The Journal of Biological Chemistry
© 1991 by The American Society for Biochemistry and Molecular Biology, Inc.

Identification of Phosphorylation Sites in the Repetitive Carboxyl-terminal Domain of the Mouse RNA Polymerase II Largest Subunit*

Jie Zhang and Jeffry L. Corden†
From the Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

The largest subunit of eukaryotic RNA polymerase II contains a carboxyl-terminal domain (CTD) which is comprised of repetitive heptapeptides with a consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. We demonstrate here that the mouse CTD expressed in and purified from Escherichia coli can be phosphorylated in vitro by a p34cdc2/CDC28-containing CTD kinase from mouse ascites tumor cells. The product of this reaction, a phosphorylated form of the CTD, contains phosphoserine and phosphothreonine, but not phosphotyrosine. The same phosphoamino acid content is observed in the in vivo phosphorylated CTD from a mouse cell line. Synthetic peptides with naturally occurring non-consensus heptapeptide sequences can also be phosphorylated by CTD kinase in vitro. Phosphoamino acid analysis of these non-consensus heptapeptides together with direct sequencing of a phosphorylated heptapeptide reveals that serines (or threonines) at positions two and five are the sites phosphorylated by mouse CTD kinase. Thus, the -Ser(Thr)-Pro- motif common to the consensus heptapeptide sequence is a consensus recognition site for mouse CTD kinase.

RNA polymerase II contains an unusual repetitive domain at the carboxyl terminus of its largest subunit. This carboxyl-terminal domain (CTD) is comprised of tandem repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser and is present in RNA polymerase II but not in RNA polymerases I or III in prokaryotic RNA polymerase (for review see Corden, 1990; Woychik and Young, 1990). Three different forms of RNA polymerase II have been described, each differing in the apparent molecular weight of the largest subunits (for review see Sawadogo and Sentenac, 1990). Form IIa contains the primary translation product of the largest subunit (IIa) and form IIb contains two proteolytically degraded large subunits (IIb) which has lost the CTD but retains catalytic activity. Form IIb contains a highly phosphorylated form of the largest subunit (IIb). Forms IIa and IIb have been observed in vivo (Kim and Dahmus, 1986), while form IIb appears to be an artifact produced during enzyme purification.

Genetic studies (Allison et al., 1988; Bartolomei et al., 1988; Nonet et al., 1987; Zehring et al., 1988) have indicated that the CTD plays an essential, albeit unknown role in RNA polymerase II function. The CTD, rich in serine and threonine, is known to be highly phosphorylated in vivo (Cadena and Dahmus, 1987) and this modification is thought to be important in CTD function. In vitro transcription experiments have shown that RNA polymerase IIA becomes phosphorylated after formation of the preinitiation complex (Kim and Dahmus, 1989) but before the initiation of transcription (Laybourn and Dahmus, 1990). Thus, the shift from form IIA to II may be involved in the transition between initiation and elongation.

We have reported the isolation of a CTD kinase from mouse ascites cells that contains two subunits (58 and 34 kDa), the smaller of which is a mouse p34cdc2/CDC28 homolog and the larger (58 kDa) is as yet uncharacterized (Cisek and Corden, 1989). More recently, we have identified a second mouse CTD kinase that contains, in addition to the p34cdc2/CDC28 subunit, a 62-kDa subunit that is recognized by anti-cyclin B antibodies (Cisek and Corden, 1990). This second enzyme elutes first from the DEAE column that fractionates the two activities, and we have therefore designated it CTD kinase E1. The previously identified enzyme elutes second from DEAE and has been designated CTD kinase E2. Preliminary experiments have indicated that these two mouse CTD kinases have identical substrate specificities; in particular, their activities on CTD substrates are indistinguishable. Because the E2 enzyme is available in a more pure form and is more stable we have concentrated our studies on this enzyme.

We show here that mouse CTD kinase E2 is able to phosphorylate a complete mouse CTD substrate. The end product of this phosphorylation reaction has an electrophoretic mobility in SDS gels similar to the CTD excised from the in vivo phosphorylated IIa subunit. Phosphoserine and phosphothreonine are detected in the in vivo phosphorylated CTD in the same ratio as seen in the in vivo phosphorylated CTD. Finally, we show that the target site of the mouse CTD kinase E2 is serine or threonine followed by a proline residue. This observation is consistent with the known specificity of p34cdc2/CDC28 kinases (Moreno and Nurse, 1990). In the accompanying paper we show that CTD phosphorylation produces a conformational change that greatly extends the CTD.

EXPERIMENTAL PROCEDURES

Polyacrylamide Gel Electrophoresis and Immunoblots—Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out according to the method of Laemmli (1970) with a modification by Dreyfuss et al. (1984). Electrophoretic transfer of proteins on SDS gels to nitrocellulose filter (Schleicher & Schuell) was accomplished by the procedure of Towbin et al. (1979) as modified by Gerace and Schiafino. The abbreviations used are: CTD, carboxyl-terminal domain; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; CNBr, cyanogen bromide; HPLC, high performance liquid chromatography.

* 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.
‡ The abbreviations used are: CTD, carboxyl-terminal domain; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; CNBr, cyanogen bromide; HPLC, high performance liquid chromatography.

2 p34cdc2/CDC28 is the catalytic subunit of a class of cell-cycle regulated kinases. This 34-kDa subunit is encoded by the cd2" gene in Schizosaccharomyces pombe, and by CDC28 in Saccharomyces cerevisiae. See Nurse, 1990 for review.
et al. (1982). We followed the immunoblot procedure by Snow et al. (1987). The blotted nitrocellulose filter was initially incubated for 20 min at room temperature in 0.5% (v/v) Triton X-100, 2% bovine serum albumin (Sigma) in phosphate-buffered saline (Triton/bovine serum albumin/PBS). The blots were sequentially incubated in each of the buffer solutions with 5 µg/ml of SWG16, a monoclonal antibody (IgG) directed against the heptapeptide repeats in CTD (Thompson et al., 1989; 2 µg/ml goat anti-mouse IgG (Cappel, Malvern, PA), and finally against 10^6 cpm/ml [γ-32P]ATP (Du Pont-New England Nuclear). Between each of the incubations, blots were washed four times with 10 min with 0.5% Triton X-100 in PBS. The immunoblots were briefly air-dried and exposed to x-ray film.

**Construction of CTDa Expression Plasmid—**Cloning procedures were performed according to Maniatis et al. (1982). The HpaII-HindIII DNA fragment encompassing 28% of the largest subunit gene of mouse RNA polymerase II (Ahearn et al., 1987), was filled in to blunt the HpaII-recessed end. This fragment was then cloned into the filled EcoRI site of pGEM-2 vector (Stratagene, La Jolla, CA) by blunt-end ligation to generate pGEM-CTDa. The EcoRI DNA fragment coding for the CTD was then cut out with EcoRI and subcloned into the EcoRI site of pBR-1 (Romm and Pollard, 1988) to generate pRX-CTDa (Fig. 1a). An mRNA whose expression is under the control of the trpE promoter was produced after transforming pRX-CTDa into Escherichia coli strain CAG-456 (inc am amp thm supC rpsL* lacI*) (Baker et al., 1984) by CaCl2 transformation. Translation of this mRNA results in 30 amino acid trp E peptide that terminates at a stop codon in the 5' end of the CTD fragment. Translation apparently reinitiates at a methionine residue immediately preceding the heptapeptide repeats, generating a protein containing the entire 52 heptapeptide repeats of the CTD (Fig. 1b).

**Expression and Purification of CTDa—**Growth and induction of CAG-456 cells containing pRX-CTDa plasmid were according to the procedure of Rimm and Pollard (1988) with the following modifications. Briefly, freshly transformed CAG-456 cells (slow in growth) were grown to stationary phase overnight at 30°C in 0.5 liter of LB CA minimal medium (Maniatis et al., 1982) at a methionine residue immediately preceding the heptapeptide repeats.

At the end of growth and induction described above, the cells were harvested. The rest of the purification steps, except those that were carried out at 4°C. The pellets were transfixed to 50 ml of lysis buffer consisting of 10 mM Tris-HCl, pH 8.0, 5 mM EDTA, 8% (v/v) sucrose, 0.5% (v/v) Triton X-100 and 0.1 mM of phenylmethylsulfonyl fluoride, and 2 µg/ml each of chymostatin, leupeptin, pepstatin, and trypsin inhibitor (all from Sigma). The cells were lysed by incubating on ice for 15 min followed by sonication for 3 min (Sonifier 450, VWR Scientific, Philadelphia, PA). The sonicated cell extract was centrifuged at 22,500 x g for 30 min at 4°C (SS34 rotor) to remove insoluble material. Into the supernatant, 12.5 ml of saturated ammonium sulfate (NH4)2SO4 solution was added dropwise to a final concentration of 20%. Protein fractionation was achieved by collecting the pellet of the (NH4)2SO4 mixture at 22,350 x g for 25 min. The CTD-containing precipitate was resuspended into 1.6 ml of 0.2% SDS, 50 mM HCl, and transferred to a 4-ml glass vial (VWR Scientific) for cyanogen bromide (CNBr) cleavage (Cadenas and Dahmus, 1987; Jay, 1984). Solid CNBr (0.2 g, Pierce Chemical Co.) was added to the vial, and the sample was slowly vortexed at room temperature for 16 h. The CNBr cleavage reaction was terminated by the addition of 0.2 ml of 1 M HCl. This CNBr-treated mixture was spun at 5,000 x g for 10 min (Microspin 245, Sorvall Instruments) and the CNBr supernatant was immediately chromatographed on two Superose 12 HR 10/30 columns in series at a rate of 0.3 ml/min with 100 mM ammonium bicarbonate (NH4HCO3, pH 7.7) as elution buffer. Aliquots of each Superose 12 fraction and of the previous purification steps were assayed by immunoblot (see above) to determine the presence of the CTD (Fig. 2a and b). Positive fractions were pooled and further purified by reverse-phase chromatography on a C4 column (4.6 x 220 mm, 10 µm, Pheric Chemical Co.). A 30-min multiphasic gradient using 0.06% trifluoroacetic acid as buffer A and 0.06% trifluoroacetic acid, 80% acetonitrile as buffer B was conducted at a rate of 0.6 ml/min and an oven temperature of 40°C while monitoring absorbance at 274 nm. The gradient is started with 99.8% buffer A, 0.2% buffer B followed by a step increase to 50% buffer B at 5 min, a linear increase of 3% buffer B/min from 5 to 15 min, 20% buffer B/min from 15 to 16 min, 100% buffer B from 16 to 20 min, and 99.8% buffer A, 0.2% buffer B at 20 min. The concentration of the CTD in the 10% of the hydrolysate was performed on PICO-TAG amino acid analysis system (Millipore, Waters, Taunton, MA).

**In Vitro Phosphorylation of CTDa—**A typical 10-µl radioactive labeling reaction consisted of 60 mM KCl, 50 mM Tris-HCl, pH 7.8, 10 mM CaCl2, 0.5 mM ATP, 1.6 µg/ml capillary peptide (Sigma), 50 µCi/mmol, Du Pont-New England Nuclear), 50 µg of CTKase E2 complex (purified by Dr. L. Cisek), and 2 pmol of purified CTD. The reaction mixture was incubated at 30°C for 4 h and stopped by 10 µl of 2 × SDS sample buffer containing 20 mM EDTA. Samples were boiled for 3 min and loaded onto SDS gels. After SDS-PAGE separation, gels were dried on two sheets of 3MM paper (Whatman, Maidstone, United Kingdom) and exposed to Kodak x-ray film. The intensities of bands on the autoradiogram were quantitated using LKB Ultrascan XL laser densitometer.

**Dephosphorylation of Phosphoamino Acid Content—**Synthetic heptapeptide substrates were synthesized by an Applied Biosystems model 430A peptide synthesizer using t-Boc chemistry. Peptide concentration was determined by A216, as described above for the CTD. Radioactive labeling of synthetic peptides was essentially the same as described above for CTD phosphorylation, except that the specific activity of [γ-32P]ATP was 10 Ci/mmol with a final concentration of 4 mM and a peptide concentration of 1 mM. To determine the phosphoamino acid content, either one- or two-dimensional paper electrophoresis was carried out according to Bylund and Huang (1976). An aliquot of [32P]-labeled peptide or CTDa, was lyophilized, rehydrated in 20 µl of 6 M HCl in an Eppendorf tube and hydrolyzed at 110°C for 3 h. The hydrolysate was dried under vacuum over NaOH and dissolved in 5 µl of pH 1.9 buffer (1:10:89 of formic acid/ acetic acid/water) containing 1 µg each of phosphoserine, phosphothreonine, and phosphotyrosine (all from Sigma). 2 µl of sample was applied to a 20 × 20-cm thin layer cellulose plate (Kodak) and electrophoresed according to Hunter and Sefton (1979). Some samples were electrophoresed on a second dimension in pH 3.5 buffer (1:10:189 of pyridine/acetic acid/water). The air-dried plates were sprayed with 0.2% (w/v) ninhydrin in ethanol and exposed to x-ray film.

**Phosphopeptide Sequencing—**Synthetic peptide T (see Fig. 5a for sequence) was 32P-labeled as described above. After labeling, 0.5 µg of sequencing grade trypsin (Boehringer Mannheim) was added for a 4-h 37°C digestion. The trypsin-cleaved [32P]phosphopeptide T was separated from free [γ-32P]ATP over a 5 × 1000-mm Bio-Gel P2 column with 1 M HAc as running buffer. The phosphopeptide T fragment was dried and rehydrated in 20 µl of 6 M HCl and hydrolyzed at 110°C for three times. Equal amounts of [32P]phosphopeptide (about 200 cpm/pmole, 400 pmol) were loaded onto each of six precut sequencing filters. Sequencing was performed according to Wang et al. (1988) on an Applied Biosystems model 475A sequenator employing a proline program at cycles 3 and 6. Phenylthiohydantoin amino acids were detected with a model 120A phenylthiohydantoin analyzer. After each cycle, the sequenator was interrupted, a piece of filter was taken out, a piece of blank filter was replaced, and the sequencing was resumed. Free phosphates generated during sequencing and the phosphopeptide T fragment were eluted from each piece of filter and quantitated by immersion of the filter in 200 µl of deionized distilled water and gently vortexing for 3 h. Approximately equal counts of extract for each cycle were loaded onto a thin layer cellulose plate and electrophoresed at pH 1.9 as described above with [32P]orthophosphoric acid (Du Pont-New England Nuclear) as a standard marker. After autoradiography, the free phosphate and phosphopeptide areas on the plate were cut out using the autoradigram as a template. Samples were counted in a scintillation counter to quantitate the percentage of free phosphate released at each cycle.
besides the reinitiation methionine just prior to the repetitive structure (see Fig. 1b). This property has been exploited in the purification of the repetitive domain by using CNBr cleavage, which degrades any other methionine-containing proteins in the extract to smaller peptide fragments. Low recovery of the CTD from the CNBr reaction due to poor solubility is compensated by the substantial purification achieved when the supernatant of the CNBr cleavage reaction is fractionated by Superose 12 gel filtration. The CTDs with a calculated molecular mass of 39,883 is well separated from most of the CNBr-degradation products that are in the included volume during gel filtration (Fig. 2b). In the final purification step, the CTDs is purified to homogeneity by HPLC reverse-phase chromatography (Fig. 2c).

The purity of the CTD is based on several lines of evidence. First, CTDs is eluted as a single peak on the C8 column with an absorbance maximum of 274 nm, as expected for a protein that contains tyrosine as the only aromatic amino acid (Fig. 2c). Second, the material in the C8 peak has an observed amino acid composition quite close to the calculated value for the mouse CTD (Table I) with a standard deviation, \( \sigma^2 \) value\(^4\) of 0.61 for all of the 15 amino acids present in CTD. Finally, conventional silver staining (Morrissey, 1981) reveals no bands when several \( \mu \)g of pure CTD is electrophoresed on a 12.5% SDS gel (data not shown). Failure of the CTD to stain is likely due to its unusual amino acid composition. The failure to detect contaminant bands indicates that the CTD is highly pure.

Purified mouse CTDs has an aberrant mobility on SDS-PAGE, probably due to its high content of proline. Furthermore, CTDs migrates at a different position compared with standard \( M_r \) markers on a different percentage SDS gels. The apparent \( M_r \) of 95 kDa on a 10% SDS gel for mouse CTDs is in close agreement with the mobility of the HeLa cell and calf-thymus CTDs excised from RNA polymerase IIa by CNBr cleavage (Cadena and Dahmus, 1987).

**Phosphorylation of CTDs by CTD Kinase E2 Causes a Shift to CTDs—Incubation of CTD kinase E2 with CTDs in the presence of [\( \gamma^{32}P \)]ATP results in a phosphorylated form of CTD that migrates at a position of 95 kDa on a 10% SDS gel (Fig. 3). The mobility shift from unphosphorylated CTDs to phosphorylated CTDs (CTD\(_{\text{p}}\)) is similar to the mobility difference between IIa and IIo forms of RNA polymerase II. Cadena and Dahmus (1987) have reported that CNBr cleavage of in vitro \( ^{32}P \)-labeled calf-thymus RNA polymerase IIo or in vivo \( ^{32}P \)-labeled HeLa cell RNA polymerase IIo releases a phosphopeptide ranging from 75 to 90 kDa. The single band at 95 kDa we observe here is probably the completely phosphorylated form of the CTD. When Ltk\(^{+}\) cells are \( ^{32}P \)-labeled in vivo, SDS-PAGE of \( ^{32}P \)-CTD purified with carrier S9 cells infected by a polyhedrin-CTD recombinant baculovirus (see under "Experimental Procedures") shows an apparent \( M_r \) of 95 kDa band on 10% gel (data not shown). Taken together, these results indicate that CTD kinase E2 produces the same shift in CTD mobility seen in vivo. This point is addressed in more detail in the accompanying paper.

**Phosphoamino Acid Contents of in Vitro and in Vivo Labeled CTD—As a first step in assessing the specificity of in vitro phosphorylation by mouse CTD kinase E2, we determined the phosphoamino acid content of the CTD labeled in vivo or in vitro. Two-dimensional paper electrophoresis was performed to determine if tyrosine, serine, or threonine are phosphorylated in CTD after in vivo labeling of mouse tissue culture cells or in vitro \( ^{32}P \)-labeling with CTD kinase E2. The

---

\(^4\) \( \sigma^2 = \frac{\sum(E_i - \bar{E})^2}{n}, \) where \( n = 15; \bar{E} \) and \( O \) are the expected and the observed percentage of amino acid \( i \) in the CTD.
FIG. 2. Purification of the CTD. a, samples of early purification steps were electrophoresed on a 10% SDS gel and transferred to nitrocellulose for immunoblotting by a monoclonal antibody, 8WG16. The autoradiogram of the immunoblot is shown here with $M_r$ values indicated in kDa at the left and the position for CTD indicated at the right. Lane 1 contained cell lysate (50 µl), lane 2 lysate supernatant (50 µl), lane 3 CNBr-cleavage mixture (5 µl), and lane 4 CNBr-cleavage supernatant (5 µl). b, chromatography profile of fast protein liquid chromatography Superose 12 fractionation is shown by the absorbance at 280 nm as a function of fraction number. Two ml of the 2-mercaptoethanol neutralized CNBr-cleavage supernatant was loaded on two Superose 12 columns in series and eluted with 100 mM NH$_4$HCO$_3$, pH 7.7. 30 µl of each of the 1.5-ml fractions were electrophoresed on 10% SDS gel for immunoblotting by 8WG16 antibody. The immunoblot autoradiogram of CTD-containing fractions is shown in the inset. c, chromatography profile of the HPLC reverse-phase C$_{18}$ column is shown by the absorbance of 274 nm as a function of elution time. The CTD fraction from Superose 12 chromatography was loaded on a C$_{18}$ column (4.6 × 230 mm) pre-equilibrated with 0.06% trifluoroacetic acid and the proteins were eluted with a trifluoroacetic acid/acetonitrile gradient as described under “Experimental Procedures.” The ultraviolet absorption spectrum of the major peak (12.3 min) is shown in the inset.
TABLE I
Amino acid composition analysis of CTDa

| Amino acid | Expected | Observed |
|------------|----------|----------|
| Ser        | 33.9     | 33.2     |
| Pro        | 27.6     | 26.3     |
| Thr        | 15.8     | 16.4     |
| Tyr        | 13.7     | 11.9     |

content of phosphoserine and phosphothreonine is quite similar for in vivo and in vitro phosphorylated CTD (compare Fig. 4, a and b). Neither in vivo nor in vitro phosphorylated CTD has any detectable amount of phosphotyrosine; the autoradiogram of the in vivo labeled sample has been intentionally over-exposed to demonstrate the lack of phosphotyrosine. In both in vitro and in vivo cases, serine is the major phosphorylation site, while threonine a minor one. Dahmus (1981) reported a similar pattern of phosphoamino acid composition for in vivo labeled HeLa RNA polymerase II.

Phosphorylation of Non-consensus Heptapeptides at Positions 2 and 5—It has been shown that CTD kinase E2 phosphorolyses only serines in a 28-amino acid synthetic peptide with consensus repeats, (Ser5-Pro3-Thr4-Ser5-Pro6-Ser7- Tyr7)7 (Cisek and Corden, 1989). Threonine in this consensus peptide is not phosphorylated. Therefore, phosphothreonines in the phosphorylated CTD of mouse are least likely derived from threonine at position 4. The mouse CTD contains several non-consensus heptapeptide repeats in which threonine replaces serine in positions 2 or 5. To determine which threonines (positions 2, 5, or 7) can act as phosphate acceptor, a series of 28 amino acid peptides were synthesized, all of which correspond to non-consensus variants occurring naturally in the mouse CTD, especially at its carboxyl terminus (Fig. 5a). These synthetic substrates were phosphorylated with CTD kinase E2 in the presence of [γ-32P]ATP. The 32P-labeled phosphopeptides were then subjected to acid hydrolysis and phosphoserine and phosphothreonine were separated by one-dimensional paper electrophoresis at pH 1.9. As shown in Fig. 5b, there is no detectable phosphothreonine in phosphopeptide T even after prolonged time of exposure, indicating that...
the non-consensus threonine at position 7 is not phosphorylated. All of the rest of the phosphopeptides shown in Fig. 5 contain phosphothreonine. Phosphothreonine in peptides K and V can only derive from the threonine at position 5 since the possibility of phosphorylation of threonine 7 has been ruled out. For the same reason, phosphothreoninines in phosphopeptides N and R are derived from threonine at position 2. The weak phosphorylation of threonine in phosphopeptide K is probably due to its location in the last repeat of the peptide making it a less favorable phosphorylation site. Taken together, these results are consistent with phosphorylation at both the second and fifth positions in the heptapeptide repeat.

-Ser(Thr)-Pro- Is the Recognition Site for CTD Kinase E2—

The above phosphoamino acid composition analysis demonstrates that threonines at positions 2 and 5, but not 7, can be phosphorylated, strongly suggesting that serines 2 and 5, but not 7 are the phosphorylation sites for CTD kinase E2. In order to directly confirm this conjecture, [32P]phosphopeptide T has been directly sequenced using a method developed by Wang et al. (1988). Because the first and last repeats in peptide T tend to be poorly phosphorylated (see for example phosphopeptide K in Fig. 5b), the [32P]-labeled peptide T was first cut with trypsin to generate two phosphopeptides with exactly the same sequence. Trypsin cleavage exposes the third repeat in peptide T directly at the amino terminus for sequencing. At each of the six cycles of sequencing, one piece of precut sample filter was taken out of the sequenator. The phosphate and phosphopeptides on the filter were extracted and separated by electrophoresis. Fig. 6 shows the percentage of released phosphate at each sequencing cycle. Prior to sequencing and after the first cycle there is only a background of nonspecific phosphate release. Edman degradation at the second position serine causes a release of 9.4% of the total phosphate as compared with a 1.6% background, indicating that this serine is phosphorylated. The minor increase of phosphate released at cycles three and four (to 11.2 and 11.7%, respectively) is mainly due to carry-over since neither phosphoserine nor proline are cleaved completely. The lack of major phosphate release at cycle four is consistent with the earlier result that no threonine at position 4 is phosphorylated. The next significant increase of phosphate release, from 11.7% at cycle four to 20.4% at cycle five, comes when serine at position 5 is degraded, indicating that serine 5 is another phosphorylation site for CTD kinase E2. These results directly demonstrate that positions 2 and 5 are the phosphorylation sites for CTD kinase E2.

-Ser-Pro-Lys(Arg)- Is a Better Substrate for Phosphorylation Than -Ser-Pro-Ser—

The identification of serines at positions 2 and 5 as the recognition sites of CTD kinase E2 is consistent with the known preference of \( \text{p34}^{cdk2/\text{CDC28}} \) kinases for this sequence (Moreno and Nurse, 1990). To see if residues other than the dipeptide -Ser-Pro- are important in the recognition of the heptapeptide by CTD kinase E2, we compared the different peptides described in Fig. 5 for their abilities to compete with the full-length CTDa in mixed substrate reactions. It is clear from the results shown in Fig. 7 that the basic peptides (R and K) compete more effectively than the consensus (peptide S). Peptide N does not seem to compete in this reaction. These results indicate that among the heptapeptides present in the CTD, mouse CTD kinase E2 prefers the nine that contain non-consensus basic residues. Such a preference is consistent with the observation that several of the known \( \text{p34}^{cdk2/\text{CDC28}} \) recognition sites contain basic residues (Moreno and Nurse, 1990).

**DISCUSSION**

We have previously reported the identification and purification of two protein kinases that phosphorylate CTD peptides in vitro (Cisek and Corden, 1989, 1990). In this study we show that a mouse CTD kinase (E2) phosphorylates a complete CTD substrate resulting in a shift in electrophoretic mobility of the CTD characteristic of the in vivo shift from \( \text{II}_a \) to \( \text{II}_b \). This result not only supports the argument that CTD kinase E2 participates in generating \( \text{II}_b \) in vivo, but also demonstrates that the generation of form \( \text{II}_a \), which has been used by others (Guilloy, 1989; Lee and Greenleaf, 1989; Payne et al., 1989) to identify and purify CTD kinases, is also a property of mouse CTD kinase E2. In the accompanying paper we show that the phosphorylation of the mouse CTD by CTD kinase E2 is associated with a profound conformational change.

One criterion for judging the in vivo relevance of a protein kinase purified by an in vitro phosphorylation assay is to show that the same substrate residues are phosphorylated in vivo and in vitro. As a first step in this analysis we have shown here that the ratio of phosphoserine to phosphothreonine in the in vivo and in vitro labeled CTD is similar (Fig. 4). We have identified the sites of phosphorylation in vitro both by examination of the phosphoamino acids in labeled non-consensus peptides, and through direct sequencing of a labeled phosphopeptide. These results demonstrate that mouse CTD kinase E2 phosphorylates positions 2 and 5 in the consensus sequence.
sequence Tyr^5-Ser^7-Pro^7-Thr^7-Ser^7-Pro^7-Ser^7. While we have not sequenced the in vivo labeled CTD, several pieces of evidence suggest that positions 2 and 5 are phosphorylated in vivo. The presence of threonine at positions 2 or 5 in some of the 52 repeats in the mouse CTD (Corden et al., 1985) leads to the prediction that the ratio of phosphoserine to phosphothreonine should be ~10:1 (there are 94 serines and 9 threonines at positions 2 and 5). This approximate ratio observed both in the in vivo and in vitro labeled CTD is consistent with the idea that positions 2 and 5 are phosphorylated in vivo.

A second argument that mouse CTD kinase E2 phosphorylates the CTD in vivo comes from comparing the recognition sites of other p34^cd2/CDC28-containing kinases: a serine (or threonine) residue followed by a proline residue. -Ser(Thr)-Pro- sites in histone H1, laminin, and several other substrates have been clearly demonstrated to be phosphorylated by p34^cd2/CDC28 kinases (Moreno and Nurse, 1990). The similarity of the CTD recognition site, and the identification of p34^cd2/CDC28 as a component of CTD kinase E2 in mouse strongly support the idea that CTD kinase E2 phosphorylates the CTD in vivo.

The mouse CTD kinase E2 used in these studies consists of p34^cd2/CDC28 complexes with a 58-kDa subunit. Recently, we (Cisek and Corden, 1990) have described a second CTD kinase E1 activity from mouse cells. This enzyme contains p34^cd2/CDC28 complexes with cyclin B to form a complex similar to M-phase-promoting factor (Nurse, 1990). Thus, in mouse cells there are at least two activities that may phosphorylate the CTD. Lee and Greenleaf have described a yeast CTD kinase, designated CTK1, that contains subunits of 58, 38, and 32 kDa. The 58 kDa subunit contains homologies to the p34^cd2/CDC28 family of kinase catalytic subunits while the 32 and 38 kDa subunits are not recognized by anti-CDC28 antibodies. This result indicates that, at least in yeast, another class of CTD kinases exists. This apparent multiplicity of CTD kinases is consistent with the idea that CTD kinase E2 may phosphorylate the CTD in vivo.

One interesting property of p34^cd2/CDC28 kinases is their variation in activity during the cell cycle (Nurse, 1990). Resulting changes in the phosphorylation state of p34^cd2/CDC28 substrates are thought to be important for the transition between G2 and M-phase and between G1 and S-phase. How phosphorylation of RNA polymerase II might affect transcription and how such transcriptional regulation might contribute to cell-cycle transitions is currently under study.

Acknowledgments—We thank Dr. David Rimro for the expression vector pHX-1, Dr. Murthy Shenbagamurthi for sequencing [32P]phosphopeptide, Dr. Clark Riley and Van Vogel for peptide synthesis, Cathy Cullen for subcloning in the initial stage of the project, Dr. Lars Cisek for CTD kinase E2, Dr. Nancy Thompson, University of Wisconsin-Madison, for monoclonal antibody, 8WG16, and Dr. Michael Dahmus (University of California at Davis), Dr. Richard Young (Massachusetts Institute of Technology), and Dr. Arno Greenleaf (Duke University) for communicating results prior to publication.

REFERENCES

Ahearn, J. M., Jr., Bartolomei, M. S., West, M. L., Cisek, L. J., and Corden, J. L. (1987) J. Biol. Chem. 262, 10695-10705
Allison, L. A., Wong, J. K.-C., Fitzpatrick, V. D., Moyle, M., Shales, M., and Ingles, C. J. (1988) J. Mol. Cell. Biol. 8, 321-329
Baker, T. A., Grossman, A. D., and Gross, C. A. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 6779-6783
Bartolomei, M. S., Halden, N. F., Cullen, C. R., and Corden, J. L. (1988) J. Mol. Cell. Biol. 8, 330-339
Bylund, D. B., and Huang, T.-S. (1976) Anal. Biochem. 73, 477-485
Cadena, D. L., and Dahmus, M. E. (1987) J. Biol. Chem. 262, 12468-12474
Cisek, L. J., and Corden, J. L. (1989) Nature 339, 679-684
Cisek, L. J., and Corden, J. L. (1991) Methods Enzymol., in press
Corden, J. L. (1990) Trends Biochem. Sci. 15, 383-387
Corden, J. L., Cadena, D. L., Ahearn, J. M., and Dahmus, M. E. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 7924-7938
Dahmus, M. E. (1981) J. Biol. Chem. 256, 3319-3325
Dreyfuss, G., Adam, S. A., and Choi, Y. D. (1984) Mol. Cell. Biol. 4, 415-423
Gerace, L., Ottaviano, Y., and Kondor-Koch, C. (1982) J. Cell Biol. 95, 826-837
Guilfoyle, T. J. (1989) Plant Cell 1, 827-836
Hunter, T., and Selton, B. M. (1979) Proc. Natl. Acad. Sci. U. S. A. 77, 1311-1315
Jay, D. G. (1984) J. Biol. Chem. 259, 15572-15578
Kim, W.-Y., and Dahmus, M. E. (1986) J. Biol. Chem. 261, 14219-14225
Kim, W.-Y., and Dahmus, M. E. (1989) J. Biol. Chem. 264, 3169-3176
Kolodziej, P., Woychik, N., Liao, S.-M., and Young, R. A. (1990) Mol. Cell. Biol. 10, 1915-1929
Laemmli, U. K. (1970) Nature 227, 680-685
Laybourn, P. J., and Dahmus, M. E. (1990) J. Biol. Chem. 265, 13165-13173
Lee, J. M., and Greenleaf, A. L. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 3264-3268
Maniatis, T., Fristach, E. F., and Sambrook, J. (1984) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Moreno, S., and Nurse, P. (1990) Cell 61, 549-551
Morrissey, J. H. (1981) Anal. Biochem. 117, 307-310
Nonet, M., Sweetser, D., and Young, R. A. (1987) Cell 50, 909-915
Nurse, P. (1990) Nature 344, 503-508
Payne, J. M., Laybourn, P. J., and Dahmus, M. E. (1989) J. Biol. Chem. 264, 19621-19629
Rimm, D. L., and Pollard, T. D. (1988) Gene (Amst.) 75, 323-327
Sawadogo, M., and Sentenac, A. (1990) Annu. Rev. Biochem. 59, 711-754
Snow, C. M., Senior, A., and Gerace, L. (1987) J. Cell Biol. 104, 143-156
Thompson, N. E., Steinberg, T. H., Aronson, D. B., and Burgess, R. R. (1989) J. Biol. Chem. 264, 11511-11520
Towbin, H., Staehelin, T., and Gordon, J. (1979) Proc. Natl. Acad. Sci. U. S. A. 76, 4350-4354
Wang, Y., Fiol, C. J., DePaoli-Roach, A. A., Bell, A. W., Hermodson, M. A., and Roach, P. J. (1988) Anal. Biochem. 174, 537-547
Woychik, N., and Young, R. A. (1990) Trends Biochem. Sci. 15, 347-351
Zehring, W. A., Lee, J. M., Weeks, J. R., Jokser, R. S., and Greenleaf, A. C. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3698-3702

A. Greenleaf, personal communication.