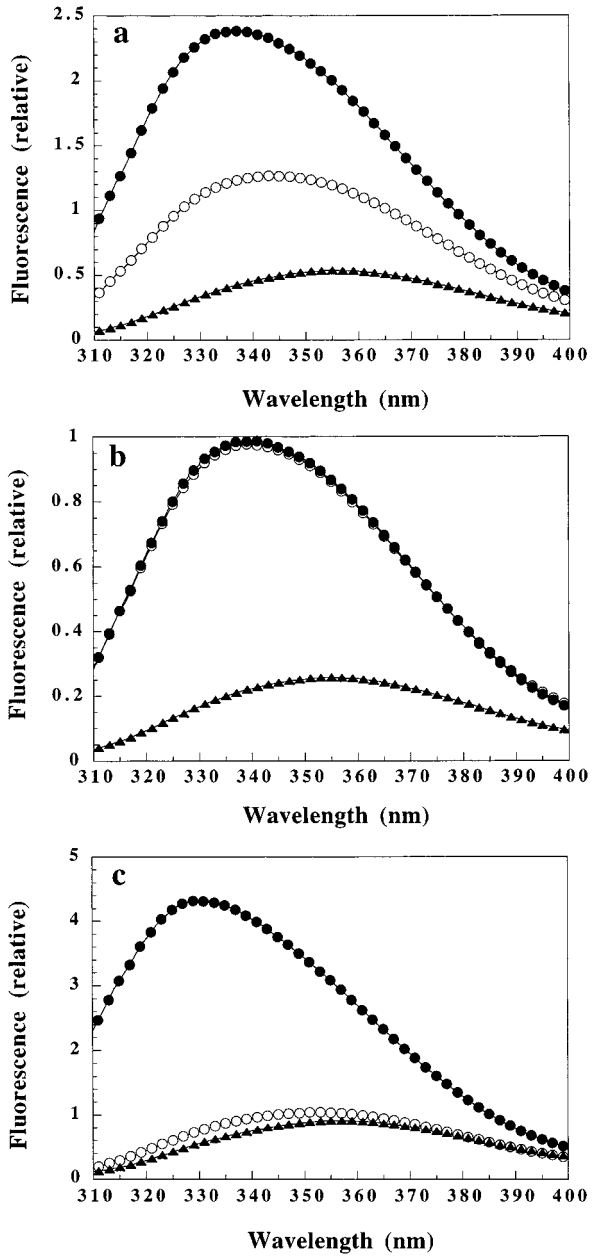

FIGURE 2: Fluorescence emission spectra of 5 μM solutions of the WW domains in buffer A with and without PY-ligand and in 6.0 M GdnHCl in buffer A. (a) Wild-type WW domain in buffer A (○), with 50 μM PY-peptide ligand added (buffer A) (●), and in 6.0 M GdnHCl (▲). Strictly analogous experiments were carried out with W39F (b) and W17F (c). Emission spectra were collected using an excitation wavelength of 295 nm and corrected by subtracting a buffer blank.

Fluorescence studies on W39F (buffer A, pH 6.0) reveal an emission spectrum which is similar to that of the wild-type domain, with the exception of a slight blue shift in the λ_{max} from 342 (wild-type) to 339 nm (Figure 2b). In 6.0 M GdnHCl, the λ_{max} of the W39F WW domain red shifts to 354 nm, similar to the response of the wild-type domain. However, the reduction in the quantum yield for the W39F variant is much more pronounced (74%), relative to the 57% reduction for the wild-type WW domain (Figure 2b). Addition of the PY-ligand to the W39F WW domain under native conditions (buffer A, pH 6.0) did not alter the fluorescence characteristics of this variant at all (Figure 2b). Neither a change in the quantum yield nor a shift in λ_{max}

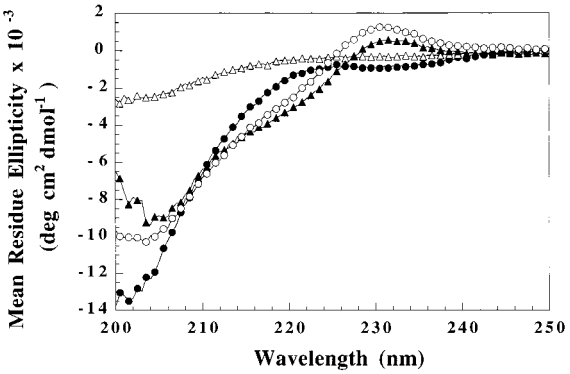

FIGURE 3: Far-UV circular dichroism spectra of 50 μM samples of wild-type (○), W39F (▲), W17F (●), and W17F/W39F (△) WW domains in buffer A (pH 6.0) acquired at 25 °C using a 0.1 cm path length quartz cell. Each spectrum is the average of three consecutive scans which have been corrected by subtracting a buffer blank.

was detected, suggesting the inability of the W39F WW domain to bind the PY-ligand. This experiment does not prove the inability of the W39F WW domain to bind the PY-ligand, but the interpretation is consistent with more conclusive NMR, isothermal titration calorimetry, CD, and ultracentrifuge data, all indicating a lack of binding (vide infra).

Application of the aforementioned experiments to the W17F WW variant revealed some unique fluorescence responses, in both native and chaotrope-containing solutions, as well as in the presence of the PY-ligand (Figure 2c). A 5 μM solution of the W17F WW domain in buffer A (pH 6.0) exhibits a single broad Trp39 emission band with a λ_{max} centered at 353 nm. When the fluorescence spectra were analyzed in 6.0 M GdnHCl, the λ_{max} of this variant red shifted only slightly to 355 nm and the quantum yield decreased by only 13%. These data are consistent with a W17F WW domain that is largely if not completely unfolded under solution conditions that yield a folded wild-type WW domain (see additional data below). Addition of the PY-ligand to the W17F WW domain in buffer A (pH 6.0) induces dramatic changes in the fluorescence emission spectrum in less than 20 s (dead time of manual mixing) (Figure 2c). In the presence of 50 μM PY-ligand, the fluorescence quantum yield of the W17F domain increased by more than a factor of 4 and the λ_{max} blue shifted 24 nm from 353 to 329 nm (Figure 2c), consistent with ligand-induced WW domain folding (vide infra).

Analysis of the W17F/W39F WW domain by Trp fluorescence spectroscopy was not possible, as it lacks Trp. However, excitation at 280 nm allowed Tyr fluorescence at 304 nm to be monitored, which was not significantly affected by either the absence or presence of 6.0 M GdnHCl (Figure 1 of the Supporting Information), consistent with an unfolded structure.

Secondary Structure of the WT and the Trp to Phe WW Domain Variants As Determined by Far-UV CD. The far-UV CD spectrum of the wild-type WW domain presented in Figure 3 exhibits two characteristic features that are associated with the three-stranded β-sheet of this fold: a weak maximum centered at 230 nm with a mean residue ellipticity of 1.2×10^{-3} deg cm^2 dmol^{-1} and a much more pronounced minimum (-10.1×10^{-3} deg cm^2 dmol^{-1}) at 202 nm (*20*). The far-UV CD of the W39F variant resembles the overall spectrum exhibited by the wild-type WW domain, although small variations in the intensities and positions of the spectral minimum and maximum are evident (Figure 3) and probably result from the loss of Trp39 which contributes to the far-UV CD spectrum. In contrast, the far-UV CD spectrum of the W17F domain lacks the maximum near 230 nm characteristic of a folded WW domain, and most closely resembles a CD spectrum representative of an ensemble of unordered conformations with a minimum near 200 nm (Figure 3) (*27*, *28*). Importantly, the 230 nm maximum is recovered upon ligand-induced folding (vide infra), demonstrating that the loss of Trp17 is not directly responsible for the loss of this signal. The construct devoid of both Trp residues, the W17F/W39F WW domain, produces a far-UV CD spectrum which is characterized by only a single broad and very shallow negative peak with a minimum around 200 nm (Figure 3). Clearly, the Trp residues contribute significantly to the far-UV CD spectrum of the WW domain.

PY-Ligand Binding to the WT and Variant WW Domains As Determined by Far-UV CD. Addition of 100 μM PY-ligand to a 50 μM solution of wild-type WW domain results in a dramatic change in the far-UV CD spectrum, especially near the 230 nm maximum characteristic of the WW domain (Figure 4a). The minimum also shifts from 205 to 208 nm. The PY-ligand in the absence of the wild-type WW domain exhibits a significant CD signal with positive and negative absorptions centered near 228 and 208 nm, respectively, consistent with a polyproline type II helical conformation (*29*). It is important to note that the simple linear sum of the WW spectrum with the spectrum of the PY-ligand (Figure 4a) does not account for the CD spectrum generated by the WW−PY-ligand complex. Deviations in λ and in the intensity of the complex are observed at both the maximum (232 nm) and the minimum (208 nm) relative to the values for WW itself or the sum of the WW and PY spectra.

In contrast to the previous example (wild-type) and the W17F case discussed below, addition of 100 μM PY-ligand to a 50 μM solution of W39F generates spectra not exhibiting deviations in λ or intensity above 208 nm relative to the spectra generated from a simple linear combination of the W39F and PY-ligand spectra (Figure 4b). This is consistent with the inability of the W39F WW domain to bind to the PY-ligand and to form a complex exhibiting differential absorption of circularly polarized light.

Addition of 100 μM PY-ligand to a 50 μM solution of the W17F WW domain results in the re-emergence of positive ellipticity in the 225−230 nm range characteristic of a folded WW domain, along with a shift in the position of the minimum from approximately 202 to 208 nm (Figure 4c). Significant is the fact that the simple linear combination of the W17F spectrum with the spectrum of the PY-ligand (Figure 4c) does not account for the CD spectrum generated by the W17F−PY-ligand complex. Clearly, additional conformational changes are induced when binding occurs between the W17F WW domain and the PY-ligand as noted above, producing the spectral characteristics of a folded WW domain, with additional broadening of the minimum as a result of the of the polyproline helical contribution.

Probing the Trp Asymmetric Environment by Near-UV CD. The wild-type WW domain exhibits a near-UV CD spectrum that is characterized by double minima at 282 and

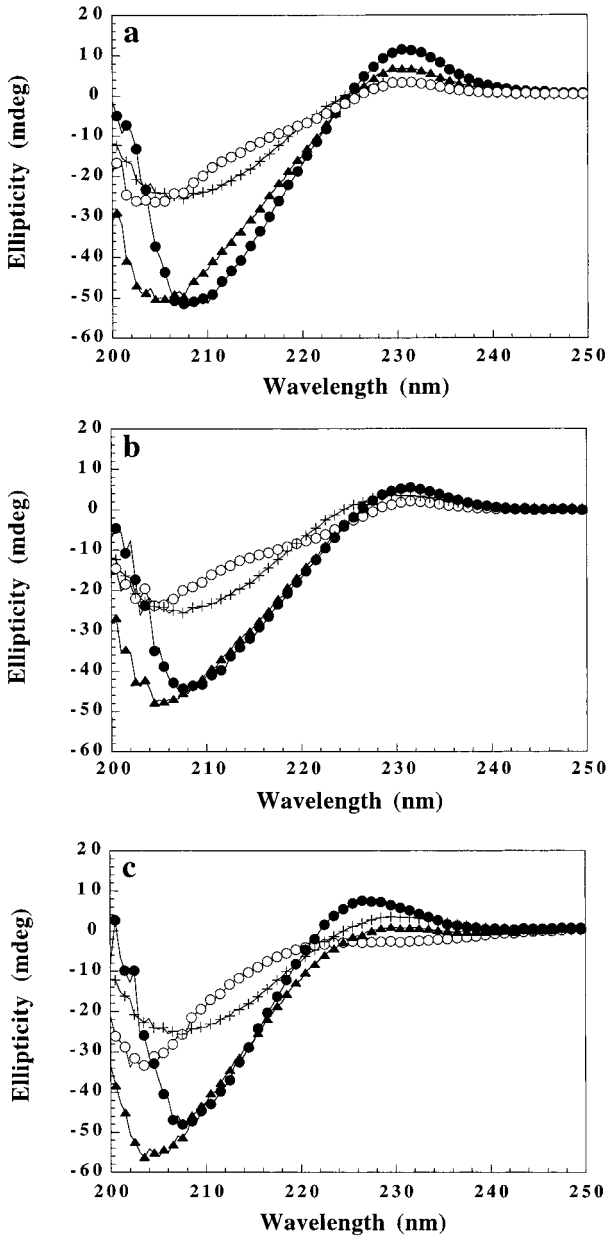

FIGURE 4: Addition of PY-ligand (EYPPYPPPPYPSG) to the wild-type and variant WW domains monitored by far-UV circular dichroism spectroscopy. (a) Far-UV CD spectra of 50 μM wild-type WW (○), 100 μM PY-ligand (+), and a combination of the wild-type WW domain (50 μM) and the PY-ligand (100 μM) (●), all in buffer A (pH 6.0). The black triangles (▲) represent the hypothetical spectrum resulting from summation of the apo wild-type WW and the apo PY-ligand CD spectra. (b) Strictly analogous experiment carried out with the W39F WW domain. (c) Strictly analogous study carried out with the W17F variant.

290 nm and by a positive absorbance centered near 260 nm (Figure 5) (*20*). The near-UV CD spectrum of the W39F WW domain exhibits the 282 and 290 nm minima, but lacks the maximum centered near 260 nm characteristic of the wild-type protein (Figure 5), suggesting that the double minimum originates from Trp17. The near-UV CD spectrum of the W17F WW domain differs significantly from the wild-type and W39F spectra and is characterized by a single, very broad, positive absorption extending over the 250−300 nm region, suggesting Trp39 is the main contributor to the maximum observed in the wild-type spectra. The fourth variant in this series, the W17F/W39F WW domain, produces

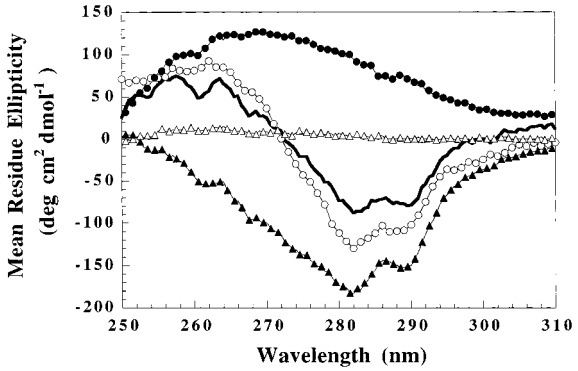

FIGURE 5: Near-UV circular dichroism spectra of wild-type (○), W39F (▲), W17F (●), and W17F/W39F (△) WW domains (100 μM) in buffer A (pH 6.0). All spectra were acquired at 25 °C using a 1.0 cm path length quartz cell. The solid line represents the sum of the W17F and W39F near-UV CD spectra.

virtually no near-UV CD signal and is indistinguishable from a buffer blank (Figure 5), demonstrating that the Trp residues are the predominant contributors to the near-UV CD spectrum of the wild-type WW domain. Summation of the W39F and W17F near-UV CD spectra yields a resultant spectrum that is similar to the wild-type spectra, implying that the major contributors to the near-UV CD maximum (Trp39) and minimum (Trp17) have been elucidated.

PY-Ligand Binding to the WT and Variant WW Domains As Determined by Near-UV CD. Addition of 2 equiv of the PY-ligand to 100 μM wild-type WW domain results in an enhancement of the intensity of the double minimum and the maximum (Figure 6a). The spectrum of the complex is dramatically different from the sum of the near-UV CD spectra of the wild-type WW domain and the PY-ligand, indicating complex formation with a unique spectral signature. A small but very reproducible near-UV CD signal is seen for the free PY-ligand, providing evidence that this peptide is at least partially structured on its own in solution.

On the contrary, the analogous addition of PY-ligand to the W39F WW domain yields a near-UV CD spectrum that is identical with that afforded by a linear sum of the apo W39F signal with the PY-ligand signal over the near-UV CD range (Figure 6b). This observation is inconsistent with W39F−PY-ligand complex formation.

Addition of 2 molar equiv of PY-ligand to a 100 μM solution of the W17F WW domain results in a dramatic increase in the near-UV CD maximum intensity over the range of 287−250 nm (Figure 6c). The linear combination of the apo W17F and the PY-ligand signal over the near-UV CD range is very different from the experimental near-UV CD spectrum of the W17F−PY-ligand complex which exhibits unique spectral features, analogous to what was observed with the far-UV CD experiments and consistent with binding-induced conformational changes.

^{1}H NMR Spectroscopy. The ^{1}H NMR spectrum of the 57-residue wild-type WW domain acquired at 5 °C exhibits the spectral characteristics of a folded structure, including good dispersion in the downfield amide proton region (Figure 7a), especially in the region from 6.4 to 9.6 ppm and in the upfield aliphatic region from 2.0 to 0.5 ppm (Figure 7b), as well as generally good resonance dispersion throughout the entire spectrum (*6*, *20*, *30*) (Figure 2 of the Supporting Information). The indole N−H proton of Trp17 in the wild-type

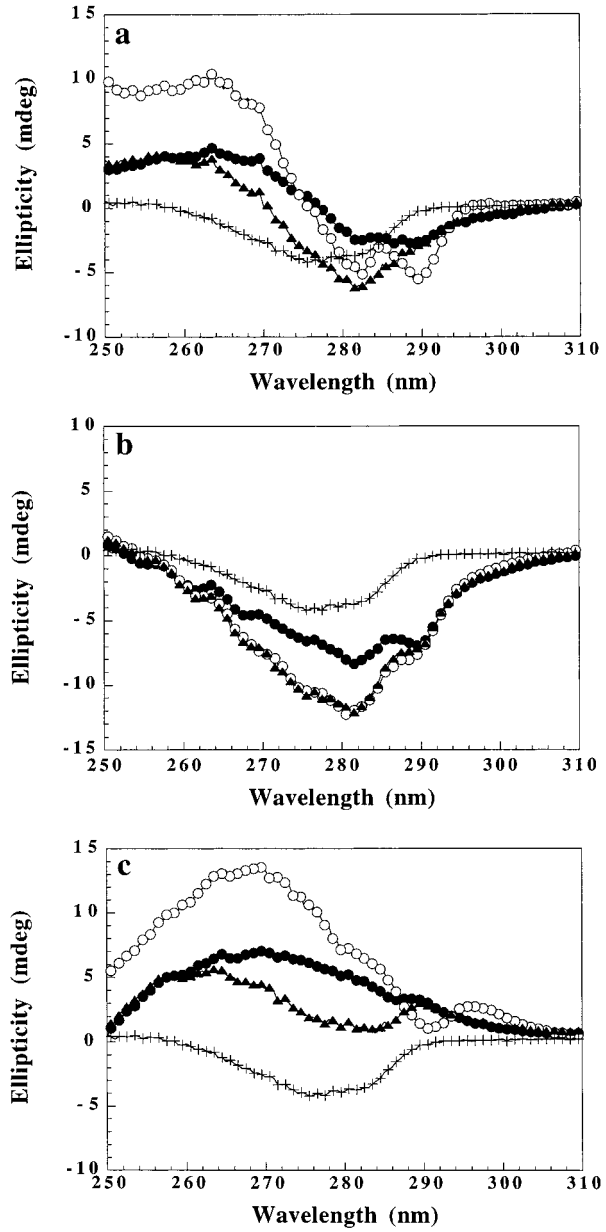

FIGURE 6: (a) Near-UV circular dichroism spectra of the PY-ligand (200 μM) alone (+) and of the wild-type WW domain (100 μM) in the absence (●) and presence (○) of the PY-ligand (200 μM). The black triangles (▲) represent the hypothetical spectrum resulting from summation of the near-UV CD signal of the apo PY-ligand with that of the apo wild-type WW domain. (b) Strictly analogous experiment with the W39F WW domain (c) Strictly analogous experiment with the W17F WW variant.

domain has a chemical shift of 10.50 ppm, while the indole N−H proton of Trp39 resonates at 10.15 ppm, consistent with a folded structure.

The ^{1}H NMR chemical shift dispersion exhibited by the W39F WW domain is very similar to the spectrum exhibited by the wild-type protein with good resonance dispersion in both the downfield (Figure 7a) and upfield (Figure 7b) regions and with the characteristic indole N−H resonance of Trp17 shifted slightly to 10.41 ppm (Figure 7a).

In contrast, the spectrum of the W17F WW domain exhibits no resonances downfield of 8.9 ppm (Figure 7a) or upfield of 0.9 ppm (Figure 7b), and the N−H indole proton of Trp39 in this domain variant exhibits resonances at 10.24 (shoulder peak) and 10.14 ppm (Figure 7a). The poor

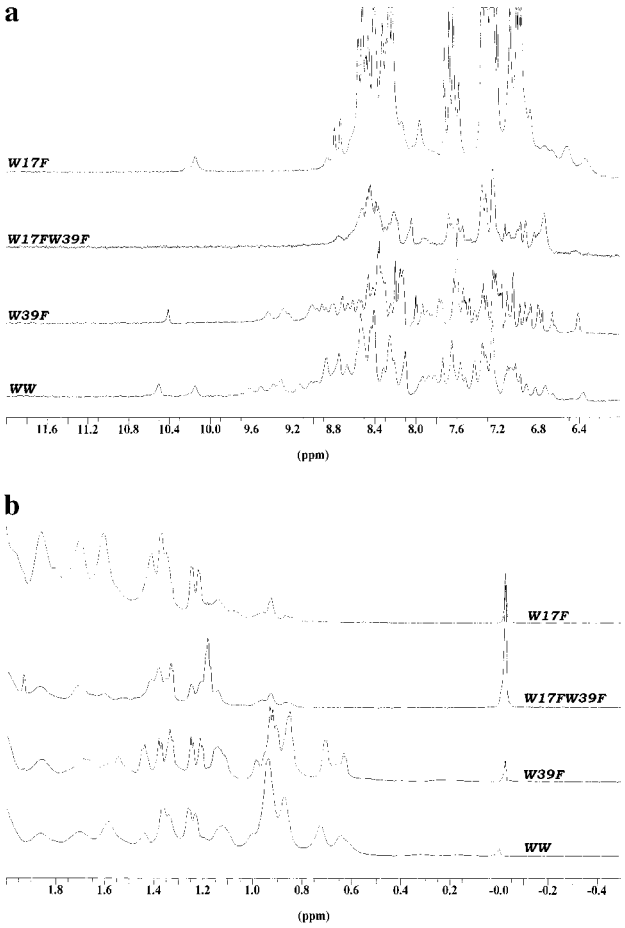

FIGURE 7: Downfield (a) and upfield (b) regions of the 750 MHz ^{1}H NMR spectra of the wild-type, W39F, W17F, and W17F/W39F WW domains acquired at 5 °C in buffer A (pH 6.0). The complete spectrum of each WW protein is presented in Figure 2 of the Supporting Information.

chemical shift dispersion exhibited by the W17F WW domain is consistent with what one observes for a polypeptide adopting an ensemble of largely unordered conformations in aqueous solution (the resonance at ≈10.24 ppm may represent a minor conformational family wherein Trp39 is involved in hydrophobic cluster formation). Like the W17F ^{1}H NMR spectrum, the spectrum of the W17F/W39F double mutant variant provides no evidence for a folded conformation (Figure 7a,b).

PY-Ligand Binding to the WT and Variant WW Domains As Determined by ^{1}H NMR Spectroscopy. Titration of the PY-ligand into a 600 μM solution of the WW domain induces dramatic changes in both the downfield (Figure 8a) and the upfield (Figure 8b) regions of the spectrum of the complex, affording readily noticeable increases in the extent of chemical shift dispersion associated with the unique spectral properties of the complex. Under saturating PY-ligand conditions (2 molar equiv), significant downfield shifts are observed for both Trp resonances, with the limiting shifts observed at 10.81 (Trp39) and 10.43 ppm (Trp17). A number of resonances upfield of 0.7 ppm that were not present in the spectrum of the apo WW domain are also evident upon PY-ligand binding. The ^{1}H NMR spectrum of the PY-ligand does not exhibit resonances upfield of 1.0 ppm or downfield of 9.0 ppm (Figure 3 of the Supporting Information).

Contrary to the results described above, titration of 2 equiv of PY-ligand into a 600 μM solution of the W39F WW

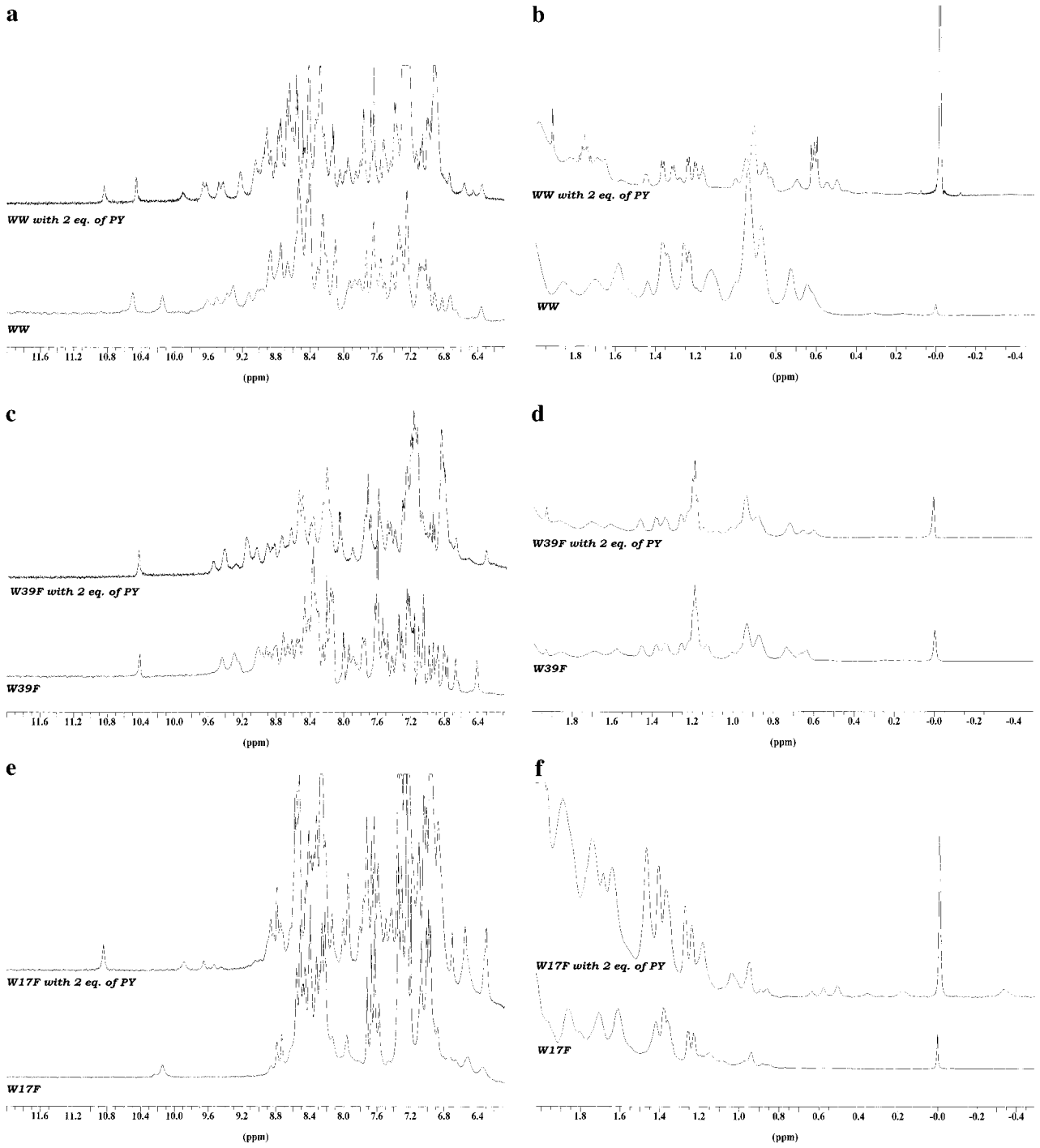

FIGURE 8: Addition of the PY-ligand to samples of the wild-type, W39F, and W17F WW domains monitored by ^{1}H NMR spectroscopy. Changes in the 750 MHz downfield (a) and upfield (b) regions of the ^{1}H NMR spectrum of the wild-type WW domain upon addition of 2 molar equiv of the PY-ligand. Analogous downfield (c) and upfield (d) regions of the W39F variant. The downfield (e) and upfield (f) regions of the ^{1}H NMR spectrum of the W17F WW domain variant monitored during the course of a PY-ligand titration (2 equiv).

domain does not produce the downfield- and upfield-shifted resonances expected for W39F−PY-ligand complex formation (Figure 8c,d). Furthermore, the Trp17 residue which is sensitive to ligand binding in the wild-type WW domain does not change its resonance position upon addition of 2 equiv of the PY-ligand. The lack of a dramatically increased extent of chemical shift dispersion beyond that expected by PY-peptide addition further supports the inability of W39F WW to bind the PY-ligand.

Titration of the PY-ligand into a 600 μM solution of the W17F WW domain induces a number of changes in downfield and upfield regions of the W17F ^{1}H NMR spectrum (panels e and f of Figure 8, respectively). Under saturating PY-ligand conditions (2 equiv), we exclusively see the indole N−H proton resonance at 10.84 ppm (Figure 8e), nearly the same resonance position exhibited by Trp39 in the wild-type WW−PY-ligand complex (10.81 ppm). The W17F WW−PY-ligand complex also exhibits multiple resonances upfield of 0.8 ppm that are not present in the spectrum of the apo form of the W17F variant (Figure 8f).

Monitoring Ligand Binding by Analytical Ultracentrifugation. Detection of the dabcylated PY-ligand over the λ

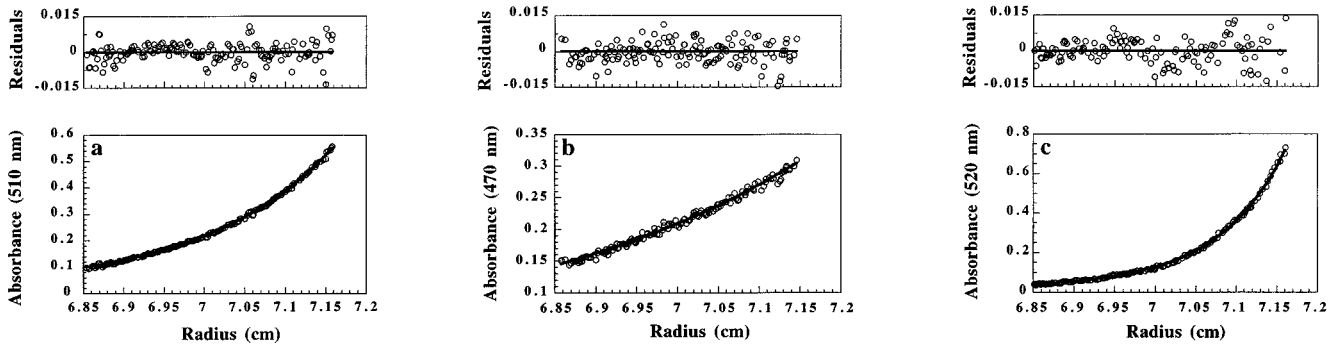

FIGURE 9: Sedimentation equilibrium analysis of 29 μM wild-type (a), 5 μM W39F (b), and 29 μM W17F (c) WW domains in buffer A (pH 6.0) with added dabcylated PY-ligand (14.5 μM to the wild-type and W17F samples and 5 μM to the W39F sample). The solid line drawn through each data set was obtained by fitting the absorbance vs radial position data (○) to eq 1 for a single ideal species, yielding molecular masses of 7924 ± 122 and 8273 ± 72 Da for the wild-type and W17F complexes, respectively. The solution of 5 μM W39F WW domain with 5 μM dabcylated PY-peptide yielded a molecular mass of 1823.5 ± 21.9 Da (470 nm), demonstrating the inability of the W39F WW domain variant to bind the PY-ligand. The residuals between the fit and the experimental data are illustrated in the top panel of each figure. The free dabcylated PY-peptide (14.5 μM) exhibited single-ideal species behavior with an apparent molecular mass of 1798 ± 14 Da (Figure 4 of the Supporting Information).

range of 470−520 nm was employed to observe a change in molecular mass from 1711 Da for the labeled PY-ligand alone to >8000 Da for the WW−PY-ligand complex (>4.7-fold increase). Monitoring the molecular mass of the WW domains (molecular mass of 6372−7095 Da; see Table 1) in the UV region with unlabeled PY-ligand (molecular mass of 1460 Da) only leads to a molecular mass increase of 1.2-fold (molecular mass of 7832−8555 Da) upon PY-ligand binding, which is difficult to analyze with confidence on the basis of our experiments.

Sedimentation equilibrium analysis of the wild-type WW domain (29 μM) in complex with the dabcylated PY-ligand (14.5 μM) clearly demonstrates binary complex formation in solution (Figure 9a). The dabcylated PY-ligand−WW domain complex was observed to come to sedimentation equilibrium as a single ideal species (eq 1) with a molecular mass of 7924 ± 122 Da, which is in excellent agreement with the expected mass of the complex (8082.9 Da). The goodness of the fit was evaluated on the basis of the randomness and magnitude of the residuals (upper panel), expressed as the difference between the theoretical curve and the experimental data, and by evaluating the reasonability of the fit parameters. The dabcylated PY-ligand (14.5 μM) itself was also studied by equilibrium analysis, demonstrating that it behaves as a single ideal species with a molecular mass of 1798 ± 14 Da, in good agreement with the calculated mass of 1710.88 Da (Figure 4 of the Supporting Information).

Analogous equilibrium studies on the W39F WW variant binding to the dabcylated PY-ligand fit best to a two-species model (eq 2), 90% free dabcylated PY-ligand, 10% dabcylated PY-ligand−W39F complex (Figure 5 of the Supporting Information). The small amount of unexpected binding is not surprising given the hydrophobicity of the dabcyl fluorophore label which apparently gives rise to nonspecific binding (see ITC data). To test this hypothesis, a reduction in the concentration of both the W39F WW domain and the dabcylated PY-ligand to 5 μM results in an excellent fit to a single-ideal species model of uncomplexed dabcylated PY-ligand as expected (Figure 9b), further supporting the inability of W39F to specifically bind the PY-ligand.

Equilibrium ultracentrifuge studies on a 29 μM solution of the W17F WW variant added to 14.5 μM dabcylated PY-ligand show the data fit to a single-ideal species model, consistent with binary complex formation (detection $\lambda = 520$ nm) (Figure 9c). Equation 1 yields a solution molecular mass of 8273 ± 72 Da (calculated mass of 8805.7 Da), where the fit to binary complex formation is good as discerned from the randomness of the residuals.

Determination of Binding Constants by Isothermal Titration Calorimetry. The heat evolved from each 5 μL injection of a 1.0 mM PY-ligand solution into a 50 μM wild-type WW domain solution is shown in Figure 10a (top panel). The normalized peak area data calculated in kilocalories per mole of injectant are displayed in the lower panel of Figure 10a. The binding isotherm shows an immediate decrease in the amount of heat that is evolved for each consecutive injection, resulting in the absence of a well-defined plateau region prior to reaching the transition area of the titration (Figure 10a, bottom panel). Nonlinear least-squares analysis of the dilution-corrected ITC data yields a dissociation constant K_d of 5.9 ± 0.33 μM (mean ± standard deviation, $n = 5$). The best fit to the data using a single set of identical binding sites model (eq 3) is achieved when the binding stoichiometry is allowed to vary during fitting. This method of data analysis produces a binding stoichiometry of 0.92. When the stoichiometry of binding is held constant at 1.0, the best fit to the data matches quite well with the experimental data, with the exception of the data at very low titrant concentrations (not shown). Irrespective of whether the stoichiometry was fixed at 1.0 or allowed to vary, the binding constants differed by only 18%. The enthalpy of association ΔH_a of the wild-type WW domain with the PY-ligand in buffer A (pH 6.0) at 25 °C is −14.0 ± 0.61 kcal mol^{-1}, and the entropy of association is −23 cal K^{-1} mol^{-1}.

The isothermal titration data recorded during the addition of the PY-ligand (1.0 mM) to 50 μM W39F WW (buffer A, pH 6.0, 25 °C) are strikingly different from those of the wild type discussed above in that no heat in excess of the heat of dilution of the ligand was evolved (Figure 10b). These data combined with the spectroscopic data outlined above demonstrate the inability of the W39F WW domain to bind the PY-ligand.

The isothermal titration data exhibited by the W17F variant (44.5 μM) being titrated with the PY-ligand (1.0 mM) in buffer A (pH 6.0) at 25 °C were similar to the data discussed above for the wild-type WW domain (Figure 6 of the

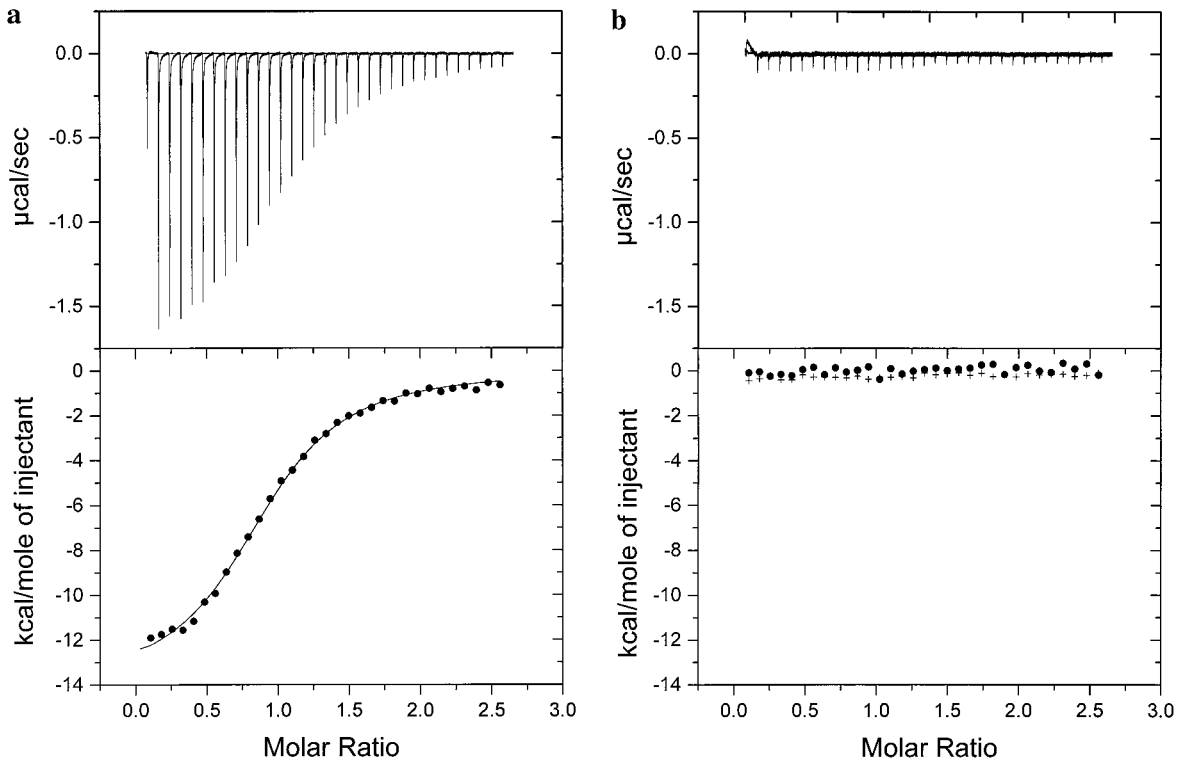

FIGURE 10: Analysis of PY-ligand binding to the wild-type and W39F WW domains using isothermal titration calorimetry. (a) The raw heat response of each 5 μL injection of PY-ligand into a 1.3407 mL ITC cell containing a 50 μM sample of the wild-type WW domain in buffer A (pH 6.0) at 25 °C is presented in the top panel (ignore the first injection), while the bottom panel reflects the integrated peak areas normalized to moles of PY-ligand added and corrected for dilution effects (●). The nonlinear least-squares fit to the binding isotherm is represented by the solid line. The average of five titration experiments yielded a dissociation constant of 5.9 ± 0.33 μM and an enthalpy of association of −14.0 ± 0.61 kcal mol^{-1}. (b) Analogous titration experiment conducted with the W39F WW domain variant. The data presented in the bottom panel are the integrated peak area of each injection (●) not corrected for the heat of dilution of the PY-ligand (+) or the heat of dilution of the W39F protein (not significant). The heat of ligand dilution (+) is equivalent to the heat evolved for the addition of the PY-ligand to the W39F WW variant (●), consistent with the inability of this WW variant to bind the PY-ligand.

Supporting Information). Fitting these data to a single-binding site model with a varying binding stoichiometry yields a dissociation constant of 15.1 ± 1.2 μM (mean ± standard deviation, $n = 4$), a binding stoichiometry of 0.91, an enthalpy of association of −20.4 ± 1.2 kcal mol^{-1}, and an entropy of association of −49 cal K^{-1} mol^{-1}. Surprisingly, the K_d values associated with PY-ligand binding by the folded wild-type WW domain and the largely unfolded W17F WW domain differ by less than 3-fold.

DISCUSSION

Signal transduction mediated by protein−protein interactions between a WW domain and its proline-rich ligand partners is important for normal physiology (*5*, *31*). Mutations that interfere with this interaction have been implicated in several diseases (*4*, *32*). In this study, the influence of the two Trp residues on WW β-sheet stability and PY-ligand binding was investigated by making Trp to Phe substitutions at the two conserved residues after which the domain was named (Trp17 and Trp39). As is often the case in investigations of this type, surprising behavior was observed for the W17F variant regarding PY-ligand binding. Even though the W17F mutation resulted in a largely unfolded WW domain, this domain was capable of binding the PY-ligand with affinity within 3-fold of that exhibited by the wild-type WW domain. The W39F variant does not significantly affect β-sheet stability (*vide infra*), but does abrogate PY-ligand binding which is not surprising in light of the prominent role that Trp39 plays in mediating interactions with a xPPxY-ligand (see the introductory section).

Evaluating the Structure of the W39F WW Domain Variant. Recent laser temperature jump studies on the kinetics of W39F WW domain folding required that we characterize the folding equilibria of this variant as a function of temperature by both fluorescence and far-UV CD (*41*). The parallel changes in fluorescence and CD as a function of temperature allowed a simultaneous two-state model fit to both data sets, revealing that the W39F variant is 1 kJ mol^{-1} more stable than wild-type WW (*41*). The substitution of Trp39 with Phe does not noticeably alter the reversible thermal transitions exhibited by this variant WW domain (Figure 7 of the Supporting Information). The two-state thermal folding and unfolding behavior of W39F ($T_m = 53.1$ °C) is nearly identical to that exhibited by wild-type WW ($T_m = 50.3$ °C) (*20*, *41*). The environment of the Trp17 residue in W39F appears to be nearly identical to that in the wild-type protein as discerned from its emission at 339 nm (Figure 2b). The far-UV CD spectrum exhibited by the W39F variant closely resembles that of the wild-type WW domain (Figure 3), the reduced intensity being attributed to the loss of the far-UV CD contributions of Trp39. The W39F near-UV CD spectrum exhibiting double minima is similar to the wild-type spectrum with the exception of the absence of the maximum (Figure 5). In fact, if one adds the near-UV CD spectra of W39F and W17F, the resulting hypothetical

spectrum is nearly identical to the wild-type spectrum (Figure 5). This interesting observation strongly suggests that the near-UV CD spectral contribution from Trp39 in the unfolded W17F variant is very similar to the Trp39 contribution in the folded wild-type WW spectrum. The one-dimensional ^{1}H NMR spectrum of the W39F variant exhibits excellent resonance dispersion in both the downfield and upfield portions of the spectra and exhibits a Trp17 indole proton resonance at 10.41 ppm, consistent with a folded three-stranded β-sheet structure (Figure 7a,b).

Evaluating W39F WW−PY-Ligand Binding. The addition of the PY-ligand to the folded W39F WW domain variant results in no changes in fluorescence intensity or wavelength (Figure 2b). Even though Trp17 is on the side of the β-sheet opposite the PY-ligand binding face, it is hard to conceive how Trp17 fluorescence would not be sensitive to binding due to its proximity and the suggested conformational changes exhibited by the wild-type WW domain on binding (see the Results, especially Figures 4, 6, and 8). The inability of W39F to bind the PY-ligand is strongly supported by the addition of the PY-ligand to the W39F WW domain which results in a far-UV CD spectrum that is not different than the sum of the PY and W39F spectra (Figure 4b). This is very different from the behavior exhibited by the wild-type and W17F WW domains which do bind the PY-ligand (Figure 4a,c) and exhibit spectral contributions unique to complex formation. Near-UV CD binding data also demonstrate the inability of the W39F WW variant to bind the PY-ligand, as the resulting spectrum is simply that expected from a sum of the two titration components, unlike the spectroscopically unique complexes formed by the wild-type− and W17F−PY-ligand complexes (Figure 6). The addition of 2 equiv of the PY-ligand to an NMR sample of the W39F variant (Figure 8c,d) reveals none of the characteristic upfield or downfield shifts clearly observed when wild-type WW binds the PY-ligand (Figure 8a,b). In addition, no chemical shift dispersion beyond that created by the addition of PY-ligand was evident.

Analytical equilibrium ultracentrifuge studies on the W39F WW (29 μM) variant binding to the dabcylated PY-ligand (14.5 μM) fit best to a two-species model, 90% free dabcylated PY-ligand, 10% dabcylated PY-ligand−W39F complex (Figure 5 of the Supporting Information). Due to the hydrophobicity of the dabcyl fluorophore, a small extent of nonspecific binding is not too surprising. Reducing the concentration of both the W39F WW domain and the dabcylated PY-ligand to 5 μM results in an excellent fit to a single-ideal species model of uncomplexed dabcylated PY-ligand (Figure 9b) (470 nm). The isothermal titration data exhibited by the addition of the PY-ligand to the W39F variant are strikingly different from those observed for the wild-type WW domain. No heat in excess of the ligand dilution heat was observed (Figure 10b). These data combined with the spectroscopic data outlined above clearly demonstrate the inability of the W39F WW domain to bind the PY-ligand.

The lack of PY-ligand binding to the W39F variant is not unexpected given the numerous important interactions that were identified between Trp39 of the WW domain and the Pro4 and Pro5 residues of the GTPPPPYTVG ligand in the NMR structure of the complex (*6*). In addition, the GST−W39F fusion protein was also not active with respect to binding the GST−proline-rich ligand having the sequence PPPYTVG (*11*). Interestingly, there are several native WW domains that have the equivalent of the W39F mutation, strongly implying that these sequences have ligand binding preferences other than the xPPxY motif (*3*, *42*, *43*; www.bork.embl-heidelberg.de/Modules/auto_ident_ww.html).

Evaluating the Structure of the W17F WW Domain Variant. Replacement of Trp17 with Phe in the W17F WW domain results in Trp39 fluorescence at 353 nm, similar to the emission observed in 6.0 M GdnHCl, consistent with a largely unfolded structure (Figure 2c). However, this piece of evidence alone is not conclusive because exposure of the Trp39 residue to solvent is significant. The lack of the characteristic far-UV CD maximum near 230 nm and the absence of a minimum near 202 nm also suggest the absence of β-sheet tertiary structure that is normally observed in folded WW domains (Figure 3). Interestingly, the near-UV CD spectrum exhibits a broad maximum over the wavelength range of 250−300 nm, suggesting the presence of a residual hydrophobic cluster involving Trp39 in the ensemble of disordered states (Figure 5). The fact that the near-UV CD spectra of the W17F and W39F variants can be added to produce a spectrum (Figure 5, solid line) similar to that of wild-type WW domain suggests that the Trp39 environment is similar in the unfolded and folded states. Since this residue is largely exposed to solvent, this result is not unexpected. The ^{1}H NMR spectrum of the W17F WW domain does not exhibit the upfield or downfield shifts or the dispersion that is characteristic of those of the folded wild-type and W39F WW domains (Figure 7a,b). Furthermore, the lone indole N−H proton of Trp39 in the W17F variant exhibits heterogeneity in its resonance position (peaks at 10.24 and 10.14 ppm), consistent with multiple environments for the indole side chain, possibly including a hydrophobic cluster (Figure 7a).

Evaluating W17F WW−PY-Ligand Binding. Addition of the PY-ligand to the ensemble of largely unfolded states characterizing this variant results in a disorder to order folding transition of the W17F WW domain. Monitoring this addition by fluorescence results in a large increase in the quantum yield (4-fold) of the Trp39 emission as well as a striking 24 nm blue shift in the wavelength of maximal emission upon PY binding (Figure 2c). These changes not only reflect a significant increase in the hydrophobicity of the Trp39 environment (blue shift) but also strongly suggest protection of the excited state from dynamic quenching. A wild-type-like far-UV CD spectrum is afforded upon addition of the PY-ligand to the W17F WW domain (Figure 4c). The maximum at 230 nm as well as the broad minimum centered at 205 nm is characteristic of a folded WW domain bound to a polyproline type II helix. The differences observed between the spectra of the complex and the hypothetical spectra derived from summing the PY-ligand and W17F spectra reveal a unique spectroscopic signature for the complex. The prominent near-UV CD changes mediated by ligand binding and the differences between the hypothetical summed spectra and the experimental spectra of the complex also demonstrate binding (Figure 6c). Addition of the PY-ligand to a W17F NMR sample is accompanied by a dramatic increase in the extent of chemical shift dispersion, a number of upfield amide and downfield aliphatic resonances, and a shift in the indole proton resonance of Trp39 to 10.84 ppm,

consistent with thermodynamically linked folding and ligand binding (Figure 8e,f).

Sedimentation equilibrium experiments using a dabcylated PY-ligand directly demonstrated the formation of an intermolecular complex (molecular mass of 8273 ± 72 Da), confirming a binding stoichiometry of 1 W17F WW domain to 1 PY-peptide ligand (Figure 9c). The dissociation constant measured by isothermal titration calorimetry is 15.1 ± 1.2 μM, only about 3-fold higher than the K_d observed for the wild-type WW domain (Figure 6 of the Supporting Information). The enthalpy of association was greater for the W17F variant (20.4 ± 1.2 kcal mol^{-1}) than that observed for the wild-type WW protein (14.0 ± 0.6 kcal mol^{-1}), perhaps as a result of the enthalpy gained from thermodynamically linked folding and binding in the case of the W17F WW domain. If this were the case, the entropy of binding of the PY-ligand would be expected to be more unfavorable for the W17F variant (−49 cal K^{-1} mol^{-1}) than for the wild-type WW domain (−23 cal K^{-1} mol^{-1}), which is observed. This is probably so because the chain entropy penalty associated with the linked folding of the W17F WW domain is not outweighed by the entropy gained by the return of ordered water molecules to the bulk solvent, because of the limited extent of tertiary structural interactions in the WW domain fold.

It is not clear whether the PY-ligand simply alters the thermodynamics of W17F WW domain folding, shifting the equilibrium to the folded state by binding to and lowering the free energy of the folded state (implying that there is a small population of folded W17F protein in the ensemble of states). Alternatively, the PY-peptide could bind to the transition state, thereby lowering the activation barrier for folding in conjunction with binding to the ground state. Kinetic and thermodynamic measurements will be made in an effort to obtain a better understanding of this effect. None of the more than 170 WW domain sequences observed to date have the equivalent of the W17F mutation, possibly because the unfolded WW domain is susceptible to proteolysis, which we observe experimentally during recombinant expression (www.bork.embl-heidelberg.de/Modules/auto_ident_ww.html).

Structural Characteristics and Binding Properties of the W17F/W39F WW Domain. The W17F/W39F WW variant appears to be both unfolded and unable to bind to the PY-ligand. Tyrosine fluorescence analysis reveals similar emission properties under native conditions (wild-type) and in 6.0 M GdnHCl (Figure 1 of the Supporting Information), consistent with an unfolded structure. The NMR dispersion in both the upfield and downfield regions of the spectra is particularly poor, consistent with an unfolded structure (Figure 7). The far-UV CD spectrum of the W17F/W39F WW domain under native conditions is inconsistent with a folded WW β-sheet structure (Figure 3). The nonexistent near-UV CD spectrum does not reveal much about the structure of the W17F/W39F double mutant, other than it is likely unfolded because of the lack of any Tyr-derived near-UV CD signal (Figure 5).

General Discussion. Replacement of Trp17 with a Phe residue in the WW domain generates a natively unfolded protein that retains ligand binding activity. To the best of our knowledge, this is the only known example where a normally folded and functional protein has been converted to an unfolded protein that retains near-native ligand binding activity by incorporation of a single conservative amino acid change. However, there are a number of proteins which have been identified as natively unfolded, some of which undergo disorder to order folding transitions induced by ligand binding. All of these examples are involved in protein−protein recognition and include the C-terminus of FlgM (a protein component of bacterial flagella), a 130-residue fragment of the *Staphylococcus aureus* fibronectin-binding protein, the cyclin-dependent kinase inhibitor p21$^{Waf1/Cip1/Sid1}$, the microtubule-associated protein MAP2, the tau protein and α-synuclein associated with Alzheimer's amyloid fibrils, the antitermination protein N of Phage λ, the transcriptional activation domain of the herpes virus protein VP16, and prothymosin α from the thymus (*12*−*16*, *33*−*39*). Of these examples, FlgM, the cyclin-dependent kinase inhibitor p21$^{Waf1/Cip1/Sid1}$, the antitermination protein N of Phage λ, and the transcriptional activation domain of the herpes virus protein VP16 all exhibit disorder to order conformational transitions upon ligand binding. It is likely that additional natively unfolded proteins will be identified. The challenge will be to determine why these proteins are largely unfolded, and what biological advantage is gained from utilizing an unfolded protein as a ligand binding target. Perhaps this type of protein−protein recognition event is used to engineer exquisite selectivity by linking the folding and binding equilibria, especially if the ligand is selected to bind to the transition state associated with folding.

Overall Summary. In summary, substitution of Trp39 with Phe results in a folded WW domain; however, it is unable to bind the PY-ligand. Substitution of Trp17 with Phe results in a largely unfolded polypeptide, yet surprisingly, it is capable of PY-ligand binding. The K_d for the latter complex is less than 3-fold greater than that exhibited by the wild-type WW domain, suggesting that enthalpy−entropy compensation allows ligand binding to occur by coupling the PY binding event to the W17F WW folding transition. Future kinetic and thermodynamic measurements will lead to a better understanding of this observation.

ACKNOWLEDGMENT

We thank Dr. Prakash Raman for collecting all the mass spectrometry data, Dr. Jennifer Kowalski for help with the NMR experiments and Hans Purkey for technical assistance with the isothermal titration calorimetry experiments.

SUPPORTING INFORMATION AVAILABLE

Molecular biology details for variant construction (W39F and W17F/W39F), W17F/W39F Tyr fluorescence (Figure 1), complete ^{1}H NMR spectra for all the WW domains (Figure 2), ^{1}H NMR spectrum of the PY peptide (Figure 3), equilibrium ultracentrifuge analysis of the dabcylated PY-ligand (Figure 4), equilibrium analysis of W39F (29 μM) combined with dabcylated PY-ligand (14.5 μM) (Figure 5), isothermal titration calorimetry of W17F binding to the PY-ligand (Figure 6), and thermal unfolding of W39F monitored by CD and fluorescence (Figure 7). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bork, P., and Sudol, M. (1994) *Trends Biochem. Sci. 19*, 531−533.

2. Sudol, M., Chen, H. I., Bougeret, C., Einbond, A., and Bork, P. (1995) *FEBS Lett. 369*, 67−71.
3. Barinaga, M. (1999) *Science 283*, 1247−1249.
4. Sudol, M. (1998) *Oncogene 17*, 1469−1474.
5. Yagi, R., Chen, L.-F., Shigesada, K., Murakami, Y., and Ito, Y. (1999) *EMBO J. 18*, 2551−2562.
6. Macias, M. J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996) *Nature 382*, 646−649.
7. Ranganathan, R., Lu, K. P., Hunter, T., and Noel, J. P. (1997) *Cell 89*, 875−886.
8. Kanelis, V., Farrow, N. A., Kay, L. E., Rotin, D., and Forman-Kay, J. D. (1998) *Biochem. Cell Biol. 76*, 341−350.
9. Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J., and Schreiber, S. L. (1992) *Science 258*, 1655−1668.
10. Morton, C. J., Pugh, D. J., Brown, E. L., Kahmann, J. D., Renzoni, D. A., and Campbell, I. D. (1996) *Structure 4*, 705−714.
11. Chen, H. I., Einbond, A., Kwak, S.-J., Linn, H., Koepf, E. K., Peterson, S., Kelly, J. W., and Sudol, M. (1997) *J. Biol. Chem. 272*, 17070−17077.
12. Kriwacki, R. W., Hengst, L., Tennant, L., Reed, S. I., and Wright, P. E. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 11504−11509.
13. Daughdrill, G. W., Chadsey, M. S., Karlinsey, J. E., Hughes, K. T., and Dahlquist, F. W. (1997) *Nat. Struct. Biol. 4*, 285−291.
14. Van Gilst, M. R., Rees, W. A., Das, A., and von Hippel, P. H. (1997) *Biochemistry 36*, 1514−1524.
15. Mogridge, J., Legault, P., Li, J., Van Oene, M. D., Kay, L. E., and Greenblatt, J. (1998) *Mol. Cell 1*, 265−275.
16. Penkett, C. J., Redfield, C., Jones, J. A., Dodd, I., Hubbard, J., Smith, R. A. G., Smith, L. J., and Dobson, C. M. (1998) *Biochemistry 37*, 17054−17067.
17. Shortle, D. (1993) *Curr. Opin. Struct. Biol. 3*, 66−74.
18. Miller, D. W., and Dill, K. A. (1997) *Protein Sci. 6*, 2166−2179.
19. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature 395*, 817−819.
20. Koepf, E. K., Petrassi, H. M., Sudol, M., and Kelly, J. W. (1999) *Protein Sci. 8*, 841−853.
21. Mach, H., Middaugh, C. R., and Lewis, R. V. (1992) *Anal. Biochem. 200*, 74−80.
22. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
23. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
24. Laue, T. M., Shad, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) *Computer-aided interpretation of analytical sedimentation data in biochemistry and polymer science*, The Royal Society of Chemistry, Cambridge, England.
25. Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.
26. Eftink, M. R. (1997) *Methods Enzymol. 278*, 221−257.
27. Johnson, W. C., Jr. (1990) *Proteins: Struct., Funct., Genet. 7*, 205−214.
28. Fasman, G. D. (1996) *Circular Dichroism and the conformational analysis of biomolecules*, Plenum Press, New York.
29. Woody, R. W. (1985) in *The Peptides, Analysis, Synthesis, and Biology* (Hruby, V. J., Ed.) Academic Press, New York.
30. Wuthrich, K. (1986) *NMR of Proteins and Nucleic acids*, John Wiley & Sons, New York.
31. Sudol, M. (1996) *Prog. Biophys. Mol. Biol. 65*, 113−132.
32. Sudol, M. (1996) *Trends Biochem. Sci. 21*, 161−163.
33. Penkett, C. J., Redfield, C., Dodd, I., Hubbard, J., McBay, D. L., Mossakowska, D. E., Smith, R. A. G., Dobson, C. M., and Smith, L. J. (1997) *J. Mol. Biol. 274*, 152−159.
34. Hernandez, M. A., Avila, J., and Andreu, J. M. (1986) *Eur. J. Biochem. 154*, 41−48.
35. Schweers, O., Schonbrunn-Hanebeck, E., and Mandelkow, E. (1994) *J. Biol. Chem. 269*, 24290−24297.
36. Weinreb, P. H., Zhen, W., Poon, A. W., Conway, K. A., and Lansbury, P. T., Jr. (1996) *Biochemistry 35*, 13709−13715.
37. Shen, F., Triezenberg, S. J., Hensley, P., Porter, D., and Knutson, J. R. (1996) *J. Biol. Chem. 271*, 4819−4826.
38. Shen, F., Triezenberg, S. J., Hensley, P., Porter, D., and Knutson, J. R. (1996) *J. Biol. Chem. 271*, 4827−4837.
39. Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H. R., Muller-Frohne, M., Zirwer, D., Czarnecki, J., and Damaschun, G. (1995) *Biochemistry 34*, 13211−13218.
40. Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.
41. Crane, J. C., Koepf, E. K., Kelly, J. W., and Gruebele, M. (1999) *J. Mol. Biol.* (submitted for publication).
42. Lu, P.-J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) *Science 283*, 1325−1328.
43. Bedford, M. T., Chan, D. C., and Leder, P. (1997) *EMBO J. 16*, 2376−2383.

BI991105L