Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I

Stefan Gross and Claire Moore*

Department of Molecular Biology and Microbiology, Sackler School of Graduate Biomedical Sciences, Tufts University, Stearns 509, 136 Harrison Avenue, Boston, MA 02111

Edited by Phillip A. Sharp, Massachusetts Institute of Technology, Cambridge, MA, and approved March 21, 2001 (received for review January 29, 2001)

Cleavage and polyadenylation of mRNA 3′ ends in *Saccharomyces cerevisiae* requires several factors, one of which is cleavage factor I (CF I). Purification of CF I activity from yeast extract has implicated numerous proteins as functioning in both cleavage and/or polyadenylation. Through reconstitution of active CF I from separately expressed and purified proteins, we show that CF I contains five subunits, Rna14, Rna15, Pcf11, Clp1, and Hrp1. These five are necessary and sufficient for reconstitution of cleavage activity *in vitro* when mixed with CF II, and for specific polyadenylation when mixed with polyadenylation factor I, purified poly(A) polymerase, and poly(A) binding protein. Analysis of the individual protein–protein interactions supports an architectural model for CF I in which Pcf11 simultaneously interacts with Rna14, Rna15, and Clp1, whereas Rna14 bridges Rna15 and Hrp1.

The posttranscriptional maturation of eukaryotic mRNA 3′ ends is an essential step in gene expression. This maturation occurs in two steps that are tightly coupled *in vivo*, but can be experimentally uncoupled *in vitro*. First, the nascent transcript is cleaved at a specific site downstream of the translational stop codon, and then a polyadenylate tail is added. This processing is effected by a multisubunit complex that provides the nuclease activity and also confers specificity to a template-independent poly(A) polymerase (PAP) (reviewed in ref. 1). The development of *in vitro* assays for the study of both mammalian (2) and yeast (3) mRNA 3′ end processing systems has led to the purification and characterization of the subunits of this complex from cell extracts. These investigations have revealed a significant degree of conservation both at the level of individual peptides as well as in the scheme of their arrangement into various subcomplexes (1, 4).

In the mammalian system, six components have been defined by biochemical separation as essential for cleavage and polyadenylation of the mRNA precursor. The multisubunit mammalian cleavage factors I and II (CF I_M and CF II_M), cleavage/polyadenylation specificity factor (CPSF), cleavage stimulatory factor (CstF), and PAP define the nascent 3′ end and perform the cleavage, whereas addition of the poly(A) tail uses CPSF, PAP, and poly(A) binding protein II.

In the yeast system, four components, also defined by protein chromatography, are required for processing of mRNA 3′ ends (5). CF I and CF II recognize the processing signals of the RNA and perform the endonucleolytic cleavage, whereas CF I, polyadenylation factor I (PF I), and the single-polypeptide PAP (6, 7) are required for the polyadenylation step (5). Further purification of these factors has led toward a determination of the number and identity of their constituent proteins. CF II contains the four peptides Cft1, Cft2, Brr5, and Pta1 (8). PF I contains the CF II subunits and the additional peptides Pfs1, Fip1, Pfs2, and Yth1 (9). As a consequence of this biochemical characterization, current work in this area is now focusing on dissection of the actual mechanisms through which these factors function (10–13).

There is still some controversy concerning the identity of the proteins of CF I. Kessler *et al.* (14) purified a CF I activity of five peptides through reconstitution of cleavage activity when mixed with a crude CF II fraction. During the purification process, CF I split into two separate subfractions, CF IA and CF IB. CF IA contained four proteins, three of which, Rna14, Rna15, and Pcf11, have homologs in mammalian factors and had been previously implicated in mRNA 3′ end formation (4, 15). The fourth protein identified, EF1α, has no known counterpart in mammalian processing. In the absence of any other evidence connecting EF1α to mRNA 3′ end formation, the significance of its copurification with CF IA activity remains unknown. CF IB was determined to be a single protein, Hrp1, and shown to be absolutely required for both cleavage and polyadenylation.

These results differ from work conducted by Minvielle-Sebastia *et al.* (16), who purified CF IA through activity complementation of a processing-deficient *rna15* mutant extract to a fraction containing five peptides (16). Three of the five proteins in this preparation were the Rna14, Rna15, and Pcf11 CF IA subunits purified by Kessler *et al.* (14). The fourth peptide was named Clp1, for cleavage/polyadenylation protein 1. A fifth, the major yeast poly(A) binding protein Pab1, also copurified with the CF I activity and was shown to participate in the control of poly(A) tail length (16–18). In contrast to Kessler *et al.*, Hrp1 was found to be required for polyadenylation, but only to have a stimulatory effect on the accuracy and efficiency of CF I-mediated cleavage (19).

To resolve the identities and necessity of the various CF I components, we undertook reconstitution of CF I from purified recombinant proteins expressed in nonyeast systems. This approach has the advantage over purification from extract in that the exact composition of each component is known. Our results show that five proteins in combination (Rna14, Rna15, Pcf11, Clp1, and Hrp1) are necessary and sufficient for reconstitution of cleavage activity in combination with CF II and for polyadenylation activity in combination with PF I, PAP, and Pab1. Direct protein–protein interaction studies indicate an arrangement of these proteins in the CF I complex that further illustrates the conservation of the yeast and mammalian machineries.

Materials and Methods

All purifications were conducted at 4°C, and reagents were obtained from Sigma, except where specifically indicated. Recombinant PAP was a gift from A. Zhelkovsky, Tufts University, recombinant Pab1 from A. Sachs, University of California,

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CF, cleavage factor; CPSF, cleavage/polyadenylation specificity factor; CstF, cleavage stimulatory factor; PAP, poly(A) polymerase; PF, polyadenylation factor; GST, glutathione *S*-transferase.

*To whom reprint requests should be addressed. E-mail: claire.moore@tufts.edu.

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

www.pnas.org/cgi/doi/10.1073/pnas.101046598

Berkeley, and recombinant glutathione *S*-transferase (GST)-Hrp1 from M. Henry, University of Medicine and Dentistry of New Jersey.

Expression and Purification of his_6-Rna14 and his_6-Clp1 from Insect Cells. The *CLP1* and *RNA14* ORFs were cloned into pFASTBAC-HTa and pFASTBAC-HTc, respectively, of the Bac-to-Bac Kit (GIBCO). These plasmids then were used to generate recombinant baculovirus according to the manufacturer's instructions. After several rounds of amplification in SF9 cells (Promega), these viruses were used for protein expression in Hi5 cells (Invitrogen).

Approximately 6×10^7 cells were harvested, washed once in PBS (PBS) + 1 mM PMSF, and resuspended in 5 ml sonication buffer (200 mM KCl/50 mM Tris·Cl, pH 7.9/10% glycerol/5 mM β-mercaptoethanol/0.1% Nonidet P-40) plus protease inhibitors (1 mM PMSF, 2 μM pepstatin A, 0.6 μM leupeptin). Cells were disrupted by sonication until complete lysis was confirmed by microscopy, and then debris was removed by centrifugation for 30 min at $11,000 \times g$. The lysate then was clarified by centrifugation at $90,000 \times g$ for an additional 30 min, frozen by immersion in liquid nitrogen, and stored at −80°C.

For affinity purification of the his_6-tagged proteins, 500 μl of Talon resin (CLONTECH) was first equilibrated in 15 ml Talon buffer (20 M Tris·Cl, pH 7.0/5 mM β-mercaptoethanol/10% glycerol/plus protease inhibitors) + 500 mM KCl and then mixed with the Hi5 lysate as prepared above. Volume was brought to 15 ml with Talon buffer + 500 mM KCl, and proteins were adsorbed for 20 min under gentle agitation. The resin was washed for 10 min with 15 ml Talon buffer + 500 mM KCl, 5 mM imidazole, pH 7.0 and 0.1% Nonidet P-40, then for 10 min with 15 ml Talon buffer + 500 mM KCl, 5 mM imidazole, pH 7.0, and finally twice for 10 min with 15 ml Talon buffer + 125 mM KCl before transferring to a 10-ml PolyPrep column (Bio-Rad). This column was washed with 10 ml Talon buffer + 125 mM KCl, and proteins were eluted with 5 ml Talon buffer + 125 mM KCl, 100 mM imidazole, pH 7.0, collecting 0.5-ml fractions. Peak fractions were identified by silver stain (Bio-Rad), pooled, and dialyzed twice for 2 h against 1 liter buffer D (20 mM Tris·Cl, pH 8.0/0.5 mM DTT/0.2 mM EDTA/10% glycerol/plus protease inhibitors) + 50 mM KCl.

Expression of myc-Tagged Pcf11 in Insect Cells. The *PCF11* ORF was amplified from yeast genomic DNA by using a 5′ primer that added the MEQKLISEEDL Flag M5 epitope before the Pcf11 start codon, cloned into pFASTBAC1 (GIBCO), and mobilized to create a recombinant baculovirus by using the Bac-to-Bac Kit. Protein was expressed as described for his_6-Rna14 and his_6-Clp1 above, except clarified lysate was dialyzed directly against buffer D.

Expression and Batch Purification of Rna15-his_6 from *Escherichia coli*. One liter of BL21 plys(S) cells containing plasmid pET21b-RNA15 (a gift from M. Kessler, Tufts University) was grown to an OD_{600} of 0.7 at 37°C and then induced for 2 h with 1 mM isopropyl β-D-thiogalactoside. Cells were collected, washed once with 30 ml PBS + 1 mM PMSF, and then resuspended in 7.5 ml Talon buffer. Cells were disrupted by a freeze/thaw cycle in liquid nitrogen, and the lysate was clarified by centrifugation for 30 min at $90,000 \times g$ before being processed as described above.

Expression and Purification of GST and Pcf11-GST from *E. coli*. The *PCF11-GST* ORF was amplified from plasmid Yeplac33-PCF11GST (a gift of C. Hammell, Dartmouth, Hanover, NH), and cloned into pET21-b (Novagen), replacing the his_6 epitope tag, and expressed as above. Cells were collected, washed with 30 ml PBS + 1 mM PMSF, and resuspended in 7.5 ml lysis buffer (250 mM KCl/50 mM Tris, pH 8.0/1 mM EDTA/0.5 mM DTT/10% glycerol/plus protease inhibitors), then lysed and clarified as described for Rna15-his_6. The supernatant was diluted with 1.5 vol buffer D and applied to a 15-ml DEAE-cellulose column (Bio-Rad) equilibrated with buffer D + 100 mM KCl. Unbound proteins were removed by extensive washing, and Pcf11-GST was eluted with 50 ml of buffer D + 250 mM KCl. This fraction then was diluted with 1.5 vol buffer D and reapplied to a 5-ml DEAE-cellulose column prepared as above. Proteins were eluted with a 100-ml gradient of buffer D + 100 mM KCl to buffer D + 250 mM KCl, collecting 5-ml fractions. Individual fractions were brought to 15 ml with 10 ml of buffer D + 125 mM KCl and 0.1% Nonidet P-40 and bound in batches to 0.5-ml glutathione-Sepharose (AP Biotech, Uppsala) for 1 h. Beads were washed three times with buffer D + 125 mM KCl, and Pcf11-GST was eluted in 0.5 ml buffer D + 125 mM KCl and 50 mM glutathione before dialyzing twice against buffer D + 50 mM KCl.

For expression and purification of GST, the pGEX-6P-2 (AP Biotech) vector was transformed in BL21 plys(S) cells and induced as described. After clarification, lysate was diluted 1:2 with buffer D and bound to 1 ml gluthathione-Sepharose beads for 30 min. The beads were washed four times with 15 ml PBS, and GST was eluted in 1 ml PBS + 50 mM glutathione and dialyzed twice against buffer D + 50 mM KCl.

Preparation of Yeast CF I, CF II, and PF I. Three hundred microliters of yeast whole-cell extract prepared according to Chen and Moore (5) was diluted 1:2 with buffer D + 50 mM KCl and clarified for 15 min at $90,000 \times g$, then applied to a 1-ml Hi-Trap Q-Sepharose FPLC column (AP Biotech). Bound proteins were eluted with a 20-ml gradient of buffer D + 50 mM KCl to buffer D + 500 mM KCl, followed by 5 ml of buffer D + 500 mM KCl. Fractions (0.5 ml) were collected and dialyzed against buffer D + 50 mM KCl. CF I-containing fractions were identified with an anti-Rna15 antibody. CF II-containing fractions were identified through reconstitution of cleavage activity with CF I in the RNA processing assay. Fractions containing both CF II and PF I activities were similarly identified through reconstitution of specific polyadenylation activity.

RNA Processing Assays. Capped, ^{32}P-labeled mRNAs were prepared by run-off transcription from plasmids pJCGAL7–1 for full-length and pJCGAL7–9 for precleaved substrates as described by Chen and Moore (5). RNA containing ≈250,000 to 280,000 cpm was used per reaction, equivalent to a final concentration of 10 nM.

Each reaction was done in a volume of 20 μl, containing 1 mM ATP, 10 mM creatine phosphate, 1 mM magnesium acetate, 75 mM potassium acetate, 2% PEG 8000, 1 mM DTT, 0.1 mg/ml BSA (New England Biolabs), 0.4 units of RNasin (Promega), and 10 nM radioactive RNA precursor. Processing assays used 1 μl of yeast whole-cell extract, 4 μl of CF I, CF II, or PF I crude fractions, 100 ng each of recombinant Rna15, Rna14, Pcf11, and Clp1, 80 ng Hrp1, and 20 ng PAP. For reconstitution of polyadenylation, 80 ng of poly(A) binding protein was also added to promote compact poly(A) tails (16–18). Reactions were assembled on ice and incubated at 4°C for 10 min, and then at 30°C for 20 min. For complementation assays using whole-cell extract made from temperature-sensitive cells, extract was preincubated for 5 min at 37°C before addition to the reaction. Reactions were stopped by the addition of proteinase K and SDS as described (5), brought to a volume of 30 μl with Tris-EDTA, pH 7.5, and extracted once with phenol/chloroform/isoamyl alcohol (25:24:1, vol/vol/vol). One-tenth of the reaction was resolved on a 5% acrylamide-8.3 M urea gel and visualized by using a Storm 960 PhosphorImager.

GST Pulldowns and Coimmunoprecipitations. To study interactions between Pcf11 or Hrp1 and other components of CF IA, 1 μg of

purified GST fusion protein was bound to 20 μl of glutathione-agarose beads in 200 μl IP-150 buffer (150 mM KCl/20 mM Tris·Cl, pH 8.0/0.1% Nonidet P-40) for 1 h. To minimize nonspecific interactions between components, the bead/GST-protein complexes then were blocked for an additional hour with 200 μl IP-150 plus 10% (vol/vol) FCS. Unbound proteins were removed with three 5-min washes with 1 ml IP-150 buffer, and 10 μg each of the additional recombinant CF I proteins were added in 200 μl IP-150 buffer for 1 h. After washing with IP-150 buffer as described above, assembled complexes were eluted with IP-150 + 50 mM glutathione and resolved by SDS/PAGE.

For immunoprecipitations, 0.5 μl anti-Rna15 polyclonal antiserum was bound to 10 μl protein-A agarose beads (GIBCO) for 1 h in 1 ml IP-150, blocked with FCS, and loaded with recombinant Rna15-his$_6$ for an additional hour. Assembly of complexes *in vitro* was conducted as described for GST fusions, and complexes were eluted by resuspending the beads in 20 μl SDS buffer and heating at 65°C for 5 min.

Antibodies to Rna15 have been described (14), penta-HIS antibody was purchased from Qiagen (Chatsworth, CA), anti-GST polyclonal serum was from Pharmagen, and the anti-myc monoclonal 9E10 and anti-Clp1 antibodies were gifts from B. Schaffhausen (Tufts University) and C. Hammell, respectively. Protein markers were from NEB, Beverly, MA.

Results

Expression and Purification of CF IA Components. The four subunits of yeast CF IA, Rna14, Rna15, Pcf11, and Clp1, were expressed heterologously either in *E. coli* or by using recombinant baculovirus infection of insect cells and purified as described in *Materials and Methods*. An SDS/PAGE analysis of the proteins used in this study is presented in Fig. 1*A*. For assays using myc-Pcf11, lysate from insect cells was used directly (Fig. 1*B*).

The functionality of Rna14, Rna15, Pcf11-GST, and myc-Pcf11 was confirmed by a complementation assay (Fig. 1*C*). Yeast whole-cell extracts made from the temperature-sensitive strains *rna15–2*, *rna14–1* (20), or *pcf11–2* (21) do not polyadenylate mRNA precursor *in vitro* when the assay is performed under nonpermissive conditions. However, they are capable of complementing each other when two different mutant extracts are mixed together, presumably by supplying a functional protein in trans. Addition of the appropriate recombinant protein also rescued the *in vitro*-specific polyadenylation activity of these defective extracts to the same levels observed through the mixture of that mutant extract with one mutant in another gene. Mutant extract can be rescued only with the appropriate recombinant protein (Fig. 1*D*). These results indicate that the wild-type recombinant protein can assemble into CF IA complexes in the extracts in place of the defective protein to form functional complexes. Because there are no published temperature-sensitive mutants in *CLP1*, we were not able to assay its functionality by using this method.

Analysis of the Subunit Interactions Necessary for Complex Assembly. The arrangement of the four subunits of CF IA was investigated through immunoprecipitations and GST pulldowns, where capture of a protein was taken as evidence of a direct protein–protein interaction. To ensure saturation of the bound protein, a 10-fold excess of soluble factors was used in each experiment.

To discover which proteins directly interacted with Pcf11, and whether these interactions were mediated via secondary interactions with other members of the CF IA complex, we used GST pulldowns with purified Pcf11-GST bound to beads and presented with a mixture of the other CF IA components either singly or in various combinations. GST alone showed no interaction with Rna14, Rna15, or Clp1 (Fig. 2*A*, lane 8). Pcf11-GST capture of Rna14 and Rna15 (Fig. 2*A*, lanes 1 and 2) confirm previously reported two-hybrid interactions and copurification

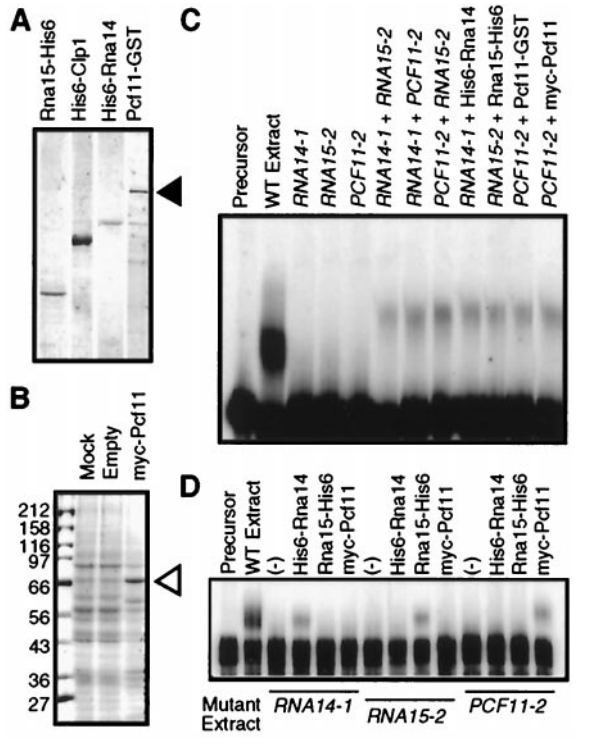

Fig. 1. Purification of active CF IA subunits. (*A*) Purity of expressed proteins used in reconstitutions was evaluated after silver staining SDS/PAGE gels; Pcf11-GST is indicated by a closed arrow. (*B*) High levels of expression of myc-pcf11 were obtained in baculovirus-infected insect cells. Mock, clarified lysates from Hi5 cells mock-infected. Empty, cells infected with baculovirus not containing myc-PCF11 ORF. Myc-Pcf11 is indicated by an open arrow. (*C*) Pure proteins or myc-Pcf11 in lysate can restore processing activity to defective extracts. Whole-cell extracts made from *rna14–1, rna15–2,* or *pcf11–2* mutant cells are defective for specific polyadenylation *in vitro*. Activity can be rescued by complementation with extract made from a different mutant. The activity of temperature-sensitive extracts also can be restored through the addition of purified recombinant protein. (*D*) Mutant extracts are rescued only by addition of the appropriate recombinant protein. The addition of buffer D alone is indicated by (−).

(21). We also observed a direct interaction between Pcf11 and Clp1 (Fig. 2*A*, lane 3). Combinations of the individual subunits did not change the efficiency of individual proteins binding to Pcf11 (Fig. 2*A*, lanes 4–7). These results prove that Pcf11 interacts directly with Rna14, Rna15, and Clp1 and also suggest that each protein probably has an independent binding site on Pcf11.

With recombinant proteins, we can demonstrate the *de novo* formation of an Rna14-Rna15 heterodimer by coimmunoprecipitation of Rna14 by Rna15 bound to beads with Rna15 antiserum (Fig. 2*B*). Kessler *et al.* (14) showed that the same Rna15 antiserum could immunoprecipiate a dimer of Rna14 and Rna15 from CF I-containing fractions isolated from yeast extract. A small amount of Rna14 binds nonspecifically to antibody coupled to beads. However, the retention of Rna14 on the beads is greatly increased with the addition of Rna15 (Fig. 2*B*, compare lanes 1 and 2). This result indicates that the other CF I components are not necessary for the assembly of the Rna14-Rna15 heterodimer.

To confirm the Pcf11-Clp1 interaction identified above and investigate the possibility of direct interactions between Clp1 and Rna14 or Rna15, we conducted immunoprecipitations using Clp1 bound to beads with anti-Clp1 antiserum and attempted to capture other CF IA components. None of the CF IA subunits were brought down in the absence of Clp1 (Fig. 2*C*, lane 6).

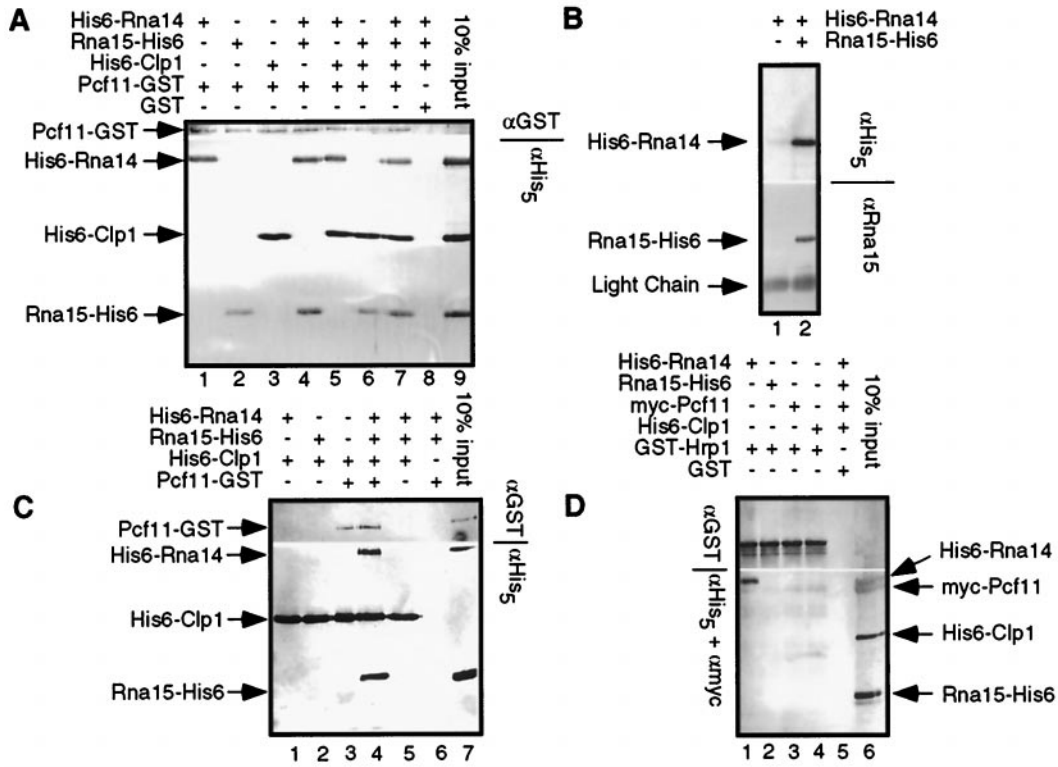

Fig. 2. Protein–protein interactions of CF I components. Complexes were assembled and recovered as described in *Materials and Methods*, resolved by SDS/PAGE, and analyzed by Western blot. A plus sign indicates the presence of a particular protein in the assembly reaction. As a negative control for the GST pulldowns, GST alone was used. After transfer, the membrane was cut at the horizontal line and probed separately with the antibodies indicated. (*A*) Pulldowns using GST-Pcf11 illustrate direct interactions with Rna14, Rna15, and Clp1. (*B*) Coimmunoprecipitation of Rna14 and Rna15. Rna15 bound to beads using polyclonal antibodies to Rna15 was able to capture Rna14. (*C*) Clp1 interacts directly with Pcf11, but not Rna14 or Rna15. Clp1 was bound to beads using Clp1 antiserum and mixed with Rna14, Rna15, or GST-Pcf11. Precipitation of Pcf11-GST depended on the presence of bound Clp1, and precipitation of Rna14 and Rna15 depended on the presence of Pcf11 bound to Clp1. (*D*) Rna14 bridges CF IB and CF IA. Pulldown with GST-Hrp1 captures Rna14, but not myc-Pcf11, Rna15, or Clp1; no proteins are precipitated with GST alone.

Precipitation of Pcf11 did not require the presence of any other protein (Fig. 2*C*, lane 3), but capture of Rna14 and Rna15 depended on the presence of Pcf11 bound to Clp1 (Fig. 2*C*, lanes 4–5), suggesting that Pcf11 bridges Clp1 and these proteins through multiple protein–protein interaction sites.

Yeast CF I can be split into two components, CF IA and CF IB (14). CF IB is a single peptide, Hrp1, that shuttles between the nucleus and the cytoplasm (18), and that, like Rna15, contains RRM-type RNA-binding motifs (22). To determine which component of CF IA bridges these two processing factors, we used a GST-Hrp1 fusion that had previously been shown to supply CF IB activity in reconstitution experiments (18). This fusion construct also fully covers a disruption of the *HRP1* gene in *Saccharomyces cerevisiae*, indicating that it must be able to make all of the contacts critical for its role in mRNA 3′ end processing. GST-Hrp1 was bound to beads and tested for capture of the CF IA components Rna14, Rna15, Pcf11, and Clp1. None of the input proteins bound when GST was used (Fig. 2*D*, lane 5). In the presence of GST-Hrp1, only Rna14 was retained on the beads (Fig. 2*D*, lanes 2–5).

All Four Components of CF IA Plus CF IB Are Required for Reconstitution of the Cleavage and Polyadenylation Activity of CF I. To determine the minimal cohort required for CF I cleavage activity, combinations of purified peptides were mixed and allowed to assemble before addition to reactions containing the Hi Trap Q CF II fraction and radioactive RNA containing sequences flanking the *GAL7* poly(A) site. This precursor is efficiently cleaved and polyadenylated in yeast extract (Fig. 3, lane 1). No processing is observed in the presence of either CF I or CF II containing fractions (Fig. 3, lanes 2 and 3). When these two factors are combined, the RNA is cleaved but not polyadenylated because of the absence of PAP and PF I-specific subunits. When recombinant proteins are used in place of CF I, cleavage is observed only when all four CF IA subunits and Hrp1 are present (Fig. 3, lane 20). Although Minvielle-Sebastia *et al.* (19) observed cleavage in the absence of Hrp1 within 3 min of starting the reaction, we observed no cleavage without Hrp1, even if the reaction was carried out to 60 min (data not shown).

A similar experiment was conducted to determine the necessity of various CF IA proteins, as well as Hrp1, for reconstitution of the polyadenylation reaction. Precursor containing only sequences upstream of the *GAL7* poly(A) site is readily polyadenylated in yeast extract (Fig. 4, lane 1). Neither CF I from yeast nor a fraction containing both CF II and PF I activity can polyadenylate this substrate when supplemented with PAP and Pab1, but efficient processing is obtained with CF I and CF II/PF I in the same reaction (Fig. 4, lanes 2–4). In the absence of the CF I fraction, reconstitution of specific polyadenylation required the presence of all four peptides of CF IA plus CF IB (Fig. 4, lane 20).

Discussion

A model of CF I architecture derived from our interaction studies is presented in Fig. 5*A*. The individual interactions between Pcf11 and Rna14, Rna15, or Clp1 did not require the presence of other CF I proteins. The experiments presented here also demonstrate that it is possible to reassemble the complex

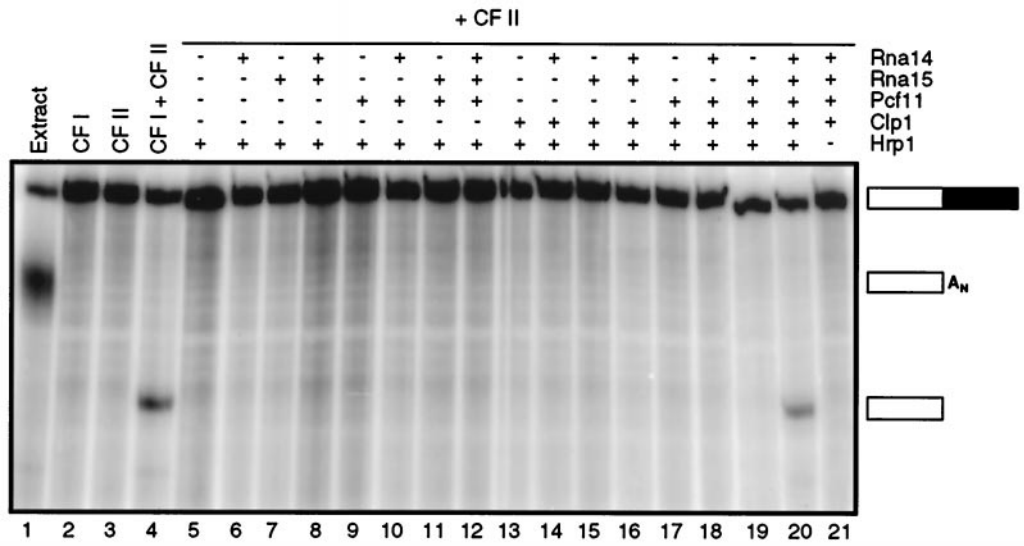

Fig. 3. Reconstitution of CF I cleavage activity requires the addition of all five peptides. Whole-cell extract from wild-type yeast strain FY86 before fractionation on a Hi-Trap Q column cleaves the precursor and polyadenylates the upstream cleavage product. CF I and CF II are sufficient in the absence of PAP and PF I to reconstitute cleavage activity when mixed. The addition of all five peptides of CF I reconstitutes cleavage activity when mixed with CF II (lane 20). The downstream cleavage product is rapidly degraded and is not detectable under these conditions.

from separately expressed and purified components, suggesting that assembly is not mediated by other cellular factors.

In our model, the Rna14 bridge between CF IB and CF IA connects Rna15 and Hrp1, two proteins with predicted RNA-binding domains. This subassembly may contribute to the overall binding strength or specificity of CF I by facilitating the cooperative binding of the two RRM-containing proteins. A similar function may be conserved in other organisms, such as *Drosophila melanogaster*, where mutations in the *RNA14* homolog, *suppressor of forked*, allow readthrough of weak polyadenylation signals (23), perhaps due to attenuated recognition of cis elements.

At least two interactions are preserved between homologues of yeast and mammals. We reassembled the Rna14-Rna15 heterodimer (14), which echoes the mammalian CstF77-CstF64 pair (24). The direct interaction we observed between Pcf11 and Clp1 confirms a reported two-hybrid interaction (25) and suggests that the coimmunoprecipitation and copurification (4) of the human homologs of these proteins also may be caused by a direct interaction.

Other interactions between sequence homologs are either not preserved from yeast to mammals or occur with different affinities. Whereas the yeast proteins Rna14, Rna15, Pcf11, and Clp1 remain tightly associated throughout the purification process as the CF I complex, their mammalian homologues do not (Fig. 5*B*). CstF64 and CstF77, the Rna15 and Rna14 homologs, respectively, are subunits of CstF (24). The human homologs of Clp1 and Pcf11, hClp1 and hPcf11, are components of CF II_M (4). CstF and CF II_M separate fairly early in the purification process (26), suggesting that any interaction between these two factors is easily disrupted. The most likely counterpart to the CstF50 subunit is the WD repeat protein Pfs2, a subunit of PF I (11), which does not copurify with CF IA (14), but does exhibit an *in vitro* interaction with Rna14.

Functional reconstitution of CF I resolves questions about the identity and necessity of its protein constitutents. In contrast to

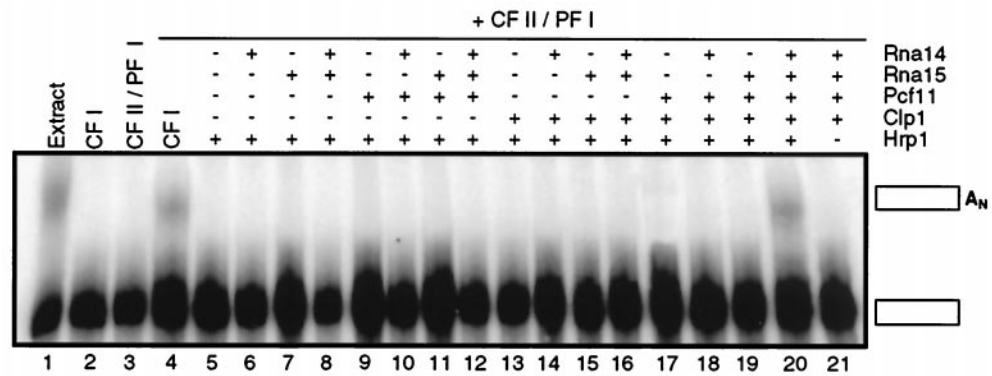

Fig. 4. Reconstitution of CF I polyadenylation activity requires the addition of all five peptides. Whole-cell extract from yeast strain FY86 polyadenylates precleaved precursor before fractionation on a Hi-Trap Q column. Reactions examined in lanes 2–21 were supplemented with the purified recombinant proteins Pab1 and PAP. A fraction containing both CF II and PF I activity can reconstitute polyadenylation activity in the presence of PAP and Pab1 proteins when mixed with either CF I obtained by fractionation of yeast whole-cell extract or all five peptide components of CF I (lanes 4 and 21). The CF I and CF II/PF I fractions are not sufficient for specific polyadenylation in the presence of PAP and Pab1 (lanes 2 and 3).

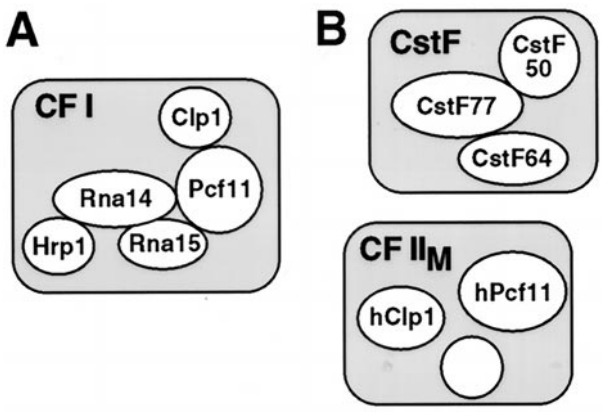

Fig. 5. Architectural model for CF I. (*A*) The proteins of CF I from *S. cerevisiae* are arranged based on direct protein–protein interactions observed in Fig. 1. (*B*) Organization of the mammalian homologs to yeast CF I subunits. CstF77 corresponds to Rna14 and CstF64 corresponds to Rna15. CstF50 is most likely homologous to Pfs2 of yeast PF I. No Hrp1 homolog has been found in mammals.

the results of Minvielle-Sebastia *et al.* (19), but in agreement with Kessler *et al.* (18), we find that CF IB is absolutely required not only for proper selection of the cleavage site, but also for reconstitution of even minimal cleavage activity upon the *GAL7* 3′ untranslated region. We confirmed that CF IB was required for reconstitution of polyadenylation activity upon this substrate, in agreement with published reports (18, 19).

We also have identified the full set of proteins comprising CF IA as being Rna14, Rna15, Pcf11, and Clp1. All four proteins are required for both the cleavage and polyadenylation activities of CF IA. The stringent recognition mechanism that defines and cleaves the nascent 3′ end in yeast is thus also required for polyadenylation, perhaps because yeast PF I may have decreased mRNA binding ability compared with CPSF and the signals specifying yeast 3′ end formation are much more divergent than their mammalian counterparts. Both CstF, which contains homologs to Rna14 and Rna15, and CF II_M, which contains homologs to Pcf11 and Clp1, are required only for cleavage, but not for polyadenylation. In mammals, CPSF possesses enough binding strength to tether PAP to the nascent 3′ end, but yeast requires the cooperativity of CF I and PF I. The association of CF I with PF I may strengthen the affinity of PF I for the mRNA. Yeast PF I contains the protein Yth1, which is a homologue of mammalian CPSF30 that preferentially binds poly(U) (27). However, Yth1 lacks a zinc knuckle present in CPSF30 that enhances its binding to mRNA. The requirement for CF IA and an additional RNA binding protein, Hrp1, in both cleavage and polyadenylation may be a compensation for the weaker mRNA binding of the Yth1 subunit of PF I. The necessity of complete CF I in both steps is a striking illustration of the functional divergence the polyadenylation machinery has undergone while maintaining significant homology at the level of individual proteins.

Pab1, although it is present in partially purified CF I (17), is not explicitly required for cleavage. EF1α, which was present in one CF I preparation (14), was not added in our reconstitution studies, but may be present in the CF II or PF I fractions that were used for reconstitution. However, given the lack of any evidence to date implicating this protein in mRNA processing, its cytoplasmic localization (28), and large abundance (29), this protein was probably a contaminant in the CF IA fraction and not a true component of CF IA.

A determination of the identity of the endonuclease as well as investigation of the actual mechanism of cleavage of the RNA precursor has been hampered by the inability to engineer components with phenotypic lethality. Future experiments with our reconstituted system will permit mapping of targeted mutations to the disruption of specific steps in 3′ end formation. Once reconstituted CF II and PF I become available, we anticipate elucidation of this machinery as well as a fuller exploration of its connection with other cellular processes.

We thank A. Zhelkovsky, D. Dawson, D. Pincus, L. Willis, C. Squires, and K. Sparks for critical reading of this manuscript. Also we are greatly indebted to A. E. Gross for her love and support at all points in this work. This work was supported by National Institutes of Health Grant GM41752 awarded to C.L.M.

1. Zhao, J., Hyman, L. & Moore, C. (1999) *Microbiol. Mol. Biol. Rev.* **63,** 405–445.
2. Moore, C. L. & Sharp, P. A. (1985) *Cell* **48,** 845–855.
3. Butler, J. S., Sadhale, P. & Platt, T. (1990) *Mol. Cell. Biol.* **10,** 2599–2605.
4. de Vries, H., Ruegsegger, U., Hubner, W., Friedlein, A., Lengen, H. & Keller, W. (2000) *EMBO J.* **19,** 5895–5904.
5. Chen, J. & Moore, C. L. (1992) *Mol. Cell. Biol.* **12,** 3470–3481.
6. Lingner, J., Radtke, I., Wahle, E. & Keller, W. (1991) *J. Biol. Chem.* **266,** 8741–8746.
7. Lingner, J., Kellermann, J. & Keller, W. (1991) *Nature (London)* **354,** 496–498.
8. Zhao, J., Kessler, M. & Moore, C. (1997) *J. Biol. Chem.* **272,** 10831–10838.
9. Preker, P., Ohnacker, M., Minvielle-Sebastia, L. & Keller, W. (1997) *EMBO J.* **16,** 4727–4737.
10. Barabino, S., Ohnacker, M. & Keller, W. (2000) *EMBO J.* **19,** 3778–3787.
11. Ohnacker, M., Barabino, S., Preker, P. & Keller, W. (2000) *EMBO J.* **19,** 37–47.
12. Zhelkovsky, A., Helmling, S. & Moore, C. (1998) *Mol. Cell. Biol.* **18,** 5942–5951.
13. Helmling, S., Zhelkovsky, A. & Moore, C. (2001) *Mol. Cell Biol.* **6,** 2026–2037.
14. Kessler, M. M., Zhao, J. & Moore, C. L. (1996) *J. Biol. Chem.* **271,** 27167–27175.
15. Minvielle-Sebastia, L., Preker, P. J. & Keller, W. (1994) *Science* **266,** 1702–1705.
16. Minvielle-Sebastia, L., Preker, P., Wiederkehr, T., Strahm, Y. & Keller, W. (1997) *Proc. Nat. Acad. Sci. USA* **94,** 7897–7902.
17. Amrani, N., Minet, M., LeGouar, M., Lacroute, F. & Wyers, F. (1997) *Mol. Cell. Biol.* **17,** 3694–3701.
18. Kessler, M., Henry, M., Gross, S., Shen, E., Zhao, J., Silver, P. & Moore, C. (1997) *Genes Dev.* **11,** 2545–2556.
19. Minvielle-Sebastia, L., Beyer, K., Krecic, A. M., Hector, R. E., Swanson, M. S. & Keller, W. (1998) *EMBO J.* **17,** 7454–7468.
20. Minvielle-Sebastia, L., Winsor, B., Bonneaud, N. & Lacroute, F. (1991) *Mol. Cell. Biol.* **11,** 3075–3087.
21. Amrani, N., Minet, M., Wyers, F., Dufour, M. E., Aggerbeck, L. P. & Lacroute, F. (1997) *Mol. Cell. Biol.* **17,** 1102–1109.
22. Henry, M., Borland, C., Bossie, M. & Silver, P. (1996) *Genetics* **142,** 103–115.
23. Audibert, A. & Simonelig, M. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 14302–14307.
24. Takagaki, Y., Manley, J. L., MacDonald, C. C., Wilusz, J. & Shenk, T. (1990) *Genes Dev.* **4,** 2112–2120.
25. Uetz, P., Giot, L., Cagney, G., Mansfield, T., Judson, R., Knight, J., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., *et al*. (2000) *Nature (London)* **403,** 623–627.
26. Takagaki, Y., Ryner, L. C. & Manley, J. L. (1989) *Genes Dev.* **3,** 1711–1724.
27. Barabino, S. M., Hubner, W., Jenny, A., Minvielle-Sebastia, L. & Keller, W. (1997) *Genes Dev.* **11,** 1703–1716.
28. Herrera, F., Correia, H., Triana, L. & Fraile, G. (1991) *Eur. J. Biochem.* **200,** 321–327.
29. Norbeck, J. & Blomberg, A. (1997) *Yeast* **13,** 1519–1534.