Requirements of the RNA Polymerase II C-Terminal Domain for Reconstituting Pre-mRNA 3' Cleavage

Kevin Ryan, Kanneganti G. K. Murthy,[†] Syuzo Kaneko, and James L. Manley*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 31 October 2001/Returned for modification 12 December 2001/Accepted 18 December 2001

RNA polymerase II (RNAP II) has previously been shown to be required for the pre-mRNA polyadenylation cleavage reaction in vitro. This activity was found to reside solely in the C-terminal domain (CTD) of the enzyme's largest subunit. Using a deletion analysis of glutathione S-transferase-CTD fusion proteins, we searched among the CTD's 52 imperfectly repetitive heptapeptides for the minimal subset that possesses this property. We found that heptads in the vicinity of 30 to 37 contribute modestly more than other sections, but that no specific subsection of the CTD is necessary or sufficient for cleavage. To investigate further the heptad requirements for cleavage, we constructed a series of all-consensus CTDs having 13, 26, 39, and 52 YSPTSPS repeats. We found that the nonconsensus CTD heptads are together responsible for only 20% of the wild-type cleavage activity. Analysis of the all-consensus CTD series revealed that the remaining 80% of the CTD-dependent cleavage activity directly correlates with CTD length, with significant activity requiring ≈ 26 or more repeats. These results are consistent with a scaffolding role for the RNAP II CTD in the pre-mRNA cleavage reaction.

Polyadenylation of nearly all eukaryotic pre-mRNAs is an important step in the flow of genetic information from DNA to functional protein (8, 26, 37, 43). During mRNA transcription, RNA polymerase II (RNAP II) does not stop at the end of a gene's last exon. Rather, it transcribes well beyond the stop codon and poly(A) site. Although on the one hand this seems wasteful and unnecessary, it appears on the other to be required for the orderly termination of transcription and the precise definition of the 3' end of the mRNA. The overall process of 3'-end formation occurs in two discrete steps that are tightly coupled in vivo but which can be uncoupled in vitro. First, the pre-mRNA undergoes endonucleolytic cleavage 10 to 30 nucleotides downstream of the AAUAAA site in the 3' untranslated region. The scissile phosphate departs with the 5' end of the downstream fragment, leaving the 3' OH of the upstream fragment free to be extended by poly(A) polymerase (PAP) in the second step. The poly(A) tail, which ultimately reaches a length of around 250 nucleotides in mammals, functions in the regulation of translation and mRNA stability (23, 27).

Two monofunctional polypeptides could, in principle, be sufficient to carry out the basic enzymatic activities of cleavage and poly(A) synthesis. However, it is now well known that the process of 3'-end formation, in yeast as well as in higher eukaryotes, requires at least 12, undoubtedly reflecting the importance and complexity of its regulation (Fig. 1A) (34). Essentially all of the factors required for the process have been identified and cloned. Two multisubunit proteins, cleavage polyadenylation specificity factor (CPSF), consisting of four subunits, and cleavage stimulation factor (CstF), consisting

of three subunits, serve to recognize and define the region surrounding the poly(A) site by binding cooperatively to this region. The CPSF-160 subunit binds specifically to the AAUAAA sequence just upstream of the poly(A) site, while CstF-64 binds a downstream G/U-rich sequence. PAP is also required for the cleavage reaction for most but not all substrates studied thus far and, in cooperation with CPSF, it synthesizes the poly(A) tail. Two other factors, CF I (two subunits) (33) and CF II (two or more subunits) (13) are also required for the cleavage reaction and likely function in pre-cleavage complex assembly and stability. Despite years of detailed study into the function of the various factors and their subunits, the endonuclease itself has not been identified.

Our lab recently made the surprising discovery that RNAP II itself, and in particular the C-terminal domain (CTD) of the largest subunit, can be an essential factor in the reconstituted in vitro cleavage reaction (17). Either RNAP II or a glutathione S-transferase (GST)-CTD fusion protein, which function equivalently when combined with CPSF, CstF, CF I, CF II, and recombinant PAP, was found sufficient to stimulate cleavage of substrate mRNA precursors. This cleavage activity was found to be independent of the CTD phosphorylation status and was detected in reaction mixtures lacking creatine phosphate, a low-molecular-weight compound that can also induce cleavage in vitro (16).

With the additional complexity of multiprotein processes comes the need for the organization and coordination of the various proteins. In recent years, the CTD has emerged as the factor most likely to play this role in pre-mRNA processing. The CTD has now been shown to be involved not only in cleavage and polyadenylation (17, 25) but also in initiation (20, 36), promoter clearance (11), 5'-end capping (7, 24, 40), and splicing (18, 25). It has become clear that mRNA synthesis and processing are coordinated events (a sort of mRNA factory [25]) and that the CTD is involved at each stage.

The RNAP II CTD is a highly repetitive domain unique

* Corresponding author. Mailing address: Department of Biological Sciences, Columbia University, 1117A Fairchild Building, New York, NY 10027. Phone: (212) 854-4647. Fax: (212) 865-8246. E-mail: jlm2@columbia.edu.

[†] Present address: Inotek Corporation, Beverly, MA 01915.

FIG. 1. Reconstituted 3' cleavage reaction activated by RNAP II CTD. (A) Schematic depiction of the hypothetical pre-mRNA in vitro cleavage complex. CPSF and CstF bind to the pre-mRNA as well as the RNAP II CTD, irrespective of CTD phosphorylation state. PAP, CFI, and CF II are also involved in the cleavage step. Arrow indicates

among the RNA polymerases (Fig. 1B) (4, 9). In mammals it consists of 52 heptapeptides having the consensus sequence YSPTSPS (Fig. 1B). Of the 52 heptads, only 21 match this sequence exactly, while the remaining 31, occurring mostly in the C-terminal half of the CTD, consist of the related sequence YX₂PX₄SPX₇, where X₂, X₄, and especially X₇ can vary considerably (Fig. 1). While the CTD is essential for viability in cultured rodent (5, 21) and yeast (1, 30, 38) cells, significant numbers of the heptads are dispensable. Studies of partially deleted and rearranged forms of the CTD have shown that a minimum of about 30 heptads are needed for viability in mouse cells, whereas mutants carrying between 31 and 39 heptads exhibit a slow-growth phenotype. These studies also concluded that CTD function depends on both its sequence and its length, and they exemplified the difficulty of separating these variables when assessing the effect of changes made to an imperfectly repetitive domain like the CTD.

The role of the CTD in pre-mRNA 3' cleavage seems likely to be related, at least in part, to the requirement for the CTD in proper transcription and viability. In order to understand which portions of the CTD are responsible for its role in cleavage and, ultimately, to understand how it brings about the cleavage reaction, we made fusion proteins between GST and a series of truncated and deleted mouse CTDs, as well as a series of all-consensus GST-CTD constructs ranging in size from 13 to 52 heptads. Using these with the other general polyadenylation factors, we then assessed the various mutant CTDs for their ability to activate the reconstituted 3' cleavage reaction. These studies have enabled us to evaluate the contributions of the nonconsensus CTD residues and the variable of CTD length—apart from any consideration of sequence—to CTD activity in the cleavage reaction.

MATERIALS AND METHODS

Protein purification. Nuclear extract was prepared from HeLa cells by the method of Dignam et al. (14) with minor modifications (35). A DEAE-Sepharose Fast Flow column (Amersham Pharmacia) was equilibrated with buffer A (20 mM Tris-HCl [pH 7.9], 0.2 mM EDTA, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, and 10% glycerol) containing 40 mM (NH₄)₂SO₄. Nuclear extract, dialyzed to 40 mM (NH₄)₂SO₄ in buffer A, was loaded onto the column. After washing with two column volumes of the same buffer, proteins were eluted with a linear gradient of (NH₄)₂SO₄ (from 40 to 250 mM). Flowthrough fractions contained CstF activity, whereas CF and CPSF activities were eluted at 80 to 100 mM and 150 to 200 mM (NH₄)₂SO₄, respectively. CstF activity in the flowthrough fraction was further purified by Superose 6, Mono Q, heparin-Sepharose, and Mono S. The following columns were run to purify CPSF activity: Superose 6, DEAE 5PW, and heparin-Sepharose. The DEAE 80-to-100 fraction containing CF activity was further purified by chromatography on columns of Superose 6 and Mono S. Mono S separates CF I and CF II activities, but the CF II prepared in this way lost activity upon storage at -80°C. The CF II used in the processing reactions was therefore not passed through the Mono S column, but the CF I was Mono S purified. The chromatographic columns were loaded and developed as described previously (35). Pol II O was isolated from HeLa cell nuclear pellets according to the method of Lu et al. (22).

cleavage site. (B) Primary structure of the mouse RNAP II CTD. (C) 3' cleavage of ³²P-labeled adenovirus L3 (Ad L3) pre-mRNA (lane 1) activated by partially purified general cleavage factors (lane 2), purified Pol II O (≈4 ng of the largest subunit; lane 3), and GST-CTD (10 ng). Background cleavage in lane 2 is attributable to traces of RNAP II in the cleavage factor preparations (see text and Materials and Methods). Each reaction mixture contained CPSF, CstF, recombinant PAP II, CF I, and CF II.

FIG. 2. C- and N-terminal GST-CTD deletion mutants used in this study. (A) Each unshaded box represents one consensus heptapeptide, YSPSTPS, while each shaded box represents one nonconsensus heptapeptide. In H(1-37) and H(1-35), the last residue of the last CTD heptapeptide is not present. (B) SDS-PAGE gel of the purified deletion mutants (150 ng), silver stained.

Quantitative Western blotting with the anti-CTD antibody 8WG16 (Babco) and the RNAP II O purified from HeLa cells (22) was used to estimate the amount of contaminating RNAP II largest subunit contributed by each of the factor preparations per 12.5- μ l reaction volume: CPSF, 0.16 ng; CF I, 0.02 ng; CF I/CF II mixture, 0.1 ng. The sum of these is approximately equal to the amount of exogenous RNAP II largest subunit needed to detectably increase the amount of cleavage over background (e.g., Fig. 1, lane 2). It is therefore reasonable to assume that the background cleavage in our experiments is due to these trace amounts of contaminating RNAP II largest subunit in our factor preparations.

C- and N-terminally deleted GST-CTD fusion proteins. The C- and N-terminal GST-CTD deletion mutants were prepared by PCR using the original GST-CTD plasmid, pGCTD, as the template (31). In the C-terminal deletion series (Fig. 2A), the forward primer was, in each case 5'-TTGTCTGGATCCGTAGG TGGTCTATGTC. The reverse primers (5' to 3') were H(1-48) GGTGGGAC TGGTGGGCGAG, H(1-44) GACTGGAGAGGTTGGTGAGTA, H(1-37) CTCTGGTGAGCTGGGACTGTA, H(1-35) GGGTGAAGTGGGGCTGTAG CT, H(1-29) CGGAATTCAGTGGGTGAGGTTGGGAATAG, and H(1-21) CGGAATTCAGTGGGAGAGTTGGTGAGTAG. This last primer was designed to prime at the 20th heptad repeat in order to produce the same number of heptads as H(33-52) but, due to the repetitive nature of the CTD cDNA, was found upon sequencing to have primed at the adjacent heptad 21 instead. In the N-terminal deletion series the reverse primer was, in each case, GGAATTCCT

CGCCCTGTTCGC. The forward primers for the N-terminal deletion series were H(24-52) CGGGATCCTATACCCCGACATCACCCAG, H(33-52) CGG GATCCTACACACCAAGCTCACCAAG, and H(41-52) CGGGATCCTATTC CTCTACCAGCCCCAC. PCRs were performed on 10 ng of template DNA in a 50- μ l volume using Vent DNA polymerase (New England Biolabs), the supplied buffer, and 0.5 μ M concentrations of each primer for 30 cycles of 95°C for 30 s, 62°C for 30 s, and 75°C for 66 s. The products in the C-terminal deletion series were digested with *Bam*HI, agarose gel purified, and ligated into the pGEX-2T plasmid (Amersham Pharmacia) after the plasmid had been cut with *Bam*HI and *Sma*I and agarose gel purified. In the case of the N-terminal deletion series, the PCR product from H(41-52) was digested with *Eco*RI and *Bam*HI and ligated into the *Eco*RI-*Bam*HI-digested pGEX-2T vector. The two longer N-terminal deletions were prepared differently. First, both the PCR products and pGCTD were separately digested with *Bam*HI and *Spe*I and gel purified. The PCR fragment was then ligated into the *Bam*HI-*Spe*I-digested pGCTD fragment. The ligation products were transformed into *Escherichia coli* DH5 α cells. H(1-29), H(1-21), H(24-52), H(33-52), and H(41-52) were sequenced entirely. The other clones were partially sequenced. The plasmids corresponding to the foregoing constructs are designated pH(N-C), where N is the number of the N-terminal heptad and C is the number of the C-terminal heptad. For expression, the strain JM101 was used and the proteins were purified on glutathione-Sepharose and Superose 12 chromatography columns as described elsewhere for GST-CTD (17). The final proteins were checked by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized with silver staining (Fig. 2).

The all-consensus CTD series of GST fusion proteins. The cDNAs for the fusion proteins were constructed as follows. First, the CON 13 cDNA was produced using PCR on pH(1-21) as the template instead of pGCTD. The following primers were used: forward 5'-pACATCTCCTAGCTACTCGCCAC CTC; reverse 5'-pGGGTGAGTAGCTGGGAGATGTCGGT. PCR was performed as described above except that the annealing temperature was 74°C. The resulting CON 13 DNA from 3 ml of PCR mixtures was combined, precipitated, and purified on a 1.5% agarose gel. In order to multimerize the CON 13, the DNA was ligated to itself (7 μ g in 53 μ l) using T4 DNA ligase (New England Biolabs) for 2 h at 22°C. The DNA was phenol-chloroform extracted, precipitated, and then double digested with *Rsa*I and *Sma*I (New England Biolabs). These restriction endonucleases catalyze the exact reverse of the blunt-end ligation reaction on all ligation products having either of the undesired orientations, thereby providing CON 13 with another chance to be ligated in the correct orientation. The ligation-restriction sequence was repeated two more times, and the DNA products were separated on a 1.2% agarose gel. The desired tandem dimer (CON 26), trimer (CON 39), and tetramer (CON 52) bands were cut from the gel and isolated using a Qiaquick gel extraction kit (Qiagen). These DNAs were then ligated into the *Sma*I-digested pBluescript II SK(+) plasmid (Stratagene). The presence of the correctly oriented fragments was detected by a *Sma*I/*Spe*I double digest. After sequencing, the cDNA coding for the CTD portion of the mutants was excised by digestion with *Bam*HI and *Sma*I and transferred to a modified version of plasmid pGEX-6P-1 (Amersham Pharmacia) that lacked the -CCGGGTC- sequence surrounding the vector's *Sma*I site. This modified vector was prepared for ligation by sequential treatment with the following enzymes: *Eco*RI, Klenow fragment fill-in, *Bam*HI, and calf intestinal phosphatase. All cloning and subcloning was done in *E. coli* DH5 α , while protein expression was done in JM101. The resulting series of fusion proteins coding for GST-[TSPSYSP] $_n$, where $n = 13, 26, 39$, or 52, was purified identically to GST-CTD and the deletion mutants described above.

The cDNAs for CON 13 and CON 26 were completely sequenced. That of CON 39 was partially sequenced (93%), while only 30% of CON 52 could be read. To verify these sequences, the GST tag of the expressed fusion protein was removed by the PreScission protease (Amersham Pharmacia), and matrix-assisted desorption/ionization-time of flight (MALDI-TOF) mass spectrometry was performed on a Perseptive Biosystems Voyager-DE Pro instrument. Values all fell within 0.16% of the calculated molecular weight, verifying that all multimerization junction ligations were in the correct orientation. All cDNAs were of the expected length, and Western blotting with the anti-RNAP II LS CTD antibody 8WG16 gave expected results on the entire series of fusion proteins (results not shown). The CTD is well known for its uncharacteristically slow mobility on SDS-PAGE (42). Substituting consensus for nonconsensus heptads exaggerates this trait.

Internal deletion mutant D30/37. PCR-mediated gene fusion (39) was used to delete heptads 30 through 37 from the wild-type GST-CTD. The portion upstream of the deleted segment was amplified by PCR using the above conditions (with 68°C annealing temperature) from pH(1-35) with the following primers: forward 5'-TTGTCTGGATCCGTAGGTTGGTGCTATGTC; reverse 5'-TTTGG

GAGAAGCTGGGGTGTAACTGGGTGAGGTTGGGGAATA. The portion downstream of the deleted segment was amplified from pH(33-52) using the following primers: forward 5'TATTCCCAACCTCACCCAGTTACACCCCAGCTTCTCCCAAA; reverse 5'GGAATTCTTCGCCCTGTTCGC. The PCR products were mixed, annealed, and amplified with the two extreme upstream and downstream primers, trimmed with *Bam*HI and *Eco*RI, gel purified, and cloned into the modified pGEX-6P-1 described above. The sequence was verified by partial sequencing (95%, including the new heptad 29-heptad 38 junction) and MALDI-TOF mass spectrometry (within 0.16%) after protease treatment to release the GST tag.

Mutant H(30-37). Heptads 30 to 37 were amplified from pH(24-52) using the following primers: forward 5'CACGCGGATCTATTCCCCACCTCACCAA GC; reverse 5'CCGGAATTCTCACTCTGGTGAGCTGGGACTGTA. The resulting fragment was trimmed with *Bam*HI and *Eco*RI, gel purified, cloned into the modified pGEX-6P-1 as described above, and completely sequenced. Where indicated, GST tag was cleaved as described above.

3' cleavage assays. The 5'-capped, internally ³²P-labeled pre-mRNA substrate used in this study was transcribed in vitro from the template plasmid pG3L3-A (35) after digestion with *Bam*HI. Cleavage assays were performed as previously described (17) except for the following differences: 2'-dATP was substituted for 3'-dATP, the final MgCl₂ concentration was 0.5 mM, and reactions were incubated for 2 h at 30°C instead of for 1.5 h. The extent of cleavage was quantitated using a Molecular Dynamics phosphorimager. The background-subtracted ratios of the 5' cleavage products to the uncleaved precursor RNA were calculated. Ratios for control reactions employing the cleavage factors alone (no GST-CTD or mutant CTD proteins) were set arbitrarily to one, and all others were reported relative to this value.

RESULTS

CTD-activated pre-mRNA cleavage is a function of both CTD length and sequence. To begin analysis of CTD requirements for 3' cleavage, we prepared from HeLa cells purified or partially purified CPSF, CstF, CF I, and CF II and from *E. coli* His-tagged recombinant PAP (see Materials and Methods). In the presence of creatine phosphate (CP), these factors are sufficient to bring about efficient 3' cleavage of numerous poly(A) signal-containing substrates, such as the model adenovirus L3 pre-mRNA (16, 35), but in the absence of CP, little if any cleavage is observed (e.g., Fig. 1C, lane 2 versus lane 1). However, cleavage can be restored under these conditions by addition of purified RNAP II (lane 3) or recombinant GST-CTD (lane 4). Our previous study (17) suggested that under these conditions the CTD was in fact necessary for cleavage. The inefficient processing observed here in the absence of RNAP II, CTD, or CP could reflect a low basal activity (lane 2). However, quantitative Western blotting of the factors used in these experiments with an anti-RNAP II largest subunit antibody indicated that the CPSF, CF I, and CF II fractions used contained low levels of RNAP II, and the amounts were consistent with the background activity observed (see Materials and Methods). Thus, our results are consistent with the view that the CTD is indeed required for cleavage in the reconstituted reactions.

To initiate our analysis of CTD requirements for 3' cleavage, we first constructed a series of mouse GST-CTD proteins lacking up to 31 C-terminal or 40 N-terminal CTD heptapeptides [Fig. 2; H(1-N) indicates a GST-CTD protein containing heptads 1-N, inclusive, except in H(1-37) and H(1-35), where the seventh residue in the last heptad is omitted]. Fusion proteins were expressed in *E. coli* and purified as described above (see Materials and Methods). The shortened forms of the CTD were then compared to full-length GST-CTD with respect to their ability to activate cleavage of the adenovirus L3 substrate with the same factors utilized in Fig. 1C. Two concentrations of

CTD fusion proteins were analyzed, one near the high end of the linear range as previously determined with full-length GST-CTD (17) and the other well above this. Figure 3A shows an example of the data obtained at the higher concentration tested.

The average cleavage efficiency observed in three to four experiments at both CTD concentrations is presented graphically in Fig. 3B. At the lower GST-CTD concentrations (10 ng of fusion protein per 12.5-μl reaction volume), progressive deletion of heptads from the C terminus resulted in small but progressive reductions in activity, indicating that heptads 36 to 52 are not critical for efficient cleavage. A more significant loss occurred when heptads 30 to 35 were removed from H(1-35) to produce H(1-29), possibly indicating that a critical length of more than 29 but fewer than 35 heptads is required to bring about efficient pre-mRNA cleavage by a C-terminal-deleted CTD, or that specific heptads in this region are important. Deletion of heptads from the N terminus at this lower concentration also showed a general correlation of CTD length with cleavage activity. In this series, however, the CTD could be shortened beyond 36 total heptads [e.g., H(33-52)] while still retaining significant, albeit reduced, cleavage activity. Overall, a similar trend was evident at the higher GST-CTD concentrations (100 ng of GST fusion protein per 12.5-μl reaction volume), with the exception that, in the C-terminal deletion series, no reduction in activity was detected until deletions extended beyond heptad 35. Taken together, these data indicate that a CTD longer than 29, and up to perhaps 34, heptads is necessary for efficient cleavage, and that heptads 1 to 23 and 36 to 52 are largely dispensable for this activity. They also suggest a general correlation of CTD length with cleavage activity, especially when the CTD contains 35 or fewer heptads.

Deletion analysis with the CTD can lead to ambiguous results because each deletion simultaneously changes the two potentially important variables of length and sequence. It is therefore important to control for length when searching for important sequences among the various CTD heptads. As shown in Fig. 2A, H(1-29) and H(24-52) both contain 29 heptads, while H(1-21) and H(33-52) differ by only one consensus repeat, having 21 and 20 heptads, respectively. When the data in Fig. 3B are interpreted among only those mutants having the same or nearly the same numbers of heptads, it becomes clear that heptads in the C-terminal half of the CTD contributed more to cleavage than did those in the N-terminal half, in the context of these shortened CTDs. More precisely, heptads 24 to 52 contributed more than 1 to 29, while heptads 33 to 52 contributed more than 1 to 21. Since heptads 36 to 52 were largely dispensable for cleavage in the context of the C-terminal deletion series, it appeared that heptads in the approximate region of 30 to 37 might have special significance for the cleavage reaction.

Heptads 30 to 37 are slightly more important than other heptads to CTD cleavage activity. To test directly the importance of heptads 30 to 37, these heptads were deleted using PCR-mediated gene fusion (Fig. 4A; see Materials and Methods). This deletion, D30/37, resulted in a CTD of 44 total heptads. The fusion protein H(1-44) also contains 44 heptads and can be used to control for length since, as described above, heptads in its deleted region (45 to 52) have little or no effect on CTD-activated cleavage. In addition, we made another fu-

FIG. 3. Reconstituted 3' cleavage reaction of adenovirus L3 RNA activated by C- and N-terminally truncated GST-CTD mutants. (A) Representative 3' cleavage reaction showing 5' cleavage products and uncleaved pre-mRNA resolved on 5% denaturing PAGE. One hundred nanograms each of GST-CTD or indicated mutant were included per 12.5- μ l reaction volume. H(1-29) and H(1-21) are not shown. (B) Cleavage activation efficiency of the GST-CTD mutants at 10 ng (white) and 100 ng (black) per 12.5- μ l reaction volume, expressed relative to a control where no recombinant GST fusion protein was added (i.e., factors alone). Each column represents an average of three or four separate experiments. Error bars indicate 1 standard deviation.

sion protein, H(30-37), which contains only heptads 30 to 37, to test whether these heptads alone may be sufficient to activate cleavage. As shown in Fig. 4B, deletion of heptads 30 to 37 resulted in loss of about 20% of activity relative to full-length GST-CTD, while H(1-44) led to a loss of only 6%. H(30-37) did not activate cleavage above background, with or without the GST moiety (data not shown; see Materials and Methods). These results demonstrated that, in the context of a CTD shortened to 44 heptapeptide repeats, heptads 30 to 37 are slightly more important to cleavage than others known to be unimportant (i.e., heptads 45 to 52). An examination of heptads 30 to 37 (Fig. 4C) shows that this region of the CTD is notable for its high number of nonconsensus residues at positions 2 and 4 in the heptad repeat, as well as for its Ser 4-Ser

5 dipeptides. Whether or not these contribute to this modest effect is not clear.

The foregoing results underscore how the variables of CTD sequence and length may work together in a way that is not readily deciphered: while heptads 30 to 37 appeared to be important in the context of a truncated CTD, deleting them did not result in a large loss of activity, and alone they were completely inactive. Taken together, though, our results indicate that a minimal length in the range of 29 to 35 heptads is the more important CTD variable, but that sequence may also contribute. The two—length and sequence—are inextricable when using a simple deletion analysis on the wild-type CTD sequence.

Construction of an all-consensus CTD to separate the contributions of length and sequence. The lack of conclusive results implicating specific CTD heptads in the cleavage reaction, along with the possibility of heptad redundancy, discouraged us from continuing with a targeted mutational strategy. Instead, we decided to create a nonnatural, full-length CTD consisting entirely of the consensus repeat (CON 52; Fig. 5A). In this way we would be able, in one experiment, to evaluate the aggregate contribution of all nonconsensus residues, and to do so completely independent of the variable of length. We anticipated that there would remain at least some residual activity in such a CTD construct because of the dependence of the cleavage reaction, at least in part, on length. If that proved true, we would then be in a position to evaluate the variable of length independent of sequence by constructing truncated versions of CON 52. As diagrammed in Fig. 5A, in addition to CON 52, we made truncated consensus CTDs of 13 (CON 13), 26 (CON 26), and 39 (CON 39) heptads and expressed each as a GST fusion protein (Fig. 5C).

Creating the cDNAs for these proteins was not straightforward. All-consensus forms of the yeast CTD, which consists of only 26 heptads, have been made before by ligating synthetic oligonucleotides (28, 38). We anticipated that this would fail for the much longer mouse CTD. In our approach, we used PCR to amplify a 13-heptad all-consensus CTD (CON 13) from heptads 7 through 20 (Fig. 5B; see Materials and Methods). The ends of the PCR product were phosphorylated, and CON 13 was multimerized by blunt-end ligation. Statistically, only one out of eight full-length cDNAs should have the correct orientation at each of the three ligations required to attain a 52-heptad cDNA. However, recognizing that each incorrect ligation must necessarily contain a palindromic sequence, we devised a selection strategy whereby all incorrectly ligated junctions would be destroyed by a suitable restriction endonuclease, leaving only products containing the desired orientation. This method led successfully to the series of all-consensus GST fusion proteins shown in Fig. 5A and C.

We first tested the full-length all-consensus protein (CON 52) in the reconstituted processing reaction (40 ng per 12.5- μ l reaction volume). Strikingly, substituting consensus residues for all of the nonconsensus residues led to only a $\approx$ 20% reduction in cleavage activity compared to the activity of GST-CTD (Fig. 5D). That is, in the context of a CTD of natural length, all nonconsensus residues together account for only 20% of the cleavage-activating property of GST-CTD.

The fact that CON 52 retained almost all of the GST-CTD cleavage activity permitted us to test cleavage as a function of

FIG. 4. Reconstituted 3' cleavage activation ability of CTD heptads 30 to 37. (A) Schematic representation of mutants used in this experiment. (B) Cleavage activation efficiency of the indicated GST-CTD mutants at 40 ng per 12.5- μ l reaction volume, as described in the legend to Fig. 4. (C) Comparison of the primary structure of heptads 30-37 with 45-52. Nonconsensus residues are underlined.

CTD length while keeping sequence constant. We determined the cleavage properties of the all-consensus series at intermediate concentrations (Fig. 6A, 40 ng per 12.5- μ l reaction volume; concentrations from 50 to 120 nM) and at a lower, equimolar concentration (Fig. 6B, 5 nM, or ≈ 1.7 to ≈ 5 ng of GST fusion protein). The results are summarized graphically in Fig. 6C. Progressively reducing the number of consensus heptads from 52 to 13 led to corresponding decreases in cleavage of from 80 to 85% of wild-type GST-CTD activity to 0%. At both low and intermediate concentrations, the largest reduction in activity occurred upon reducing the number of heptads from 39 to 26. This was similar to the trend observed in the natural CTD deletion series (Fig. 3B). At 5 nM CON 26, no cleavage above background was observed, though a small amount could be restored by increasing the concentration approximately 10-fold to 40 ng per reaction mixture. No cleavage activity was observed for CON 13 at either concentration.

FIG. 5. Construction and cleavage properties of the all-consensus CTD heptad series. (A) Schematic representation of CON 52, the all-consensus analog of the wild-type GST-CTD, and its truncation mutants. (B) Depiction of the strategy used to construct the all-consensus CTD series cDNA. PCR of the heptad-7-to-heptad-20 fragment was followed by blunt-end ligation and selection by restriction endonuclease digestion of the products containing the undesired ligation orientations. See text and Materials and Methods. (C) SDS-PAGE of the purified mutants (≈ 150 ng), silver stained. (D) Reconstituted adenovirus L3 3' cleavage activation efficiency of CON 52 versus the wild-type GST-CTD at 40 ng per 12.5- μ l reaction volume, expressed relative to a control where no recombinant GST fusion protein was added (i.e., factors alone). The wild-type CTD and CON 52 columns represent the average of three separate experiments. Error bars indicate 1 standard deviation.

FIG. 6. Reconstituted 3' cleavage reaction activated by the all-consensus CTD series. (A and B) 3' cleavage reaction showing 5' cleavage products and uncleaved pre-mRNA resolved on 5% denaturing PAGE; activation was with 40 ng of the indicated mutant per 12.5- μ l reaction volume (A) or using equimolar amounts (5 nM) of the mutants (B). (C) Graphical representation of the data shown in panels A and B.

Consistent with this, no activity was observed (data not shown) for a four-heptad synthetic consensus peptide (19). Therefore, the physical length of the CTD, formed as a consequence of the repeated YSPTSPS motif, accounts for approximately 80% of the *in vitro* cleavage activity of GST-CTD. We conclude that the principle determinant of CTD activity is length, not sequence, and that a length equivalent to roughly half the full-length mammalian CTD, i.e., about 26 repeats, is required to activate 3' cleavage.

DISCUSSION

Polyadenylation of mRNA is an important step in the maturation of RNAP II transcripts. As such, it can serve as a regulation point in the overall process of gene expression (2, 8). Although most of the polyadenylation factors have been identified, the functions of several remain to be elucidated. Our lab recently found that RNAP II and, in particular, the CTD of the largest subunit, is required for the *in vitro* reconstituted cleavage reaction (17). To understand how the CTD activates cleavage we analyzed a series of mutant GST-CTD

fusion proteins for their ability to activate the 3' cleavage *in vitro*. Previous studies by other labs have shown that not all of the CTD heptads are functionally equivalent (5, 21). This led us to search for the minimal subset of CTD heptads that might possess cleavage activity. Our results have shown that cleavage activity does not depend to any great extent on a specific subset of CTD repeats, but instead depends mostly on CTD length.

While most mutational studies of the CTD have been carried out in yeast (1, 30, 38), one detailed study has been done on heptad requirements of the mouse CTD. Corden and co-workers (5) made extensive deletions and rearrangements in the wild-type CTD, attached them to an α -amanitin-resistant RNAP II largest subunit, and tested them for their ability to confer amanitin resistance to cultured rat myoblast cells. Although the CTD was found to be essential, a significant number of heptads could be removed without preventing the transfer of amanitin resistance. With respect to total CTD length, it was found that a C-terminally deleted CTD consisting of 36 N-terminal heptads (Del 36-Ter) resulted in wild-type behavior. The minimum number of heptads required for viability was 25, and a gradual loss in the ability to confer the resistant phenotype was found in the range from 36 to 25 heptads. Intriguingly, we found that H(1-35), which is nearly identical to Del 36-Ter, supported nearly full-strength cleavage activity. Furthermore, we also detected intermediate 3' cleavage activity in the same range as they did for viability and a lack of significant activity in CTDs having fewer than 29 heptads. When heptads 23 to 36 were deleted, amanitin resistance was reduced to 64%, while we found mutant D30/37 (44 total heptads) reduced activity similarly, to about 80%. Similar results were obtained in a more recent study by the same group, in which they produced mice homozygous for a 13-heptad deletion from Ser-5, heptad 22 through Thr-4, heptad 36 (21). These mice were viable, but they were smaller than normal and showed a high degree of neonatal lethality. Neither rodent study addressed specific functional roles of the CTD, but rather they assayed the CTD for all its parts essential for viability. Nevertheless, our findings regarding CTD-activated 3' processing are surprisingly similar to the profile of CTD essential requirements found in the two *in vivo* studies. All three studies indicate that heptads in the region of heptads 23 to 36, while not essential for viability and/or for 3' cleavage, might be especially important, and that a length greater than ≈ 26 is necessary. This similarity supports the possibility that the defective mutants analyzed by Corden and colleagues were inactive at least in part due to impaired 3' processing.

Replacing each of the nonconsensus heptads one at a time by their consensus counterparts might be expected to reveal the identity of those heptads having critical importance to the cleavage reaction in the context of a full-length CTD. However, the prospect of functional redundancy among the heptads might render even this thorough an analysis uninterpretable. Mindful of this potential problem, we instead removed all of the nonconsensus residues simultaneously in order to assess their collective importance. The resulting all-consensus CTD, CON 52, showed that the nonconsensus residues contribute 20% of the wild-type cleavage activity in the context of a 52-heptad CTD. This approach completely separated the two variables of CTD length and sequence. The result was surprising not only because it was a relatively small effect but also

because it was approximately the same amount lost upon deletion of heptads 30 to 37 (mutant D30/37). In view of the cleavage results with other mutants [H(1-35), H(1-37), and H(24-56)], it is quite possible that wild-type heptads 30 to 37 contain all or nearly all of the nonconsensus heptad contribution to cleavage. Since this effect is small, the most important aspect of the CTD in cleavage must be length.

Recently, using transient-transfection experiments with α -amanitin-resistant fusion proteins, Fong and Bentley reported that a mutant CTD consisting of heptads 27 to 52 is as effective as the full-length wild-type CTD at promoting cleavage/polyadenylation of a cotransfected reporter transcript, whereas a CTD containing only heptads 1 to 25 is comparable to having no CTD at all (15). They concluded that the C-terminal half of the CTD is "sufficient to enhance 3' processing," whereas the N-terminal half is not. The results of our initial deletion analysis are largely in agreement with these findings, as H(24-52) produced on average from 72 to 88% of wild-type cleavage activity and H(1-21) and H(1-29) produced almost none. It is reasonable to conclude that the C-terminal half more effectively activates cleavage than the N-terminal half, as assayed by two different systems, when in the context of a CTD truncated to approximately half the natural length. However, the all-consensus series of CTDs permitted us to perform the more relevant experiment where the CTD was kept at its natural length, which revealed that the identity of the C-terminal heptads makes only a minor contribution. We conclude that the unique sequence of the C-terminal half of the wild-type CTD is not necessary for cleavage. For unknown reasons, it appears that the C-terminal half of the CTD assumes greater than normal activity in the context of a CTD artificially shortened to half its natural length. It also may be that it plays a more significant role in the context of intact nuclei, which contain numerous other proteins that may compete for binding the CTD. Nevertheless, our experiments show that the nonconsensus heptads play only a minimal role in the reconstituted cleavage reaction.

It is now known that, in addition to 3' processing, the CTD also participates directly in the other processing events of capping (7, 24, 40), cap methylation (32), and pre-mRNA splicing (18, 25, 41). The role of the CTD in capping has been studied in the most detail. The CTD (24), and in particular the hyperphosphorylated CTD (7), stimulates 5' capping by recruiting the capping enzyme to the nascent transcript. It was later found *in vitro* that synthetic CTD peptides as short as two heptads bearing a phosphate on Ser-5 can stimulate the capping reaction, while the corresponding Ser-2 phosphorylated peptides could mediate binding without stimulating capping (19). The nonphosphorylated consensus peptides do neither. Thus, in both mammals (19) and yeast (6) the phosphorylated CTD recruits the capping enzyme to the nascent RNA and allosterically stimulates the reaction by increasing the affinity of the capping enzyme for GTP. Our data for the CTD-activated cleavage reaction show both contrasts and similarities to the capping reaction. In contrast to capping, the phosphorylation state of the CTD appears to have no effect on cleavage (17). Also, CTD peptides from 4 to ≈ 26 repeats fail to stimulate cleavage. On the other hand, neither process requires the presence of any nonconsensus residues, and increasing numbers of tandem heptads increasingly stimulate both capping

(19) and cleavage (this work). In view of the contrasts and length requirement, it is unlikely that the CTD plays an allosteric role in cleavage. Since the endonuclease has not yet been identified, it is impossible to say whether the CTD specifically recruits that activity to the pre-mRNA or participates in the actual catalysis of the cleavage reaction. However, the results of the all-consensus series do rule out the possibility that the nonconsensus residues function catalytically during cleavage.

The CTD has been described as a "landing pad" for factors that process pre-mRNA, because many of these factors have been shown to bind to it. Among the mammalian 3' processing factors, CPSF (25) and CstF (15, 25), which have long been known to bind sequence-specifically to the pre-mRNA itself, bind to the CTD irrespective of its phosphorylation status. In yeast a third factor, PCF IIp, has recently been shown to bind the phospho-CTD (3). The other factors may in turn bind indirectly through these factors. Having such a platform on the polymerase allows the factors to be recruited, even as early as during the initiation process (12), and carried along during transcription elongation, which is postulated to facilitate the channeling of the nascent transcript to the processing machinery (25). The recently reported crystal structure of RNAP II is consistent with this hypothesis (10). Although the CTD is too disordered to be visible, the last discernible residue is located near the RNA exit channel, implying that the CTD is able to survey the nascent RNA as it leaves the enzyme.

Our *in vitro* cleavage reactions are performed in the absence of transcription and, hence, must reflect a role for the CTD in 3' processing distinct from channeling the nascent transcript, one in which the CTD participates either directly in the cleavage event or indirectly by preparing those factors that do participate directly. One possibility is that the staging area becomes a scaffold on which the factors align in a specific array. This differs from the coordination role described above in that the CTD here is postulated to remain with the factors and the poly(A) sequences (Fig. 1A) rather than depositing the factors on the RNA while passing. This type of role would require a minimum number of heptads to accommodate all of the multisubunit factors along with other non-CTD-binding subunits, and indeed our data indicate that more than 26 heptads are necessary. We speculate that CPSF (perhaps with PAP), CstF, CF I, and CF II all must bind the CTD in order to activate cleavage and that binding to the CTD either enhances the factor-factor interaction or prepares them conformationally to cleave the RNA. A CTD shorter than 26 to 29 heptads, we suggest, cannot sterically accommodate all of the factors simultaneously, and hence cleavage efficiency is reduced. Our results support the idea that the physical size of the CTD platform, or scaffold, is critical for CTD-dependent cleavage activity, apart from a catalytic or allosteric role and, presumably, in addition to a coordinating function. Our data rule out what was at the beginning of this study a realistic and interesting possibility: that the nonconsensus sequences might act to position in the proper alignment those cleavage factors that must occupy the CTD simultaneously. Instead, our results with the all-consensus CTDs indicate that the proper alignment must be coded for within the protein-protein interactions between the factors themselves, and not between the CTD and the factors. One such interaction, between CPSF-160 and Cstf-77 (29), is already known.

Finally, any explanation concerning how the RNAP II largest subunit CTD activates cleavage should also explain how the small molecule creatine phosphate and related phosphoamino acids produce the same 3' processing results in the reconstituted system (16, 17). We can only speculate that creatine phosphate may circumvent the need for the CTD by enhancing the protein-protein recognition among the factors so they can attain an active cleavage complex with the RNA. In any event, the important conclusion from our studies is that all-consensus heptad repeats can suffice for activation of 3' cleavage but that a large number, roughly half the length of the mammalian CTD, is necessary.

ACKNOWLEDGMENTS

We thank Yutaka Hirose for helpful discussions and the CF II fraction used in the 3' cleavage assays. We also thank Stewart Shuman and C. Kiong Ho for providing the four-heptad synthetic CTD peptide.

This work was supported by National Institutes of Health grant RO1-GM28983 and a National Research Service Award (5F32AI09655-03) to K.R.

REFERENCES

- Allison, L. A., J. K. Wong, V. D. Fitzpatrick, M. Moyle, and C. J. Ingles. 1988. The C-terminal domain of the largest subunit of RNA polymerase II of *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and mammals: a conserved structure with an essential function. *Mol. Cell. Biol.* **8**:321–329.
- Barabino, S. M., and W. Keller. 1999. Last but not least: regulated poly(A) tail formation. *Cell* **99**:9–11.
- Barilla, D., B. A. Lee, and N. J. Proudfoot. 2001. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **98**:445–450.
- Barron-Casella, E., and J. L. Corden. 1992. Conservation of the mammalian RNA polymerase II largest-subunit C-terminal domain. *J. Mol. Evol.* **35**:405–410.
- Bartolomei, M. S., N. F. Halden, C. R. Cullen, and J. L. Corden. 1988. Genetic analysis of the repetitive carboxyl-terminal domain of the largest subunit of mouse RNA polymerase II. *Mol. Cell. Biol.* **8**:330–339.
- Cho, E. J., C. R. Rodriguez, T. Takagi, and S. Buratowski. 1998. Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **12**:3482–3487.
- Cho, E. J., T. Takagi, C. R. Moore, and S. Buratowski. 1997. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **11**:3319–3326.
- Colgan, D. F., and J. L. Manley. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes Dev.* **11**:2755–2766.
- Corden, J. L., D. L. Cadena, J. M. Ahearn, Jr., and M. E. Dahmus. 1985. A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **82**:7934–7938.
- Cramer, P., D. A. Bushnell, and R. D. Kornberg. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**:1863–1876.
- Dahmus, M. E. 1996. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271**:19009–19012.
- Dantoni, J. C., K. G. Murthy, J. L. Manley, and L. Tora. 1997. Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. *Nature* **389**:399–402.
- de Vries, H., U. Rueggsegger, W. Hubner, A. Friedlein, H. Langen, and W. Keller. 2000. Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J.* **19**:5895–5904.
- Dignam, J. D., R. M. Lebovitz, and R. G. Roeder. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**:1475–1489.
- Fong, N., and D. L. Bentley. 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev.* **15**:1783–1795.
- Hirose, Y., and J. L. Manley. 1997. Creatine phosphate, not ATP, is required for 3' end cleavage of mammalian pre-mRNA in vitro. *J. Biol. Chem.* **272**:29636–29642.
- Hirose, Y., and J. L. Manley. 1998. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395**:93–96.
- Hirose, Y., R. Tacke, and J. L. Manley. 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev.* **13**:1234–1239.
- Ho, C. K., and S. Shuman. 1999. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* **3**:405–411.
- Liao, S. M., I. C. Taylor, R. E. Kingston, and R. A. Young. 1991. RNA polymerase II carboxy-terminal domain contributes to the response to multiple acidic activators in vitro. *Genes Dev.* **5**:2431–2440.
- Litingtung, Y., A. M. Lawler, S. M. Sebal, E. Lee, J. D. Gearhart, H. Westphal, and J. L. Corden. 1999. Growth retardation and neonatal lethality in mice with a homozygous deletion in the C-terminal domain of RNA polymerase II. *Mol. Gen. Genet.* **261**:100–105.
- Lu, H., O. Flores, R. Weinmann, and D. Reinberg. 1991. The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl. Acad. Sci. USA* **88**:10004–10008.
- Maquat, L. E., and G. G. Carmichael. 2001. Quality control of mRNA function. *Cell* **104**:173–176.
- McCracken, S., N. Fong, E. Rosonina, K. Yankulov, G. Brothers, D. Siderovski, A. Hessel, S. Foster, S. Shuman, and D. L. Bentley. 1997. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**:3306–3318.
- McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S. D. Patterson, M. Wickens, and D. L. Bentley. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**:357–361.
- Minvielle-Sebastia, L., and W. Keller. 1999. mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. *Curr. Opin. Cell Biol.* **11**:352–357.
- Mitchell, P., and D. Tollervey. 2001. mRNA turnover. *Curr. Opin. Cell Biol.* **13**:320–325.
- Morris, D. P., J. M. Lee, D. E. Sterner, W. J. Brickey, and A. L. Greenleaf. 1997. Assaying CTD kinases in vitro and phosphorylation-modulated properties of RNA polymerase II in vivo. *Methods* **12**:264–275.
- Murthy, K. G., and J. L. Manley. 1995. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev.* **9**:2672–2683.
- Nonet, M., D. Sweetser, and R. A. Young. 1987. Functional redundancy and structural polymorphism in the large subunit of RNA polymerase II. *Cell* **50**:909–915.
- Peterson, S. R., A. Dvir, C. W. Anderson, and W. S. Dynan. 1992. DNA binding provides a signal for phosphorylation of the RNA polymerase II heptapeptide repeats. *Genes Dev.* **6**:426–438.
- Pillutla, R. C., Z. Yue, E. Maldonado, and A. J. Shatkin. 1998. Recombinant human RNA cap methyltransferase binds capping enzyme/RNA polymerase II complexes. *J. Biol. Chem.* **273**:21443–21446.
- Rueggsegger, U., D. Blank, and W. Keller. 1998. Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol. Cell* **1**:243–253.
- Shatkin, A. J., and J. L. Manley. 2000. The ends of the affair: capping and polyadenylation. *Nat. Struct. Biol.* **7**:838–842.
- Takagaki, Y., L. C. Ryner, and J. L. Manley. 1988. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* **52**:731–742.
- Thompson, C. M., A. J. Koleske, D. M. Chao, and R. A. Young. 1993. A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**:1361–1375.
- Wahle, E., and U. Rueggsegger. 1999. 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol. Rev.* **23**:277–295.
- West, M. L., and J. L. Corden. 1995. Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics* **140**:1223–1233.
- Yon, J., and M. Fried. 1989. Precise gene fusion by PCR. *Nucleic Acids Res.* **17**:4895.
- Yue, Z., E. Maldonado, R. Pillutla, H. Cho, D. Reinberg, and A. J. Shatkin. 1997. Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongating form of RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **94**:12898–12903.
- Zeng, C., and S. M. Berget. 2000. Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol. Cell. Biol.* **20**:8290–8301.
- Zhang, J., and J. L. Corden. 1991. Phosphorylation causes a conformational change in the carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J. Biol. Chem.* **266**:2297–2302.
- Zhao, J., L. Hyman, and C. Moore. 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**:405–445.