Phosphorylation of RNA Polymerase II CTD Fragments Results in Tight Binding to the WW Domain from the Yeast Prolyl Isomerase Ess1[†]

Jeffrey K. Myers, Daniel P. Morris, Arno L. Greenleaf, and Terrence G. Oas*

Department of Biochemistry, Box 3711, Duke University Medical Center, Durham, North Carolina 27710

Received December 7, 2000; Revised Manuscript Received April 13, 2001

ABSTRACT: The yeast prolyl isomerase, Ess1, has recently been shown to interact via its WW domain with the hyperphosphorylated form of the RNA polymerase II C-terminal domain (CTD). We have investigated folding of the Ess1 WW domain and its binding to peptides representing the CTD by circular dichroism and fluorescence. Ess1 WW folds and unfolds reversibly, but in the absence of ligand is only marginally stable with a melting temperature of 19 °C. The WW domain is stabilized by the addition of anionic ligands, namely, chloride, inorganic phosphate, phosphoserine, and phosphorylated CTD peptides. Dissociation constants were measured to be 70−100 μM for CTD peptides phosphorylated at one serine, and 16−21 μM for peptides with two or more phosphorylated serines. Weaker or no affinity was observed for nonphosphorylated CTD peptides. There is surprisingly little difference in the affinity for peptides phosphorylated at Ser 2 or Ser 5 of the consensus repeat, or for peptides with different patterns of multiple phosphorylation. The binding of Ess1 to phosphorylated CTD peptides is consistent with a model wherein the WW domain positions Ess1 to catalyze isomerization of the many pSer−Pro peptide bonds in the phosphorylated CTD. We suggest that cis/trans isomerization of prolyl peptide bonds plays a crucial role in CTD function during eukaryotic transcription.

The carboxy-terminal domain of the largest subunit of RNA polymerase II (CTD)[1] plays a dominant role in regulating mRNA production in eukaryotes (*1*−*6*). The CTD is an essential domain that consists of multiple repeats of a highly conserved seven amino acid consensus sequence: YSPTSPS. The yeast CTD consists of 26 repeats, while higher organisms have up to 52 repeats. The CTD is subject to extensive phosphorylation during transcription, particularly on serines at positions 2 and 5 of the repeat (*7*). CTD phosphorylation correlates with the transition between transcription initiation and elongation, and thus may be involved in transcriptional regulation. During elongation, the hyperphosphorylated CTD (P-CTD) facilitates and controls pre-mRNA processing (*3, 4, 6, 8*).

An essential yeast peptidyl prolyl isomerase, Ess1, has been shown to bind the phosphorylated CTD (P-CTD) via its WW domain (*9*). Binding to the P-CTD would position Ess1 appropriately to influence P-CTD-mediated events, such as pre-mRNA processing. Localization of Ess1 to the P-CTD suggests that the many pSer−Pro peptide bonds are substrates for Ess1, and the protein catalyzes isomerization of these bonds. It is known from studies on small peptides that phosphorylation of serine increases the ratio of cis/trans isomers in Ser−Pro peptide bonds and slows down the rate of isomerization (*10*). Ess1 and its human homologue, Pin1, are specific for phosphorylated Ser−Pro or Thr−Pro bonds in model peptides (*11, 12*). Prolyl bond isomerization could obviously affect the binding of other proteins to the P-CTD. Ess1 may also catalyze the isomerization of prolyl bonds within other proteins bound to the P-CTD, for example, phosphatases (*13*−*16*) or pre-mRNA processing factors (*17*−*21*). Genetic evidence strongly implicates Ess1 in pre-mRNA 3′-end formation and other transcription-related events (*12, 22*). It is becoming generally accepted that most RNA processing occurs co-transcriptionally, with the P-CTD serving to coordinate both processes (*3, 4, 6*).

Two crystal structures of Pin1, a human isomerase highly homologous to Ess1, have been published by Noel and co-workers (*23, 24*) (Figure 1A). The structure features 2 domains: a 120-residue isomerase domain containing the active site, and a small, 40-residue WW domain. The WW and the isomerase domains are connected by a flexible linker, which is disordered in the crystal structure (*23, 24*). WW domains are small, β-sheet proteins that bind proline-rich targets and mediate protein−protein interactions. Named for two conserved tryptophan resides, WW domains are found in many proteins of diverse function (*25, 26*). Previous studies have investigated the binding of WW domains to various proline-rich ligands (*11, 27*−*32*), and the WW domain from human YAP protein has been studied as a model for β-sheet formation (*33*−*35*). We have investigated the folding and unfolding of the isolated yeast Ess1 WW domain by using thermal denaturation monitored by circular dichroism. We have also studied the binding of CTD-related ligands by circular dichroism and fluorescence. We find that multiple phosphorylation sites on CTD peptides are required

[†] Supported by National Institutes of Health Grants GM40505 (A.L.G.) and GM45322 (T.G.O.).

* Corresponding author. Phone: 919-684-4363. Fax: 919-681-8862. E-mail: oas@biochem.duke.edu.

[1] Abbreviations: CD, circular dichroism; CTD, C-terminal domain of RNA polymerase II; P-CTD, hyperphosphorylated form of CTD; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; PAGE, polyacrylamide gel electrophoresis; T_m, midpoint of thermal denaturation curve.

10.1021/bi0027884 CCC: $20.00
Published on Web 06/21/2001

for tight binding, and observe a surprisingly low degree of discrimination between different phosphorylation patterns.

MATERIALS AND METHODS

The sequence encoding the Ess1 WW domain (residues 1−49) was cloned into the pTCLEH expression vector (*36*) by PCR. This vector has a leader TrpLE peptide that causes the overexpressed product to accumulate in inclusion bodies. A single methionine residue between the WW domain and TrpLE allows cleavage of the protein from the leader by cyanogen bromide. The TrpLE leader has a His_6 tag, which allows separation of the leader and the WW domain on a nickel column. The pTCLEH/WW plasmid was transformed into BL21(DE3) *E. coli*, and cultures were grown in LB media to an OD_{600} of 0.8 at 37 °C with shaking, and then induced with IPTG and grown a further 4 h. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris, pH 7.8, 1 mM EDTA, 25% sucrose) so that the volume of buffer added was 20% (w/v) of the cells. Lysozyme was added to 1 mg/mL, and the suspension was mixed and placed on ice for 30 min. To this suspension were added $MgCl_2$ (to 10 mM), $MnCl_2$ (to 1 mM), and DNase I (to 50 μg/mL); again, the suspension was mixed and placed on ice 30 min. The addition of 2 volumes of ice-cold detergent buffer (200 mM NaCl, 1% deoxycholate, 1% Nonidet P40, 20 mM Tris, pH 7.8, 2 mM EDTA) followed. The solution was mixed and centrifuged at 5000*g* for 15 min at 4 °C to collect the inclusion bodies. The pelleted inclusion bodies were then washed by resuspension in the following solutions followed by centrifugation at 5000*g* for 10 min: once with 0.5% Triton X-100 + 1 mM EDTA, once with 1 mg/mL deoxycholate + 1 mM EDTA + 1 mg/mL lysozyme, twice with 1 mg/mL deoxycholate + 1 mM EDTA, and 4 times with water.

Washed inclusion bodies were dissolved in 10 mL of 70% formic acid. The leader sequence was cleaved from the WW domain by adding 500 mg of CNBr (Sigma) (dissolved in 70% formic acid). The solution was incubated at room temperature for 30 min. The formic acid and CNBr were then drawn off with a rotary evaporator located in a fume hood and fitted with a liquid nitrogen trap.

Dried product from the cleavage reaction was dissolved in 6 M GuHCl/20 mM Tris, pH 8.0, adjusted to pH 8.0 with NaOH, and applied to a 15 mL Ni-NTA-agarose (Qiagen) column previously equilibrated with the same buffer. The flow-through, which contained the WW domain, was collected. The leader was washed off the column with 6 M GuHCl/20 mM Tris/200 mM imidazole, pH 8, which also prepared the column for subsequent reuse. The WW-containing flow-through (ca. 30 mL) was loaded directly onto a 400 mL Sephadex G10 (Pharmacia) column and eluted with 5% acetic acid. This serves primarily as a desalting step to remove GuHCl. Lyophilization of the protein followed, and the resulting solid was dissolved in 3 mL of 5% acetic acid and loaded onto a 300 mL Sephadex G50 (Pharmacia) column that had been equilibrated with 5% acetic acid. Fractions containing the WW domain were combined and lyophilized. Purity was checked using analytical reverse-phase HPLC and SDS−PAGE. The molecular mass was determined to be 5170 Da by electrospray mass spectrometry, in agreement with the expected molecular mass.

CTD peptides were ordered from SynPep, except for CTD_{21} and $CTD_{21}^{5,5,5p}$ which were synthesized by the Peptide Synthesis Facility, New York State Department of Health, and were a generous gift from Steven Hanes (SUNY, Albany). Purity was checked by analytical reverse-phase HPLC, and the peptides were either purified by reverse-phase HPLC on a Vydac C_{18} semi-prep column using a water/acetonitrile solvent system or used without further purification. All peptides except the three-repeat peptides had acetylated N-termini and amidated C-termini. Peptide molecular weights were confirmed by electrospray mass spectrometry.

Circular Dichroism. WW domain stock solutions were made in water (500 μM to 1 mM). Protein concentrations were measured in 6 M GuHCl with an extinction coefficient of 13 940 M^{-1} cm^{-1} at 280 nm (*37*). Peptide stock solutions were made in water, and the concentrations were measured in 6 M GuHCl at 276 nm using tyrosine absorbance.

CD spectra were collected on an Aviv model 202 circular dichroism spectropolarimeter, equipped with a Peltier temperature controller, with 1.0 nm bandwidth and an averaging time of 20 s at each wavelength. One millimeter path length quartz cells (Hellma) were used, with 50 μM protein, and a buffer of 20 mM potassium phosphate, pH 7.0, 100 mM KCl. Near-UV CD spectra were collected at a protein concentration of 1 mg/mL in 1 cm quartz cells, with a buffer of 20 mM cacodylate, pH 7.0, 100 mM NaCl. Buffer blanks were subtracted from each spectrum. Thermal melts were performed in a 1 cm path length quartz cell with a magnetic stir bar; protein concentration was 5 μM. A buffer of 20 mM cacodylate, pH 7.0, was used. Centrifuged and degassed samples were equilibrated at low temperature for 15 min before each run, then for 1 min at each subsequent temperature followed by 30 s of data collection. Thermal melts were fit to the equation:

$$Y = \frac{(a + bT) + (c + dT)\exp[\Delta H(1/T - 1/T_m)/R]}{1 + \exp[\Delta H(1/T - 1/T_m)/R]} \quad (1)$$

where Y is the observed ellipticity, a and b are the intercept and slope of the pre-transition baseline, c and d are the intercept and slope of the post-transition baseline, T_m is the melting temperature (transition midpoint), T is the absolute temperature, and R is the gas constant (*38*). For thermal melts of the protein in the presence of CTD peptides, a melt of the CTD peptide alone was subtracted from the data in order to control for the contribution of the CD signal of the peptides to the curve.

Fluorescence. Fluorescence measurements were made on an SLM-Aminco Bowman Series 2 Luminescence Spectrometer. Square 1 cm path length quartz cuvettes were used with a protein concentration of 5 μM. Emission scans were recorded using an excitation wavelength of 278 nm, in a buffer of 20 mM potassium phosphate, pH 7.0, 100 mM NaCl. For titrations, a buffer of 20 mM cacodylate, pH 7.0, 100 mM NaCl was used. Samples were thermostated at 25 °C with an attached water bath. Aliquots of a stock solution of titrant were added to the stirred solution, equilibrated for 5 min, and the fluorescence emission was monitored at 340 nm for 1 min. The average signal was then corrected for dilution. Titrations were fit to the equation:

$$F = F_d + \{((1 + [L]/K_d)/K_u)(F_n - F_d)\}/\{1 + ((1 + [L]/K_d)/K_u)\} \quad (2)$$

where F is the observed fluorescence, F_d is the fluorescence of the denatured protein, F_n is the final fluorescence, [L] is the free ligand concentration, K_u is the equilibrium unfolding constant, and K_d is the dissociation constant of the ligand (*39*). The equilibrium unfolding constant, K_u, was fixed in the fits to a value determined from the CD thermal denaturation curve in the absence of ligand. The free ligand concentration [L] is found from

$$[NL] = \{[P]_{tot} + [L]_{tot} + (K_d + K_u K_d) - \{([P]_{tot} + [L_{tot}] + (K_d + K_u K_d))^2 - 4[P]_{tot}[L]_{tot}\}^{1/2}\}/2$$

where

$$[L] = [L]_{tot} - [NL] \quad (3)$$

where $[P]_{tot}$ is the total protein concentration and $[L]_{tot}$ is the total ligand concentration.

RESULTS

Initially we focused on the folding and stability of the Ess1 WW domain monitored by circular dichroism (CD). The CD spectrum of native Ess1 WW features a minimum at ~200 nm and a maximum at 228 nm (Figure 1B), consistent with other WW domain proteins (*34*). WW domains are typically three-stranded β-sheets (Figure 1A); the lack of a characteristic β-sheet signal is probably due to the abundance of aromatic residues, which can make a strong contribution to the CD spectrum in the range of 215−235 nm (*40*, *41*). Since the maximum at 228 nm disappears when the protein is denatured at high temperatures (Figure 1B), it is likely the maximum arises from asymmetric environments of the tryptophan or tyrosine residues in the native state of the protein. Lowering the temperature after the high-temperature scan resulted in a spectrum that overlays the original low-temperature CD spectrum, indicating that the folding and unfolding are highly reversible. This reversibility is in contrast to many other β-sheet proteins, which are generally prone to aggregate when thermally denatured.

The temperature dependence of the CD signal at 228 nm is shown in Figure 2. This thermal denaturation curve is sigmoidal, indicating a cooperative unfolding transition. The fit to a two-state native (N) ⇆ denatured (D) unfolding model is shown; however, rigorous evidence for two-state unfolding has not yet been obtained (see below). The domain is partially native and partially denatured at room temperature, with a fitted thermal transition midpoint in cacodylate buffer, pH 7.0, of 19.1 °C. The enthalpy change upon denaturation, ΔH, is small (18 600 cal/mol), typical of small protein domains (*42*). This marginal stability under our experimental conditions may not be significant functionally, as the WW domain could be stabilized in vivo by interactions with the isomerase domain of Ess1, or by binding of anionic ligands (see below). Future investigations of the full-length Ess1 protein will reveal the extent of thermodynamic and functional interaction between the two domains.

Koepf et al. observed a residual near-UV CD signal when the YAP WW domain was unfolded in denaturants (*34*).

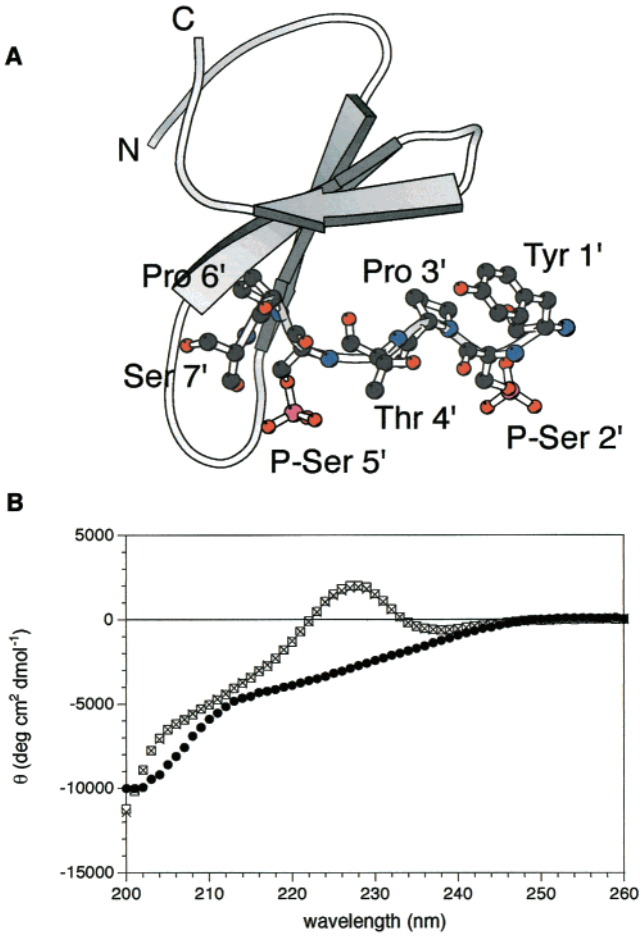

FIGURE 1: (A) Structure of the Pin1 WW domain bound to a P-CTD peptide (*23*); the WW domain is shown as a ribbon diagram, and the peptide ligand is shown in ball-and-stick representation. The figure was generated using MOLSCRIPT (*50*). (B) Circular dichroism (CD) spectra of Ess1 WW domain: folded at 0 °C (open squares), unfolded at 80 °C (closed circles), and rescanned at 0 °C after being unfolded and scanned at 80 °C (crosses). Greater than 95% reversibility is observed.

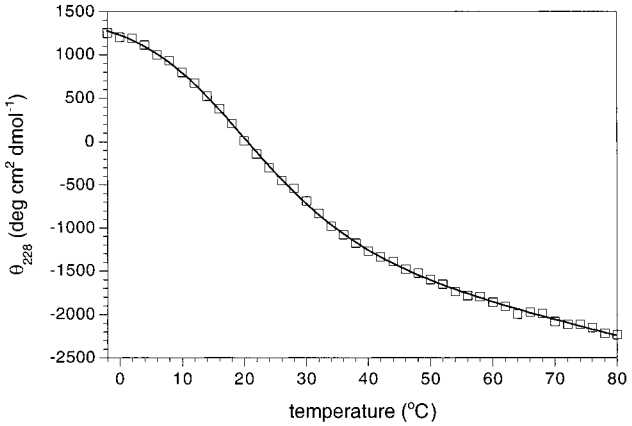

FIGURE 2: Thermal denaturation of Ess1 WW domain followed by CD at 228 nm. The solid line is fit to eq 2, giving a melting temperature of 19.1 °C and an enthalpy of unfolding of 18.6 kcal/mol.

They suggested the presence of an aromatic cluster in the denatured state of YAP WW, and that this cluster could be a general mechanism by which WW domains avoid aggregation. We used near-UV CD, which is sensitive to the environment of aromatic residues in proteins, to probe the structure of Ess1 WW domain. Figure 3 shows scans of 1

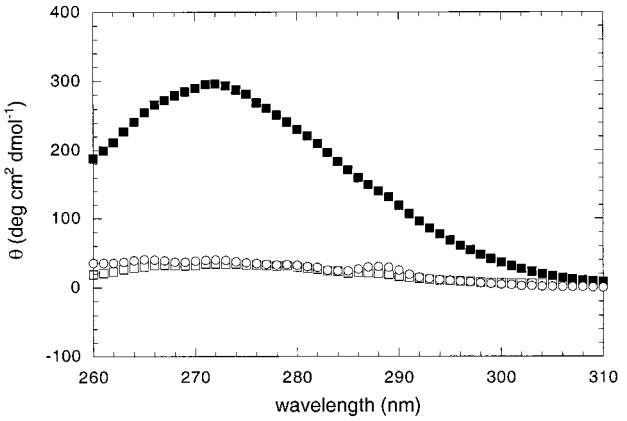

FIGURE 3: Near-UV CD scans of the WW domain: folded at 0 °C (closed squares), unfolded at 80 °C (open squares), and unfolded in 5 M guanidine hydrochloride (open circles).

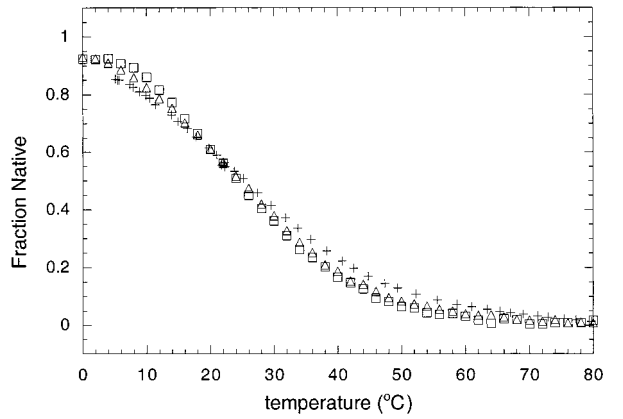

FIGURE 4: Fraction of native protein versus temperature followed by three different probes: far-UV CD at 228 nm (open squares), near-UV CD at 272 nm (open triangles), fluorescence emission at 340 nm (plus signs).

mg/mL WW domain in cacodylate buffer, folded at low temperature and unfolded in 5 M guanidine hydrochloride or by high temperature (80 °C). The native protein shows a pronounced maximum around 272 nm, which disappears upon unfolding with temperature or guanidine hydrochloride. No evidence of residual structure is observed, and therefore aromatic clustering is apparently not present in the denatured state of Ess1 WW domain.

Ess1 WW domain has a strong fluorescence signature resulting from the two tryptophan residues. Fluorescence emission scans show a maximum at 340 nm (excitation at 278 nm), while scans of the unfolded protein show a reduced fluorescence signal, with a shift of the emission maximum to 350 nm (data not shown). We followed thermal denaturation using fluorescence emission at 340 nm; the resulting curve was nearly superimposable on those derived from near- and far-UV CD (Figure 4). It is important to determine whether a two-state model can adequately describe the unfolding of the WW domain. Coincidence of multiple spectroscopic probes of denaturation is usually taken as evidence of a two-state transition (*43*). The curves in Figure 4 have similar midpoints; however, the fluorescence curve deviates somewhat from the other two at high and low temperature, possibly a sign of the presence of intermediate states. In any case, the tryptophan residues are likely dominating all three probes, and therefore even exact coincidence of the denaturation curves would be weak evidence for two-state folding. Further studies by NMR will be carried out to address this question.

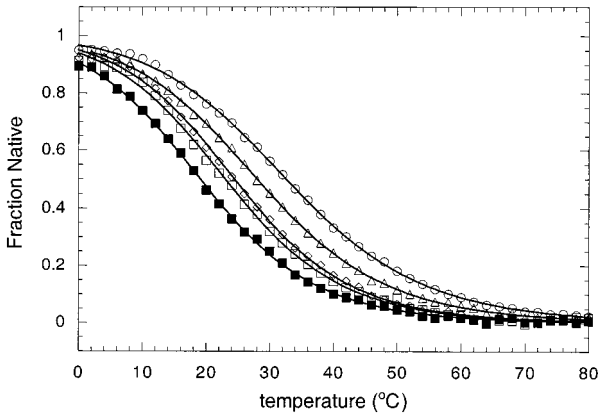

FIGURE 5: Stabilization of the native WW domain by small anions: thermal denaturation monitored by CD at 228 nm in 20 mM cacodylate buffer (closed squares), plus 1 mM phosphoserine (open squares), 100 mM NaCl (open diamonds), 20 mM sodium phosphate (open triangles), or 100 mM sodium phosphate (open circles). Solid lines are fits to eq 1 (Materials and Methods).

Table 1: Binding of CTD Peptides to Ess1 WW Domain[a]

peptide	sequence	ΔT_m (°C)[b]	K_d (μM)[c]
	1 2 3 4 5 6 7		
CTD_{21}	YSPTSPSYSPTSPSYSPTSPS	6.6±2	N.D.
CTD_9^{2p}	TSPSYSPTS	8.0±0.6	76±4
CTD_9^{5p}	SPTSPSYSP	5.9±1	98±18
CTD_{11}	SPTSPSYSPTS	0.7±0.7	N.D.
CTD_{11}^{2p}	SPTSPSYSPTS	9.8±0.6	67±11
CTD_{11}^{5p}	SPTSPSYSPTS	6.4±0.6	79±13
$CTD_{11}^{5,2p}$	SPTSPSYSPTS	25.8±0.5	21±3
$CTD_{14}^{2,5p}$	TSPSYSPTSPSYSP	25.2±0.5	16±2
$CTD_{21}^{5,5,5p}$	YSPTSPSYSPTSPSYSPTSPS	25.9±0.6	17±2

[a] Peptides labeled according to length (subscript) and phosphorylation state (superscript). Peptide sequences in single-letter amino acid code. The numbers above the CTD_{21} sequence indicate the traditional numbering of residues in the CTD repeat. Outlined S denotes phosphorylated serine. 21-residue peptides had free N- and C-termini; all other peptides had termini blocked with acetyl and amide groups, respectively. [b] Difference in fitted melting temperature from CD thermal denaturation curves [T_m(50 μM peptide) − T_m(0 μM peptide)]. [c] Fitted dissociation constant from titrations followed by fluorescence.

Ligands that bind the native state of a protein more tightly than the denatured state stabilize the native state relative to the denatured state (*44*). Thermal denaturation curves of the WW domain obtained in the presence of NaCl, phosphate, and phospho-Ser are shown in Figure 5. Binding of these anions resulted in higher melting temperatures, although the apparent enthalpy of the transition was still low, and the curves remain broad. WW domain binding to CTD peptide fragments, corresponding to sequences of varying length and phosphorylation state (see Table 1), was investigated by performing CD thermal denaturation of the WW domain in the presence of these peptides. The CD spectrum of each of the peptides alone was consistent with random coil structure (data not shown), although the peptides had some CD signal at 228 nm [see also (*45*)]. This signal as a function of temperature was subtracted from the respective thermal denaturation curves. Thermal denaturation curves of the WW domain, free and in the presence of 50 μM CTD peptides,

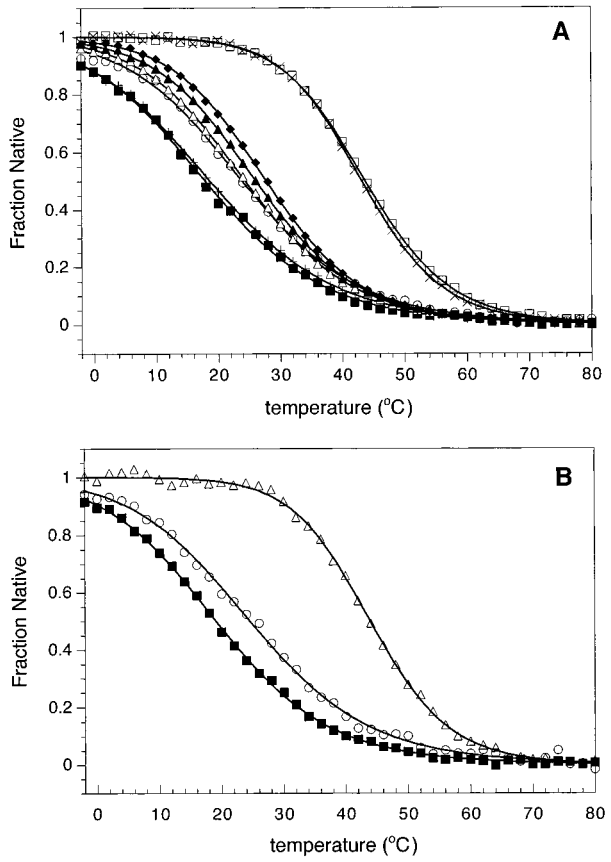

FIGURE 6: Stabilization of the native WW domain by (A) 9−14-residue CTD peptides: thermal denaturation monitored by CD at 228 nm in 20 mM cacodylate buffer (closed squares), plus 50 μM CTD_{11} (plus signs), CTD_9^{5p} (open circles), CTD_{11}^{5p} (open triangles), CTD_9^{2p} (closed triangles), CTD_{11}^{2p} (closed diamonds), $CTD_{11}^{5,2p}$ (open squares), or $CTD_{14}^{2,5p}$ (crosses). (B) 21-residue, 3-repeat CTD peptides: thermal denaturation of the WW domain in 20 mM cacodylate (closed squares), plus 50 μM CTD_{21}(open circles), or 50 μM $CTD_{21}^{5,5,5p}$ (open triangles). Solid lines are fits to eq 1.

are shown is Figure 6A,B. The native WW domain is significantly stabilized by micromolar concentrations of phosphorylated peptides (Table 1). The four singly phosphorylated peptides increase the melting temperature by 6−10 °C. There is only a slight preference for binding to a phospho-Ser at position 2 of the repeat over position 5. Multiple phosphorylations further increase the binding affinity; the three peptides with more than one phosphate increase the melting temperature by >25 °C. Surprisingly, no difference in affinity is observed among the three, despite differences in number and location of phosphorylation. The nonphosphorylated peptides, CTD_{11} and CTD_{21}, show considerably weaker binding than the equivalent-length phosphorylated peptides; in fact, the denaturation curve in the presence of the 11-mer is indistinguishable from the buffer-only curve.

Direct titration of ligands into solutions containing the WW domain was carried out using fluorescence as a probe. Fluorescence emission intensity at 340 nm was followed as a function of ligand concentration. Information from structural studies and mutagenesis indicates that one of the tryptophans is buried in the hydrophobic core and is required for folding, while the other is required for ligand binding (*23*, *33*). Thus, it is likely that both folding and binding of the ligand contribute to the change in fluorescence signal.

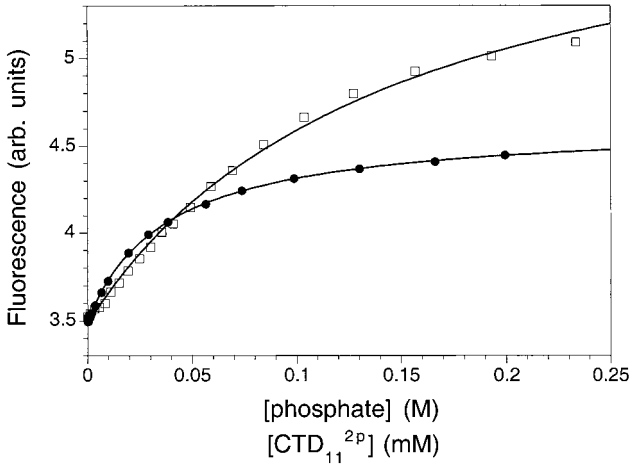

FIGURE 7: Direct titration of the WW domain monitored by fluorescence. Representative titration curves with inorganic phosphate (closed circles) and CTD_{11}^{2p} peptide (open squares). Solid lines are fits to eq 2.

Since at 25 °C the domain is 70% denatured, titration with ligands at room temperature results in linked folding and binding phenomena:

$$N{\cdot}L \overset{K_d}{\rightleftharpoons} N + L \overset{K_u}{\rightleftharpoons} D + L$$

where N represents native WW domain, D represents denatured WW domain, and N·L represents native WW domain bound to ligand. Therefore, we fit titration curves to an equation that takes both folding and binding into account (*39*). The apparent equilibrium constant for the reaction written above is $1/K_{app}$, with

$$\begin{aligned} K_{app} &= ([N] + [N{\cdot}L])/[D] \\ &= \{[N]/[D](1 + [L]/K_d)\} \\ &= [(1 + [L]/K_d)/K_u] \end{aligned} \quad (4)$$

where [L] is the free ligand concentration, K_u is the unfolding equilibrium constant, and K_d is the ligand dissociation constant.

For this analysis, we have assumed that (1) ligands do not bind to D, (2) the binding occurs with 1:1 stoichiometry, and (3) there is a two-state unfolding transition (see above). We have fit the titration data to a standard binding equation using the above definition of K_{app} (see Materials and Methods). It is impossible to decouple K_u and K_d and obtain independent values by fitting a single titration curve. However, it was possible to use the thermal denaturation curve in the absence of ligand (Figure 2) to determine the unfolding equilibrium constant at 25 °C ($K_u = 1.39$), and use this value as a fixed parameter in the titration fits.

Representative fluorescence titration data are shown in Figure 7 for inorganic phosphate and CTD_{11}^{2p} peptide. The fitted dissociation constant for phosphate is 16 mM. Fitting these data with a stoichiometry of $n = 2$ or higher results in residuals which increase in magnitude and are nonrandom; therefore, we have assumed a 1:1 stoichiometry for phosphate and the other ligands as well. Fitted dissociation constants are tabulated in Table 1 for the peptides studied. The data are generally consistent with the CD denaturation curves. The four singly phosphorylated peptides have K_d values

between 67 and 98 μM. There is a slight preference for binding to peptides with phospho-Ser 2 over peptides with phospho-Ser 5, although the difference is not dramatic. The peptides with two or three phosphates all have higher affinity, in the range of 16−21 μM. Curiously, titration with the nonphosphorylated peptides CTD_{11} and CTD_{21} causes a decrease rather than an increase in the fluorescence signal (not shown). This decrease could be indicative of a different mode of binding that quenches the fluorescence of Trp 38, and cancels out the fluorescence enhancement resulting from folding. The weak affinity and small signal change prevent precise measurements of the binding constants in this case. In our CD measurements, CTD_{21} apparently increased the T_m by 6.6 °C, indicating a higher affinity than suggested by fluorescence, whereas CTD_{11} showed no evidence of binding by CD. The CTD_{21} and $CTD_{21}^{5,5,5p}$ peptides have relatively large intrinsic CD signals at 228 nm. The shift in the CD thermal denaturation curve in the presence of CTD_{21} may result from an artifact caused by subtracting this large intrinsic signal. Alternatively, the negative charge on the unblocked C-terminus of CTD_{21} may mimic a phosphorylation and allow some affinity for the WW domain. In any case, both the fluorescence and CD results point to a significant increase in CTD affinity for Ess1 WW domain upon CTD phosphorylation.

DISCUSSION

Ess1, the only prolyl isomerase known to be essential in yeast, has recently been identified as a protein that specifically binds the phosphorylated CTD via its WW domain by affinity-matrix and far-western analysis (*9*). The results presented here for purified Ess1 WW domain are consistent with these observations. WW domain-mediated targeting of Ess1 to the P-CTD would position Ess1 effectively to catalyze the cis/trans isomerization of the many pSer−Pro peptide bonds in the P-CTD. Because factors that bind to the P-CTD may discriminate between cis and trans proline bonds, catalysis of the normally slow isomerization reaction may be necessary to expedite factor binding. Another possibility is that Ess1 catalyzes proline isomerization in other P-CTD binding factors.

In addition to biochemical data, genetic evidence supports a role for Ess1 in transcription-related events. For example, mutations in Ess1 cause defects in pre-mRNA 3′-end processing (*12*). Moreover, six suppressors of Ess1 temperature-sensitive mutants were recently isolated in yeast (*22*). Five of the six are transcription-related, and the sixth protein is cyclophilin A, another prolyl isomerase. Thus, genetic evidence strongly links not only Ess1 but also isomerase activity to transcription. Mutations in the WW domain can lead to defective Ess1 proteins (*22*), perhaps by impairing the binding of Ess1 to P-CTD.

In our in vitro binding studies, the 11-residue peptide, SPTSPSYSPTS (CTD_{11}), showed little evidence of interacting with the WW domain unless serine residues preceding proline residues were phosphorylated. The introduction of phospho-Ser at the fourth or eighth amino acid of SPTSPSYSPTS, corresponding to Ser 5′ and Ser 2′ of the consensus repeat YSPTSPS (the prime indicates conventional repeat numbering), produces peptides that clearly bound the WW domain. It may be significant that CTD_{11}^{2p} peptide bound somewhat more strongly than CTD_{11}^{5p}, particularly as this result was replicated with two similar peptides: CTD_9^{2p} and CTD_9^{5p}. On the other hand, it is probably more significant that phosphorylation at either Ser 2′ or Ser 5′ yielded peptides with about the same affinity, indicating a lack of strong discrimination. This same conclusion can be drawn from the similar binding behaviors of the peptides with two or three phosphates, which have very different patterns of phosphorylation. Thus, our data suggest that the presence of multiple phosphates is more important for tight binding than is either selection between Ser 2′ and Ser 5′ phosphorylation locations, or a specific pattern of phosphorylation. As a consequence, Ess1 can bind the P-CTD effectively despite the heterogeneous or changing phosphorylation that occurs in vivo. The WW domain of Ess1 may have evolved to bind multiple sites along the hyperphosphorylated CTD with similar affinity. In this case, each consensus CTD repeat would have two possible pSer-Pro binding sites.

The structure of Pin1 (the human homologue of Ess1) bound to a single repeat, doubly phosphorylated CTD peptide (YpSPTpSPS), has been published (*23*). Inspection of the crystal structure (Figure 1B) does not provide an obvious explanation of the enhancement of binding affinity resulting from multiple sites of phosphorylation. The peptide adopts an extended conformation, and both Ser−Pro bonds are trans. The most extensive peptide/protein contacts are between the WW domain and the C-terminal half of the peptide (there are no contacts between the peptide and the isomerase domain of Pin1). Trp 34 (Pin1 numbering) and Tyr 23 of the WW domain both contact Pro 6′ (prime numbering indicates conventional CTD repeat numbering) and provide a hydrophobic binding pocket. The phosphate on Ser 5′ forms a hydrogen bond/salt bridge with Arg 17, and also hydrogen bonds to a buried water molecule, which is in turn hydrogen bonded to the side-chain hydroxyl of Tyr 23. In contrast, the phosphate on Ser 2′ does not contact the protein, and the closest basic groups, Arg 14 and Arg 17, are 4.7 and 6.6 Å away, respectively. From inspecting the structure alone, one would predict that peptides phosphorylated at Ser 5′ would have the same affinity as one phosphorylated at Ser 5′ and Ser 2′. Indeed, in solution studies of the Pin1 WW domain, the peptide YSPTpSPS had the same affinity as YpSPTpSPS [~30 μM, (*23*)], suggesting that phospho-Ser 2′ was not important in the binding interaction. This result is consistent with the crystal structure but differs from our results. There may be several possible explanations. The CTD peptide sequence used in the Pin1 study is moderately different from our peptides. The studies of Verdecia et al. utilized heptapeptides with the single tyrosine residue at the N-terminus. Our studies utilized peptides with tyrosine residues and phosphorylated serines in more central positions (Table 1), containing a -SPSY- sequence not present in the Tyr-terminated heptapeptide. This sequence, similar to the -PPXY- recognition motif in other WW domains (*29*), may be involved in the binding of our peptides. Another possibility is that the structure of the Ess1 WW binding site is substantially different than the WW domain of Pin1. However, the identity of the residues in the apparent binding site is conserved between Ess1 WW and Pin1 WW, with one minor exception: Arg17 of Pin1 is a Lys in Ess1. Certainly our data imply that the doubly phosphorylated

peptides $CTD_{11}^{5,2p}$ and $CTD_{14}^{2,5p}$ interact with the Ess1 WW domain in a more complex manner than suggested by the Pin1 crystal structure. Nor can we eliminate the possibility that peptides bind to Ess1 WW in the opposite N- to C-terminal orientation, as exemplified by the binding of a peptide to the WW domain of dystrophin (*46*, *47*).

Our unexpected binding results may indicate that the Ess1 WW domain could bind more than one phospho-Ser simultaneously in ways not shown by the Pin1 crystal structure. Granted, the measured stoichiometry of inorganic phosphate appears to be equal to 1. However, the second phosphate site could be considerably weaker than the first and still contribute significantly to the observed peptide affinity, since by binding to the first site the peptide already has paid much of the entropic cost [see (*48*) and the discussion below]. Alternatively, each phospho-Ser could represent an individual binding site, so that the three peptides with multiple phosphates (and the full-length P-CTD) present the WW domain with multiple binding sites. A combination of these two explanations is also possible. The ability of Ess1 WW domain to bind different phospho-Ser CTD sites with comparable affinity would allow Ess1 to bind at a manifold of possible sites along the P-CTD (two for each consensus repeat), localizing the isomerase domain sufficiently to catalyze bond isomerization. Ess1 may even processively catalyze cis/trans isomerization along the length of the P-CTD, with the WW domain assisting binding and movement along the molecule. This situation may be akin to the well-known one-dimensional "sliding" of DNA binding proteins along a DNA duplex.

While the binding constants for short P-CTD peptides to the WW domain are fairly modest ($\sim$20−100 μM), the affinity of intact Ess1 for the full-length P-CTD should be stronger. Both the WW domain and the isomerase domain of Ess1 have binding sites for P-CTD. The isomerase domain alone has weak affinity for P-CTD heptapeptides (*23*); however, since the entropic penalty for binding is largely paid in the initial binding event to the WW domain, the effective affinity of the isomerase site is enhanced (*48*). The full length P-CTD would contain many binding sites, so the local concentration of sites is quite high. The situation with Ess1 bound to two different places on the P-CTD using both the isomerase domain and WW domain is similar to the A−B binding scenario discussed by Jencks (*48*). The presence of the isomerase domain, and the many potential binding sites on the full-length P-CTD, would lead to an interaction between Ess1 and P-CTD that is tighter than indicated by our measured binding constants. This may explain why Ess1 bound very tightly to a P-CTD affinity column, eluting only at 1 M NaCl (*9*).

Two other WW-containing proteins, yeast Prp 40 and human CA150, interact with phosphorylated CTD (*8*, *49*). The WW domains of Rsp5 and YAP have been shown to bind unphosphorylated CTD (*32*). Thus, WW domains are protein modules that frequently mediate protein/CTD interactions. Further studies of the WW domain interaction with the P-CTD will illuminate the mechanism and consequences of nuclear factor interactions with the two forms of CTD. In the case of Ess1, hyperphosphorylation of the CTD results in binding that does not depend on the precise locations or arrangements of phosphorylated serines. This lack of discrimination presumably allows Ess1 to bind multiple sites along the P-CTD, either to catalyze pSer−Pro bond isomerization or to act on other P-CTD-bound proteins.

ACKNOWLEDGMENT

We thank Steven Hanes for the three-repeat CTD peptides and Homme Hellinga for the use of his fluorometer.

REFERENCES

1. Corden, J. L. (1990) *Trends Biochem. Sci. 15*, 383−387.
2. Conaway, J. W., Shilatifard, A., Dvir, A., and Conaway, R. C. (2000) *Trends Biochem. Sci. 25*, 375−380.
3. Bentley, D. (1999) *Curr. Opin. Cell Biol. 11*, 347−351.
4. Greenleaf, A. L. (1993) *Trends Biochem. Sci. 18*, 117−119.
5. Koleske, A. J., and Young, R. A. (1995) *Trends Biochem. Sci. 20*, 113−116.
6. Steinmetz, E. J. (1997) *Cell 89*, 491−494.
7. Dahmus, M. E. (1996) *J. Biol. Chem. 271*, 19009−19012.
8. Morris, D. P., and Greenleaf, A. L. (2000) *J. Biol. Chem. 275*, 39935−39943.
9. Morris, D. P., Phatnani, H. P., and Greenleaf, A. L. (1999) *J. Biol. Chem. 274*, 31583−31587.
10. Schutkowski, M., Bernhardt, A., Zhou, X. Z., Shen, M., Reimer, U., Rahfeld, J. U., Lu, K. P., and Fischer, G. (1998) *Biochemistry 37*, 5566−5575.
11. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997) *Science 278*, 1957−1960.
12. Hani, J., Schelbert, B., Bernhardt, A., Domdey, H., Fischer, G., Wiebauer, K., and Rahfeld, J. U. (1999) *J. Biol. Chem. 274*, 108−116.
13. Cho, H., Kim, T. K., Mancebo, H., Lane, W. S., Flores, O., and Reinberg, D. (1999) *Genes Dev. 13*, 1540−1552.
14. Chambers, R. S., and Dahmus, M. E. (1994) *J. Biol. Chem. 269*, 26243−26248.
15. Chambers, R. S., and Kane, C. M. (1996) *J. Biol. Chem. 271*, 24498−24504.
16. Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999) *Mol. Cell 4*, 55−62.
17. McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M., and Bentley, D. L. (1997) *Nature 385*, 357−361.
18. Hirose, Y., Tacke, R., and Manley, J. L. (1999) *Genes Dev. 13*, 1234−1239.
19. Hirose, Y., and Manley, J. L. (1998) *Nature 395*, 93−96.
20. Cho, E. J., Takagi, T., Moore, C. R., and Buratowski, S. (1997) *Genes Dev. 11*, 3319−3326.
21. Corden, J. L., and Patturajan, M. (1997) *Trends Biochem. Sci. 22*, 413−416.
22. Wu, X., Wilcox, C. B., Devasahayam, G., Hackett, R. L., Arevalo-Rodriguez, M., Cardenas, M. E., Heitman, J., and Hanes, S. D. (2000) *EMBO J. 19*, 3727−3738.
23. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T., and Noel, J. P. (2000) *Nat. Struct. Biol. 7*, 639−643.
24. Ranganathan, R., Lu, K. P., Hunter, T., and Noel, J. P. (1997) *Cell 89*, 875−886.
25. Sudol, M. (1996) *Prog. Biophys. Mol. Biol. 65*, 113−132.
26. Sudol, M., and Hunter, T. (2000) *Cell 103*, 1001−1004.
27. Shen, M., Stukenberg, P. T., Kirschner, M. W., and Lu, K. P. (1998) *Genes Dev. 12*, 706−720.
28. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) *Science 283*, 1325−1328.
29. Linn, H., Ermekova, K. S., Rentschler, S., Sparks, A. B., Kay, B. K., and Sudol, M. (1997) *Biol. Chem. 378*, 531−537.
30. Kanelis, V., Farrow, N. A., Kay, L. E., Rotin, D., and Forman-Kay, J. D. (1998) *Biochem. Cell Biol. 76*, 341−350.
31. Chen, H. I., Einbond, A., Kwak, S. J., Linn, H., Koepf, E., Peterson, S., Kelly, J. W., and Sudol, M. (1997) *J. Biol. Chem. 272*, 17070−17077.

32. Chang, A., Cheang, S., Espanel, X., and Sudol, M. (2000) *J. Biol. Chem. 275*, 20562−20571.
33. Koepf, E. K., Petrassi, H. M., Ratnaswamy, G., Huff, M. E., Sudol, M., and Kelly, J. W. (1999) *Biochemistry 38*, 14338−14351.
34. Koepf, E. K., Petrassi, H. M., Sudol, M., and Kelly, J. W. (1999) *Protein Sci. 8*, 841−853.
35. Crane, J. C., Koepf, E. K., Kelly, J. W., and Gruebele, M. (2000) *J. Mol. Biol. 298*, 283−292.
36. Calderone, T. L., Stevens, R. D., and Oas, T. G. (1996) *J. Mol. Biol. 262*, 407−412.
37. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) *Protein Sci. 4*, 2411−2423.
38. Swint, L., and Robertson, A. D. (1993) *Protein Sci. 2*, 2037−2049.
39. Henkels, C. H., Kurz, J., Fierke, C. A., and Oas, T. G. (2001) *Biochemistry 40*, 2777−2789.
40. Sreerama, N., Manning, M. C., Powers, M. E., Zhang, J. X., Goldenberg, D. P., and Woody, R. W. (1999) *Biochemistry 38*, 10814−10822.
41. Woody, R. W. (1994) *Eur. Biophys. J. 23*, 253−262.
42. Robertson, A. D., and Murphy, K. P. (1997) *Chem. Rev. 97*, 1251−1267.
43. Brandts, J. F. (1969) in *Structure and Stability of Biological Macromolecules* (Timasheff, S. N., and Fasman, G. D., Eds.) pp 213−290, Marcell Dekker, New York.
44. Schellman, J. A. (1975) *Biopolymers 14*, 999−1018.
45. Bienkiewicz, E. A., Moon Woody, A., and Woody, R. W. (2000) *J. Mol. Biol. 297*, 119−133.
46. Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000) *Nat. Struct. Biol. 7*, 634−638.
47. Zarrinpar, A., and Lim, W. A. (2000) *Nat. Struct. Biol. 7*, 611−613.
48. Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4046−4050.
49. Carty, S. M., Goldstrohm, A. C., Sune, C., Garcia-Blanco, M. A., and Greenleaf, A. L. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 9015−9020.
50. Kraulis, P. K. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI0027884