Phospho-Carboxyl-Terminal Domain Binding and the Role of a Prolyl Isomerase in Pre-mRNA 3'-End Formation*  
(Received for publication, July 28, 1999, and in revised form August 24, 1999)  
Daniel P. Morris, Hemali P. Phatnani, and Arno L. Greenleaf*  
From the Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710  

A phospho-carboxyl-terminal domain (CTD) affinity column created with yeast CTD kinase I and the CTD of RNA polymerase II was used to identify Ess1/Pin1 as a phospho-CTD-binding protein. Ess1/Pin1 is a peptidyl prolyl isomerase involved in both mitotic regulation and pre-mRNA 3'-end formation. Like native Ess1, a GST-Ess1 fusion protein associates specifically with the phosphorylated but not with the unphosphorylated CTD. Further, hyperphosphorylated RNA polymerase II appears to be the dominant Ess1 binding protein in total yeast extracts. We demonstrate that phospho-CTD binding is mediated by the small WW domain of Ess1 rather than the isomerase domain. These findings suggest a mechanism in which the WW domain binds the phosphorylated CTD of elongating RNA polymerase II and the isomerase domain reconfigures the CTD though isomerization of proline residues perhaps by a processive mechanism. This process may be linked to a variety of pre-mRNA maturation events that use the phosphorylated CTD, including the coupled processes of pre-mRNA 3'-end formation and transcription termination.

The carboxyl-terminal domain (CTD) of the largest subunit of RNA polymerase II plays a central role in controlling mRNA production in eukaryotes. In most eukaryotes the CTD consists of between 26 and 52 repeats with the consensus sequence, YSPTSPS (1). This unusual domain is not found in eukaryotic RNA Pol I or Pol III or in prokaryotic RNA polymerase. By mechanisms that are still largely not understood, RNA Pol II behavior is controlled by CTD kinases and phosphatases that modulate the phosphorylation state of the CTD (2). A body of evidence has accumulated indicating that the unphosphorylated CTD is present in RNA Pol II holoenzyme and is required during the initial stages of transcription (3–6) but that a hyperphosphorylated CTD is present in elongating polymerase (7–10).

The hypothesis that the phospho-CTD functions in pre-mRNA processing (11) has received firm experimental backing with the recognition that several components of the 5' mRNA capping enzyme complex are phospho-CTD-binding proteins (12–14). In addition, indirect evidence suggests that the phospho-CTD also binds splicing components (15–18). Further, both the phosphorylated and the unphosphorylated CTD can interact with pre-mRNA 3'-end processing proteins (19, 20). The combined evidence suggests that pre-mRNA processing may in general be cotranscriptional, with the (phospho)-CTD serving as a master coordinator of the events involved.

Recently, mutations in the peptidyl prolyl isomerase, ESS1/PTFI, were found to be responsible for the temperature-sensitive phenotypes of two yeast mutants that displayed a pre-mRNA 3'-end formation defect (21). Yeast ESS1 was originally identified as essential gene (22), and is in fact the only essential peptidyl prolyl isomerase (PPIase) in Saccharomyces cerevisiae (23). The human protein, Pin1, and the Drosophila protein, Dodo, are close homologues of Ess1 and both can functionally replace Ess1 in yeast (24, 25). All three proteins have a small 31-residue WW domain amino-terminal to the isomerase domain. Studies of the peptidyl prolyl isomerase activity of recombinant Ess1 (21) and Pin1 (26) indicate these enzymes display more than a 1000-fold higher specificity toward substrates in which the other amino acid at the immobilizable bond is negatively charged (Glu, phospho-Ser, and phospho-Thr). In higher eukaryotes a variety of techniques have been used to show that Pin1 stably associates with phosphorylated but not unphosphorylated mitotic regulators that are targets of Cdc2/cyclinB (26–29). The present study demonstrates that the hyperphosphorylated CTD of RNA Pol II contains high affinity binding sites for the WW domain of Ess1.

MATERIALS AND METHODS

Identification of Ess1 as a Phospho-CTD-binding Protein—To make the phospho-CTD affinity column, CTD kinase I was purified using polyclonal antibodies against a peptide from CKT1 as described (30). Phosphorylation of a β-galactosidase/yeast CTD fusion protein, βGal-CTD (FFY1 in Ref. 30), was done as described (31) with sufficient CTD kinase I to cause a complete shift of βGalCTD to the hyperphosphorylated form as observed on a stained gel. A duplicate blank reaction was done without kinase, and both reactions were desalted into PBS using Quick Spin Protein Columns (Roche Molecular Biochemicals). The AminoLink Plus Immobilization kit (Pierce) was used to make 250 µl each of phospho-CTD and CTD matrix (pH 10 coupling protocol). The yeast strain, DY6 (a leu2 trp1 ura3–52 prb1–1122 pep4–3 prc1–407) was grown at 30 °C in minimal medium to an A600 of 1 to 2. Cells were collected and disrupted by glass beads using a protocol (39) modified for a mini bead beater and substituting a protease inhibitor mix from Roche Molecular Biochemicals (catalog 1206 8934). The only other modification was the direct recovery of the extract from the disruption tube by centrifugation (15,000 RCF for 1 min) followed by pipette removal of as much supernatant from the beads as was possible. Prior to affinity chromatography, extracts were thawed and then clarified by centrifugation at 12,000 RCF for 30 s. The supernatant was diluted with 2 volumes of buffer (30 µl Tris, pH 7.8, 3 µl PMSF) and reclarified by centrifugation. Equal amounts of on-put (2.5 ml) were loaded onto paired 250-µl columns containing identical amounts of phosphorylated or unphosphorylated βGal-CTD. Each column was washed five times with 400-µl column equilibration buffer (10 mM Tris, pH 7.8, 100 mM NaCl, and 1 mM PMSF) and eluted with sequential steps consisting of 1 column volume of 110, 125, 140, 160, 190, 220, 250, 300, 350, 425, 500, 600, 700, 850, 1000, 1000, and 1000 µM NaCl in 10 mM Tris, pH 7.8, with 1 µl PMSF. Initial gel analysis using silver staining revealed a 21-kDa protein eluting specifically from the phospho-CTD column at 1

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. E-mail: arno@biochem.duke.edu.

§ The abbreviations used are: CTD, carboxyl-terminal domain; Pol I, II, III, polymerase I, II, and III, respectively; PMSF, phenylmethylsulfonyl fluoride; MALDI-MS, matrix-assisted laser desorption/ionization-mass spectrometry; PBS, phosphate-buffered saline; DTT, dithiothreitol; GST, glutathione S-transferase; RCF, relative centrifugal force.
m NaCl. The corresponding high salt fractions from both columns were brought to 0.1% SDS, concentrated, and desalted with Microcon 10 units (Amicon) and separated on a 4–15% SDS precast gel (Bio-Rad). Following a protocol from HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory at Yale University, the gel containing the 21-kDa band was excised and destained for 3 h in 10% acetic acid and 50% methanol. The band was frozen and sent to the biotechnology facility at Yale for in-gel trypsinization and MALDI-MS analysis.

Expression of GSTEss1 Fusion Proteins—Fusion proteins were expressed using plasmids created by ligating Ess1 polymerase chain reaction (PCR) amplification products into the EcoRI/BamHI sites of the polylinker of pGEX2TK (Amersham Pharmacia Biotech). The primers CGGATCCCATGATCATTGTGACTTATTCC and GGAATTCCTAGTTGGTGCCCTCAGGCTC (WW) were used to amplify the fragments containing the WW domain (GST-Ess1WW) and isomerase domain (GST-Ess1ISO) of pGEX2TK-GSTEss1 using the primers CGGATCCATGATCATTGTGACTTATTCC and GGAATTCCTAACCTACCGCGTTGGTGCCCTCAGGCTC (WW) and GGAATTCCTAACCTACCGCGTTGGTGCCCTCAGGCTC (ISO) inserted into the EcoRI/BamHI sites of pGEX2TK (Amersham Pharmacia Biotech) and sequenced. Protein expression and purification procedures according to the protocol of the manufacturer (Amersham Pharmacia Biotech, XX-058-00-01) with minor modification. BL21 cells containing the expression plasmids were grown in LB medium with 50 μg/ml ampicillin, and the fusion proteins were purified by batch adsorption. Complete elution of GST-Ess1 or GSTEss1WW required an overnight incubation with modified elution buffer (50 mM Tris, 10 mM glutathione, and 100 mM NaCl) prepared essentially as in Ref. 33 for 1 h, then washed 5 × 12 min with PBS-Tw prior to addition of the 2′ antibody and treatment as for the far-Western analyses.

RESULTS

 Yeast CTD kinase I can quantitatively convert fusion protein substrates containing the CTD to a hyperphosphorylated form, with the accompanying characteristic decrease in electrophoretic mobility (31). To identify proteins that bind specifically to the hyperphosphorylated CTD, βGal-yCTD was hyperphosphorylated and cross-linked to agarose beads. Total yeast extracts were loaded in parallel onto this and an identical column prepared with unphosphorylated βGal-yCTD and eluted using a step gradient of NaCl. While many proteins eluted from both columns, a protein of about 21 kDa appeared to elute only from the phospho-CTD column at 1 m NaCl (Fig. 1A). This protein band was excised and subjected to in-gel trypsinization followed by MALDI mass spectrometric analysis of the tryptic fragments (Fig. 1B). Data base searches using the masses of the fragments gave an unambiguous identification of the protein as Ess1.

To confirm the ability of Ess1 to bind the phosphorylated CTD, a GSTEss1 fusion protein was purified and linked to agarose beads. As shown in Fig. 2, GSTEss1 beads were able to specifically deplete hyperphosphorylated GST-yCTD fusion protein from a solution containing both phosphorylated and unphosphorylated GSTyCTD. In addition to the unphosphorylated fusion protein, some forms of phosphorylated GSTyCTD were not effectively adsorbed, including a partially shifted intermediate clearly visible in Fig. 2. Extending the incubation an additional 30 min did not result in significant depletion of the remaining unbound material (data not shown). This may suggest that not all phosphorylations create a recognition site for Ess1.

The hyperphosphorylated CTD produced by CTD kinase I can also be used as an effective radiolabeled probe to identify phospho-CTD-binding proteins in blots of cellular extracts. Preliminary far-Western analysis of proteolyzed Ess1 suggested the WW domain could bind the phospho-CTD. 3 To test this hypothesis we created expression constructs for GST fusion proteins containing only the WW (GSTEss1WW) or isomerase (GSTEss1ISO) domains. Far-Western analysis of the fusion proteins (Fig. 3) shows similar association of the hyperphosphorylated CTD probe with GSTEss1 (lanes 1–3) and GSTEss1WW (lanes 4–6) but no association with GSTEss1ISO (lanes 7–9). Clearly, under these conditions the phosphorylated CTD appears to associate only with fusion proteins including the WW domain. As a control to demonstrate specificity, total Escherichia coli proteins were run in lane 10 of Fig. 3. There should be no natural phospho-CTD-binding proteins in E. coli, and indeed far-Western analysis reveals no phospho-CTD-binding proteins.

To demonstrate that in vivo hyperphosphorylated RNA polymerase II from yeast can associate with Ess1 and to determine whether the WW domain could mediate this interaction, we analyzed total cellular extracts of asynchronously grown yeast by far-Western using the fusion proteins as probes (Fig. 4). In these experiments a protein of >200 kDa was the dominant interacting species, bound equivalently by GSTEss1 (lanes 5 and 6) and GSTEss1WW (lanes 8 and 9). This protein is almost certainly the hyperphosphorylated form of the largest

2 A. L. Greenleaf, unpublished data.

3 D. P. Morris, unpublished observation.
subunit of RNA Pol II (subunit IIo), as it has an identical electrophoretic mobility to IIo detected using affinity-purified anti-phospho-CTD polyclonal antibodies (lanes 11 and 12).

Note that neither the control lanes probed with GST (lanes 2 and 3) nor the lanes probed with GSTEss1ISO (lanes 14 and 15) gave any detectable signal.

DISCUSSION

Our results effectively demonstrate that the phosphorylated CTD binds directly to Ess1 and that it is the WW domain that mediates this binding. Although a direct association between the phospho-CTD and the human Pin1 has not been demonstrated, Albert et al., have recently shown that beads containing Pin1 can precipitate a fraction of the hyperphosphorylated RNA Pol II from HeLa cell extracts (34). Given our results with Ess1, this finding strongly suggests that Pin1 also interacts directly with the phosphorylated CTD.

Also recently, it has been shown that the WW and not the isomerase domain of Pin1 is responsible for Pin1 binding to proteins containing mitosis-specific phospho-epitopes (29). Further, the WW domain of Ess1 was also able to bind most of these mammalian proteins. At present, the connection between Ess1/Pin1 association with the phospho-CTD and with mitotic regulatory proteins remains to be determined.

It has been demonstrated that isomerase activity of both Pin1 (27) and Ess1 (21) has a strong preference for proline
A Phospho-CTD Binding Prolyl Isomerase

Duplicate analysis. About 500, 50, and 5 ng of each of the full-sized fusion proteins 4–15% SDS gels were run for Coomassie staining and far-Western proteolyzed species although the smaller species are abundant only for GSTEss1ISO. Lane 10 contains 5 μl of E. coli cells (BL21) from a culture grown to saturation. Lane 11 contains prestained markers with the indicated masses.

**FIG. 3.** A phosphorylated CTD probe associates only with fusion proteins containing the WW domain of Ess1. Duplicate 4–15% SDS gels were run for Coomassie staining and far-Western analysis. About 500, 50, and 5 ng of each of the full-sized fusion proteins were loaded as indicated (lanes 1–9). Note that all the proteins include proteolyzed species although the smaller species are abundant only for GSTEss1ISO. Lane 10 contains 5 μl of E. coli cells (BL21) from a culture grown to saturation. Lane 11 contains prestained markers with the indicated masses.

INTERACTION BLOTS

**FIG. 4.** The predominant Ess1 binding protein is the hyperphosphorylated form of the largest subunit of RNA Pol II. Multiple identical sets of lanes containing 1 and 5 μl of total yeast extract separated by prestained markers (lanes 1, 4, 7, 10, and 13) were run on a 4–15% SDS gel and transferred to nitrocellulose. Strips were cut and probed with GST, GSTEss1, GSTEss1WW, affinity-purified rabbit anti-GST IgG, and all strips were developed using alkaline phosphatase-coupled goat anti-rabbit IgG as a 2nd antibody. The strips with GST and with GSTEss1ISO were developed twice as long as the others.

bonds in which the proline is preceded by an acidic residue (Glu, phospho-Ser, and phospho-Thr). These functional similarities, the conservation of amino acid sequence and domain organization, and the ability of Pin1 to substitute for Ess1 in yeast all indicate a very high level of functional conservation between these two proteins.

Beyond the preference of the isomerase domain for an acidic residue prior to proline, the specificity requirements for the domains of Ess1/Pin1 are not well understood. The initial report examining the specificity of Pin1 found that its ability to bind peptides could be used as a basis for designing efficient isomerase substrates (26). This seemed to suggest the peptide binding site and prolyl isomerase site were the same; however, our study and the study mentioned above (29) strongly suggest it was the specificity of the WW domain that was responsible for peptide binding. Relevant to our work, a sequence found to be a preferred binding site, YpSP, represents a sequence potentially present in the phosphorylated CTD repeats. Peptide substrates with this sequence were also the best isomerase substrates. While the binding sites of the WW and isomerase domains are clearly distinct, they may still have similar recognition determinants.

WW domain associations with Pol II are not limited to the phosphorylated CTD. It has been shown that several WW domains, including one found in the mouse ubiquitin ligase, NEDD4, one found in human YAP (35), and at least one found in yeast Rsp5 (36), can associate with the unphosphorylated CTD. In agreement with our results, it has also been shown that the WW domain of Ess1 fails to interact with the unphosphorylated CTD (35). Interestingly, two studies have looked at the sequence requirements for peptide binding to the WW domain of human YAP, and both identified a sequence present in the CTD, PxY, as the most important for the binding site motif (37, 38). The earlier and more thorough study found the position prior to proline was best occupied by another proline but was tolerant of Leu, Ser, Val, or Ala (37). The sequence SPxY is present in the CTD sequence, probably explaining the ability of human YAP to bind the unphosphorylated CTD.

Previous work has revealed a direct physical connection between 3′-end processing and the CTD. Mammalian cleavage stimulation factor (CstF) can be eluted from either a phosphorylated or an unphosphorylated CTD column, and the CstF p50 and p77 components have been shown to bind directly to the unphosphorylated CTD (19). Recently, both the unphosphorylated and the phosphorylated CTD have been found to be potent stimulators of the 3′ cleavage reaction (20). A functional link between Ess1 and 3′-end processing was established by the isolation of mutations causing a 3′-end formation defect which mapped to the isomerase domain of Ess1 (21). Our demonstration that Ess1 stably associates with the phosphorylated CTD immediately suggests a mechanistic explanation for these observations; phospho-CTD binding by the WW domain localizes Ess1 to elongating Pol II, where it is positioned to isomerize pSP bonds present in the phospho-CTD. In addition, the separation of the binding and isomerase activities of Ess1 into independent domains creates the potential for processive isomerase function. Changes in CTD structure resulting from the activity of Ess1 would be expected to affect 3′-end formation by altering interactions between the CTD and cleavage/polyadenylation components. Involvement of CTD kinase I in this process is consistent with the finding that a suppressor of CTDK-I deficiency-caused cold sensitivity contains several regions homologous to components of the 3′-end processing machinery in yeast and mammals.4 The possibility that Ess1 reconfigures the phosphorylated CTD though processive isomerization of its proline residues certainly adds an unexpected twist to the investigation of eukaryotic pre-mRNA transcription and 3′-end formation.

REFERENCES

1. Corden, J. L. (1990) Trends. Biochem. Sci. 15, 383–387
2. Dahmus, M. (1996) J. Biol. Chem. 271, 19009–19012
3. Thompson, C. M., Koleske, A. J., Chao, D. M., and Young, R. A. (1997) Cell 73, 1361–1375
4. Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H., and Kornberg, R. D. (1994) Cell 77, 599–608
5. Myer, V. E., and Young, R. A. (1998) J. Biol. Chem. 273, 27757–27760
6. Hampsey, M., and Reinberg, D. (1999) Curr. Opin. Genet. Dev. 9, 132–139
7. Cadena, D. L., and Dahmus, M. E. (1987) J. Biol. Chem. 262, 12468–12474
8. Weeks, J. R., Hardin, S. E., Shen, J., Lee, J. M., and Greenleaf, A. L. (1993) Genes Dev. 7, 2329–2344
9. O’Brien, T., Hardin, S., Greenleaf, A. L., and Lis, J. T. (1994) Nature 370, 75–77
10. Greenblatt, J. (1997) Curr. Opin. Cell Biol. 9, 310–319
11. Greenleaf, A. L. (1993) Trends Biochem. Sci. 18, 117–119
12. Cho, E. J., Takagi, T., Moore, C. R., and Buratowski, S. (1997) Genes Dev. 11.

---

4 D. Skaar and A. L. Greenleaf, manuscript in preparation.
13. McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D. L. (1997) *Genes Dev.* 11, 3306–3318

14. Ho, C. K., and Shuman, S. (1999) *Mol. Cell* 3, 405–411

15. Du, L., and Warren, S. L. (1997) *J. Cell Biol.* 136, 5–18

16. Kim, E., Du, L., Bregman, D. B., and Warren, S. L. (1997) *J. Cell Biol.* 136, 19–28

17. Carden, J. L., and Patturajan, M. (1997) *Trends Biochem. Sci.* 22, 413–416

18. Hirose, Y., Tacke, R., and Manley, J. L. (1999) *Genes Dev.* 13, 1234–1239

19. 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

20. Hirose, Y., and Manley, J. L. (1998) *Nature* 395, 93–96

21. Hani, J., Schelbert, B., Bernhardt, A., Domdey, H., Fischer, G., Wiehauer, K., and Rahfeld, J. U. (1999) *J. Biol. Chem.* 274, 108–116

22. Hanes, S. D., Shank, P. R., and Bostian, K. A. (1989) *Yeast* 5, 55–72

23. Dolinski, K., Muir, S., Cardenas, M., and Heitman, J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 13093–13098

24. Malessza, K., Hanes, S. D., Hackett, R. L., de Couet, H. G., and Miklos, G. L. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 447–451

25. Lu, K. P., Hanes, S. D., and Hunter, T. (1996) *Nature* 380, 544–547

26. 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

27. Shen, M., Stukenberg, P. T., Kirschner, M. W., and Lu, K. P. (1998) *Genes Dev.* 12, 706–720

28. Crenshaw, D. G., Yang, J., Means, A. R., and Kornbluth, S. (1998) *EMBO J.* 17, 1315–1327

29. Lu, P. J., Zhou, X. Z., Shen, M., and Lu, K. P. (1999) *Science* 283, 1325–1328

30. Sterner, D. E., Lee, J. M., Hardin, S. E., and Greenleaf, A. L. (1995) *Mol. Cell. Biol.* 15, 5716–5724

31. Lee, J. M., and Greenleaf, A. L. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 3624–3628

32. Morris, D. P., Lee, J. M., Sterner, D. E., Brickey, W. J., and Greenleaf, A. L. (1997) *Methods* 12, 264–275

33. Lee, J. M., and Greenleaf, A. L. (1991) *Gene Expr.* 1, 149–167

34. Albert, A., Lavoie, S., and Vincent, M. (1999) *J. Cell Sci.* 112, 2493–2500

35. Gavva, N. R., Gavva, R., Ermekova, K., Sudol, M., and Lu, K. P. (1997) *J. Biol. Chem.* 272, 24105–24108

36. Wang, G., Yang, J., and Huibregtse, J. M. (1999) *Mol. Cell. Biol.* 19, 342–352

37. Linn, H., Ermekova, K. S., Rentschler, S., Sparks, A. B., Kay, B. K., and Sudol, M. (1997) *Biol. Chem.* 378, 531–537

38. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N., and Lim, W. A. (1998) *Science* 282, 2088–2092

39. Ausubel, F. M., et al. (1994) *Current Protocols in Molecular Biology*, Section 13.13.4, John Wiley & Sons, Inc., New York