

Organization and Function of APT, a Subcomplex of the Yeast Cleavage and Polyadenylation Factor Involved in the Formation of mRNA and Small Nucleolar RNA 3'-Ends*

Received for publication, April 29, 2003, and in revised form, June 19, 2003
Published, JBC Papers in Press, June 20, 2003, DOI 10.1074/jbc.M304454200

Eduard Nedeá‡§, Xiaoyuan He¶, Minkyu Kim||, Jeff Pootoolal‡, Guoqing Zhong‡,
Veronica Canadien**, Timothy Hughes‡, Stephen Buratowski||‡‡, Claire L. Moore¶,
and Jack Greenblatt‡**§§

From the ‡Banting and Best Department of Medical Research and Department of Molecular and Medical Genetics, University of Toronto, Toronto, Ontario M5G 1L6, Canada, the ¶Department of Molecular Biology and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111, the ||Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and **Affinium Pharmaceuticals, Toronto, Ontario M5J 1V6, Canada

Messenger RNA 3'-end formation is functionally coupled to transcription by RNA polymerase II. By tagging and purifying Ref2, a non-essential protein previously implicated in mRNA cleavage and termination, we isolated a multiprotein complex, holo-CPF, containing the yeast cleavage and polyadenylation factor (CPF) and six additional polypeptides. The latter can form a distinct complex, APT, in which Pti1, Swd2, a type I protein phosphatase (Glc7), Ssu72 (a TFIIB and RNA polymerase II-associated factor), Ref2, and Syc1 are associated with the Pta1 subunit of CPF. Systematic tagging and purification of holo-CPF subunits revealed that yeast extracts contain similar amounts of CPF and holo-CPF. By purifying holo-CPF from strains lacking Ref2 or containing truncated subunits, subcomplexes were isolated that revealed additional aspects of the architecture of APT and holo-CPF. Chromatin immunoprecipitation was used to localize Ref2, Ssu72, Pta1, and other APT subunits on small nucleolar RNA (snoRNA) genes and primarily near the polyadenylation signals of the constitutively expressed *PYK1* and *PMA1* genes. Use of mutant components of APT revealed that Ssu72 is important for preventing readthrough-dependent expression of downstream genes for both snoRNAs and polyadenylated transcripts. Ref2 and Pta1 similarly affect at least one snoRNA transcript.

Formation of the 3'-ends of mature mRNAs involves two tightly coupled biochemical modifications, an initial endonucleolytic cleavage step followed by addition of a polyadenosine tail at the 3'-end of the upstream cleavage product. An *in vitro* assay (1) has been used to guide the purification of the yeast

factors required for one or both of the steps of 3'-end formation: cleavage factor (CF)¹ IA, CFIB, CFII, polyadenylation factor (PF) I, and poly(A) polymerase (Pap1) (2, 3). However, a larger complex, cleavage and polyadenylation factor (CPF), which contains PFI, CFII, and Pap1, was subsequently described (4). Additionally, genetic evidence indicated that the *REF2* gene product is involved in 3'-end formation (5), but the connection of Ref2 to the biochemically defined cleavage apparatus had been obscure. In parallel, DNA sequences required for efficient mRNA 3'-end processing were identified. In particular, three elements governing efficiency, positioning, and the actual site of the *in vivo* mRNA 3'-end processing reaction have been described in yeast (reviewed in Ref. 6). Recently, *in silico* studies performed with the 3'-untranslated regions (UTR) of a large number of *Saccharomyces cerevisiae* genes (7, 8) added two more U-rich elements positioned on either side of the cleavage site as regulatory elements, and their involvement in the cleavage reaction has already been confirmed (9).

Transcription by RNAPII has been functionally linked with maturation of the newly synthesized mRNA in both lower and higher eukaryotes (10–12). In particular, mutations in either *cis*-elements required for cleavage (13) or in genes encoding cleavage factors (14, 15) result in defective termination by RNAPII, indicating that termination of transcription by yeast RNAPII is coupled with 3'-end formation. Direct associations between the C-terminal domain (CTD) of the largest subunit (Rpb1) of RNAPII and CFIA (16) or CPF (17–19) have been described. Rna15, a component of CFIA, has been found to interact with Sub1 and Mbp1, two factors that have roles in initiation of transcription by RNAPII (20, 21). Consistent with these observations, chromatin immunoprecipitation (ChIP) assays have revealed that the recruitment of yeast 3'-end processing proteins in CFI and CPF is co-transcriptional (12, 22) and requires the RNAPII CTD (12).

In this study we used a systematic proteomic approach to identify components of the yeast cleavage and polyadenylation machinery and factors that link 3'-end processing to the tran-

* This work was supported in part by grants from the Canadian Institutes of Health Research, the National Cancer Institute of Canada, the Canadian Cancer Society (to J. G.), by National Institutes of Health Grants GM41752 (to C. M.) and GM56663 (to S. B.), and by a Canadian Institutes of Health Research Operating Grant MOP 49451 (to T. H.). 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.

§ Supported by a C. H. Best doctoral fellowship.

‡‡ Scholar of the Leukemia and Lymphoma Society.

§§ To whom correspondence should be addressed: Banting and Best Dept. of Medical Research, University of Toronto, 112 College St., Rm. 210, Toronto, Ontario M5G 1L6, Canada. Tel.: 416-978-4141; Fax: 416-978-8528, E-mail: jack.greenblatt@utoronto.ca.

¹ The abbreviations used are: CF, cleavage factor; CPF, cleavage and polyadenylation factor; RNAPII, RNA polymerase II; UTR, untranslated regions; PF I, polyadenylation factor I; CTD, C-terminal domain of the largest subunit of RNA polymerase II; Pap1, poly(A) polymerase; ChIP, chromatin immunoprecipitation; TAP, tandem affinity purification; MALDI-TOF, matrix assisted laser desorption ionization-time of flight; MS, mass spectrometry; ORF, open reading frame; snoRNA, small nucleolar RNA; GST, glutathione S-transferase.

scription machinery. We present a comprehensive description of the polypeptides associated with CPF and describe various protein-protein interactions between subunits of CPF. In addition, we describe functional assays suggesting new roles for various components of CPF.

EXPERIMENTAL PROCEDURES

Yeast Strain Construction and TAP Purification—All the strains used in this study are isogenic derivatives of W303-1A (23) with the exception of the strains used in Fig. 3, C and D, which are *trp1Δ* derivatives of BY4741 (*REF2⁺* and *ref2Δ::Kan*) (Research Genetics), constructed by integration and counter-selection of the pNKY1009 *trp* deleter vector (24). TAP-tagged strains (including C-terminal truncations) were constructed as described (25, 26) and subsequently verified by Western blotting using a rabbit anti-Werner syndrome protein primary antibody and anti-rabbit horseradish peroxidase-coupled secondary antibody. Purifications of TAP-tagged proteins were performed from 4l of cells grown in YPD to an A_{600} of 1.5 essentially as described (25, 27) with minor modifications.

Recombinant DNA and *In Vitro* Protein-Protein Interaction Assays—The *SSU72* gene from a YEpl3-based plasmid that contains the *SSU72* gene and suppresses the temperature-sensitive phenotype of a strain carrying *SSU72-C-TAP* was cloned into pRS316. GST or His₆-tagged versions of recombinant proteins were obtained by cloning the full-length or truncated versions of genes into appropriate expression vectors (pGEX-2TK-Pti1, pET21b-Ssu72, -Mpe1, -Pta1, or pET28a-Sub1). Expression and purification of recombinant proteins, as well as *in vitro* binding experiments were carried out essentially as described (28, 29). Radiolabeled proteins were generated *in vitro* with the TNT rabbit reticulocyte lysate system (Promega) in the presence of [³⁵S]methionine in a total volume of 50 μ l according to the manufacturer's instructions. The eluted proteins were resolved on SDS-10% polyacrylamide gels, and detected by autoradiography or by Western blot with anti-His₆ antibody (Qiagen).

RNA Analysis—Total RNA was isolated by the hot phenol method (30). For each primer extension reaction, 10 μ g of total RNA and 200 fmol of 5'-³²P-labeled primer were denatured in 10 μ l of 1 \times First Strand Buffer (FSB) (Invitrogen) (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl) for 2 min at 90 °C and annealed for 1 h in a 41 °C water bath. The reaction was started by adding 10 μ l of 1 \times First Strand Buffer containing 20 mM dithiothreitol, 0.1 mM dNTPs, 20 units of RNasin (Promega), 100 units of SuperScriptII-RNaseH (Invitrogen) and stopped after 1 h with an equal volume RNA loading dye containing 90% formamide. Samples were separated on a 6% acrylamide, 7 M urea polyacrylamide gel. The amount of reverse-transcribed U6 loaded was adjusted to match the signal given by snoRNA cDNAs. The primers specific for sequences downstream of the mature snoRNAs have been described (31). The primer used for assessing readthrough of snR33 was 5'-GC-AATGGTGCAGATTGTGCAACTC-3', whereas the primers used for reverse transcription at the *CUP1* locus were CUP1_{INT} (5'-TCATTTCCAGAGCAGCATGACTTCTTGG-3') and CUP1_{POST 3'-END} (5'-GGATCTCTATACAGAGTTGTAAGTTAGGC-3'). Gene expression profiling using a yeast ORF oligonucleotide set was performed essentially as described (32).

β -Galactosidase Assay—Three or four independent colonies transformed with various reporter vectors (pHZ18 Δ 2-no terminator; pL101-*ADH2* terminator; pL501-*GAL7* terminator) (33, 34) were grown in synthetic medium lacking uracil containing 2% galactose to an A_{600} ~0.6 and assayed as described (35).

Chromatin Immunoprecipitation—Cells were grown in YPD to an A_{600} of 0.6–0.8 and processed essentially as described previously (36). Chromatin solution was incubated overnight at 4 °C with 20 μ l (50% slurry) of rabbit IgG-agarose beads (Sigma) pre-washed with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). The immunoprecipitate was washed stringently, and the recovered chromatin as well as the input chromatin was de-cross-linked and amplified by PCR as described (36). Various segments of a gene were amplified after ChIP as illustrated in Fig. 6. The 3'-ends of snoRNA genes amplified are as follows: +118 to +306 for *snR13* and +84 to +278 for *snR71*. Numbering is with reference to the translational start for ORFs or the 5'-end of the RNA for snoRNA genes. The most downstream primers used for PCR amplification of snR13 and snR71 3'-ends annealed up to 96 nucleotides before the translational start of TRS31 and up to 121 nucleotides inside the coding region of the convergently transcribed *LIN1* gene, respectively.

RESULTS

Proteins Involved in Transcription Initiation and Termination Associate *In Vivo* with a Holo-CPF—Ref2 was identified in a genetic screen for factors involved in mRNA 3'-end formation (termination and/or 3'-end processing) (5, 37). As the product of a non-essential gene, Ref2 is unlikely to be required for the cleavage reaction but seems to be needed for efficiently processing mRNAs that have suboptimal polyadenylation sites, indicating a possible role in the assembly of the cleavage machinery.

To test its association with the mRNA cleavage machinery, Ref2 was derivatized at its C terminus with a tandem affinity purification (TAP) tag (25), containing two IgG-binding modules of protein A and a calmodulin-binding peptide, by recombining tag-encoding DNA into the chromosomal *REF2* gene. Hence, a normal expression level of the Ref2 protein was maintained, allowing isolation of physiologically relevant protein complexes. Ref2-containing protein complexes were then purified sequentially on IgG and calmodulin columns and analyzed by SDS-PAGE and silver staining as presented in Fig. 1A (*left panel*), which shows that Ref2-Tap is specifically associated with a large multiprotein complex. Moreover, the association is not likely mediated by RNA as RNase A treatment (10 μ g/ml) had little effect on the pattern of associated proteins (Fig. 1A, lanes 3 and 4). Upon identification of the polypeptides copurified with Ref2 by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry (MS) (38) and peptide-mass fingerprinting (39), the Ref2-containing complex, holo-CPF, proved to be a larger version of the originally described CPF complex. The original CPF complex contained CFII subunits (Cft1, Cft2, Ysh1, and Pta1) and Pap1 and PF1 subunits (Fip1, Yth1, and four other unidentified polypeptides, simply designated p58, p53, p37, and p35) (4, 40). p53 was later identified as the polyadenylation factor subunit (Pfs) 2 (40), required for both mRNA cleavage and polyadenylation. The Ref2-containing complex shown in Fig. 1 contained all the originally characterized components of CPF as well as seven new components also recently identified by several other groups (17, 41–44). Three of them were similar in size to the uncharacterized p58, p37, and p35 components of PFI and were identified as Ykl059c, Ykl018w, and Glc7, respectively. In addition, our holo-CPF complex included three other polypeptides not originally shown to be associated with the yeast cleavage and polyadenylation machinery, namely Ygr156w, Ssu72, and Yor179c. We have given the name Syc1 (Similar to Ysh1 C-terminal) to Yor179c. Ygr156w and Ykl018w are also known as Pti1 and Swd2.

Very similar holo-CPF complexes have been described recently (17, 41–43). There are only small discrepancies between their findings and ours. For example, we and Gavin *et al.* (41) identified Ref2 as a band that migrates just above Pap1 rather than below it (17). On the other hand, we and Dichtl *et al.* (17) positioned Fip1 just below Pfs2 and above Pti1 rather than below (41). In addition, Dichtl *et al.* (17) did not identify Syc1 but instead identified a protein of unknown function of similar size encoded by the non-essential gene, *YDL094c*. Walsh *et al.* (42) described a CPF complex lacking Ssu72, Pti1, and Syc1. He *et al.* (43) very recently also described a complex identical in composition to ours.

Ssu72 is a factor that affects start site selection during transcription initiation. Ssu72 (suppressor of *sua7*, gene 2) interacts both genetically and physically with the general transcription factor TFIIB, encoded by the *SUA7* gene, and the second largest subunit of RNAPII (45, 46). Its presence in the holo-CPF complex was confirmed by tagging and purifying the Ssu72 protein, as shown in Fig. 1A (*lane 2*). The *ssu72-1* and

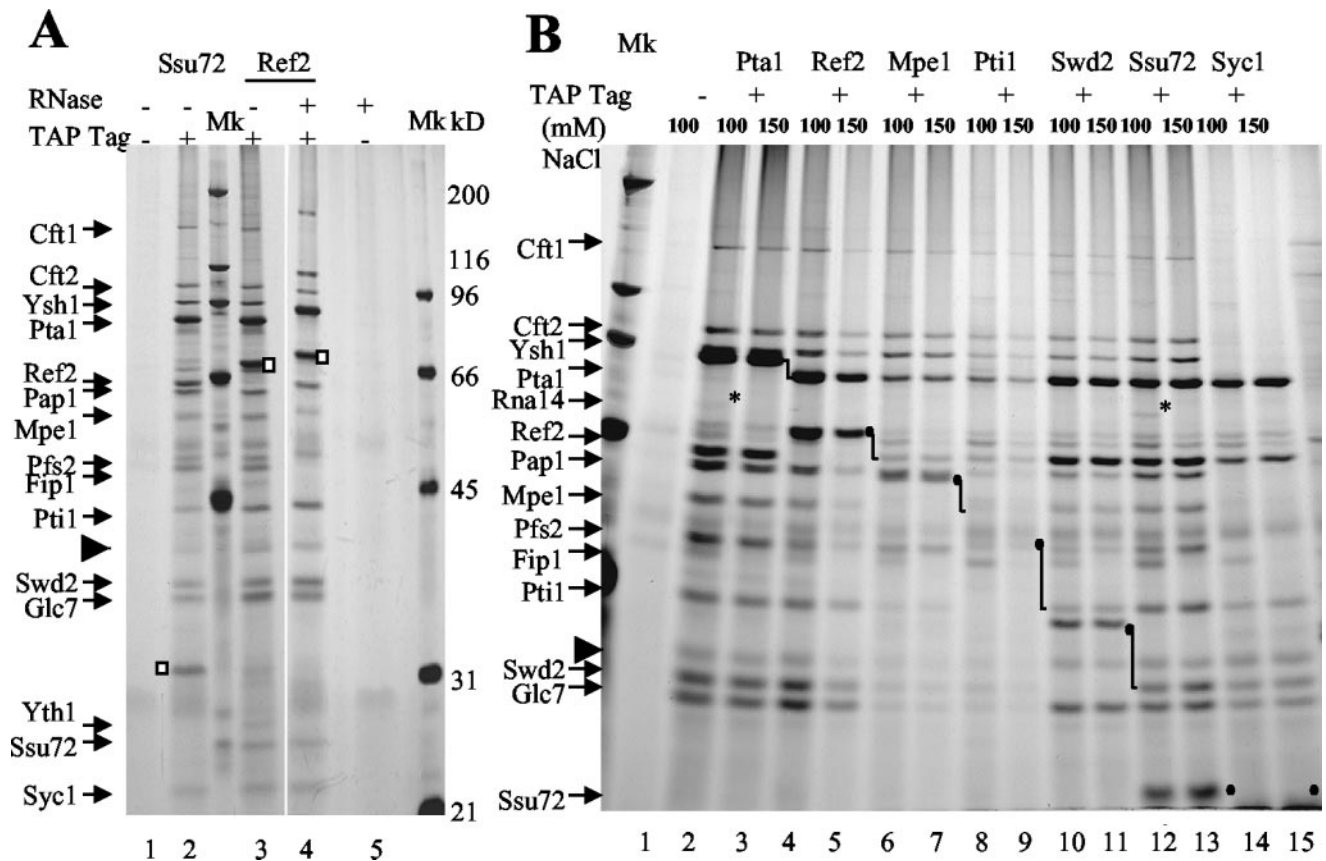


FIG. 1. Tandem affinity purification of known and putative cleavage/polyadenylation factors. Affinity-purified complexes were separated by 10% SDS-PAGE and visualized by silver staining. All the labeled proteins were identified by mass spectrometry. *A*, Ref2-Tap and Ssu72-Tap are specifically associated with a CPF-containing complex (compare the control lane 1 with lanes 2 and 3, respectively). RNA does not play an important role in mediating the association of Ref2 with the CPF complex as the treatment with RNase A (10 μ g/ml) had little effect on the band pattern (lanes 3 and 4). The open squares in this and subsequent figures indicate the tagged polypeptides. *Mk* denotes lanes containing protein molecular weight markers. *A* and *B*, the arrowheads on the left designate a degradation fragment of Pti1. *B*, the more recently identified subunits of CPF in *A* were individually tagged and purified using two different salt concentrations (100 and 150 mM NaCl) during the purification. The upper arms of the zig-zag lines point toward the TAP-tagged versions of the polypeptides, and the other ends indicate the untagged proteins. The asterisks indicate substoichiometric amounts of the CFII subunit Rna14. The protein molecular marker bands (*Mk*) correspond to 200, 116, 96, 66, 45, and 31 kDa.

sub1 Δ alleles interact genetically in an allele-specific manner with the same set of *sua7* alleles, suggesting that there could be a functional relationship between them (47). Interestingly, Sub1 was proposed to interact with Rna15, a subunit of CFII, and suppress termination by RNAPII (20).

Ygr156w (Pti1) is encoded by an essential gene (48), contains an RNA recognition motif, has weak similarity with human CSTF2 outside its RNA recognition motif, and was shown very recently to have a role in 3'-end formation (49). The N- and C-terminal regions of Pti1 and Rna15 are also weakly similar. Glc7 (GlyCogen-deficient 7) is the only essential type I Ser/Thr protein phosphatase of *S. cerevisiae* and participates in many cellular biochemical pathways by interacting with a myriad of regulators (50). Ykl059c, probably p58 in the original CPF, is an essential protein recently described as Mpe1 (mutant *PCF11* extragenic suppressor) and required for both steps of pre-mRNA 3'-end processing (51). Ykl018w (Swd2), also a component of the Set1 complex (52, 53), is indispensable for viability in yeast (48) and contains multiple WD repeats (seven, according to the Protein Structure Analysis server at bmercwww.bu.edu/psa/request.htm), a motif found in proteins involved in a variety of cellular functions and often components of multiprotein complexes (54). Yor179c (Syc1) is a non-essential protein (48), whose amino acid sequence is highly similar to that of the C-terminal region of the CFII component Ysh1 (35% identity and 52% similarity over 199 amino acids).

Pti1, Swd2, Mpe1, and Syc1 Are Genuine Components of CPF—In order to confirm our findings, polypeptides identified in holo-CPF that were not originally known to be subunits of CPF were individually tagged and purified as described above. As the salt concentration during our purifications had a significant impact on the yield of some purified polypeptides² (also see below), we carried out a series of purifications at 100 and 150 mM NaCl with tagged Pti1, Swd2, Mpe1, and Syc1 as well as Ssu72 and Ref2. A strain containing tagged Pta1, a known CFII subunit of CPF, was used as a reference. The results, shown in Fig. 1*B*, indicated that the new putative CPF-associated factors described here and elsewhere (17, 41–44) are genuine components of CPF. Full-length Pti1 and Mpe1 were purified less efficiently than the other CPF components, but their C-terminal truncated versions were purified efficiently and showed the complete holo-CPF composition² (Fig. 3), suggesting that the tags on the intact proteins might cause some steric hindrance. As also described elsewhere (42), Glc7-TAP co-purified with Pta1, a component of CFII, which we identified by MS, and many other polypeptides.²

Interestingly, the purification of tagged Syc1 yielded the six new polypeptides present in holo-CPF (Ref2, Pti1, Swd2, Glc7, Ssu72, and Syc1) (see also Refs. 17 and 41–44), but

² E. Nedea and J. Greenblatt, unpublished data.

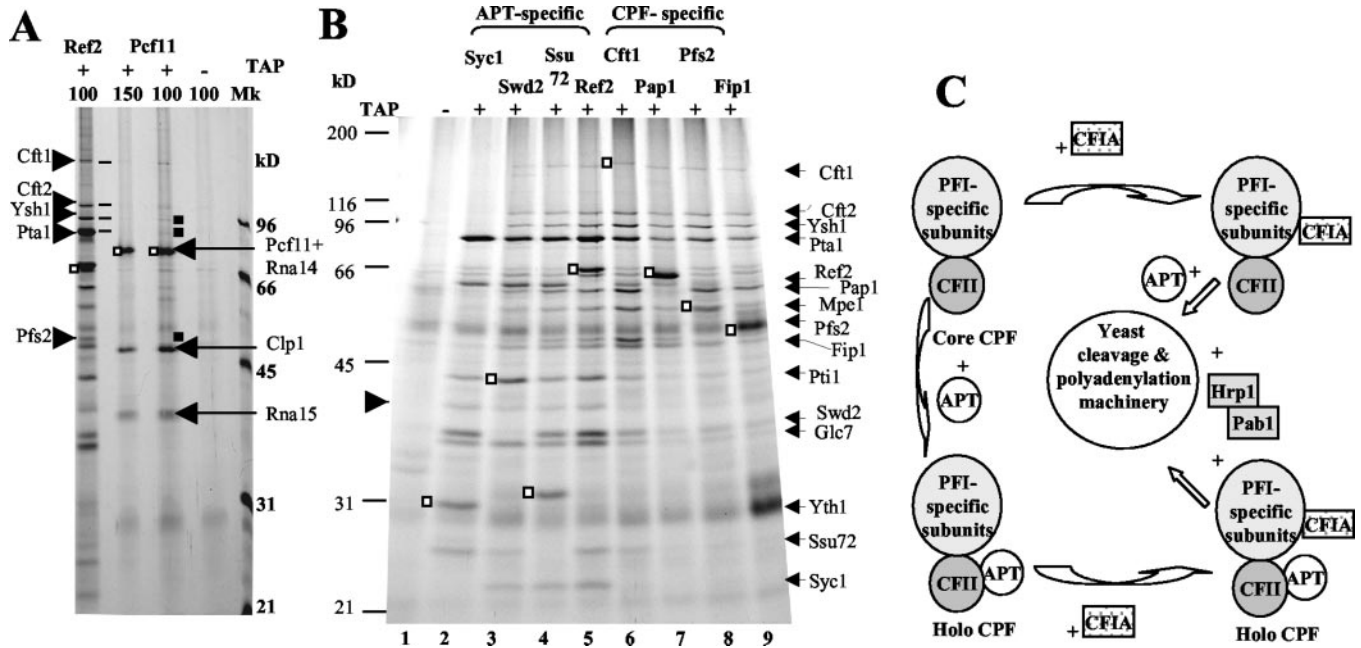


FIG. 2. Interacting modules in the cleavage and polyadenylation machinery. *A*, CFIA associates weakly with subunits of the core CPF complex. The arrowheads indicate CPF subunits found in preparations of CFIA subunit Pcf11. The closed squares represent the positively identified CPF polypeptides. *B*, evidence for distinct APT, core-CPF, and holo-CPF complexes. Core-specific (lanes 6–9) and APT-specific (lanes 2–5) polypeptides were TAP-tagged and purified at 150 mM NaCl. *C*, a model describing two alternative pathways for assembly/association of the yeast cleavage and polyadenylation machinery. See text for details.

none of the previously identified CFII or PFI components except Pta1 (Fig. 1*B*, lanes 14 and 15, and Fig. 2*B*, lane 2). Accordingly, we have named this subcomplex APT (Associated with Pta1). It is possible that there are two complexes, APT and core-CPF (containing only CFII, Pap1, and the PFI components Mpe1, Pfs2, Fip1, and Yth1) (40), that dynamically associate to form holo-CPF. Alternatively, the binding to solid supports of tagged Syc1 during its purification may lead to the dissociation of the APT complex from the other components of CFII and PFI. Regardless of which interpretation is correct, this observation indicates that the six new CPF-associated polypeptides described here and elsewhere (17, 41–44) are able to associate with each other and Pta1 in the APT subcomplex.

CFIA Associates Weakly with CPF—In the course of these experiments, we also obtained evidence that CFIA associates with CPF. First, we found that Rna14, a subunit of CFIA, was present and identified by MALDI-TOF MS, as indicated by the asterisks in lanes 2 and 12 of Fig. 1*B*, when Pta1 and Ssu72 were tagged and purified in buffers containing 100 mM NaCl, although the yield of Rna14 was clearly substoichiometric. A similar result was found for Pti1-TAP (Fig. 3*C*, lane 2). Second, when the CFIA subunits, Pcf11 (see Fig. 2*A*) and Rna15,² were tagged and purified in buffers containing 100 mM NaCl, substoichiometric amounts of various CPF subunits were visible by silver staining and three of them, Ysh1, Pta1, and Pfs2, were identified by MS along with the CFIA-specific subunits Rna14, Rna15, Pcf11, and Clp1. Based on these data, it seems likely that CFIA is weakly associated *in vivo* with core-CPF and holo-CPF and tends to be separated from them during our purification procedure. An association between CFIA and core CPF is supported by other biochemical and genetic data (40, 55). A complex containing holo-CPF and CFIA would have all the known components required for a standard, accurate cleavage and polyadenylation reaction, except CFIB (Hrp1/Nab4) and the poly(A)-binding protein Pab1 (see Fig. 2*C*). Interestingly, these latter two proteins, as well as the nuclear import protein Kap104, co-purify in

buffers containing 100 mM salt (but not 150 mM) when Hrp1/Nab4 is used as the tagged protein.³

Distinct Core-CPF and Holo-CPF Complexes—That there are distinct core-CPF and holo-CPF complexes is supported by the experiment shown in Fig. 2*B*. In this experiment, originally identified (lanes 6–9) and recently identified (lanes 2–5) components of holo-CPF were tagged and purified in parallel. It is apparent that when various components of core-CPF were tagged and purified, the purified complexes contained mainly CFII (Cft1, Cft2, Ysh1, and Pta1), Pap1, and PFI (Mpe1, Pfs2, Fip1, and Yth1), the components of core-CPF. The APT-specific components were clearly obtained in substoichiometric yield compared with the core-CPF-specific polypeptides. In contrast, when the APT-specific polypeptides were tagged and purified, all the holo-CPF polypeptides were obtained in roughly similar amounts, although the silver-staining procedure precluded a strictly quantitative analysis. In these instances, it was also noticeable that Pta1 was obtained in higher yield than the other CFII-specific subunits Cft1, Cft2, and Ysh1 (compare lanes 2–5 with 6–9), suggesting that Pta1 may be associated with both core-CPF and APT. It is striking that holo-CPF can be obtained by simply combining core CPF with the APT subcomplex. A model describing these findings is presented in Fig. 2*C*.

Ref2 Recruits Glc7 and Swd2 to the Holo-CPF—The Ref2 requirement for *in vitro* cleavage of pre-mRNA with suboptimal polyadenylation sites has been linked to its last 200 amino acids (37). We investigated the importance of the C terminus of Ref2 for the integrity of holo-CPF by constructing and purifying truncated TAP-tagged versions of Ref2 missing the last 100 amino acids (Ref2Δ100-TAPp) or last 200 amino acids (Ref2Δ200-TAPp). As seen in Fig. 3*A*, lane 6, Ref2Δ100-TAPp lost contact with all CPF subunits except Glc7, consistent with the observation that there is a two-hybrid interaction between Ref2 and Glc7 (56). An additional 100-amino acid deletion

³ X. He and C. Moore, unpublished data.

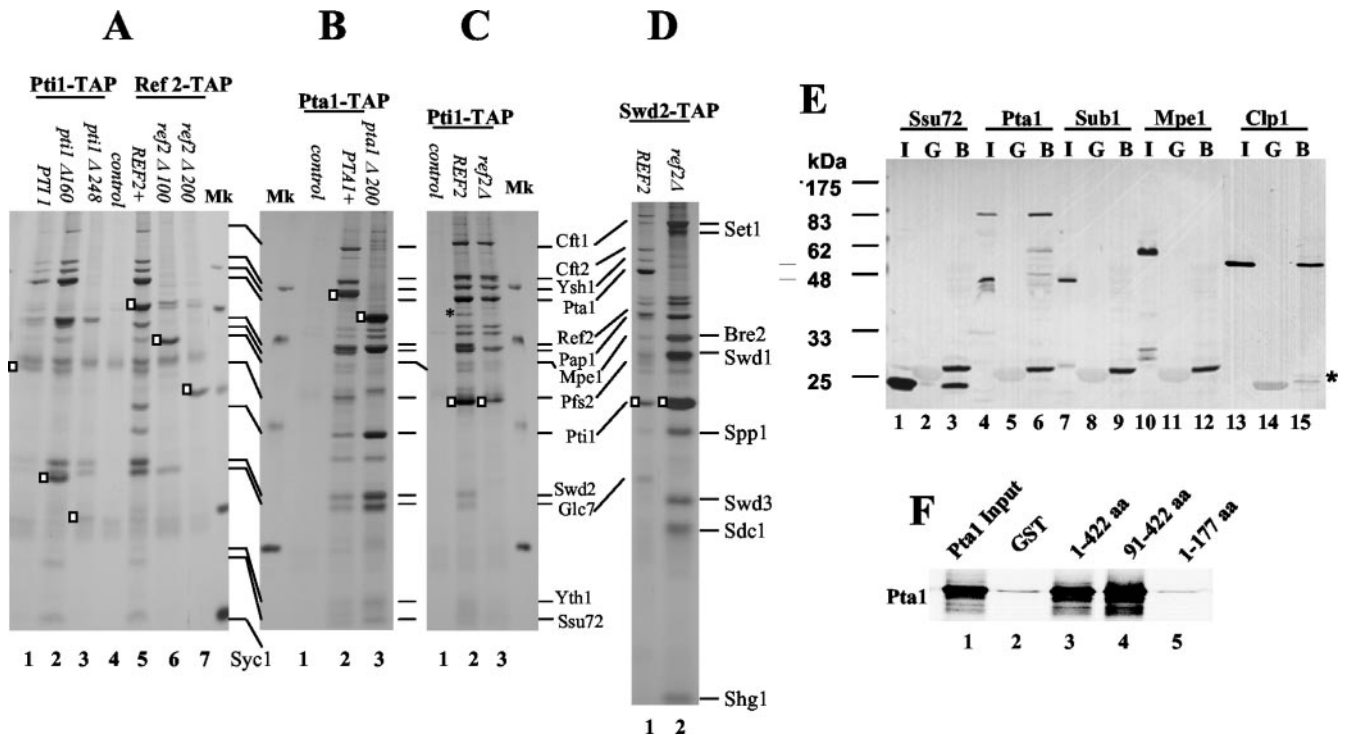


FIG. 3. **Molecular dissection of holo-CPF.** TAP-tagged C-terminally truncated forms of Pti1 and Ref2 (A) or Pta1 (B) were purified and analyzed as in Fig. 1A. Pti1-TAP (C) and Swd2-TAP (D) were similarly purified from strains containing or lacking Ref2. The protein molecular marker bands (Mk) correspond to 96, 66, 45, 31, and 21 kDa. E, GST pull-down assays were performed with full-length recombinant GST-Pti1 or GST alone and *in vitro* translated His₆-tagged Ssu72, Pta1, Sub1, Mpe1, or Clp1. After extensive washing, 10% of the input of each *in vitro* translated protein and the proteins bound to GST or GST-Pti1 were separated on a 10% polyacrylamide-SDS gel and detected by Western blot analysis using anti-His5 antibody. * indicates a nonspecific protein. I, 10% of the input of the *in vitro* translated protein; G, GST-bound protein; B, GST-Pti1 bound protein. F, GST pull-down assays were performed as in E using a full-length (lane 3) or truncated (lanes 4 and 5) GST-Pti1 and Pta1 radiolabeled by *in vitro* translation. Proteins bound to GST, GST-Pti1, or GST-Pti1 truncations were detected by autoradiography.

disrupted the interaction between Ref2 and Glc7 (Fig. 3A, lane 7), implying that residues 333–433 of Ref2 are required for its interaction with Glc7. A putative type I protein phosphatase-binding site, similar to the consensus site (KR)X (RH-QKAMNST)(VI)(RHSATMK)(FW)X{0,3}(END) (58), is found in this region (amino acids 368–376). These experiments not only show that Ref2 and Glc7 interact directly *in vivo*, but also strongly suggest that Ref2 is probably the sole interacting partner for Glc7 in CPF, as no other protein co-purifies along with full-length Glc7 and Ref2Δ100-TAP. If this were true, one would expect that holo-CPF would lack Glc7 in a strain lacking Ref2. In order to test this hypothesis, we TAP-tagged and purified Pti1 in a *ref2Δ* background. As shown in Fig. 3C, lane 3, the type I protein phosphatase was no longer present in CPF, validating our initial finding. In addition, the lack of Swd2 in this complex suggests Ref2 interacts with Swd2 and excludes a strong interaction between Pti1 and either Glc7 or Swd2.

The requirement of Ref2 for Swd2 recruitment to CPF is further supported by the fact that Swd2-TAP purification in a *ref2Δ* background yields no CPF components (Fig. 3D, lane 2). Instead, a recently described Set1-containing complex (52, 53) contained Swd2. This suggests that the two complexes compete for their common component, Swd2. As Swd2-TAP purification in a *REF2*⁺ strain yielded mainly CPF, most Swd2 is probably bound to the cleavage machinery *in vivo*.

Pti1 and Pta1 Are Important Scaffold Subunits in CPF—In order to analyze the role of Pti1 in the architecture of CPF, we used the same strategy as for Ref2. Intriguingly, deletion of the last 117² or 160 (see Fig. 3A, lane 2) amino acids of Pti1 did not result in any alteration of the composition of CPF, ruling out a structural role for the missing C-terminal end. However, when a further 88 amino acids were deleted, leaving only amino acids

1–174, a dramatic change took place, the only CPF components still co-purifying with Pti1 being Ref2, Swd2, and Glc7 (Fig. 3A, lane 3). A TAP-tagged strain with the C-terminal 248 amino acids of Pti1 deleted, leaving amino acids 1–174 of Pti1 fused to TAP, was viable. Because it was recently found (49) that deleting amino acids 185–260 of Pti1 was lethal, the TAP tag on Pti1-(1–174) may stabilize Pti1 and partially preserve its function.

An interpretation of these data is shown in Fig. 4. Our results exclude stable interactions between Ref2, Glc7, or Swd2 on the one hand, and CPF components other than Pti1. Rather, it would appear that Pti1 is a central adaptor that bridges different modules in the ~1-MDa holo-CPF complex. As Ref2, Glc7, Pti1, and Swd2 form a structural module in which Ref2 interacts with Glc7 and Pti1 interacts with neither Swd2 nor Glc7, it follows that Ref2 contacts both Swd2 and amino acids 1–174 of Pti1. This conclusion is in accord with the requirement of Ref2 for the presence of Swd2 in CPF. An interaction between Pti1 and Ref2 by two-hybrid analysis was also reported very recently (44).

Similarly, when a version of Pta1 with a 200-amino acid C-terminal truncation was purified, an important change in the CPF composition occurred (Fig. 3B). In this case, mainly Ref2, Pti1, Swd2, Glc7, Yth1, and Ssu72 co-purified with the tagged Pta1Δ200. As Ref2, Glc7, and Swd2 do not interact with Pta1 (see the explanations above), and the purification of Syc1 does not result in the co-purification of Yth1 (Fig. 1B and 2B), it seemed very likely that Pti1 either directly contacts Pta1 or uses Ssu72 as a bridge to Pta1 (see Fig. 4). As well, this result implied that the Pta1 subunit of CFII may interact with the Yth1 subunit of PFI.

The roles of Pti1 and Pta1 in our model for the organization of CPF (Fig. 4), as inferred from our data on the co-purification

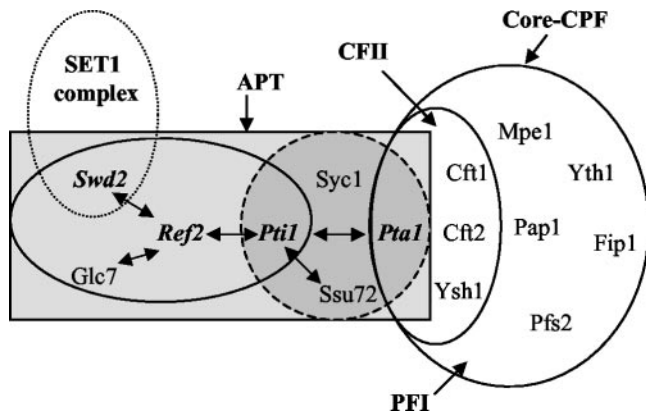


FIG. 4. A model indicating the compartmentalized architecture of holo-CPF. The solid double arrows indicate demonstrated direct interactions (visualized or inferred). The light gray rectangle represents the APT complex obtained from the purification of Syc1-TAP. The left light gray ellipse symbolizes the subcomplex purified with Pti1 Δ 270-TAP. The dark gray circle represents an inferred submodule. Pti1 is required for joining the two submodules of APT. The right vertical ellipse represents CFII, as described previously (76), which is also contained in the core-CPF. Also contained in the large ellipse on the right are the other components of core-CPF, Pap1 and PFI (Pfs2, Yth1, Fip1, and Mpe1). Pta1 is required for bridging CFII/core-CPF to APT and is a component of both complexes. Swd2 is shown as a common component of the distinct APT/holo-CPF and Set1 complexes. Ref2 is responsible for recruiting the majority of Swd2 to CPF.

of various proteins from whole cell extracts, were also examined by *in vitro* protein interaction assays using recombinant proteins. Full-length Pti1 interacted with both Ssu72 and Pta1, in good agreement with the predictions from the TAP data (Fig. 3E). In addition, we found that amino acids 1–90 of Pti1 were dispensable for its interaction with Pta1 (Fig. 3F). As a truncated form of Pti1 containing amino acids 1–262 efficiently bound both Pta1 and Ssu72 *in vivo*, and amino acids 1–174 were not sufficient for Pti1 to interact with Pta1 or Ssu72 *in vivo* or *in vitro* (Fig. 3, A and F), it seems likely that the region of Pti1 containing amino acids 178–262 is required for its interaction with both Ssu72 and Pta1.

Similar Pti1-binding experiments were also carried out with a number of other proteins. Binding interactions were found with Clp1 (Fig. 3E) and Pcf11,³ both of which are subunits of CFIA, suggesting that Pti1 interactions with Clp1 and Pcf11 may at least partly mediate the weak interaction between CFIA and CPF (see Fig. 2C). No interaction or only weak interactions, which may be nonspecific, were found for Pti1 with Rna14, Rna15, Yth1, Pap1, Brr5, Fip1,³ Mpe1, and Sub1 (Fig. 3E).

Our observations that Pti1 directly bound Pta1 and Ssu72 and that tagged Pti1 co-purified efficiently only with subunits of CPF, rather than with subunits of CFIA, suggest that Pti1 is an integral component of holo-CPF. It was recently suggested, however, that Pti1 is in an altered form of CFIA, in which it replaces Rna15, on the basis of weak sequence similarity between Pti1 and Rna15, as well as *in vitro* protein-protein interactions between Pti1 and either Pcf11 or Rna14 (49). It seems to us unlikely, however, that Pti1 is a component of CFIA, as Pti1 did not co-purify with tagged Pcf11 in amounts high enough to be detected by silver staining in our purifications² (see also Fig. 2A). Components of a tagged protein complex must be stably associated in order to survive the dilutions associated with two affinity purifications. On the other hand, binding *in vitro* to a concentrated ligand may reflect weak binding with a K_d value as high as 10^{-5} M. Therefore, we favor the conclusion that Pti1 is a stable component of CPF that associates only weakly with subunits of CFIA.

Ssu72 Is Important for Preventing Readthrough in Vivo to Downstream Genes—The association of a putative transcription initiation factor, Ssu72, with the cleavage and polyadenylation machinery was very intriguing. We asked whether Ssu72 plays a role in mRNA 3'-end formation by taking advantage of the fact that a strain containing a C-terminally TAP-tagged Ssu72 showed a mutant phenotype (slow growth at 30 °C and temperature sensitivity at 37 °C). A strain containing N-terminally TAP-tagged Ssu72 (N-TAP-Ssu72) has no growth defect (at 30 or 37 °C) and N-TAP-Ssu72 associates with the same set of polypeptides as Ssu72-C-TAP.² The growth defect induced by *ssu72-C-TAP* was rescued by the wild type gene expressed from a low copy vector (see Fig. 5A).

To test whether *SSU72* might be involved in 3'-end formation for polyadenylated transcripts, we used a previously described system (34) that allows one to evaluate the efficiency of mRNA cleavage and/or transcription termination by RNAPII *in vivo* by monitoring the levels of β -galactosidase. In this assay, briefly summarized in Fig. 5B, a vector containing the *lacZ* gene fused in-frame to an exon that follows a modified *rp51* intron is used to assess the expected levels of transcription (100% readthrough). In a second vector, a wild type polyadenylation signal is inserted within the *rp51* intron to ensure cleavage and polyadenylation, as well as termination, and prevent transcription of the downstream *lacZ* gene. Inefficient transcription termination and/or mRNA cleavage is detected when the downstream *lacZ* gene becomes transcribed and β -galactosidase is synthesized.

Our C-TAP allele of *ssu72* showed a significant readthrough (21 and 10%) at semipermissive temperature (28 °C) for two polyadenylation signals, derived from the *ADH2* and *GAL7* genes, respectively, that efficiently suppress readthrough (<2%) in a wild type strain (see Fig. 5C). An allele (*pcf11-9*) of a CFIA subunit known to be involved in cleavage and termination (14) was used as a positive control and caused substantially greater readthrough (70 and 40%). As a signal becomes apparent only after all or most of the *lacZ* gene is transcribed, it is possible that readthrough values may be underestimated (*i.e.* it is possible that a population of transcription complexes may terminate before the end of *lacZ* is reached). At this point, we cannot readily distinguish whether the role of Ssu72 in preventing readthrough is direct or else is mediated indirectly by cleavage of the transcript followed by exonucleolytic degradation of the downstream RNA. Indeed, Ssu72 was recently shown to be required for the cleavage reaction for at least some pre-mRNA substrates (43). This same study, using different mutations in Ssu72, did not detect an effect of Ssu72 on termination by RNAPII using a nuclear run-on assay, which is a more direct measure of termination by RNAPII (43). Therefore, it seems likely that the readthrough effect of our *ssu72-C-TAP* allele reflects inefficient cleavage of the transcript and consequent stabilization of the downstream readthrough RNA. It is possible, of course, that Ssu72 is required for termination downstream of some poly(A) signals but not the poly(A) signal of the *CYC1* gene studied by He *et al.* (43). Mutations we tested in other members of the CPF complex that showed various defects *in vivo* (*e.g.* growth rate) or *ex vivo* (TAP purification) (*ref2* Δ , *pta1* Δ 200TAP, and *pti1* Δ 248TAP) did not show a significant level of readthrough for either of the polyadenylation signals that we tested.²

As an additional test for the participation of Ssu72 in normal termination and/or 3'-end formation, we examined whether readthrough occurred at the *CUP1* locus, commonly used as a sensor for deficient mRNA 3'-end formation. By using reverse transcription analysis, we were able to show that transcripts extended beyond the normal cleavage site (59) to a higher

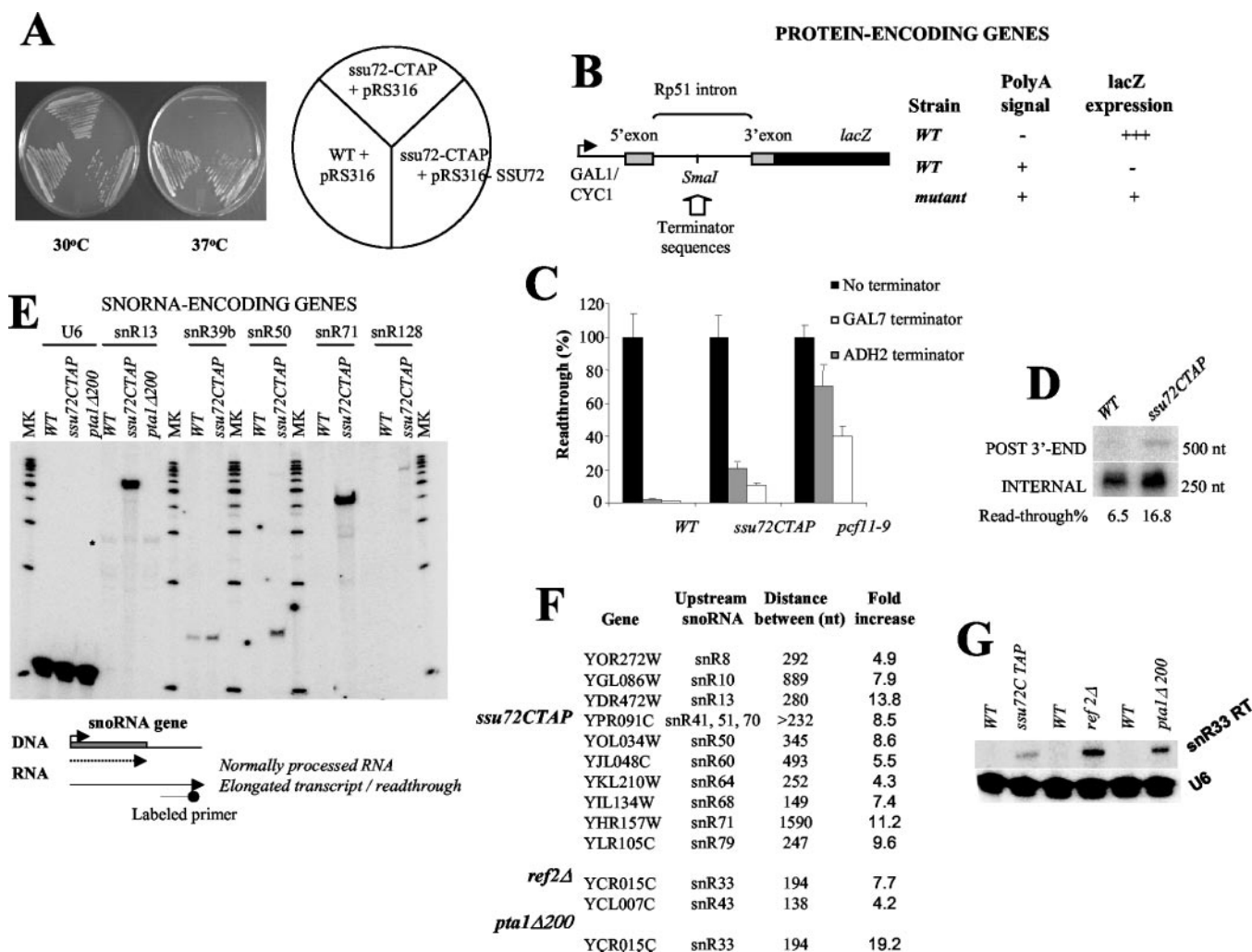


FIG. 5. *Ssu72* requirement for suppressing readthrough by RNAPII. *A*, strain containing a C-terminal TAP tag on *Ssu72* has a temperature-sensitive phenotype that is suppressed by wild type (*WT*) *Ssu72* expressed from a low copy vector. *B*, schematic showing the rationale of the system in which *lacZ* expression is used to assess the efficiency of RNA cleavage and/or transcription termination by RNAPII. See text for details. *C*, quantitative measurements of β -galactosidase activity were performed from wild type or mutant strains transformed with reporter vectors lacking or containing an efficient polyadenylation signal. Each measurement was carried out at least three times, and for each strain, the values represent the percentage of the value obtained in the absence of a cleavage and polyadenylation signal. *D*, evidence of elongated *CUP1* transcripts in a mutant *ssu72-C-TAP* strain as assessed by reverse transcription. Primers that anneal within the coding region or downstream of the cleavage site of the *CUP1* transcript were extended by reverse transcription. *E*, analysis of the 3'-ends of various snoRNAs by reverse transcription in wild type and mutant strains. The asterisk represents normal transcripts of the gene situated downstream of *snR13*, *TRS31*, which is transcribed from the same strand. *MK* lanes contain labeled 100-bp double-stranded DNA ladder. *F*, list of genes located downstream of snoRNA loci whose expression is increased more than 4-fold by mutations in APT components as a result of readthrough. *Distance between* represents the interval in nucleotides separating the 3'-end of the mature snoRNA transcript and the translational start of the downstream ORF transcribed from the same strand. *G*, extended transcripts (readthrough) downstream of the *snR33* locus, labeled as *snR33* reverse transcription, caused by various mutations and detected by primer extension. The lower panel served as a control for RNA recovery and shows similar levels of U6 transcripts estimated by primer extension were recovered from all the strains tested.

extent (more than 2-fold) in our *ssu72-C-TAP* mutant even at the semi-permissive temperature of 30 °C (see Fig. 5D).

We next investigated the possibility that *Ssu72* may also be important for 3'-end formation of another class of transcripts generated by RNAPII, namely snoRNAs. It has been shown that CFIA subunits, but not the originally identified CPF components, seem to be involved in snoRNA processing (60, 61). By using primer extension analysis to quantify the amounts of readthrough fusion transcripts, we found that our *ssu72-C-TAP* allele shows significant readthrough when compared with wild type at several snoRNA loci, *snR13*, *-39b*, *-50*, *-71* and *-128*, that we tested (see Fig. 5E), but no discernible alteration at several other loci (*snR3*, *-45*, and *-47*),² suggesting that the requirement of *Ssu72* for efficient snoRNA 3'-end formation may be gene-specific. The mutations in CPF subunits, which showed no defect in the readthrough assay for mRNA cleavage signals shown in Fig. 5, *B* and *C*, namely *ref2Δ*, *pta1Δ200TAP*,

and *pti1Δ248TAP*, were also negative for readthrough at the *snR13* locus² (Fig. 5E).

A recent study (31) suggested that genes situated downstream of snoRNA loci may appear up-regulated in an expression profiling microarray experiment if a mutation causes a deficiency in 3'-end formation for the upstream snoRNAs. In order to identify snoRNA transcripts whose correct 3'-end formation potentially requires *Ssu72* or other components of APT, we performed expression profiling experiments with *ssu72CTAP*, *pta1Δ200* (after 2 h at 38 °C), and *ref2Δ* (30 °C) mutants, and the corresponding wild type strains using an oligonucleotide set representing 6200 yeast ORFs. Consequently, our oligonucleotide array can detect putative snoRNA readthrough products that contain ORFs transcribed from the same strand only, unlike the ORF array used previously (31). It is quite likely that there is also readthrough we failed to detect downstream of other snoRNA genes in which the downstream

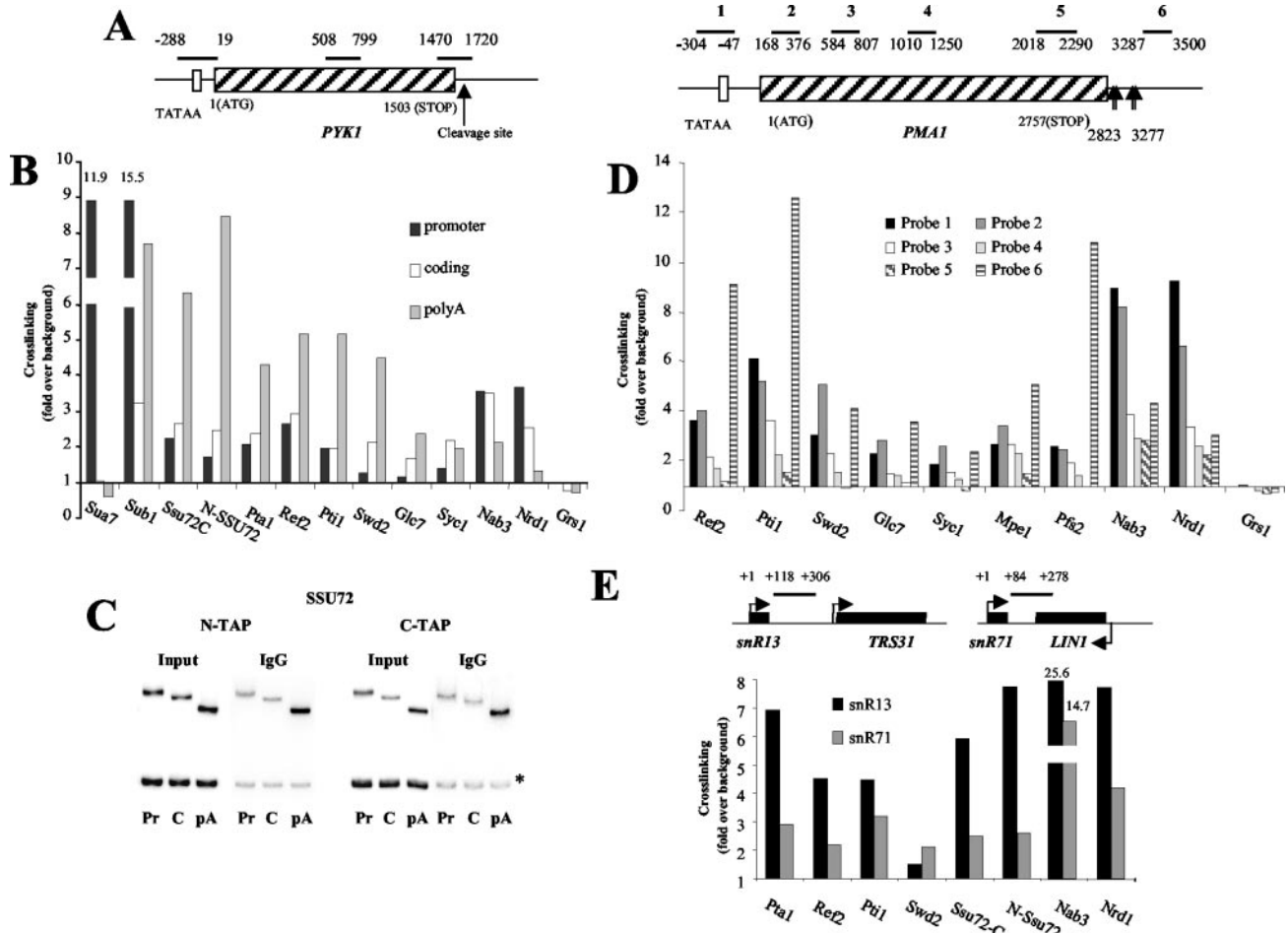


FIG. 6. ChIP localizes APT subunits to genes transcribed by RNAPII. Cross-linked chromatin from strains containing TAP-tagged subunits of APT or other factors was immunoprecipitated with IgG beads and amplified by PCR using primers specific for various regions of *PYK1* (*B* and *C*) or *PMA1* (*D*) or the 3'-ends of two snoRNA-encoding genes (*E*). A non-transcribed region of chromosome V marked by an asterisk in *C* was used as a control. *A* shows the regions of *PYK1* or *PMA1* amplified by PCR (horizontal lines), and the arrows indicate the cleavage site(s). *C*, a representative gel showing the pattern of cross-linking of wild type and mutant Ssu72 is illustrated. *Input* and *IgG* represent the signals before and after immunoprecipitation, respectively. *B*, *D*, and *E* summarize the degree of cross-linking (relative to the control region from chromosome V) for various proteins along the *PYK1* (*B*) or *PMA1* (*D*) genes or on the *snR13* and *snR71* genes (*E*). Each point is the average (S.D. <20%) of the signals of three PCRs normalized against the controls.

transcript has the opposite orientation. Interestingly, in the *ssu72* mutant, a considerable set of ORFs located downstream of snoRNA genes and transcribed in the same direction were dramatically up-regulated. Approximately half (7 of 15) of the genes whose expression was increased at least 7-fold were such genes (see Fig. 5F). Overall, 16 of 234 genes up-regulated more than 2-fold in *ssu72CTAP* were downstream of snoRNA loci. In three of these cases, readthrough from the mature snoRNA transcript to the downstream region was confirmed by reverse transcription (Fig. 5E), strongly suggesting that the increased expression of post-snoRNA ORFs is a direct result of readthrough. As well, two genes found to be up-regulated more than 2-fold in *ref2* and *pta1* mutants by microarray analysis were located downstream of snoRNA loci (snR33 and -43 and snR5 and -33, respectively). Primer extension analysis revealed that readthrough from the snR33 locus into the downstream gene was substantial in the *ref2* and *pta1* mutants and somewhat less pronounced in the *ssu72* mutant (Fig. 5G).

Ssu72 and Other Subunits of APT Associate with snoRNA Genes and Primarily with the 3'-Ends of mRNA-encoding Genes—ChIP after formaldehyde cross-linking has been used to demonstrate that some cleavage and polyadenylation factor subunits bind to transcriptionally active genes (12, 36). We investigated the possibility that Ssu72 might directly interact with 3'-ends of genes transcribed by RNAPII as do various

subunits of CFIA or core-CPF.⁴ The level of cross-linking of N- and C-terminally TAP-tagged Ssu72 (wild type and mutant, respectively) along a constitutive, highly expressed gene, *PYK1*, was analyzed as described previously (36) (see Fig. 6, *B* and *C*). The levels of cross-linking were compared with those obtained for a control region from chromosome V devoid of open reading frames as well as sn/snoRNA genes. Either tagged variant of Ssu72 cross-linked at the 5'-end, middle and 3'-end of the *PYK1* gene at levels higher than the control sequence, but cross-linking to the 3'-end region containing the polyadenylation signal (62) was substantially enriched (see Fig. 6, *B* and *C*). Other subunits of APT, namely Ref2, Pti1, Swd2, Gle7, Syc1 (weakly), as well as Pta1, showed a pattern of cross-linking to the *PYK1* gene similar to that of Ssu72 (see Fig. 6B), suggesting that the APT subunits of holo-CPF co-localize at least at some genes. The preferential localization of APT subunits to the region around the polyadenylation signal is likely to be quite general, because we have preferentially cross-linked APT subunits to the 3'-ends of the *PMA1* and *PYK1* genes (Fig. 6D), as well as the *ADH1*^{2,4} gene. Grs1, a tRNA synthetase proposed to be required for cleavage-independent transcription termination by RNAP II (63), did not cross-link

⁴ M. Kim, S. H. Ahn, and S. Buratowski, unpublished data.

above the background on the *PYK1* and *PMA1* genes (Fig. 6, *B* and *D*). TAP tagging and purification of Grs1 suggested that Grs1 is not strongly associated with any other protein.²

As mutations in *SSU72* and *SUB1* can specifically suppress alleles of *SUA7* deficient in transcription start site selection, it has been proposed that they are functionally related, at least at the level of transcription initiation. On the other hand, Sub1 has also been reported to play a negative role in transcription termination by RNAPII (20) and, more recently, a positive role in mRNA 3'-end processing (43). In CHIP experiments on the *PYK1* gene, we found that although Sua7 is detectable only at the promoter, as expected for a general initiation factor, Sub1, the yeast homolog of the human co-activator PC4 (64, 65), is substantially enriched both at the promoter and in the poly(A) signal region (Fig. 6*B*), which is in good agreement with its proposed roles. The lower level of Sub1 cross-linking to the middle of the *PYK1* gene suggests that Sub1 probably has independent roles at the promoter and 3'-end.

Consistent with the involvement of Ssu72 in snoRNA 3'-end formation, Ssu72, like Nrd1 and Nab3, which were previously shown to be involved in this process (31), also cross-linked at two snoRNA loci, *snR13* and *snR71*, which showed readthrough in an *ssu72-C-TAP* mutant strain (see Fig. 6*E*). SnoRNA loci are too small for the CHIP technique to distinguish whether Ssu72, Nrd1, and Nab3 are localized primarily at the 3'-ends. Similarly, the other APT subunits Ref2, Pti1, Swd2, and Pta1 were bound to the *snR13* and *snR71* genes (see Fig. 6*E*) even though the lack of Ref2 did not result in increased readthrough at these genes. Apparent cross-linking of APT subunits to *snR71* could, in principle, reflect cross-linking to the 3'-end of the downstream *LIN1* gene. Cross-linking of APT to *snR13*, however, should not reflect cross-linking to the promoter of the downstream *TRS31* gene, because the *TRS31* promoter is quite distant from the PCR primers and because APT cross-links poorly to promoters (Fig. 6, *B* and *D*, and data not shown). Therefore, the APT subcomplex of holo-CPF appears associated as a whole with snoRNA genes and primarily with the 3'-ends of mRNA-encoding genes even when some subunits of APT are not required to efficiently prevent readthrough. Conversely, Nrd1 and Nab3 are associated with the *PYK1*, *PMA1* (Fig. 6, *B* and *D*), and *ADH1*² genes even though these proteins are not known to participate in 3'-end formation for polyadenylated transcripts. Interestingly, Nrd1 and Nab3 are associated with the *PYK1* and *PMA1* genes primarily near their promoter regions and taper off toward the 3'-ends, which is the converse of the pattern observed for subunits of the APT complex. Our finding that Nrd1 and Nab3 are associated with mRNA-encoding genes, as well as snoRNA genes, is supported by several lines of evidence. First, Nab3 was identified as a nuclear protein that binds polyadenylated RNAs and is also required for the formation of various mRNAs (66). Second, Nrd1 binds the CTD of the largest subunit, Rpb1, of RNAPII suggesting Nrd1 might have a more general role in RNAPII transcription (67). Third, Nrd1 and Nab3 are required for maintaining normal levels of the Nrd1 mRNA and possibly other mRNAs (31).

DISCUSSION

In this study we performed a thorough analysis of the composition of the yeast cleavage and polyadenylation factor complex. Our results are in accord with recently published studies (9, 17, 41–44). In addition, we provide a more complete picture of the overall architecture of the holo-CPF complex by describing details regarding the connectivity of various structural modules and subunits of the holo-CPF (see Fig. 4).

In particular, we describe a subcomplex of holo-CPF, APT, which contains the Pta1 subunit of CFII and six additional

polypeptides (Ref2, Pti1, Swd2, Glc7, Ssu72, and Syc1) whose association with CPF was not discovered until recently (9, 17, 41–44). By the systematic tagging and purification of each subunit, we showed that five of the six polypeptides associated with Pta1 in APT exist *in vivo* mostly as members of holo-CPF-APT complexes. The exception is Glc7, which associates with many polypeptides² (42). Because Pta1 is present in both CFII and APT, dimerization of Pta1 could, conceivably, mediate the association of APT with CFII. Because only some, perhaps one-quarter to one-half, of the CPF molecules in a yeast extract are associated with APT, it is possible that APT is only engaged on some genes transcribed by RNAPII. We also found that there is a weak, rather than a stable (41), interaction between CPF and CFIA, enabling us to provide an outline for the assembly of the yeast cleavage and polyadenylation machinery (Fig. 2*C*).

There is evidence that the mammalian cleavage and polyadenylation specificity factor interacts with the general transcription factor TFIID at promoters and then is transferred to the elongating RNAPII (68). It has been proposed that the 3'-end machinery may engage the transcription machinery in a similar way in the yeast *S. cerevisiae* (12, 46). Our tandem affinity purification of Ssu72 and CHIP data argue against a stable interaction between Ssu72 and Sua7 (γ TFIIB). However, a transient or weak interaction at the promoter level is still entirely possible. Our data suggest that APT is initially recruited at promoters, as Ssu72 and the other APT components that we tested cross-linked at the *PYK1* and *PMA1* promoters significantly above the background.

Regardless of when APT is initially recruited to the transcription unit, our observation that there is enhanced cross-linking in the region that contains the poly(A) signal makes it likely that an encounter with the polyadenylation signal (at the DNA or RNA level) triggers an accumulation of APT at the 3'-UTR. The simplest hypothesis that explains our CHIP results is that a weak, unstable interaction between APT and the RNAPII transcription initiation and/or elongation complexes that first form when RNAPII is near the promoter is stabilized when the polyadenylation signal is transcribed. Interactions between RNAPII and various subunits of CFIA (16) or CPF (17–19, 46) have been described, suggesting that RNAPII *per se* could initially recruit the 3'-end processing machinery. As some of these interactions are mediated by a hyperphosphorylated CTD, and RNAPII must be hypophosphorylated in order to initiate transcription (69), it seems plausible that regulating the phosphorylation state of Rpb1 may influence 3'-end processing. The yeast CTD phosphatase, Fcp1 (70), may modulate 3'-end formation for RNAPII transcripts by opposing Ctk1, which phosphorylates Ser-2 of the CTD heptad repeats YSPTSPS (71) and was recently suggested to participate in this process (49). However, Glc7, the protein phosphatase subunit of the APT subcomplex in CPF could also potentially play a role in 3'-end formation by dephosphorylating RNAPII or some other substrate, like Pap1 (72). Intriguingly, Ssu72 was proposed recently (73, 74) to have a tyrosine phosphatase activity.

While our manuscript was in preparation, a report (17) showing that Ssu72 plays a role in suppressing pausing by RNAPII upstream of a polyadenylation signal was published. Evidence was also presented that Ssu72 contributes to transcription termination but only for the *CUP1* gene. Our results support and complement the findings of that report by suggesting that Ssu72 may be more generally important as a positive effector in suppressing readthrough by RNAPII not only at protein-encoding genes but also at snoRNA-encoding genes. Similar results were reported very recently (74). Although our experiments do not distinguish between the effects of Ssu72 on cleavage and

termination, one possibility is that Ssu72, which interacts directly with RNAPII, acts directly as a termination factor for RNAPII at some sites. The observation that the temperature sensitivity of an *ssu72-2* strain can be partly suppressed by 6-azauracil (17), which depletes UTP and GTP pools and enhances pausing by RNAPII, is consistent with a role for Ssu72 in termination by RNAPII. Alternatively, Ssu72 may only indirectly decrease the level of terminator readthrough by enhancing cleavage at the poly(A) site (43), followed by exonucleolytic digestion of the downstream RNA, or else by directly recruiting the exosome or some other RNA-processing nuclease to the transcript. This latter possibility does not agree particularly well with the apparent specificity of readthrough among snoRNA loci caused by the *ssu72-C-TAP* mutation (Fig. 5E). Our findings physically position Ssu72 (and other recent CPF components in APT) mainly at the 3'-ends of the protein-encoding genes *PYK1*, *PMA1*, and *ADH1*. Similar observations have been made for other cleavage and polyadenylation factors.⁴

It is interesting that Ssu72 is also necessary for forming the 3'-ends of many snoRNAs. In this sense, the effect of Ssu72 resembles those of Nrd1 and Nab3 (31) and subunits of the CFIA complex (60, 61). We were also able to identify one snoRNA, namely *snR33*, for which other subunits of the APT complex, Ref2 and Pta1, are important for preventing readthrough into downstream genes. Because Sen1 is involved in the snoRNA 3'-end formation (31), the apparently limited involvement of Pta1 and Ref2 in normal snoRNA 3'-end formation could possibly be explained by the ~2-fold up-regulation of the Sen1 mRNA we observed in microarray experiments. Very recently, Dheur *et al.* (44) showed that Ref2 and Pti1 are important for snoRNA 3'-end maturation and proposed a role for these proteins in uncoupling of cleavage and polyadenylation during small RNA 3'-end formation. Therefore, four of the seven APT subunits are required for the correct formation of the 3'-ends of snoRNAs which suggests a role for APT as a modulator of uncoupling cleavage from polyadenylation of the RNAPII transcripts.

Our experiments did not distinguish whether effects of Ssu72 on the expression of downstream genes were caused by transcriptional readthrough or stabilization of the downstream RNA as a consequence of failure to cleave at the upstream cleavage signal. Indeed, Ssu72 appears to be required for cleavage but not termination on the *CYC1* gene (43). Nevertheless, Ssu72 may contribute to termination on some snoRNA genes (and perhaps some protein-coding genes) by influencing pausing by RNAPII (17). How Ssu72 could suppress pausing by RNAPII upstream of a termination signal and enhance termination once the termination signal has been transcribed would require further explanation, but there is a precedent in the activity of the NusA protein of *E. coli*. Like Ssu72, NusA interacts with RNAP, and NusA normally enhances pausing and termination by RNAP, but the presence of *nut* site RNA and the bacteriophage λ N protein causes NusA to suppress termination (75). In a similar way, perhaps the presence of the polyadenylation signal in RNA and the proteins that recognize it cause the activity of Ssu72 to change from suppression of pausing to enhancement of termination.

Acknowledgments—We are grateful B. Seraphin, R. Haw, and C. Koth for providing plasmids and strains, to A. Golshani for technical help with the chromatin immunoprecipitation assay, and to B. Sheldrick and D. Richards for help with some of the mass spectrometry.

REFERENCES

- Butler, J. S., Sadhale, P. P., and Platt, T. (1990) *Mol. Cell. Biol.* **10**, 2599–2605
- Chen, J., and Moore, C. (1992) *Mol. Cell. Biol.* **12**, 3470–3481
- Kessler, M. M., Zhao, J., and Moore, C. L. (1996) *J. Biol. Chem.* **271**, 27167–27175
- Preker, P. J., Ohnacker, M., Minvielle-Sebastia, L., and Keller, W. (1997) *EMBO J.* **16**, 4727–4737
- Russnak, R., Nehrke, K. W., and Platt, T. (1995) *Mol. Cell. Biol.* **15**, 1689–1697
- Zhao, J., Hyman, L., and Moore, C. (1999) *Microbiol. Mol. Biol. Rev.* **63**, 405–445
- Graber, J. H., Cantor, C. R., Mohr, S. C., and Smith, T. F. (1999) *Nucleic Acids Res.* **27**, 888–894
- van Helden, J., del Olmo, M., and Perez-Ortin, J. E. (2000) *Nucleic Acids Res.* **28**, 1000–1010
- Dichtl, B., and Keller, W. (2001) *EMBO J.* **20**, 3197–3209
- Cho, E. J., Takagi, T., Moore, C. R., and Buratowski, S. (1997) *Genes Dev.* **11**, 3319–3326
- 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
- Licatalosi, D. D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J. B., and Bentley, D. L. (2002) *Mol. Cell* **9**, 1101–1111
- Russo, P., and Sherman, F. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8348–8352
- Birse, C. E., Minvielle-Sebastia, L., Lee, B. A., Keller, W., and Proudfoot, N. J. (1998) *Science* **280**, 298–301
- Greger, I. H., Aranda, A., and Proudfoot, N. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8415–8420
- Barilla, D., Lee, B. A., and Proudfoot, N. J. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 445–450
- Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., and Keller, W. (2002) *Mol. Cell* **10**, 1139–1150
- Dichtl, B., Blank, D., Sadowski, M., Hubner, W., Weiser, S., and Keller, W. (2002) *EMBO J.* **21**, 4125–4135
- Rodriguez, C. R., Cho, E. J., Keogh, M. C., Moore, C. L., Greenleaf, A. L., and Buratowski, S. (2000) *Mol. Cell. Biol.* **20**, 104–112
- Calvo, O., and Manley, J. L. (2001) *Mol. Cell* **7**, 1013–1023
- Aranda, A., and Proudfoot, N. (2001) *Mol. Cell* **7**, 1003–1011
- Komarnitsky, P. B., Michel, B., and Buratowski, S. (1999) *Genes Dev.* **13**, 2484–2489
- Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122**, 19–27
- Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J. D. (1998) *Yeast* **14**, 115–132
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) *Nat. Biotechnol.* **17**, 1030–1032
- Puig, O., Caspari, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001) *Methods* **24**, 218–229
- Krogan, N. J., Kim, M., Ahn, S. H., Zhong, G., Kobor, M. S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J. F. (2002) *Mol. Cell. Biol.* **22**, 6979–6992
- Helmling, S., Zhelkovsky, A., and Moore, C. L. (2001) *Mol. Cell. Biol.* **21**, 2026–2037
- Gross, S., and Moore, C. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 6080–6085
- Kohrer, K., and Domdey, H. (1991) *Methods Enzymol.* **194**, 398–405
- Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001) *Nature* **413**, 327–331
- Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., Hughes, J. E., Snesrud, E., Lee, N., and Quackenbush, J. (2000) *BioTechniques* **29**, 548–556
- Hyman, L. E., Seiler, S. H., Whoriskey, J., and Moore, C. L. (1991) *Mol. Cell. Biol.* **11**, 2004–2012
- Hyman, L. E., and Moore, C. L. (1993) *Mol. Cell. Biol.* **13**, 5159–5167
- Kippert, F. (1995) *FEMS Microbiol. Lett.* **128**, 201–206
- Komarnitsky, P., Cho, E. J., and Buratowski, S. (2000) *Genes Dev.* **14**, 2452–2460
- Russnak, R., Pereira, S., and Platt, T. (1996) *Gene Expr.* **6**, 241–258
- Karas, M., and Hillenkamp, F. (1988) *Anal. Chem.* **60**, 2299–2301
- Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., Grimley, C., and Watanabe, C. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5011–5015
- Ohnacker, M., Barabino, S. M., Preker, P. J., and Keller, W. (2000) *EMBO J.* **19**, 37–47
- Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M. A., Copley, R. R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., and Superti-Furga, G. (2002) *Nature* **415**, 141–147
- Walsh, E. P., Lamont, D. J., Beattie, K. A., and Stark, M. J. (2002) *Biochemistry* **41**, 2409–2420
- He, X., Khan, A. U., Cheng, H., Pappas, D. L., Jr., Hampsey, M., and Moore, C. L. (2003) *Genes Dev.* **17**, 1030–1042
- Dheur, S., Vo, L. T. A., Voisinet-Hakil, F., Minet, M., Schmitter, J.-M., Lacroute, F., Wyers, F., and Minvielle-Sebastia, L. (2003) *EMBO J.* **22**, 2831–2840
- Sun, Z. W., and Hampsey, M. (1996) *Mol. Cell. Biol.* **16**, 1557–1566
- Pappas, D. L., Jr., and Hampsey, M. (2000) *Mol. Cell. Biol.* **20**, 8343–8351
- Wu, W. H., Pinto, I., Chen, B. S., and Hampsey, M. (1999) *Genetics* **153**, 643–652
- Costanzo, M. C., Hogan, J. D., Cusick, M. E., Davis, B. P., Fancher, A. M., Hodges, P. E., Kondu, P., Lengieza, C., Lew-Smith, J. E., Lingner, C., Roberg-Perez, K. J., Tillberg, M., Brooks, J. E., and Garrels, J. I. (2000) *Nucleic Acids Res.* **28**, 73–76
- Skaar, D. A., and Greenleaf, A. L. (2002) *Mol. Cell* **10**, 1429–1439
- Tu, J., Song, W., and Carlson, M. (1996) *Mol. Cell. Biol.* **16**, 4199–4206
- Vo, L. T., Minet, M., Schmitter, J. M., Lacroute, F., and Wyers, F. (2001) *Mol.*

- Cell. Biol.* **21**, 8346–8356
52. Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., and Shilatifard, A. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12902–12907
53. Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W., Wilm, M., Aasland, R., and Stewart, A. F. (2001) *EMBO J.* **20**, 7137–7148
54. Smith, T. F., Gaitatzes, C., Saxena, K., and Neer, E. J. (1999) *Trends Biochem. Sci.* **24**, 181–185
55. Minvielle-Sebastia, L., Preker, P. J., and Keller, W. (1994) *Science* **266**, 1702–1705
56. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) *Nature* **403**, 623–627
57. Deleted in proof
58. Zhao, S., and Lee, E. Y. (1997) *J. Biol. Chem.* **272**, 28368–28372
59. McNeil, J. B., Agah, H., and Bentley, D. (1998) *Genes Dev.* **12**, 2510–2521
60. Fatica, A., Morlando, M., and Bozzoni, I. (2000) *EMBO J.* **19**, 6218–6229
61. Morlando, M., Greco, P., Dichtl, B., Fatica, A., Keller, W., and Bozzoni, I. (2002) *Mol. Cell. Biol.* **22**, 1379–1389
62. Graber, J. H., McAllister, G. D., and Smith, T. F. (2002) *Nucleic Acids Res.* **30**, 1851–1858
63. Magrath, C., and Hyman, L. E. (1999) *Genetics* **152**, 129–141
64. Henry, N. L., Bushnell, D. A., and Kornberg, R. D. (1996) *J. Biol. Chem.* **271**, 21842–21847
65. Knaus, R., Pollock, R., and Guarente, L. (1996) *EMBO J.* **15**, 1933–1940
66. Wilson, S. M., Datar, K. V., Paddy, M. R., Swedlow, J. R., and Swanson, M. S. (1994) *J. Cell Biol.* **127**, 1173–1184
67. Steinmetz, E. J., and Brow, D. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6699–6704
68. Dantonel, J. C., Murthy, K. G., Manley, J. L., and Tora, L. (1997) *Nature* **389**, 399–402
69. Lu, H., Zawal, L., Fisher, L., Egly, J. M., and Reinberg, D. (1992) *Nature* **358**, 641–645
70. Kobor, M. S., Archambault, J., Lester, W., Holstege, F. C., Gileadi, O., Jansma, D. B., Jennings, E. G., Kouyoumdjian, F., Davidson, A. R., Young, R. A., and Greenblatt, J. (1999) *Mol. Cell* **4**, 55–62
71. Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J., and Buratowski, S. (2001) *Genes Dev.* **15**, 3319–3329
72. Mizrahi, N., and Moore, C. (2000) *Mol. Cell. Biol.* **20**, 2794–2802
73. Meinhart, A., Silberzahn, T., and Cramer, P. (2003) *J. Biol. Chem.* **278**, 15917–15921
74. Ganem, C., Devaux, F., Torchet, C., Jacq, C., Quevillon-Cheruel, S., Labesse, G., Facca, C., and Faye, G. (2003) *EMBO J.* **22**, 1588–1598
75. Greenblatt, J., Nodwell, J. R., and Mason, S. W. (1993) *Nature* **364**, 401–406
76. Zhao, J., Kessler, M. M., and Moore, C. L. (1997) *J. Biol. Chem.* **272**, 10831–10838