WW: An isolated three-stranded antiparallel β -sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state

EDWARD K. KOEPF,¹ H. MICHAEL PETRASSI,¹ MARIUS SUDOL,² AND JEFFERY W. KELLY¹

¹Department of Chemistry and the Skaggs Institute of Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Rd., La Jolla, California 92037

²Department of Biochemistry, Mount Sinai School of Medicine, New York, New York 10029

(RECEIVED August 24, 1998; ACCEPTED December 16, 1998)

Abstract

The objective of this study was to evaluate the suitability of the WW domain as a desirable model system to understand the folding and stability of an isolated three-stranded antiparallel β -sheet structure. The WW domain was subjected to thermal and chaotropic denaturation/reconstitution utilizing a variety of biophysical methods. This three-stranded sheet folds reversibly and cooperatively utilizing both urea and GdnHCl as denaturants; however, the denatured state retains structure in the form of a hydrophobic cluster involving at least one aromatic side chain. In contrast to chaotropic denaturation, thermal denaturation appears to be more complete and may be a two state process. The suitability of the WW domain for future studies aimed at understanding the kinetics and thermodynamics of antiparallel β -sheet folding clearly emerges from this initial study. The most exciting and significant result in this manuscript is the finding that the chaotropic denatured state of WW has a hydrophobic cluster as discerned by near-UV CD evidence. The role that the denatured state plays in the folding and stability of a three-stranded β -sheets, and its capacity for preventing aggregation may be particularly important and is the subject of ongoing studies.

Keywords: β -sheet folding; hydrophobic cluster; reversible folding; WW

The WW domain is a recently discovered protein module consisting of a three-stranded antiparallel β -sheet structure that is named for the two highly conserved Trp residues separated in the sequence by 20–22 residues (Bork & Sudol, 1994; Chen & Sudol, 1995). It is a small compact domain (38–40 residues) found in species ranging from lower eukaryotes to mammals (Bork & Sudol, 1994; Chen & Sudol, 1995; Macias et al., 1996; Sudol, 1996; Linn et al., 1997). Like the SH2, SH3 (Src homology 2,3), and PTB (phospho-tyrosine binding) domains, the WW domain mediates protein–protein interactions. This domain is found in nearly

100 intracellular proteins of great functional diversity, ranging from cytoskeletal to regulatory proteins (Sudol, 1996).

The WW domain was initially identified in the 65 kD human Yes associated protein (hYAP) during a search for substrates that bind to the Yes proto-oncogene product, a nonreceptor tyrosine kinase of the Src family (Bork & Sudol, 1994; Chen & Sudol, 1995). The association of YAP with Yes is not mediated through the WW domain. Instead, binding is the result of the interaction of the SH3-binding motif of YAP (consensus sequence PxxP where P is Pro and x is any other amino acid) with the SH3 domain of Yes (Sudol, 1994). To identify ligands capable of binding directly to the WW domain of YAP, a functional screen of a cDNA expression library was conducted. Two WW binding proteins were discovered, WBP-1 and WBP-2, each having the identical proline-rich pentapeptide sequence PPPPY (Chen & Sudol, 1995). Further refined screening by Ala scanning mutagenesis identified the minimal consensus sequence for binding as xPPxY, where x is any residue and P and Y are Pro and Tyr residues, respectively. Additional characterization of ligands capable of binding to the WW domain was conducted by screening peptide libraries prepared by the “SPOTs” method of multiple peptide synthesis on derivatized cellulose (Blankenmeyer-Menge

Reprint requests to: Jeffery W. Kelly, The Scripps Research Institute, 10550 North Torrey Pines Rd., MB12, La Jolla, California 92037; e-mail: jkelly@scripps.edu.

Abbreviations: WW, Trp-Trp domain; GdnHCl, guanidinium hydrochloride; CD, circular dichroism; YAP, human Yes associated protein; WBP, WW binding protein; GST, glutathione-S-transferase; SDS-PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis; MALDI, matrix-assisted laser desorption ionization; PY, proline rich internal peptide from p53 binding protein-2 having the following sequence: EYPPYPPPPYPSG; PTB, phospho-tyrosine binding.

et al., 1990), and by screening a phage displayed peptide library biased for the presence of two consecutive prolines (Chen et al., 1997; Linn et al., 1997). Both lines of investigation confirmed the xPPxY consensus sequence, and identified that the strongest binding ligands contain the PPPPY sequence found in WBP-1, WBP-2, and in the p53 tumor suppressor binding protein-2 (a potential alternative ligand for the WW domain).

The NMR solution structure of a 57 residue recombinant protein fragment containing the WW domain of hYAP (GSMSFEIPDD VPLPAGWEMA KTSSGQRYFL NHIDQTTTWQ DPRKAMLS QM NVTAPTS) in complex with a polyproline-rich ligand (GTP-PPPYTVG) revealed a twisted and slightly bent three-stranded antiparallel β -sheet structure (Fig. 1) (Macias et al., 1996). The β -1 strand of the sheet is made up of 7 residues, Gly16 to Thr22, while the sequence of Gln26 to His32 comprises strand β -2. The smallest of the three strands, β -3, runs from Gln35 to Trp39 and contains amino acid residues that have a high β -sheet propensity (Minor & Kim, 1994a, 1994b; Smith et al., 1994). Strands β -1 and β -2 are connected by a three residue γ -turn, Ser-Ser-Gly, while the connection of strand β -2 to β -3 is made through a type II β -turn where Ile and Asp occupy the $i + 1$ and $i + 2$ positions. Located on the concave face of the β -sheet are the side chains of Tyr28, Leu30, Trp39, and a series of Thr methyl groups which define the hydrophobic surface that is involved in xPPxY-mediated ligand binding. The second conserved Trp at position 17 is located on the convex face of the bent sheet and appears to be important for making contacts with residues on the β -2 and β -3 strands. Recently, crystallographic analysis of the mitotic rotamase Pin1 produced the second structure of a WW domain, revealing a high degree of structural similarity to the WW domain of hYAP (Ranganathan et al., 1997).

We aim to use the WW domain as a model system to investigate the thermodynamic and kinetic principles governing β -sheet folding and stability. The WW domain is the smallest isolated, natu-

rally occurring three-stranded β -sheet identified to date. Although β -sheets and α -helices are both predominant forms of secondary structural elements found in proteins, considerably more effort has been put into elucidating the principles that govern α -helix formation. Investigation of small helical peptide model systems has provided significant insight into the nature of α -helical folding (Munoz & Serrano, 1994; Baldwin, 1995), but the corresponding development of a short peptide or peptidomimetic system for β -sheets has proven to be more difficult, primarily due to the propensity for isolated β -sheet structures to aggregate in aqueous environments (DeGrado & Lear, 1985; Osterman & Kaiser, 1985; Schneider & Kelly, 1995; Choo et al., 1996; Nesloney & Kelly, 1996), although recent progress is notable (Kortemme et al., 1998; Schenck & Gellman, 1998; Sharman & Searle, 1998a, 1998b).

The three-stranded antiparallel β -sheet of the WW domain of YAP adopts a context independent sheet structure that is ideal in size to be used as a model system for investigating the contributions that the hydrophobic effect, electrostatic interactions, hydrogen bonding, hydrophobic clusters in the denatured state, and β -turn propensities make to β -sheet folding and stability. These interactions appear to drive the formation of, and impart stability to, β -sheets in aqueous solutions. The growing number of naturally occurring WW sequences (~ 100) should help us sort out the importance and specificity of side-chain-side-chain interactions (Sudol, 1996). The monomeric nature of WW is critical, as this allows dissection of the energetic contributions that drive β -sheet formation in the absence of competing intermolecular interactions that lead to aggregation (Choo et al., 1996; Kelly, 1996, 1998). Through mutagenesis and other experiments it is our goal to understand the key molecular features that prevent aggregation, which is important for studying diseases based on protein misfolding and assembly (Colon & Kelly, 1992; Kelly, 1996, 1998). Three groups have recently reported de novo designed peptides capable of folding into a three-stranded antiparallel β -sheet structure (Kortemme et al., 1998; Schenck & Gellman, 1998; Sharman & Searle, 1998a, 1998b). One of these designs provides evidence that packing between side chains projecting on the same face of a β -sheet greatly influences stability (Kortemme et al., 1998), while others have invoked hydrogen bonding and turn propensities as being the predominant factors that facilitate β -sheet folding (Constantine et al., 1995). Clearly, we are still in the early stages of understanding the underlying forces and principles that govern the uniqueness and stability of a three-stranded β -sheet fold (Tsang et al., 1994). Further development and characterization of water soluble β -sheets based upon the WW amino acid sequence that are amenable to both chemical (P. Raman & J. Kelly, pers. comm.; Camarero et al., 1998) and recombinant synthesis will expand our fundamental knowledge of β -sheets, and should provide significant insight into understanding the thermodynamic and kinetic factors that influence β -sheet folding and stability. Herein, evidence is provided for the suitability of WW as an interesting β -sheet model system, in addition to demonstrating the presence of a hydrophobic cluster in the chaotrope mediated denatured state, which may be important for β -sheet folding.

Results

Expression and purification of the WW domain

The recombinant WW domain was purified from the soluble cell lysate of *Escherichia coli* BL21 (DE3) by a double affinity chro-

Fig. 1. Ribbon diagram of the polypeptide backbone of the isolated WW domain from the human yes associated protein (hYAP). The side chains or residues Trp17 and Trp39 are illustrated. This figure was prepared using MOLSCRIPT (Kraulis, 1991) based on the coordinates of Macias et al. (1996).

matographic approach, employing glutathione S-transferase and a poly-histidine affinity tag (Macias et al., 1996). A 3 L culture of LB media produced between 150–200 mg of (His)₆-GST-WW fusion protein after affinity chromatography with Glutathione Sepharose 4B resin (Pharmacia, Piscataway, New Jersey). The WW domain was cleaved from the (His)₆-GST affinity tag/carrier protein by incubation with bovine thrombin, and following immobilized metal chelation chromatography on Ni-NTA (Quiagen, Inc., Valencia, California) to sequester the (His)₆-GST portion, the WW domain was purified to homogeneity by gel permeation chromatography (HiLoad 16/60 Superdex 30 column (Amersham Pharmacia Biotech, Piscataway, New Jersey)). A yield of 20 mg of pure WW was generally attainable. Purified WW migrated as a single band on SDS-PAGE and eluted as a single peak when analyzed by reverse-phase HPLC (C18). Five cycles of N-terminal sequencing yielded the expected sequence of GSMSEF. The mass of the WW peptide determined by both electrospray and MALDI mass spectrometry was 6,372.2 Da (expected mass 6,372.24 Da).

Urea and GdnHCl-induced denaturation and reconstitution monitored by fluorescence

The fluorescence spectrum of a 1 μ M solution of the WW domain in 10 mM potassium phosphate, 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, and 0.02% NaN₃ pH 6.0 (buffer A) shows a single broad emission band with the maximum fluorescence centered at 341 nm (Fig. 2). Upon the addition of urea, the quantum yields of one or both of the tryptophans diminish and the wavelength of maximum emission red-shifts to lower energy, consistent with the exposure of the tryptophan indole groups to the bulk solvent environment. In 8.0 M urea, the emission of the WW sample decreased by approximately 45% and the wavelength of maximum

emission red shifted to 350 nm (Fig. 2). To evaluate the reversibility of denaturation and refolding of the WW domain, a concentrated protein stock solution was initially prepared in 8.0 M urea and diluted accordingly to yield 1 μ M WW solutions with urea concentrations ranging from 0.5–8 M. The 0.5 M urea WW sample, the lowest urea concentration achievable by dilution, produces a sample that yields an emission spectrum with a maximum at 341 nm (consistent with the folded structure) and one that regains nearly 95% of the native WW fluorescence signal (Fig. 2). This suggests that the WW domain, having an ensemble of conformations in 8.0 M urea (see below), folds back to a compact, native conformation upon dilution of the chaotropic agent. Upon dilution of the chaotropic agent, the tryptophans re-orient themselves (especially Trp17) into an environment that is protected from collisional quenching, resulting in a blue-shifted spectra with enhanced fluorescence.

The denaturation and reconstitution fluorescence-based unfolding/refolding curves of WW samples (1.0 μ M) monitored at 340 nm as a function of urea concentration largely superimpose upon each other suggesting a reversible denaturation-renaturation folding pathway (Fig. 2). The denaturation curve exhibits a 0–1 M pre-transition region, followed by a sigmoidal unfolding region characterized by a nonlinear reduction in the fluorescence at 340 nm with increasing urea concentration between 1.0 and 6.0 M urea. In the post-transition concentration range above 7.0 M urea, both the emission intensity and the wavelength of maximum emission remain nearly constant. Transformation of the data from the transition region (1.0–6.0 M urea) yields a linear plot of ΔG_{app} against urea concentration ($r^2 = 0.995$), from which the values of ΔG_{H_2O} and m can be calculated. The ΔG_{H_2O} for the WW domain is 2.24 kcal mol⁻¹ and the m -value (sensitivity to denaturant) is 0.58 kcal mol⁻¹ M⁻¹, assuming that a two state assumption holds (see below). At the midpoint of the unfolding transition, the concentration of urea $[D]_{1/2}$ is 3.86 M. Using m and $[D]_{1/2}$ from the linear transformation of the denaturation data along with the pre- and post-transition baseline slopes and intercepts as initial regression values for a nonlinear least-squares fit to Equation 1, we calculate a best fit to the data, which is consistent with an ideal two-state model for protein folding (Fig. 2). Comparison of the calculated best-fit (solid line) with our experimental data suggests a two-state system, although slight variations are evident between the theoretical and experimental curves at both the low and high urea concentration limits. These small but reproducible deviations for samples having different histories indicate that there may be intermediates populated, particularly at higher urea concentrations, suggesting deviations from two-state behavior, which will be further evaluated in ongoing studies (see discussion) (Pace, 1986; Dill & Shortle, 1991; Shortle, 1993, 1996).

The denaturant guanidine hydrochloride (GdnHCl) was used in a series of experiments similar to those described above with urea. The behavior of the WW fluorescence emission spectra as a function of GdnHCl concentration resembles that observed with urea, with the exception that the pretransition baseline disappears because of the increased efficacy of GdnHCl as a denaturant. The addition of GdnHCl to a WW sample is accompanied by an immediate decrease in fluorescence intensity at 340 nm (Fig. 3). The majority of this signal loss occurs over the 0–1.5 M GdnHCl concentration range, and by approximately 2.0 M GdnHCl the WW domain appears devoid of any β -sheet structure. Reconstitution of the WW domain from 5.0 M GdnHCl produces a curve that coincides well with the denaturation profile, suggesting that these

Fig. 2. Urea-induced denaturation ($\blacktriangle$) and renaturation ($\times$) curves of 1 μ M WW in 10 mM potassium phosphate (pH 6.0) 100 mM NaCl, 0.1 mM DTT, 0.1 mM EDTA, and 0.02% NaN₃ (buffer A) monitored by fluorescence at 20 °C. Sample excitation was at 295 nm while emission was monitored at 340 nm, with both the excitation and emission slitwidths set to 1 nm. The solid line represents the best fit to a two-state model for protein folding according to Equation 1. Emission spectra taken at the start and completion of the titration experiment are shown in the inset. WW in 0 (●) and 8.0 M (▼) urea during the unfolding experiment, and WW in 8.0 M (Δ) and 0.5 M (○) urea during reconstitution.

Fig. 3. GdnHCl-induced unfolding ($\blacktriangle$) and refolding ($\times$) curves of 1 μ M WW in buffer A (pH 6.0) monitored by fluorescence at 25 $^{\circ}$ C. The excitation and emission monochromator were set to 295 and 340 nm, respectively, using slit widths of 2 nm for each. The best fit to a two-state model according to Equation 1 is represented by the solid line. The inset shows selected emission spectra taken before and after the addition of GdnHCl for both the unfolding and refolding experiment; WW in 0 M (∇) and 5 M GdnHCl ($\circ$) during denaturation, and in 5 M ($\bullet$) and 0.2 M ($\blacktriangledown$) GdnHCl during reconstitution.

two events proceed via a reversible pathway(s). During the course of renaturation, the fluorescence of the unfolded WW domain shows little dependence upon the GdnHCl concentration from approximately 5.0 to 2.5 M GdnHCl. Below 2.0 M GdnHCl, the fluorescence signal increases sharply indicating the onset of WW folding. Clearly, the majority of GdnHCl-mediated unfolding and refolding of the WW domain occurs in the 0–1.5 M concentration range, while in urea this transition occurs over a much larger range (1.0–6.0 M), consistent with GdnHCl's more powerful denaturation properties. Nonlinear least-squares analysis of the unfolding data according to the two-state model described by Equation 1 produces a good fit to the experimental data (solid line), consistent with a model where WW folds reversibly from the denatured ensemble to the native state without significantly populating intermediate species (Fig. 3). However, small reproducible deviations observed at high GdnHCl concentration for samples having different histories indicate that there may be intermediates populated, particularly at high GdnHCl concentrations (Fig. 3; Dill & Shortle, 1991; Shortle, 1993, 1996).

Urea and GdnHCl-induced denaturation and reconstitution monitored by near-UV CD

Near-UV CD studies on the WW domain establish that the transitions observed by fluorescence and far-UV CD as a function of urea are not superimposable (Fig. 4), providing evidence against a two-state transition. The near-UV CD transitions are less cooperative than the fluorescence based urea denaturation and refolding curves. Even more intriguing is the discovery that the near-UV CD signal exhibited by WW does not disappear at chaotrope concentrations far into the post-transition region (5.0 M GdnHCl or 6.0 M urea) (Fig. 5). The near-UV CD spectrum of WW in 3.0 M urea is nearly as intense as that exhibited by normally folded WW. The

Fig. 4. WW unfolding transitions monitored by near-UV CD ($\circ$) (290 nm) and fluorescence emission intensity ($\bullet$) (340 nm, data from Fig. 2), as a function of urea concentration. Near-UV CD measurements were made in triplicate with independent 1.0 mg/mL WW samples in buffer A (pH 6.0) in the presence of varying urea concentrations using a 1.0 cm pathlength cell.

near-UV CD intensity decreases above 3.0 M urea up to a concentration of 6.0 M, but does not disappear, even at very high urea or GdnHCl concentrations (Fig. 5). The display of a near-UV CD spectrum by the hYAP WW domain at high denaturant concentrations suggests the presence of a hydrophobic cluster-based intermediate having a side-chain composition that is likely aromatic. The formation of a hydrophobic cluster in the “denatured state(s)” could prove to be very important for the nucleation of a unique β -sheet fold and/or for avoiding aggregation. These data imply that the urea or GdnHCl denatured state is not a random coil, instead it seems to retain residual structure that involves hydrophobic side-chain interactions. Such hydrophobic clusters have been observed previously in the conversion of a β -hairpin to a three-stranded β -meander and in the folding of β -hairpins nucleated by unnatural amino acids (Neri et al., 1992).

Fig. 5. Near-UV CD spectra of 1.0 mg/mL WW solutions in buffer A (pH 6.0) acquired in urea and GdnHCl. WW in 0 ($\bullet$), 3.0 ($\blacktriangle$), and 6.0 M ($\circ$) urea, and in 5.0 M (∇) GdnHCl. The spectra are averages of three consecutive scans that have been corrected by subtracting buffer blanks containing the identical concentration of denaturant as was present in each WW sample. All scans were made using a 1.0 cm pathlength cell with the CD sample holder thermostated to 20 $^{\circ}$ C.

The extent of hydrophobic surface area exposed by both native and chaotrope denatured WW domains was investigated by 1-anilinonaphthalene-8-sulfonic acid (1,8 ANS) fluorescence-detected binding studies. A 5 μ M solution of the folded WW domain (buffer A, pH 6.0) having a tenfold excess of 1,8 ANS (50 μ M) did not exhibit ANS binding, based on the lack of the quantum yield increase or emission blue-shift characteristic of binding to hydrophobic surfaces (Supplementary Fig. 1) (Lai et al., 1996). A WW solution adopting a hydrophobic cluster conformation(s) in 5.0 M GdnHCl or 6.0 M urea does not appear to display appreciable ANS binding either, suggesting minimal exposed hydrophobic surfaces large enough to bind to 1,8 ANS. In contrast, transthyretin, which is able to bind 2 mol of ANS in its hydrophobic binding channel, clearly exhibits the characteristics of an exposed hydrophobic surface (Lai et al., 1996).

Thermal unfolding and reconstitution of WW

The far-UV CD spectrum of the 57 residue WW domain recorded at 4 and 88 $^{\circ}$ C during the course of a thermal denaturation and a cooling induced reconstitution experiment is presented in Figure 6. The native WW far-UV CD spectrum is characterized by two peaks, a small absorption with a positive ellipticity centered around 230 nm and an intense negative peak at 202 nm. The spectrum does not exhibit the classical β -sheet signal designated by a negative peak at 217–218 nm and a positive peak at 195–197 nm (Woody, 1985, 1987, 1993; Manning & Woody, 1987) but more closely resembles the signal of an ensemble of less-ordered states (Fasman, 1996). However, the positive peak at 230 nm, originating from the aromatic contribution to the far-UV CD spectrum (Woody, 1994; Krittanaï & Johnson, 1997) as well as the negative peak above 200 nm, is consistent with the spectra exhibited by several small structurally characterized β -sheet structures, suggesting that the WW domain is folded under these conditions (the N- and C-termini are disordered in the folded structure, partially explaining the random coil-like CD spectrum) (Viguera et al., 1994; Knapp et al., 1998; Reid et al., 1998).

The thermal unfolding profile of the WW domain was monitored at 202 and 230 nm in the temperature range from 4 to 88 $^{\circ}$ C, both

Fig. 6. Far-UV CD spectra of 50 μ M solutions of the WW domain in buffer A (pH 6.0) acquired at 4 $^{\circ}$ C (●) and 88 $^{\circ}$ C (▼) during the course of a thermal denaturation and reconstitution experiment in a 0.1 cm path-length cell. A spectrum of WW recorded at 4 $^{\circ}$ C (▽) after cooling from 88 $^{\circ}$ C. A 5 min thermal equilibration period was used prior to making a CD measurement.

Fig. 7. Thermal unfolding (×) and refolding (●) of the WW domain in buffer A (pH 6.0) monitored by CD at 230 nm. A 5 min thermal equilibration period was established at each new temperature. The best fit to the data assuming a two state transition described by Equation 2 is illustrated by the solid line.

during the course of heat-induced unfolding and cooling-induced refolding. At 202 nm no significant changes in the magnitude of the CD signal were observed up to approximately 80 $^{\circ}$ C, but a slight loss in signal was noticed at 88 $^{\circ}$ C (Fig. 6). The 230 nm far-UV CD maximum provided a much better means of monitoring the thermal unfolding and refolding events of the WW domain, as a change of approximately 10 mdeg was observed when comparing the 4 and 88 $^{\circ}$ C spectra. In this temperature range, a single unfolding transition was evident by monitoring the 230 nm far-UV CD signal, with a midpoint unfolding temperature (T_m) of 48.9 ± 0.6 $^{\circ}$ C (Fig. 7). The ellipticity changes observed during the renaturation (cooling) cycle of the experiment matched well with the denaturation (heating) profile, with a recovery of CD signal that varied between 80–100% of the value measured at 4 $^{\circ}$ C. Visible inspection of the renatured sample revealed no signs of precipitation (even when signal recovery was observed to be 80%), and analysis by analytical ultracentrifugation confirmed the absence of any higher ordered oligomeric species. The variation in signal recovery is being further investigated and appears to be associated with the amount of time the sample spends around 88 $^{\circ}$ C. The solid line in Figure 7 is the nonlinear least-squares fit to the denaturation data assuming two-state denaturation behavior described by Equation 2. The calculated best fit superimposes well on the experimental data, consistent with two state thermal unfolding and refolding behavior.

The near-UV CD-based thermal unfolding and reconstitution curves (290 nm) for the WW domain are very similar (Fig. 8A). Furthermore, the far-UV CD-based denaturation and refolding curves at 230 nm occur over the same temperature range. Interestingly, the near-UV CD spectrum at 88 $^{\circ}$ C suggests nearly complete unfolding by thermal denaturation based on the absence of a near-UV CD signal at high temperature (88 $^{\circ}$ C) (Fig. 8B). This is in stark contrast to the near-UV CD spectra exhibited by WW in concentrated chaotrope solutions, where the near-UV CD signal is not lost. The similarity of the near- and far-UV CD-based thermal folding and reconstitution curves suggests that this transition may be a two-state transition; however, at this point we cannot be certain of this. Unfortunately, the far-UV CD (230 nm) and near-UV CD data (290 nm) are both likely measuring the asymmetry of aromatic side chain packing. It would be better to have a second probe that measures β -sheet structural changes, but this is not

Fig. 8. A: Thermal denaturation (○) and reconstitution (●) of 1.0 mg/mL WW monitored by near-UV CD spectroscopy at 290 nm. The sample was heated and cooled from 4 to 88 °C, with measurements taken at 4 °C intervals. A 5 min thermal equilibration period was established at each temperature prior to making a measurement. B: Near UV-CD spectra of 1.0 mg/mL WW solutions recorded during the course of a thermal denaturation/reconstitution experiment. WW at 4 °C (○) at the start of the heating cycle, at 88 °C (●), and at 4 °C (+) upon completion of the cooling cycle.

possible using the amide-based far-UV CD signal in this case. The transitions traced out in Figures 7 and 8A are very similar; the T_m s only differ by 2 °C suggesting the absence of intermediates. The NMR studies planned as a function of temperature will allow us to evaluate denaturation and folding on a residue specific basis, which is essential in the search for intermediates. Future kinetic studies of WW undergoing thermal and chaotropic denaturation/reconstitution should prove useful for detecting intermediates and for evaluating the suitability of a two-state model.

Illustrated in Figure 9 are the ^1H NMR spectra of a 2 mM WW sample recorded at 5 and 45 °C. At the lower temperature, the amide chemical shift region shows good resonance dispersion and sharp peaks throughout the spectrum, consistent with a folded structure (Wüthrich, 1986; Macias et al., 1996). Once heated to 45 °C (the temperature limit of our probe, close to the 48.9 °C T_m of the WW domain), the N-H dispersion decreases significantly and the chemical shifts of these protons begin to coalesce, ap-

proaching values observed for an ensemble of less ordered states. Further NMR studies with a probe allowing access to elevated temperatures and monitoring the whole spectrum of WW up to a temperature of 80 °C are necessary to further understand the extent of thermal denaturation as a function of temperature.

Extent of WW aggregation or lack thereof monitored by analytical ultracentrifugation

To evaluate the aggregation propensity of the WW domain, we analyzed this protein by analytical ultracentrifugation at a series of different concentrations. From sedimentation equilibrium analysis of WW (25 μM sample in buffer A), the solution MW of the WW domain was determined to be $6,502.0 \pm 42.1$ Da, a value that compares well with the theoretical mass of 6,372.24 Da (Fig. 10). The sedimentation equilibrium data are evaluated by nonlinear least-squares analysis using the Origin software package provided

Fig. 9. ^1H NMR spectra of a 2 mM solution of WW in 90% buffer A (pH 6.0), 10% D_2O at 4 °C (top) and 45 °C (bottom) taken during the course of a partial thermal unfolding and refolding experiment. The spectra have been truncated to emphasize the chemical shift dispersion of the indole and amide bond protons.

Fig. 10. Sedimentation equilibrium analysis of the recombinant WW domain (25 μM) in buffer A (pH 6.0). The solid line drawn through the data was obtained by fitting the absorbance vs. radial position data to Equation 3 for a single ideal species. The residuals between the fit and the experimental data are shown in the top panel.

by Beckman with the assumption that the WW domain behaves as a single ideal solute described by Equation 3. The validity of this assumption is supported by evaluation of the residuals, which are the calculated differences between the experimental data and the best fit model (Fig. 10). In this case, the residuals are small and distributed randomly over the length of the radial distance examined with approximately the same variance across the cell, indicating that these data are consistent with a single ideal species model. Similar equilibrium experiments were conducted with WW at a concentration of 69, 120, 220, and 440 μM , well above the concentrations used for the fluorescence and far-UV CD-based biophysical experiments during the course of this study. Up to a concentration of 440 μM , WW behaved as a single ideal solute that did not self-associate. The 2 mM WW samples analyzed by NMR produced sharp resonances with narrow linewidths as expected for a protein of this size and did not visually scatter light, even at elevated temperatures, observations consistent with an absence of aggregation. Equilibrium analysis further showed that a 21.5 μM WW solution reconstituted from 8.0 M urea back into buffer A was still monomeric, like native WW at a comparable concentration.

Equilibrium ultracentrifugation studies were also conducted to investigate the solution stoichiometry (potential aggregation) of the WW domain under conditions where the hydrophobic cluster conformation exists. The WW domain (100 μM) in 6.0 M urea behaves as a single ideal nonassociating species, as does the WW domain (100 μM) in 5.0 M GdnHCl, exhibiting weight average molecular weights of $6,123 \pm 115$ and $5,797 \pm 105$ Da, respectively (Supplemental Figs. 2, 3). Both calculated masses are slightly lower than the expected value of 6,372.24 Da, likely due to the error involved in estimating the partial specific volume under these conditions. However, it is clear that the WW domain does not self-assemble under these conditions.

A 180 μM WW domain sample was subjected to a sedimentation velocity run carried out at 50,000 rpm to further scrutinize the possible presence of oligomers. Under these experimental conditions, the g-force applied to the sample would quickly sediment any large self-assembled species to the bottom of the cell, if present. The WW domain sedimented as a single broad boundary that was fit using direct svedberg analysis, yielding a sedimentation coefficient of 0.85 s (Supplemental Fig. 4). Attempts to fit the sedimentation data to a nonmonomeric model failed, further confirming that at this concentration the WW domain exists as a monomer. Similar sedimentation velocity analysis of a WW sample that went through a heating (88 $^{\circ}\text{C}$) and cooling cycle in the CD spectrometer, but which only regained 80% of the initial CD signal upon reconstitution at 4 $^{\circ}\text{C}$, also provided no evidence for WW aggregation.

Determination of the dissociation constant (K_d) of the PY-WW domain complex

A synthetic peptide derived from the internal sequence of the p53 binding protein-2 EYPPYPPPYPSG (PY peptide, residues 742 to 754 of p53 binding protein-2), previously characterized as a second generation ligand for the WW domain (Linn et al., 1997), was used to probe the structural integrity of the recombinant WW fragment. The NMR solution structure of the WW domain complexed with the first generation ligand (GTPPPYTVG) revealed a hydrophobic-rich cleft defining the ligand binding site of the WW domain. The indole of Trp39 is part of the xPPxY binding site,

hence the binding of xPPxY ligands to WW should produce an enhancement in the fluorescence emission of the protein. Figure 11 illustrates the observed increase in the Trp fluorescence of the WW domain after the addition of increasing amounts of PY peptide, ultimately saturating the binding site. Binding of the PY peptide to the WW domain results in an increase in the fluorescence intensity and a shift in the fluorescence maximum to a shorter wavelength, indicating that Trp39 becomes buried and protected from the bulk solvent environment. A direct fit of the fluorescence titration data using a nonlinear least-squares approach to Equation 4, with the assumption that the binding stoichiometry is 1:1, yields a K_d of 6.0 ± 0.34 μM . The best fit (solid line) and the experimental data are in excellent agreement, indicating ideal ligand binding behavior (Fig. 11). A similar ligand binding titration in 6.0 M GdnHCl showed no increase in fluorescence, consistent with the expectation that the polyproline rich ligand should not bind to unfolded or partially folded (see below) WW (data not shown). The WW domain initially denatured in 6.5 M urea and refolded and dialyzed to remove the last traces of urea binds the EYPPYPPPYPSG ligand analogous to WW not undergoing a denaturation/reconstitution cycle, indicating that a functional fold is produced upon reconstitution.

Discussion

WW unfolds and refolds reversibly in urea and GdnHCl

The primary structure of the WW domain contains two tryptophan residues, one at position 17, which is located in the β -1 strand of the sheet, and the other at position 39, found at the C-terminus of strand β -3 (Macias et al., 1996). The molecular packing around Trp17 is such that this residue is constrained into an environment that gives it some protection from the bulk solvent, while Trp39 is moderately surface exposed. The changes in fluorescence observed during the course of the chaotropic unfolding and refolding experiments are likely dominated by Trp17, which interacts with residues on the β -2 and β -3 strands, while ligand binding induced fluorescence changes are probably reported primarily by Trp39.

Urea and GdnHCl appear to act by stabilizing the unfolded state of proteins (Timasheff, 1992). Unfolding of the WW domain in

Fig. 11. Titration of a 5 μM solution of the WW domain in buffer A (pH 6.0) with the PY peptide at 25 $^{\circ}\text{C}$. Excitation was at 295 nm and emission at 340 nm. The excitation and emission slits were both set to 2 mm. The solid line is the best fit to the data assuming ideal ligand binding behavior described by Equation 4.

8 M urea results in a 45% decrease in the fluorescence emission intensity at 340 nm, and red-shifts the maximum emission wavelength 9 nm from 341 to 350 nm (Fig. 2). When the denaturant was 5 M GdnHCl, the fluorescence loss at 340 nm relative to native WW was 63.2%, and the red-shift of the maximum emission wavelength was more pronounced, shifting 15 nm to 355 nm (Fig. 3). These results suggest that urea and GdnHCl interact with the WW domain "denatured" states to different degrees (Pace et al., 1990; Monera et al., 1994), and suggest that the WW domain may retain some structure in urea and GdnHCl. The non-native near-UV CD spectra exhibited by WW in the transition and post-transition regions of the chaotropic denaturation/reconstitution curves strongly suggest that structure is retained at high denaturant concentrations. Instead of a random coil denatured state (lacking side-chain-side-chain interactions), a hydrophobic cluster appears to form under denaturing conditions, a state that has been observed previously in other proteins (Dill & Shortle, 1991; Neri et al., 1992; Shortle, 1993, 1996; Logan et al., 1994; Wang & Shortle, 1995; Dill et al., 1995). The differences in the chaotrope dependence of the fluorescence- and near-UV CD-based unfolding curves also suggest that a two-state assumption is likely not valid (Fig. 4). Fluorescence- and near-UV CD spectroscopy likely detect different intermediates, hence the nonsuperimposable nature of the transitions. Significant experimental support exists for the presence of structure within proteins at high chaotrope concentrations (post-transition region). The structured denatured state of WW has significant implications for understanding ΔG_{H_2O} (see below). Structure in the denatured state of WW also explains the surprisingly fast cyclization rate of WW observed by Muir and colleagues during the preparation of cyclic WW analogs (Camarero et al., 1998). Typically the cyclization rate of folded proteins with proximal N- and C-termini (activated ester) decreases markedly under denaturing conditions, which is not the case for WW owing most likely to hydrophobic cluster conformation in urea and GdnHCl.

The equilibrium denaturation and reconstitution curves are largely superimposable in urea and GdnHCl and show single unfolding/refolding transitions, demonstrating reversibility of WW domain denaturation and reconstitution (Figs. 2, 3). In urea, the fluorescence-based unfolding/reconstitution curves can be fit to a two-state model, showing only slight deviations from the ideal case at low and high denaturant concentrations. In the case of a protein that unfolds into a random coil, the two-state approximation allows one to determine the stability of the folded state relative to an ensemble of unfolded states lacking side-chain-side-chain interactions (Shortle, 1996). In the case of WW, interpreting the data is more challenging because the denatured states are structured, hence the ΔG measurement is evaluating the free energy difference between the unique folded state and the structured denaturation state, which likely does not reflect the ΔG between the folded and unfolded states. Hence the 2.24 kcal/mol is almost surely a lower limit on the stability of the WW domain. In principle, we could get ΔG_{H_2O} by knowing the ΔG of two equilibria, folded WW to hydrophobic cluster and hydrophobic cluster to random coil; however, it has been documented that it is very difficult to denature hydrophobic clusters (Dill & Shortle, 1991; Neri et al., 1992; Shortle, 1993, 1996; Logan et al., 1994; Wang & Shortle, 1995; Choo et al., 1996). A thermodynamic analysis of CspA, another small all β -sheet protein, yielded values of 3.0 ± 0.1 kcal mol⁻¹ for ΔG_{H_2O} , similar to the lower limit of the stability of the WW domain (Reid et al., 1998).

A structured denatured state for WW was anticipated and serves as a good system to test the hypothesis that such preorganized

denatured states are important for the folding of isolated three-stranded antiparallel β -sheets. The bad news about structured denatured states, as pointed out by Shortle, is that they make it more difficult to interpret the thermodynamic effect of mutations because mutations can effect the folded state, the hydrophobic cluster states, or both (Shortle, 1996). However, the good news is that significant advancements in our understanding of β -sheet folding may result from considering structured denatured states. It is clear that both states (folded and hydrophobic cluster) must be understood to appreciate the energetics associated with the interconversion of two states. Experimental methods to probe the structure of the denatured state have quickly evolved and include multidimensional NMR studies, H/D exchange studies, small angle X-ray scattering, and kinetic methods to cite a few (Dyson & Wright, 1991; Koide et al., 1993; Shortle, 1996).

Reversibility of the thermal unfolding transition

Thermal denaturation is a mechanistically complex process that appears to act largely on the denatured state by lowering its free energy ($-T\Delta S$). The reversibility of thermal unfolding of WW and cooling induced reconstitution has been established by the 230 nm far-UV and 290 nm near-UV CD data as a function of temperature. Thermal denaturation appears to be more complete than chaotropic denaturation as discerned by the loss of the near-UV CD signal at high temperature (Fig. 8B). The thermal unfolding profile of the WW domain monitored by the far-UV CD signal at 230 nm exhibits a single transition with a melting temperature of 48.9 ± 0.6 °C (Fig. 7). Fitting these data to Equation 2, the equation for a two-state model, produces a best fit, which is in good agreement with the experimental data. Unlike the case of urea denaturation (fluorescence at 340 nm and near-UV CD at 290 nm), two spectroscopic methods (far- and near-UV CD) afford thermal transitions that have very similar temperature dependencies, consistent with a two-state model. As discussed above, more demanding experiments will be needed before definitive conclusions can be drawn regarding the appropriateness of a two-state model for thermal unfolding. During the course of refolding, the ellipticity readings at 230 nm generally produced an excellent match with the values measured during unfolding as a function of temperature, down to approximately 20 °C. For some samples, however, the signal below 20 °C observed during reconstitution decreased relative to that observed in the unfolding direction, such that the extent of reversibility varied from 80 to 100%. Selected samples that did not regain 100% of the native signal upon cooling, particularly those at the low end of the recovery range, were analyzed by equilibrium analytical ultracentrifugation, which unequivocally demonstrated that the loss of signal was not associated with protein aggregation. Further studies to understand this variation, which seems to correlate with extent of incubation at 88 °C, are underway. For example, refolding is fully reversible by far-UV CD (230 nm) when the temperature range is limited to 45 °C (data not shown).

The reversibility of thermal unfolding was also probed by ¹H NMR up to 45 °C, the temperature limit of our current probe. We collected data over the range of 5 to 45 °C in 5 °C increments for both heating and cooling cycles to evaluate reversibility. Over this temperature range we observed 100% reversibility as evaluated by the chemical shift changes of the Trp17 and Trp39 indole N-H protons (data not shown). A comparison of the proton spectra of WW taken at 5 °C to a sample subjected to a heating and cooling

cycle produced an exact match. The NMR and 230 nm far-UV CD data demonstrate that the unfolding and refolding pathways up to 45 °C are fully reversible. Unfolding of the WW domain at 45 °C, just below the T_m , is significant as judged from the NMR spectrum, consistent with a more complete thermally unfolded state. An NMR study above 45 °C should prove useful for evaluating the extent of denaturation in the post-transition region. Such experiments may also provide insight into the lack of reversibility in some cases. Since the WW domain contains five Pro residues, perhaps the loss of CD signal observed for some of our samples may be associated with a slow cis-trans proline isomerization event. Alternatively, high temperature may have induced a chemical transformation within the WW domain, perhaps some deamination, which may affect the refolding event. A detailed study should help us gain further insight into the origin of this phenomenon.

Monomeric nature of the WW domain in an aqueous environment

A major problem encountered when studying model β -sheets is their propensity to aggregate in an aqueous solvent, especially partially folded β -sheet rich intermediates. The nature of β -sheet formation requires strand-strand interactions, which can lead to aggregation when these contacts occur intermolecularly (Smith & Regan, 1997). The WW polypeptide is remarkable in that it is very resistant to aggregation, perhaps the hydrophobic cluster formed in urea and GdnHCl solutions and the more extensive denatured state at elevated temperature is how Mother Nature prevents sheet-mediated aggregation. At concentrations up to 440 μ M, we did not detect any higher ordered oligomers or aggregates by either equilibrium or velocity studies. This was also the case for a WW sample that had been re-analyzed 4 days after initial equilibrium analysis, a sample that had been reconstituted from 8.0 M urea, and one that showed approximately 80% signal recovery from a CD thermal unfolding/refolding experiment. Sedimentation equilibrium analysis showed that the protein behaves as an ideal species with a molecular mass of $6,502.0 \pm 42.1$ Da, corresponding to monomeric WW (calculated MW = 6,372.24 Da). The WW domain also remains monomeric in 6 M urea and 5 M GdnHCl, conditions populating the hydrophobic cluster conformation. Sedimentation velocity experiments further verified the absence of soluble oligomeric species or insoluble aggregates. The velocity studies demonstrated that the WW domain sediments as a single ideal species with a sedimentation coefficient of 0.85 s, consistent with a monomeric solution structure.

Ligand interaction with the PY peptide

To demonstrate that recombinant WW has a functional fold, we titrated the WW domain with the PY peptide, a ligand derived from the sequence of the p53 (tumor suppressor) binding protein-2 that contains the second generation WW binding consensus sequence identified by Sudol et al. (Linn et al., 1997). The binding of the PY peptide to the WW domain is accompanied both by an enhancement in the fluorescence emission and by a blue-shift in the maximum emission wavelength. These changes are the result of peptide binding to the hydrophobic patch located on the surface of the WW domain, thereby protecting the excited state of Trp39 from solvent induced collisional quenching. The dissociation constant for the WW-PY peptide complex is 6.0 ± 0.34 μ M, which is approximately one order of magnitude lower than the values determined

for a series of peptides with the xPPxY consensus sequence (Macias et al., 1996). The WW domain denatured in 6.5 M urea and reconstituted by dilution is capable of binding to the PY peptide, demonstrating the reformation of a functional fold.

Conclusion

This represents the first in what will likely be a series of papers utilizing the WW domain as a system to understand more about the thermodynamics and kinetics of β -sheet folding. Based on what we currently know, the WW domain appears to be an ideal polypeptide to study β -sheet folding and is the best characterized three-stranded antiparallel β -sheet reported. The WW domain from hYAP is accessible from recombinant sources (Chen & Sudol, 1995) and from chemical synthesis (Camarero et al., 1998). It has been structurally characterized by NMR, and crystallography is underway. In addition, there is a functional assay (xPPxY binding) that is useful for demonstrating proper folding and for future selection experiments. The WW domain certainly folds reversibly utilizing both urea and GdnHCl as denaturants; however, the denatured state retains structure in the form of a hydrophobic cluster involving at least one aromatic side chain. Further studies on the structured denatured state of WW may prove particularly enlightening in terms of the role the denatured state plays in the design and stability of three stranded β -sheets and the role such states may play in preventing aggregation. The disadvantage of the structured denatured state is that it makes the thermodynamic measurements hard to interpret. In contrast to chaotropic denaturation, thermal denaturation may be a two-state process, where the denatured state is largely unstructured. The variability observed in the thermal reconstitution yield is currently under evaluation by several approaches and appears to be caused by prolonged incubation at high temperature in the post-transition region.

Materials and methods

Expression and purification of recombinant WW domain

The WW domain of the human yes associated protein (hYAP), GenBank accession number X80507, was introduced into a modified glutathione S-transferase (GST) fusion vector pGAT2 and expressed in *E. coli* BL21 (DE3) in the form of a N-terminal His-tagged, GST-fusion protein (Macias et al., 1996). Protein production was carried out at 37 °C in LB media supplemented with 100 μ g/mL ampicillin following induction with 1 mM isopropyl- β -D-thiogalactopyranoside (IPTG). Initial fusion protein purification was achieved by passing the total cell lysate through Glutathione Sepharose 4B affinity resin (Amersham Pharmacia Biotech, Piscataway, New Jersey) as recommended by the manufacturer, followed by elution with a solution of 10 mM glutathione (Calbiochem, La Jolla, California) dissolved in 50 mM Tris-HCl, pH 8.0 (buffer B). The WW domain was cleaved from the His-tagged, GST-fusion protein directly in buffer B by the addition of bovine thrombin (Sigma T4648, St. Louis, Missouri) protease at a ratio of 2 U of enzyme per milligram of fusion protein. Following overnight proteolysis at room temperature (≈ 22 °C) with gentle mixing on a Ferris wheel style lab rotator, the GST fragment was removed from the target protein by passing the reaction mixture through Ni-NTA-Agarose resin (Qiagen, Inc., Valencia, California). The WW containing eluant flow through was concentrated by ultrafiltration with

a YM3 membrane (Amicon, Inc., Beverly, Massachusetts), and then purified to homogeneity by gel permeation chromatography on a HiLoad 16/60 Superdex 30 column (Amersham Pharmacia Biotech, Piscataway, New Jersey). SDS-PAGE modified for peptide and small protein analysis was used to visualize the WW-domain on polyacrylamide gels (Schagger & von Jagow, 1987). The identity of the purified WW domain was confirmed by N-terminal protein sequencing and by both MALDI and electrospray mass spectrometry. The concentration of WW was determined by measuring the absorbance at 280 nm in 6.0 M GdnHCl using a calculated absorption coefficient of $12,660 \text{ M}^{-1} \text{ cm}^{-1}$.

PY peptide synthesis and purification

The proline-rich PY peptide (EYPPYPPPYPSG) was synthesized on an ABI 433A Peptide Synthesizer (Applied Biosystems, Inc., Foster City, California) using "Fastmoc" chemistry on Rink amide resin (Advanced Chemtech, Inc. Louisville, Kentucky). After cleavage from the solid support and side-chain deprotection, the PY peptide was purified to homogeneity by reverse-phase chromatography on C18 and analyzed by electrospray mass spectrometry. Additional details on the synthesis and purification protocol of the PY peptide are available upon request.

Chemical denaturants

Electrophoresis grade urea from Fisher Biotech and ultraPURE™ guanidine hydrochloride from GibcoBRL Life Technologies (Gaithersburg, Maryland) were used without further purification. Stock solutions of the denaturants were freshly prepared at room temperature, filtered through $0.2 \mu\text{m}$ membranes, and analyzed by refractometry to determine the exact concentration (Nozaki, 1972; Pace, 1986).

Equilibrium unfolding and refolding by urea

A Milton Roy SLM-Aminco 8100 spectrofluorometer was used to measure the fluorescence emission of WW samples in the 310 to 400 nm range upon excitation at 295 nm. All measurements were made in a 1.0 cm pathlength Suprasil quartz cell (Hellma, Forest Hills, New York) with both the emission and excitation slit widths set to 1 mm, and with the temperature thermostated to 20°C by a circulating waterbath.

The intrinsic fluorescence of the WW domain in the presence of varying concentrations of urea was used to monitor the extent of protein unfolding or refolding. For the unfolding experiment, a set of independent $1 \mu\text{M}$ WW samples were prepared by dilution of a $20 \mu\text{M}$ stock WW solution that had been dialysed against 10 mM potassium phosphate (pH 6.0), 100 mM NaCl, 0.1 mM EDTA, 0.1 mM DTT, and 0.02% NaN_3 (Buffer A). Aliquots of freshly prepared 10 M urea were added to the stock protein to cover a denaturant concentration range of 0 to 8 M. The final volume of each sample assayed was 1.0 mL. Matrix effects were minimized by using the WW dialysis buffer for all dilutions. To generate the samples for the reconstitution experiment, a $34.3 \mu\text{M}$ WW domain solution initially prepared in 8.0 M urea was appropriately diluted with buffer producing samples of $1 \mu\text{M}$ WW in a series of differing urea concentrations. Reference blanks containing the identical concentration of urea as was present in each WW sample were prepared and analyzed concurrently at each interval investigated during the course of both the unfolding and refolding experiments.

Each data point presented in Figure 2 is the average taken from nine independent samples.

The combination of the linear extrapolation method with a two-state unfolding model was applied to describe the shape of the denaturation curve in Figures 2 and 3 according to Equation 1:

$$Y = \frac{Y_f + m_f[\text{urea}] + (Y_u + m_u[\text{urea}])\exp(-m(D_{1/2} + [\text{urea}])/RT)}{1 + \exp(-m(D_{1/2} + [\text{urea}])/RT)} \quad (1)$$

where Y is the measured fluorescence, m is the dependence of free energy of unfolding on the denaturant concentration, and $D_{1/2}$ is the denaturant concentration at the midpoint of unfolding (Pace, 1986; Santoro & Bolen, 1988; Myers et al., 1997). The slope and intercept of the pretransition baseline is given by m_f and Y_f , respectively, while m_u and Y_u describe the slope and intercept of the post-transition baseline.

ANS binding to probe hydrophobic surface exposure in WW

A stock solution of 8-anilino-1-naphthalene sulfonic acid (1,8 ANS) prepared in deionized water was added to samples of the WW domain ($5 \mu\text{M}$) in buffer A (pH 6.0), either in 6.0 M urea or 5.0 M GdnHCl. The final 1,8 ANS concentration present in all samples studied was $50 \mu\text{M}$ as determined by UV absorbance measurements at 372 nm using a molar extinction coefficient of $7,800 \text{ M}^{-1} \text{ cm}^{-1}$. The fluorescence emission spectra were recorded on an Aviv Model ATF105 automated titrating differential/ratio spectrofluorometer using an excitation wavelength of 375 nm, with both excitation and emission slit widths set to 2 mm. The emission spectra presented in Supplementary Figure 1 have been corrected by subtracting the fluorescence of unbound ANS measured in each of the solutions listed above in the absence of the WW domain (25°C).

Equilibrium unfolding and refolding by guanidine hydrochloride (GdnHCl)

Data for the GdnHCl-induced unfolding profile of the WW domain were recorded on an Aviv Model ATF105 Automated Titrating Differential/Ratio Spectrofluorometer (Lakewood, New Jersey) at 25°C using an excitation wavelength of 295 nm with detection at 340 nm. Both the excitation and emission slitwidths were set at 2 mm.

Samples analyzed during the course of GdnHCl-induced unfolding and refolding of WW (Fig. 3) were prepared by methods analogous to those used for the urea experiments. Aliquots of an 8 M GdnHCl solution were added to stock protein to generate a series of $1 \mu\text{M}$ WW solutions containing different denaturant concentrations. For the reconstitution portion of the GdnHCl experiment, a $100 \mu\text{M}$ stock WW domain solution in 6 M GdnHCl was diluted appropriately, producing $1 \mu\text{M}$ WW samples in GdnHCl concentrations that varied from 5 to 0.08 M. The data presented in Figure 3, an average of three independent samples, were corrected by subtracting a buffer blank containing an identical amount of GdnHCl.

Circular dichroism spectroscopy

All CD studies were performed on an AVIV Model 202SF Stopped Flow Circular Dichroism Spectrometer (Lakewood, New Jersey)

equipped with a Peltier temperature-controlled cell holder. The CD spectra represented in Figures 5, 6, and 8B are the averages of three consecutive scans, which were corrected by subtracting a buffer blank. Temperature-induced unfolding and refolding of 50 μ M WW domain samples in buffer A (pH 6.0) were monitored at 202 and 230 nm in a 0.1 cm pathlength Suprasil quartz cell (Hellma, Forest Hills, New York). The temperature was cycled from 4 to 88 °C and then back to 4 °C in increments of 4 °C. Prior to taking a reading, a 5 min thermal equilibration period was established at each newly set temperature. Experimental parameters similar to those used for thermally induced unfolding and refolding of the WW monitored in the far-UV CD spectral region were employed for the near-UV CD experiments, except that the protein concentration was increased to 1.0 mg/mL and the pathlength lengthened to 1.0 cm. The extent of reversibility of thermal denaturation varied between 80–100% for all WW samples analyzed. The shape of the thermal denaturation profile in Figure 7 was analyzed by a nonlinear least-squares fit assuming two state behavior according to Equation 2:

$$Y = \frac{(y_n + m_n T) + (y_d + m_d T) \exp(\Delta H_m / R(1/T_m - 1/T))}{1 + \exp(\Delta H_m / R(1/T_m - 1/T))} \quad (2)$$

where Y is the measured ellipticity, ΔH_m is the enthalpy at the unfolding transition, T_m is the melting temperature, T is temperature in Kelvin, and R is the universal gas constant (Santoro & Bolen, 1988; Swint & Robertson, 1993). The pretransition baseline slope and intercept is described by m_n and y_n , respectively, while m_d and y_d are the slope and intercept of the post-transition baseline.

Urea-induced unfolding of the WW domain was monitored by near-UV CD spectroscopy using a protein concentration of 1.0 mg/mL and a pathlength of 1.0 cm. Independent samples of protein made up with varying concentrations of urea were scanned three consecutive times from 320 to 260 nm and corrected by subtracting a buffer blank containing the identical concentration of denaturant as was present in each WW sample. The average of the scans collected in 0, 3, and 6 M urea is displayed in Figure 5, as is a scan acquired in 5.0 M GdnHCl. The ellipticity measurements at the fixed wavelength of 290 nm from these scans are presented in Figure 4, and are displayed with the fluorescence denaturation data from Figure 2 for comparison purposes.

¹H NMR spectroscopy

All spectra were recorded on a Varian Unity 600 spectrometer operating at a 599.1 MHz proton frequency with a sweepwidth of 8,000 Hz. Prior to data acquisition at each newly set temperature, a period of 20 min elapsed to establish thermal equilibration. The average of 64 consecutive scans was recorded at each temperature using the Watergate water suppression pulse sequence, with the chemical shifts standardized relative to the external standard 3-(trimethylsilyl)propionate-2,2,3,3-d₄ (Aldrich, Milwaukee, Wisconsin).

Sedimentation equilibrium to establish solution stoichiometry

To determine the quaternary structure, or lack thereof, of the WW domain and to evaluate the extent of protein aggregation, if any, a series of WW samples at various concentrations were subjected to

analytical ultracentrifugation experimentation. Equilibrium data were collected on a temperature-controlled Beckman XL-I analytical ultracentrifuge (Beckman Instruments, Inc., Palo Alto, California) equipped with a An-60Ti rotor and a photoelectric scanner. Samples of the WW domain that had been dialysed against buffer A (pH 6.0) were loaded into a double sector cell equipped with a 12 mm Epon centerpiece and sapphire optical windows. The sample compartment of the cell was charged with approximately 140 μ L of WW solution, while dialysis buffer was loaded into the reference compartment. All data were monitored at 280 nm (20 °C) using rotor speeds of 3,000 to 40,000 rpm with a step size of 0.001 cm. Samples came to equilibrium after 18–24 h of centrifugation, determined by overlaying duplicate scans taken at 3 h intervals. Twenty scans were taken at each 3 h time point during the course of the centrifugation run. The WW samples studied in chaotropic solutions were analyzed similarly.

Data analysis was carried out by a nonlinear least-squares approach using Origin software (Microcal Software, Inc., Northampton, Massachusetts). The data were fit to a single ideal species model described by Equation 3:

$$A_r = \exp[\ln(A_0) + (M\omega^2(1 - \bar{v}\rho)/2RT) \cdot (x^2 - x_0^2)] + E, \quad (3)$$

where A_r is the sample absorbance at radius x , A_0 is the absorbance at a reference radius x_0 (usually the meniscus), $\bar{v}$ is the partial specific volume of the protein (mL/g), ρ is the density of the solvent (g/mL), ω is the angular velocity of the rotor (radian/s), 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 volume of the WW domain (0.736 mL/g) was calculated based on its amino acid sequence according to the method of Cohn and Edsall as implemented in the XL-I software (Laue et al., 1992), while the value for ρ was 1.0 g/mL. 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 reason ability.

Determination of K_d by fluorescence

The interaction of the PY peptide (ligand) with the WW domain (acceptor) was characterized and quantified by monitoring the changes in the WW fluorescence at 340 nm upon peptide binding. Assuming a 1 to 1 stoichiometry between the ligand and acceptor, a plot of fractional acceptor saturation vs. free ligand concentration produces an ideal hyperbolic binding curve described by Equation 4, from which the K_d may be calculated:

$$F = F_f + (F_b - F_f)[\text{Ligand}_f]/(K_d + [\text{Ligand}_f]), \quad (4)$$

where F is the total fluorescence, F_f is the fluorescence of the unbound acceptor, F_b is the fluorescence of the ligand-acceptor complex, $[\text{Ligand}_f]$ is the free ligand concentration, and K_d is the dissociation constant of the complex (Viguera et al., 1994; Winzor & Sawyer, 1995). The dissociation constant of the WW-PY complex was calculated by a nonlinear least-squares fit of the binding data to Equation 4.

The peptide binding curve data were recorded on an Aviv Model ATF105 Automated Titrating Differential/Ratio Spectrofluorometer at 20 °C using an excitation wavelength of 295 nm with de-

tection at 340 nm. Both the excitation and emission slitwidths were set at 2 nm. A 1.0×1.0 cm Suprasil quartz cell (Hellma, Forest Hills, New York), charged with 2 mL of 5 μ M WW domain in buffer A (pH 6.0), was titrated with a solution of PY peptide using a Microlab 540B Dual Syringe Dispenser (Hamilton Company, Reno, Nevada). Throughout the entire titration the sample was constantly stirred, and following each aliquot of PY peptide, a 1 min mixing period elapsed before the next reading was taken. A simultaneous titration of a cell containing only buffer A with PY peptide, placed in the reference channel of the instrument, was used to correct the emission data presented in Figure 11. The PY peptide alone did not contribute to the fluorescence signal when excited 295 nm.

Supplementary material in Electronic Appendix

Figure 1. Addition of 1,8 ANS to the WW domain in buffer A alone (pH 6.0) (○), or buffer A (pH 6.0) with either 5.0 M GdnHCl (▲) or 6.0 M urea (+) added. The WW and 1,8 ANS concentrations in each sample were 5 and 50 μ M, respectively. All scans were made employing an excitation wavelength of 375 nm, and corrected by subtracting the emission spectrum of 1,8 ANS in each of the above solutions in the absence of WW (25 °C). Addition of 50 μ M 1,8 ANS to 5 μ M human transthyretin (●), a well-characterized protein that binds two equivalents of ANS, illustrating the quantum yield increase and blue-shift of the emission maximum expected upon ANS binding. Electronic file name: Koepfsf1 (KaleidaGraph).

Figure 2. Sedimentation equilibrium analysis of 100 μ M WW domain sample in 6.0 M urea. The solid line drawn through the data was obtained by fitting the absorbance vs. radial position data to Equation 3 for a single ideal species model. The residuals between the fit and the experimental data are shown in the top panel. Electronic file name: Koepfsf2 (KaleidaGraph).

Figure 3. Sedimentation equilibrium analysis of 100 μ M WW sample in 5.0 M urea. The solid line drawn through the data was obtained by fitting the absorbance vs. radial position data to Equation 3 for a single ideal species model. The residuals between the fit and the experimental data are shown in the top panel. Electronic file name: Koepfsf3 (KaleidaGraph).

Figure 4. Sedimentation velocity analysis of a 0.18 mM WW sample in buffer A (pH 6.0) carried out at 50,000 rpm (20 °C). Initial absorbance scans were made using a rotor speed of 5,000 rpm to detect any fast sedimenting species, then the rotor speed was increased to 50,000 rpm and data collected at 10 min intervals. The WW domain sedimented as a single broad boundary, which was fit using direct svedberg analysis, yielding a sedimentation coefficient of 0.85 s. The inset represents the residual difference between the experimental and fitted data. Electronic file name: Koepfsf4 (Adobe Photoshop for Mac).

Acknowledgments

We thank the NIH R01GM51105 for primary financial support, the Skaggs Institute of Chemical Biology and the Lita Annenberg Hazen Foundation for Infrastructure Support, Dr. Hartmut Oschkinat and Dr. Maria Macias for providing helpful advice throughout the course of this work and for the NMR assignments, Dr. Kwaku Dayie of the Williamson group at Scripps for helping out with the NMR experiments, Hilal Lashuel for centrifugation assistance and Dr. Nick Pace for the valuable suggestions on data analysis. We also thank Dr. L.D. Burtnick for comments on the manuscript.

References

- Baldwin RL. 1995. α -Helix formation by peptides of defined sequence. *Biophys Chem* 55:127–135.
- Blankenmeyer-Menge B, Nimtz M, Frank R. 1990. An efficient method for anchoring Fmoc-amino acids to hydroxyl-functionalised solid supports. *Tetrahedron* 31:1701–1704.
- Bork P, Sudol M. 1994. The WW domain: A signaling site in dystrophin? *TIBS* 19:531–533.
- Camarero J, Pavel J, Muir T. 1998. Chemical synthesis of a circular protein domain: Evidence for folding-assisted cyclization. *Angew Chem Int Ed* 37:347–349.
- Chen HI, Einbond A, Kwak SJ, Linn H, Koepf E, Peterson S, Kelly JW, Sudol M. 1997. Characterization of the WW domain of human Yes-associated protein and its polyproline-containing ligands. *J Biol Chem* 272:17070–17077.
- Chen HI, Sudol M. 1995. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci USA* 92:7819–7823.
- Choo DW, Schneider JP, Graciani NR, Kelly JW. 1996. Nucleated antiparallel β -sheet that folds and undergoes self-assembly: A template promoted folding strategy towards controlled molecular architectures. *Macromolecules* 29:355–366.
- Colon W, Kelly JW. 1992. Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry* 31:8654–8660.
- Constantine K, Mueller L, Andersen N, Tong H, Wandler CF, Friedrichs MS, Brucoleri RE. 1995. Structural and dynamic properties of a β -hairpin-forming linear peptide. 1. Modeling using ensemble-averaged constraints. *J Am Chem Soc* 117:10841–10854.
- DeGrado WF, Lear DJ. 1985. Induction of peptide conformation at apolar/water interfaces. 1. A study with model peptides of defined hydrophobic periodicity. *J Am Chem Soc* 107:7684–7689.
- Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, Thomas PD, Chan HS. 1995. Principles of protein folding—A perspective from simple exact models. *Protein Sci* 4:561–602.
- Dill KA, Shortle D. 1991. Denatured states of proteins. *Annu Rev Biochem* 60:795.
- Dyson HJ, Wright PE. 1991. Defining solution conformations of small linear peptides. *Annu Rev Biophys Biophys Chem* 20:519–538.
- Fasman GD, ed. 1996. *Circular dichroism and the conformational analysis of biomolecules*. New York: Plenum Press.
- Kelly JW. 1996. Alternative conformations of amyloidogenic proteins govern their behavior. *Curr Opin Struct Biol* 6:11–17.
- Kelly JW. 1998. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol* 8:101–106.
- Knapp S, Mattson P, Christova P, Berndt K, Karshifoff A, Vihinen M, Smith E, Ladenstein R. 1998. Thermal unfolding of small proteins with SH3 domain folding pattern. *Proteins Struct Funct Genet* 31:309–319.
- Koide S, Dyson HJ, Wright PE. 1993. Characterization of a folding intermediate of apoplastocyanin trapped by proline isomerization. *Biochemistry* 32:12299–12310.
- Kortemme T, Ramirez-Alvarado M, Serrano L. 1998. Design of a 20-amino acid three-stranded β -sheet protein. *Science* 281:253–256.
- Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946–950.
- Krittana C, Johnson WC. 1997. Correcting the circular dichroism of peptides for contributions of absorbing side chains. *Anal Biochem* 253:57–64.
- Lai Z, Colon W, Kelly JW. 1996. The acid-mediated denaturation pathway of transthyretin yields a conformational intermediate that can self-assemble into amyloid. *Biochemistry* 35:6470–6482.
- Laue TM, Shad BD, Ridgeway TM, Pelletier SL. 1992. Computer-aided interpretation of analytical sedimentation data for protein. In: Harding SE, Rowe AJ, Horton JC, eds. *Analytical ultracentrifugation in biochemistry and polymer science*. Cambridge, England: The Royal Society of Chemistry, pp 90–125.
- Linn H, Ermekova KS, Rentschler S, Sparks AB, Kay BK, Sudol M. 1997. Using molecular repertoires to identify high-affinity peptide ligands of the WW domain of human and mouse YAP. *Biol Chem* 378:531–537.
- Logan TM, Theriault Y, Fesik SW. 1994. Structural characterization of the FK506 binding protein unfolded in urea and GdnHCl. *J Mol Biol* 236:637–648.
- Macias MJ, Hyvonen M, Baraldi E, Schultz J, Sudol M, Saraste M, Oschkinat H. 1996. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* 382:646–649.
- Manning MCW, Woody R. 1987. Theoretical determination of the CD of proteins containing closely packed antiparallel β -sheets. *Biopolymers* 26:1731–1752.
- Minor DL Jr, Kim PS. 1994a. Context is a major determinant of β -sheet propensity. *Nature* 371:264–267.

- Minor DL Jr, Kim PS. 1994b. Measurement of the β -sheet-forming propensities of amino acids. *Nature* 367:660–663.
- Monera OD, Kay CM, Hodges RS. 1994. Protein denaturation with guanidine hydrochloride or urea provides a different estimate of the stability depending on the contributions of electrostatic interactions. *Protein Sci* 3:1984–1991.
- Munoz V, Serrano L. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct Biol* 1:399–409.
- Myers JK, Pace CN, Scholtz JM. 1997. Helix propensities are identical in proteins and peptides. *Biochemistry* 36:10923–10929.
- Neri D, Billeter M, Wider G, Wüthrich K. 1992. NMR Determination of residual structure in a urea-denatured protein, the 434 repressor. *Science* 257:1559–1563.
- Nesloney C, Kelly J. 1996. Progress towards understanding β -sheet structure. *Bioorg Med Chem* 4:739–766.
- Nozaki Y. 1972. The preparation of guanidine hydrochloride. *Methods Enzymol* 26:43–50.
- Osterman DG, Kaiser ET. 1985. Design and characterization of peptides with amphiphilic β -strand structures. *J Cell Biochem* 29:57–72.
- Pace CN. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131:266–280.
- Pace CN, Laurents DV, Thomson JA. 1990. pH Dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. *Biochemistry* 29:2564–2572.
- Ranganathan R, Lu KP, Hunter T, Noel JP. 1997. Structural and functional analysis of the mitotic rotamase Pin1 suggests substrate recognition is phosphorylation dependent. *Cell* 89:875–886.
- Reid K, Rodriguez H, Hillier B, Gregoret L. 1998. Stability and folding properties of a model β -sheet protein, *Escherichia coli* CspA. *Protein Sci* 7:470–479.
- Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α -chymotrypsin using different denaturants. *Biochemistry* 27:8063–8068.
- Schagger H, von Jagow G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 166:368–379.
- Schenck HL, Gellman SH. 1998. Use of a designed triple-stranded antiparallel β -sheet to probe cooperativity in aqueous solution. *J Am Chem Soc* 120:4869–4870.
- Schneider JP, Kelly JW. 1995. Templates that induce α -helical, β -sheet, and loop conformations. *Chem Rev* 95:2169–2187.
- Sharman GJ, Searle MS. 1998a. Cooperative interaction between the three strands of a designed antiparallel β -sheet. *J Am Chem Soc* 120:5291–5300.
- Sharman GJ, Searle MS. 1998b. Origin of β -hairpin stability in solution: Structural and thermodynamic analysis of the folding of a model peptide supports hydrophobic stabilization in water. *J Am Chem Soc* 120:1996–2007.
- Shortle D. 1993. Denatured states of proteins and their roles in folding and stability. *Curr Opin Struct Biol* 3:66.
- Shortle D. 1996. The denatured state (the other half of the folding equation) and its role in protein stability. *FASEB J* 10:27–34.
- Smith CK, Regan L. 1997. Construction and design of β -sheets. *Acc Chem Res* 30:153–161.
- Smith CK, Withka JM, Regan L. 1994. A thermodynamic scale for the β -sheet forming tendencies of the amino acids. *Biochemistry* 33:5510–5517.
- Sudol M. 1994. Yes-associated protein (YAP-65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product. *Oncogene* 9:2145–2152.
- Sudol M. 1996. Structure and function of the WW domain. *Prog Biophys Mol Biol* 65:113–132.
- Swint L, Robertson AD. 1993. Thermodynamics of unfolding for turkey ovomucoid third domain: Thermal and chemical denaturation. *Protein Sci* 2:2037–2049.
- Timasheff SN. 1992. Water as a ligand: Preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry* 31:9857–9864.
- Tsang KY, Díaz H, Graciani N, Kelly JW. 1994. Hydrophobic cluster formation is necessary for dibenzofuran-based amino acid to function as β -sheet nucleators. *J Am Chem Soc* 116:3988–4005.
- Viguera AR, Arrondo JLR, Musacchio A, Saraste M, Serrano L. 1994. Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. *Biochemistry* 33:10925–10933.
- Wang Y, Shortle D. 1995. The equilibrium folding pathway of staphylococcal nuclease: Identification of the most stable chain-chain interactions by NMR and CD spectroscopy. *Biochemistry* 34:15895–15905.
- Winzor D, Sawyer W. 1995. *Quantitative characterization of ligand binding*. New York: Wiley-Liss, Inc.
- Woody RW. 1985. Circular dichroism of peptides. In: Hruby V, eds. *The peptides, analysis, synthesis and biology*. New York: Academic Press. pp 15–104.
- Woody RW. 1987. The circular dichroism of biopolymers. *FECS Int Conf Circ Dichroism(Proc)*:270–295.
- Woody RW. 1993. The circular dichroism of oriented β sheets: Theoretical predictions. *Tetrahedron: Asymmetry* 4:529–544.
- Woody RW. 1994. Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. *Eur Biophys J* 23:253–262.
- Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: John Wiley & Sons.