A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor

Pascal J. Preker¹, Martin Ohnacker,
Lionel Minvielle-Sebastia and Walter Keller²

Department of Cell Biology, Biozentrum, University of Basel,
Klingelbergstrasse 70, CH-4056 Basel, Switzerland

¹Present address: Department of Biochemistry and Biophysics,
UCSF School of Medicine, San Francisco, CA 94143-0448, USA

²Corresponding author

Polyadenylation is the second step in 3' end formation of most eukaryotic mRNAs. In *Saccharomyces cerevisiae*, this step requires three *trans*-acting factors: poly(A) polymerase (Pap1p), cleavage factor I (CF I) and polyadenylation factor I (PF I). Here, we describe the purification and subunit composition of a multi-protein complex containing Pap1p and PF I activities. PF I–Pap1p was purified to homogeneity by complementation of extracts mutant in the Fip1p subunit of PF I. In addition to Fip1p and Pap1p, the factor comprises homologues of all four subunits of mammalian cleavage and polyadenylation specificity factor (CPSF), as well as Pta1p, which previously has been implicated in pre-tRNA processing, and several as yet uncharacterized proteins. As expected for a PF I subunit, *pta1-1* mutant extracts are deficient for polyadenylation *in vitro*. PF I also appears to be functionally related to CPSF, as it polyadenylates a substrate RNA more efficiently than Pap1p alone. Possibly, the observed interaction of the complex with RNA tethers Pap1p to its substrate.

Keywords: polyadenylation factor I/poly(A) polymerase holoenzyme/pre-messenger RNA 3' end processing/yeast

Introduction

The formation of the 3' ends of eukaryotic mRNAs occurs in two tightly coupled steps and requires a complex set of *cis*- and *trans*-acting factors. In mammals, the cleavage and polyadenylation specificity factor (CPSF) is a key component in this process, as it is required for both steps, endonucleolytic cleavage of the primary transcript and polyadenylation of the upstream cleavage product by the enzyme poly(A) polymerase (PAP). CPSF is a multimeric protein complex (Bienroth *et al.*, 1991) that binds the conserved hexanucleotide sequence AAUAAA upstream of the cleavage site, by virtue of its largest subunit (Keller *et al.*, 1991; Murthy and Manley, 1995). CPSF binding occurs cooperatively with the cleavage stimulation factor (CstF), which recognizes a less well-defined U- or GU-rich sequence downstream of the cleavage site through its 64 kDa subunit (Takagaki *et al.*, 1992; MacDonald *et al.*, 1994). In addition, cleavage requires the presence of PAP and two cleavage factors, CF I_m and CF II_m (Takagaki

et al., 1989; Rügsegger *et al.*, 1996). After cleavage, the poly(A) tail is added by PAP in a processive event that requires CPSF and a poly(A)-binding protein (PAB II) (Wahle, 1991; Bienroth *et al.*, 1993). The genes encoding PAP, PAB II and the subunits of CPSF and CstF have been cloned and sequenced (reviewed in Wahle and Keller, 1996; see also Jenny *et al.*, 1996; Barabino *et al.*, 1997).

The observation that RNA sequences directing 3' end formation are far more redundant and degenerate in yeast than they are in mammals led to the view that the factors that specify the mRNA 3' end are distinct. However, more recent findings confirm a fundamental structural and functional conservation between the 3' end processing machineries of these distantly related eukaryotes. Exhaustive analyses of the sequences required for accurate 3' end formation of the iso-1-cytochrome *c* (*CYC1*) mRNA has led to the identification of two essential elements, both located upstream of the cleavage site. These are a 'positioning element' close to the cleavage site and a distal 'efficiency element' (Guo and Sherman, 1996, and references therein). The sequence AAUAAA can function as the positioning element, but other sequences work as well. Fractionation of *trans*-acting 3' end processing factors was pioneered by Chen and Moore (1992). These authors separated yeast whole-cell extracts into four chromatographic fractions: poly(A) polymerase (Pap1p), which is encoded by *PAP1* (Lingner *et al.*, 1991a,b), CF I and CF II, and polyadenylation factor I (PF I). In reconstituted *in vitro* systems, cleavage requires CF I and CF II, whereas polyadenylation occurs upon combination of CF I, Pap1p and PF I. The first genes besides *PAP1* that were demonstrated to encode subunits of a yeast 3' end processing factor were *RNA14* and *RNA15* (L. Minvielle-Sebastia *et al.*, 1994). Rna14p and Rna15p are components of CF I, and mutations in either gene abolish both cleavage and polyadenylation activity (Minvielle-Sebastia *et al.*, 1994). Further purification revealed that CF I can be separated into two activities (CF IA and CF IB) that are both required for cleavage and polyadenylation (Kessler *et al.*, 1996). We independently have purified CF IA and found that five polypeptides co-fractionate with the activity. They include Rna14p, Rna15p, Pcf11p (Amrani *et al.*, 1997), a new protein called Clp1p (L. Minvielle-Sebastia *et al.*, unpublished data) and remarkably, the major poly(A)-binding protein Pab1p (Minvielle-Sebastia *et al.*, 1997). *PCF11* initially was found in a two-hybrid screen designed to identify proteins interacting with Rna14p and Rna15p (Amrani *et al.*, 1997). *RNA14*, *RNA15* and *PCF11* also interact genetically, as combinations of temperature-sensitive mutations in either of them are synergistically lethal (Minvielle-Sebastia *et al.*, 1994; Amrani *et al.*, 1997). The Rna15p subunit of purified CF IA contacts the pre-mRNA, albeit with unknown specificity. The protein has an RNA-binding domain with similarity to that of the

64 kDa subunit of mammalian CstF. Moreover, Rna14p is significantly related to the 77 kDa subunit of CstF (Takagaki and Manley, 1994).

A screen for proteins that interact with Pap1p *in vivo* led to the identification of Fip1p as an essential component of PF I (Preker *et al.*, 1995). Two additional subunits of PF I were identified, initially based on their homology to subunits of mammalian CPSF (Chanfreau *et al.*, 1996; Jenny *et al.*, 1996; Barabino *et al.*, 1997). Ysh1p (yeast 73 kDa homologue) and Yth1p (yeast 30 kDa homologue) are 53 and 40% identical to the 73 and 30 kDa subunits of bovine CPSF, respectively (hence their names). Surprisingly, Cft1p, a putative homologue of the largest subunit of CPSF, has been reported to be a subunit of CF II (Stumpf and Domdey, 1996). Thus, homologues of different subunits of CPSF, which is required for both cleavage and polyadenylation, appear to be associated with two separate yeast factors that are required for either of the two activities only.

Here we report the purification of a multimeric complex required for pre-mRNA polyadenylation in yeast and the identification of seven polypeptides associated with its activity. These include the putative homologues of all four subunits of CPSF, as well as Fip1p, Pap1p and Pta1p, a protein that previously has been implicated in pre-tRNA maturation (O'Connor and Peebles, 1992). All subunits characterized to date are essential, underscoring the important role of this factor in gene expression.

Results

Purification of the multimeric PF I complex

To facilitate purification of PF I, a strain (PJP14) was constructed in which a disruption of the chromosomal *FIP1* gene was rescued by expression of a tagged form of Fip1p provided on a plasmid. The tag was fused in-frame to the amino-terminus of Fip1p and consisted of the influenza haemagglutinin HA1 epitope followed by six consecutive histidine residues (Figure 1A). The tag did not adversely affect cell growth or polyadenylation activity, and, thus, PF I function, in cell-free extracts. Ammonium sulfate-fractionated extract obtained from 500 g of PJP14 cells was first subjected to chromatography on a Macro-Prep Q column. CF I- and PF I-containing fractions were identified by their ability to complement the 3' end processing defect of *rna14-1* and *fip1-1* mutant extracts, respectively (Minvielle-Sebastia *et al.*, 1994; Preker *et al.*, 1995). The two activities eluted separately at ~150 and 270 mM salt, respectively. PF I activity was purified further by chromatography on Blue-Sepharose and heparin-Sepharose columns. Western blot analysis of the complementing fractions confirmed that Fip1p always co-purified with PF I activity. The pool of activity from the heparin-Sepharose column was loaded on a Ni²⁺-nitrilotriacetic acid (NTA)-agarose column for affinity purification of the His₆-tagged protein complex. This step resulted in a 5-fold purification only (Figure 1B), possibly because the six histidine residues, located between the HA1 tag and the Fip1p sequence, are not easily accessible. PF I activity was purified further by chromatography on a Mono S column. From the Macro-Prep Q column to this step, recovery of activity was ~70% and PF I was purified ~90-fold (Figure 1B).

Fig. 1. Purification of yeast polyadenylation factor I. (A) Schematic representation of strain PJP14 made for PF I purification. The chromosomal *FIP1* gene is replaced by a double-tagged allele provided on a centromeric plasmid. The sequence of the HA1 epitope is printed in bold. Centromere sequences are represented by open circles. (B) Purification of PF I by conventional and metal chelate affinity chromatography. Protein concentrations were determined with a colorimetric assay (Bio-Rad) with bovine serum albumin (BSA) as a standard. One unit of PF I is defined as the amount of protein that polyadenylates 1 fmol of the 5' product generated by cleavage of the *CYC1* pre-mRNA in a *fip1-1* mutant extract per minute. Assays contained 20 fmol of ³²P-labelled substrate, and the amount of polyadenylated end product was calculated with IPLab Gel software (version 1.5, Signal Analytics Corporation) after scanning of gels on a PhosphorImager 425 (Molecular Dynamics). All the values obtained were corrected by subtracting the background value determined from an empty lane. Because the crude extract and the ammonium sulfate-fractionated extract are competent for polyadenylation by themselves, they are not amenable to quantification. (C) Immunoprecipitation of PF I. A sample from the Mono S pool containing partially purified PF I was pre-adsorbed to protein A-Sepharose and subsequently subjected to immunoprecipitation with a control antibody (rabbit anti-mouse immunoglobulins; lane 5), the monoclonal antibody 12CA5, specific for the HA1-peptide (lane 7), or a polyclonal antibody to Fip1p (lane 9). Bound proteins were eluted from the pre-adsorption resin (lane labelled no ab) and the antibody resins by boiling in sample buffer. The eluate was analysed by electrophoresis on an SDS-polyacrylamide (9%) gel and silver staining. A sample of each antibody was loaded in adjacent lanes. Lane 1, 20 µl of the input fraction; lane 2, 20 µl of the supernatant (sn) after precipitation with affinity-purified antibodies to Fip1p. Major polypeptides in the eluates are labelled according to their apparent molecular mass. The polyclonal antibodies contain exogenous BSA. The positions of the light and heavy chains of the immunoglobulins and of molecular mass standards (M) are indicated; sizes are in kilodaltons.

Fip1p and its associated proteins were immunoprecipitated from a PF I fraction of the Mono S column. Anti-HA1 or anti-Fip1p antibodies were immobilized on protein

A-resin and mixed with a sample of the PF I fraction. After extensive washing, bound proteins were eluted and separated by SDS-PAGE (Figure 1C, lanes 7 and 9). The anti-HA1 and anti-Fip1p antibodies immunoprecipitated a virtually identical set of proteins. None of these proteins was retained in control experiments with an unrelated antibody (lane 5) nor with antibodies to Fip1p that had been pre-adsorbed to the antigen prior to immunoprecipitation (data not shown). The presence of the immunoglobins in the eluate precluded detection of any proteins with M_r s of ~50 and 30 kDa on the silver-stained gel. Pre-treatment of the input fraction with RNase A did not affect the composition of proteins in the immunoprecipitate, suggesting that the integrity of the complex does not require an RNase A-sensitive component (data not shown). Finally, a similar set of proteins was also found in a distinct purification scheme that did not involve immunoaffinity purification (data not shown). The supernatant and the eluate of the anti-HA1 antibody precipitation were probed on blots with the anti-Fip1p antibody and vice versa. In both cases, Fip1p was only detected in the eluate, indicating that the antibodies had adsorbed the protein completely (data not shown).

To recover PF I in its native form, Mono S fractions were loaded on an anti-HA1 antibody column, and PF I was eluted with competing HA1 peptide (Field *et al.*, 1988; Keys *et al.*, 1994). No signal was seen if immunoblots of the eluate were probed with antibodies to mouse immunoglobins, indicating that the anti-HA1 antibody had not desorbed from the column in detectable amounts (data not shown). The immunopurified fractions were analysed for their protein content by gel electrophoresis on an SDS-polyacrylamide gel and silver staining (Figure 2B, lane 2). The five largest proteins with M_r s of 150, 105, 100, 85 and 64 kDa were similar to proteins observed in the previous immunoprecipitation experiments (Figure 1C). Other major polypeptides had apparent M_r s of 58, 55, 53, 36 and 35 kDa. To confirm that the immunoprecipitated proteins co-migrate with PF I activity, a fraction of the eluate was subjected to chromatography on a Mini Q column. The composition of the complex did not change due to chromatography after immunopurification, indicating that the factor has been purified to near homogeneity. All polypeptides were recovered at apparently equimolar ratios with the exception of the 55 (see below) and the 58 kDa subunits. Fractions were assayed for their ability to restore polyadenylation of *fip1-1* mutant extract (Figure 2C). Polyadenylation activity co-eluted with the protein peak in fractions 24–26 (Figure 2B, lanes 5 and 6 and C, lanes 9 and 10). In addition, polyadenylation of the *CYC1* pre-cleaved precursor could be obtained upon combination of partially purified CF I (Chen and Moore, 1992) and purified PF I from the same Mini Q column fractions (Figure 2D, lanes 4 and 7–16). No exogenous poly(A) polymerase was required. The presence of the poly(A) polymerase in PF I fractions explained this surprising result (see below).

Although the immunopurification step only recovered a fraction of the PF I activity in the eluate, this step may afford another 20- to 50-fold purification as judged by polyacrylamide gel electrophoresis (Figure 1C).

Fip1p, Pap1p, Pta1p and homologues of the four subunits of CPSF are associated with PF I activity

Protein fractions of the final Mini Q column were tested for cross-reactivity with several antibodies. Antibodies to Fip1p detected a protein that corresponded to a polypeptide of 55 kDa. The apparent absence of stoichiometry of Fip1p is reproducible and may be due to the inherently poor staining of Fip1p with silver or partial degradation of the protein.

Affinity-purified antibodies to Pap1p detected a protein of 64 kDa in active fractions (Figure 2A). This is surprising because previous studies have indicated that Pap1p and PF I elute from a Mono Q column as separate factors (Chen and Moore, 1992). That the 64 kDa protein detected by Western blot analysis is Pap1p was confirmed by several experiments. First, immunoprecipitation with the anti-Pap1p antibodies precipitated the same set of proteins as the anti-Fip1p antibodies (data not shown). Second, Pap1p purified from recombinant *Escherichia coli* (Lingner *et al.*, 1991a) had an electrophoretic mobility indistinguishable from that of the 64 kDa subunit of PF I (data not shown). Third, in the absence of other factors, the PF I fractions unspecifically polyadenylated the *CYC1* full-length and pre-cleaved pre-mRNA substrates (Figure 2C and D, lanes 5, respectively) or a poly(A) primer (data not shown). As observed with the authentic enzyme, poly(A) is elongated much more efficiently in the presence of manganese ions than in the presence of magnesium ions (data not shown). Finally, the notion that Pap1p is a component of PF I is consistent with the observation that specific polyadenylation of a pre-cleaved pre-mRNA substrate can be reconstituted from purified PF I–Pap1p and partially purified CF I from a Mono Q column (Chen and Moore, 1992) in the absence of exogenous Pap1p (see Figure 2D). Remarkably, small amounts of the 64 kDa subunit began to elute from fraction 10 of the Mini Q column, ahead of the PF I peak (Figure 2B, lane 3; and data not shown). However, the Pap1p peak coincided with that of the other PF I subunits and thus with PF I activity.

A yeast homologue of the 73 kDa subunit of the mammalian CPSF recently has been cloned and shown to be a subunit of PF I (Jenny *et al.*, 1996). Ysh1p has a coding capacity for a 87 kDa protein (note that the same gene was cloned independently as *BRR5* by Chanfreau *et al.*, 1996). Antibodies directed against Ysh1p recognized a polypeptide of 100 kDa in fractions containing purified PF I (Figure 2A).

In the same way that we characterized Ysh1p, the cloning of the 30 kDa subunit of CPSF allowed us to identify a yeast protein of 26 kDa which showed 40% identity with the mammalian protein. It was therefore called Yth1p (Barabino *et al.*, 1977). Extracts prepared from a *yth1* mutant strain show normal cleavage but are deficient in polyadenylation (Barabino *et al.*, 1977). Although proteins smaller than 28 kDa are not shown on the protein gel in Figure 2B, no additional major polypeptides were observed on gels that resolved proteins as small as ~15 kDa (data not shown). However, antibodies against Yth1p clearly recognized a band of the expected size (26 kDa) in fractions containing PF I activity (Figure 3B, lanes 2 and 3; Barabino *et al.*, 1997). As observed with Fip1p and the 58 kDa protein, it is possible that Yth1p does not stain well with silver. Nevertheless, results

obtained with *yth1* mutant extracts clearly showed that this polypeptide behaves as an actual subunit of PF I.

Microsequences of proteolytic fragments of the 85, 105 and 150 kDa subunits of PF I were obtained (see Materials and methods). Five peptides derived from the 150 kDa subunit are identical to a predicted protein of 153 kDa encoded by an open reading frame on chromosome IV (systematic name: YDR301w, SwissProt accession No.

S61187). Three peptides from the 105 kDa protein match a predicted protein of 96 kDa (YLR115w, S64952). Strikingly, these proteins are 23.5 and 24.4% identical to the 160 and 100 kDa subunits of CPSF (Jenny *et al.*, 1994; Jenny and Keller, 1995; Murthy and Manley, 1995), respectively. Although less obvious than for Ysh1p and Yth1p, these homologies are statistically significant [error probabilities $P \leq 6.3 \times 10^{-23}$ (Altschul *et al.*, 1994)] and lend strong support to the hypothesis that the two genes encode actual subunits of PF I. The 105 kDa protein is also significantly related to Ysh1p and to CPSF-73 (Jenny *et al.*, 1996). The two genes were therefore tentatively named *YHH1* (yeast 160 kDa homologue 1) and *YDH1* (yeast double homologue 1). While this manuscript was in preparation, a report appeared showing that Yhh1p is required for pre-mRNA 3' end formation in yeast (Stumpf and Domdey, 1996). Based on immunodepletion experiments, the authors concluded that the protein is a subunit of CF II and thus named it cleavage factor two 1 protein (Cft1p; see Discussion). Therefore, we will hereafter refer to the yeast homologue of CPSF-160 as 'Yhh1p/Cft1p'. As expected, a polyclonal antibody raised against the carboxy-terminus of Yhh1p/Cft1p (Stumpf and Domdey, 1996) recognized the 150 kDa subunit of PF I in peak fractions of the Mini Q column (see Figure 3B).

Standard genetic analyses demonstrated that *YDH1* and *YHH1/CFT1* are essential for cell viability (see Materials and methods). *YDH1*, cloned from an *S.cerevisiae* genomic plasmid bank, could rescue a disruption of that gene when provided on a single-copy plasmid. We repeatedly failed to clone *YHH1/CFT1* from the same library by using a probe corresponding to amino acids 1–379 of the protein. Possibly, this gene is detrimental to bacterial growth.

Fig. 2. PF I is a multiprotein complex containing Pap1p, Pta1p, Fip1p and homologues to CPSF subunits. (A) Western blot analysis across the PF I activity peak of the Mini Q column. Proteins were detected with polyclonal antibodies to Ysh1p, Pap1p and Fip1p. (B) Subunit composition of PF I. Aliquots (20 µl) of the Mini Q column fractions indicated at the top of the panel were separated on an SDS-polyacrylamide (9%) gel and visualized by silver staining. Proteins that correspond in size to proteins detected by immunoblotting (A) or that have been identified by peptide sequencing are indicated. Other major polypeptides are designated according to their apparent molecular mass. (C) PF I activity of the Mini Q fractions shown in (B). Of the indicated Mini Q fractions, 4 µl of each were mixed with 3 µl of an ammonium sulfate-fractionated *fip1-1* extract and 32 P-labelled *CYC1* pre-mRNA under standard reaction conditions at 30°C for 45 min. Hatched and open boxes represent the 5' and 3' portions of the precursor, respectively. Control reactions were done in parallel with the following components: lane 1, dialysis buffer (G-50); lane 2, 3 µl of a wild-type extract; lane 3, 3 µl of *fip1-1* extract; lane 4, as lane 3, but containing 3 µl of a crude PF I fraction from the Macro-Prep Q column; lane 5, 4 µl of the peak fraction (# 24) from the Mini Q column; lane 6, 3 µl of the load fraction and 3 µl of *fip1-1* extract. (D) Reconstitution of specific polyadenylation activity with partially purified CF I and PF I fractions of the Mini Q column. For assays, 4 µl of the PF I fractions shown in (B) and 3 µl of partially purified CF I were incubated with the *CYC1* pre-cleaved transcript (represented by a hatched box) under standard reaction conditions. Lane 1, unreacted *CYC1* pre-cleaved; lane 2, 3 µl of a wild-type extract; lane 3, 3 µl of CF I only; lane 4, 3 µl of CF I and 4 µl of Mini Q fraction 24; lane 5, 4 µl of Mini Q fraction 24 only; lane 6, 3 µl of CF I and 4 µl of the load fraction. In (C) and (D) the reaction products were resolved on a 5% polyacrylamide–8.3 M urea gel and exposed to X-ray film for 12 h. Markers (M) are *Hpa*III-digested pBR322 DNA. Their sizes, in number of nucleotides, and positions are indicated on the right.

Fig. 3. Mutant *pta1-1* extracts are deficient for polyadenylation and can be complemented by PF I fractions containing Fip1p and homologues to CPSF subunits. (A) The *in vitro* phenotype of *pta1-1* mutant extracts is similar to that of *fip1-1* mutant extracts. Three to six μ l of extract or buffer and 1 μ l of a Mono S fraction containing PF I activity were combined with 32 P-labelled *CYC1* substrate RNA as indicated at the top of each lane, and incubated under standard reaction conditions for 40 min. Reaction products were separated by denaturing gel electrophoresis, and the gel was exposed to film, as described in the legend to Figure 2. Lower case italicized labels at the top refer to mutant extracts. PF I refers to the partially purified protein. Markers (M) are as in Figure 2. (B) The complementing PF I fractions contain Fip1p, Yhh1p/Cft1p, Ysh1p and Yth1p. PF I fractions (lanes 2 and 3) and CF II/CF IB fractions (lanes 1 and 4; hydroxyapatite column, Minvielle-Sebastia *et al.*, 1997) were electrophoresed on an SDS-polyacrylamide (10%) gel, blotted to membrane and probed with antibodies to Yhh1p/Cft1p, Ysh1p and Yth1p (lanes 1 and 2). After stripping, the blot was reprobed with antibodies to Ysh1p, Yth1p and Fip1p (lanes 3 and 4). A polypeptide with a higher electrophoretic mobility was reproducibly detected with antibodies to Fip1p. This protein is probably a proteolytic form of Fip1p and was therefore denoted Fip1p* (see text).

Peptide microsequencing of the 85 kDa subunit of PF I revealed that this protein is the product of the essential gene *PTA1* (coding capacity 88 kDa). *PTA1* initially was defined by a conditional growth mutation, *pta1-1*, that causes the accumulation of unspliced pre-tRNAs *in vivo* (O'Connor and Peebles, 1992). All 10 intron-containing tRNA families are affected. Surprisingly, extracts prepared from *pta1-1* cells have normal pre-tRNA splicing endo-

nuclease activity (O'Connor and Peebles, 1992). To test whether the *PTA1* gene product is also involved in 3' end formation, extracts from a *pta1-1* mutant strain were assayed for their ability to cleave and polyadenylate *CYC1* pre-mRNA substrates *in vitro*. Whereas cleavage activity was normal, the extracts failed to polyadenylate both the upstream cleavage product (Figure 3A, lane 6) and a pre-cleaved *CYC1* RNA (data not shown). In contrast, extracts from the *pta1-1* strain transformed with the wild-type gene borne on a single-copy plasmid polyadenylated the substrate RNA with efficiency comparable with the wild-type extract (Figure 3A, compare lanes 2 and 3). Processing activity could be restored efficiently by addition of extracts mutant in CF I, which on their own neither cleave nor polyadenylate (compare lanes 5 and 8), but not by extracts of mutants in the Fip1p subunit of PF I (lane 7). It thus appears that Pta1p and Fip1p are tightly associated in a complex required for polyadenylation. Most importantly, polyadenylation activity could be restored to *pta1-1* mutant extracts by addition of PF I from the Mono S column (Figure 3A, lane 11) or the Mini Q column (data not shown). Differences in the length of the polyadenylated products after complementation of *fip1-1* and *pta1-1* extracts were not reproducible (lanes 8, 10 and 11).

The complementing fractions contained Fip1p, Yth1p, Ysh1p, Yhh1p/Cft1p (Figure 3B, lanes 2 and 3) and Pap1p (not shown), as determined by immunoblotting. In contrast, CF II/CF IB fractions do not contain any of these proteins (lanes 1 and 4; see Discussion). In addition, the anti-Fip1p antibody detected a faster migrating protein species (lane 3), referred to here as Fip1p*. Fip1p* had been observed throughout the entire purification, but its abundance relative to the full-length protein varied. Because this species is also recognized by the anti-HA1 antibody, it probably represents a C-terminally truncated form of the protein.

PF I modulates the polyadenylation activity of Pap1p associated with it

The purified PF I–Pap1p complex polyadenylates RNA in the absence of other factors and without specificity for a genuine pre-mRNA 3' end (Figures 2C, lane 5, and 3A, lane 9). We compared the kinetics of poly(A) addition with *CYC1* pre-mRNA by either the PF I–Pap1p complex or Pap1p alone. Time course experiments were done with immunoaffinity-purified PF I, containing ~10 ng of Pap1p (final concentration ~3 nM), or the same amount of the recombinant enzyme (Figure 4). Recombinant Pap1p elongated the primer at a slow and uniform rate to a final average tail length of ~200 adenosine residues after 1 h. In stark contrast, polyadenylation by Pap1p in the PF I complex was fast. After 4 min, poly(A) tails were ~400 residues long. After 32 min, no further elongation could be detected owing to the limited resolution capacity of the gel. Another remarkable difference is that, in contrast to what was observed with Pap1p alone, only a minor portion of the RNA was used as a substrate by the poly(A) polymerase holoenzyme (Figure 4, compare lanes 6 and 12). Two possible explanations may account for this result. First, the RNA is polyadenylated more processively when Pap1p is in a complex with the other components of PF I. Second, a substrate RNA is elongated preferentially once it has received a minimal number of adenosine residues,

Fig. 4. Poly(A) polymerase activity is modulated by other components of PF I. Pap1p activity was assessed in standard reaction mixtures containing (in 50 μ l) 100 fmol of 32 P-labelled *CYC1* pre-mRNA and either 200 fmol of immunoaffinity-purified PF I-Pap1p (lanes 1–6) or 200 fmol (13 U) of recombinant Pap1p (lanes 7–12). Aliquots (8 μ l) were withdrawn from each tube (0 min time point), and the remainder was pre-warmed briefly to 30°C. The reaction was started by addition of ATP to a final concentration of 0.5 mM. Samples (8 μ l) were removed at the time points indicated at the top and diluted into proteinase K digestion buffer (Minvielle-Sebastia *et al.*, 1994). The RNA was extracted from the aliquots and analysed on a denaturing 5% polyacrylamide gel. The size of DNA markers (M) in number of nucleotides is indicated on the left.

as in mammals where CPSF and PAB II stimulate the activity of the poly(A) polymerase after oligoadenylation of the precursor (Wahle, 1991; Bienroth *et al.*, 1993). Because the authentic poly(A) polymerase and the recombinant enzyme purified from *E.coli* have indistinguishable biochemical and enzymatic properties (Lingner *et al.*, 1991a), we consider it unlikely that the differences are attributable to an intrinsic property of the recombinant enzyme. When combined with purified CF IA and CF IB/CF II fractions, the PF I-Pap1p complex was fully active in the specific cleavage and polyadenylation reaction with *CYC1* pre-mRNA (Minvielle-Sebastia *et al.*, 1997). It also specifically polyadenylated a pre-cleaved RNA substrate when combined with partially purified CF I (see Figure 2D).

Discussion

Complexity of yeast 3' end processing factors

PF I was purified from a strain expressing epitope-tagged Fip1p. In addition to Fip1p, PF I comprises Pap1p, Ptalp and, at a minimum, eight other protein subunits. The finding that Fip1p is stably associated in a complex agrees with the observations that recombinant Fip1p was neither able to substitute for PF I in the reconstituted *in vitro* system nor to complement polyadenylation-deficient *fip1* extracts (P.J.Preker, unpublished results).

The purified PF I-Pap1p complex can specifically

polyadenylate a pre-cleaved pre-mRNA substrate when combined with partially purified CF I, with no additional poly(A) polymerase. Our results differ from those of earlier work indicating that PF I and Pap1p are two separable factors, each one required individually to restore specific polyadenylation activity (Chen and Moore, 1992). We do not know if these discrepancies are attributable to variations in the protocols used for extract preparation and biochemical fractionation or might reflect the different genetic backgrounds of the strains used in these studies. Remarkably, the previously reported purification of Pap1p from yeast did not result in the identification of any additional polypeptides (Lingner *et al.*, 1991b). Because this purification relied solely on the unspecific polyadenylation activity of the enzyme, any polyadenylation activity associated with PF I may have escaped detection. Early work of Haff and Keller (1975) demonstrated that Pap1p activity of yeast cell extracts can be separated into distinct peaks by chromatography on DEAE-cellulose. A minor portion of the activity eluted at high salt concentrations after the major peak of activity. Possibly, this portion corresponded to Pap1p in a complex with PF I.

Of the other polypeptides that co-purified with PF I activity, one was identified by Western blot analysis as Ysh1p, a protein that is highly homologous to mammalian CPSF-73 (Chanfreau *et al.*, 1996; Jenny *et al.*, 1996). Also, Western blot analysis confirmed the presence of Yth1p (the homologue of CPSF 30 kDa subunit) in this factor even though no polypeptide with an M_r <35 kDa was visible in highly purified PF I preparations. The biochemical characteristics of Ysh1p and Yth1p are consistent with these proteins being genuine subunits of PF I. Antibodies directed against the recombinant proteins depleted PF I activity from whole-cell extracts or PF I fractions. In both cases, activity could be restored by adding back purified PF I (Jenny *et al.*, 1996; Barabino *et al.*, 1997). Moreover, extracts from conditional *ysh1/brr5* or *yth1* mutants are deficient for polyadenylation but not for cleavage (Chanfreau *et al.*, 1996; Barabino *et al.*, 1997).

The two largest subunits associated with PF I activity are significantly related over the entire sequence length to the 160 and 105 kDa subunits of CPSF, respectively. Therefore, it is highly likely that these proteins play an essential role in 3' end formation as well. Yhh1p/Cft1p, the CPSF-160 homologue, recently has been reported to be a subunit of CF II (Stumpf and Domdey, 1996). The authors showed that antibodies to Cft1p immunodeplete cleavage and polyadenylation activity from wild-type extracts. The two activities could be complemented by addition of Mono Q fractions containing CF II and PF I activity, respectively. In Western blot experiments, Cft1p was found associated with CF II activity. However, the two activities largely overlapped in the column fractions used for complementation (Stumpf and Domdey, 1996). Although we cannot rigorously rule out the possibility that the CPSF-related polypeptides are involved in cleavage *in vivo*, we found no evidence that this is the case in our reconstituted *in vitro* system. The fractions required for cleavage, purified CF IA and partially purified CF II/CF IB (Minvielle-Sebastia *et al.*, 1997), do not contain detectable amounts of Yth1p, Ysh1p or Yhh1p (Figure 3B). Moreover, cleavage can be reconstituted efficiently

without PF I. In any case, the observation that immunodepletion with antibodies to Pap1p or PF I subunits, such as Fip1p and Ysh1p/Brr5p, also impairs cleavage to various degrees (Minvielle-Sebastia *et al.*, 1994; Preker *et al.*, 1995; Chanfreau *et al.*, 1996) suggests that cleavage factors and PF I may be directly associated and that some aspects of this association may be maintained *in vitro*.

The finding that the 85 kDa subunit of PF I is identical to Pta1p (O'Connor and Peebles, 1992) adds yet another layer of complexity to the mechanism of yeast 3' end formation. *In vitro*, *pta1-1* mutants have a 3' end processing defect very similar to that of *fip1-1*, *ysh1/brr5* or *yth1-1* mutants. Pta1p has been implicated in pre-tRNA splicing, because *pta1-1* mutants accumulate all 10 end-trimmed, intron-containing pre-tRNA families *in vivo* (O'Connor and Peebles, 1992). However, the mutant exhibits no pre-tRNA splicing defect *in vitro*. Similar phenotypes are common to a number of other mutants, such as nucleoporin mutants (Simos *et al.*, 1996) and the pleiotropic *rnal-1* mutant (Hopper *et al.*, 1978). It has been speculated that pre-tRNA processing might be coupled to mRNA export (Simos *et al.*, 1996). Likewise, Pta1p (and so, possibly, other components of PF I) may have overlapping functions in both tRNA and mRNA maturation. Alternatively, the accumulation of unspliced pre-tRNAs in *pta1-1* mutants might be a secondary effect of a reduced 3' end processing efficiency in these mutants. To this end, we have used an *in vivo* assay based on loss of suppressor tRNA activity (Simos *et al.*, 1996) to test whether mutants in 3' end processing factors are generally impaired in tRNA processing. However, we failed to detect any significant decrease in suppressor tRNA activity in an *rnal4-1*, *rnal5-1* or *fip1-1* mutant background (P.J.Preker, unpublished data).

We showed that purified PF I exhibits non-specific polyadenylation activity on its own that is significantly more efficient than that of Pap1p alone. A plausible explanation of how the other subunits of PF I stimulate the activity of Pap1p could be that the holoenzyme contains an RNA-binding component that promotes its interaction with the RNA substrate. In fact, immunoaffinity-purified PF I can form a complex with the *CYC1* pre-mRNA with a lower electrophoretic mobility in non-denaturing gels (results not shown). In contrast, free Pap1p does not form a shifted complex under the same conditions. The peak of the RNA-binding activity coincided with that of PF I activity when fractions of the final Mini Q column were assayed for complex formation. This indicates that PF I contains at least one RNA-binding component. In this respect, recombinant Yth1p has been shown to have unspecific RNA-binding activity *in vitro* (Barabino *et al.*, 1997). Whether or not additional subunits of PF I are binding to RNA remains to be determined. However, the binding of PF I–Pap1p to the RNA is not specific for 3' end processing-competent substrates because an RNA–protein complex was also formed with a *cyc1-512* mutant RNA, which is not processed *in vitro* (results not shown). Similarly, CF IA interacts with wild-type and mutant pre-mRNAs (Kessler *et al.*, 1996; L.Minvielle-Sebastia, P.J.Preker, Y.Strahm and W.Keller, unpublished data). Thus, sequence-specific RNA binding may require different conditions and/or additional factors.

Evolutionary conservation of 3' end formation in eukaryotes

Recently, Murthy and Manley (1995) have reported that CPSF-160 interacts with both bovine PAP and the 77 kDa subunit of CstF. In view of the finding that the Fip1p subunit of yeast PF I tethers Pap1p to the Rna14p subunit of CF I (Preker *et al.*, 1995), these authors speculated that Fip1p might be the functional homologue of the CPSF 160 kDa subunit. An extension of this idea is that PF I and CPSF may have similar functions in the two organisms. The sequence homologies of PF I components to the four subunits of mammalian CPSF strengthens this hypothesis. Sequence similarities have also been reported between the Rna14p and Rna15p subunits of yeast CF I and, respectively, the 77 and 64 kDa subunits of mammalian CstF (Takagaki and Manley, 1994). Thus, 3' end processing factors appear to be conserved between yeast and higher eukaryotes, at least at the amino acid sequence level. Obvious differences exist in the function of related genes in either system. Whereas mammalian CstF only participates in the cleavage reaction, CF I is essential for both cleavage and polyadenylation in yeast. On the other hand, CPSF is required for both steps in 3' end formation, whereas PF I is dispensable for the initial cleavage step. Similarly, poly(A) polymerase is required for cleavage in the mammalian, but not in the yeast system. The apparent discrepancies between sequence and functional homology might be rationalized in part by the fact that the inventory of the factors is still incomplete and that the detailed molecular mechanism of how the known protein factors act in the reaction is largely unknown. In addition, in contrast to the mammalian situation, it is unclear how CF I and PF I are arranged on the RNA substrate. Both factors bind RNA *in vitro* but their sites of interaction have not yet been mapped. It could be that the interaction of PF I–Pap1p with the pre-mRNA is not specific or strong enough and that only the simultaneous binding of the factor to the RNA and to the already bound CF I leads to the formation of a stable and specific polyadenylation complex via the Fip1p–Rna14p protein–protein contact. In a model for the assembly of a yeast polyadenylation complex, Pap1p is tethered to the RNA primer by at least two distinct recognition events (see Figure 5). First, direct binding of PF I–Pap1p to the efficiency element of the upstream cleavage fragment involves Yth1p and possibly other components of PF I. Second, the Fip1p subunit of PF I acts as a bridge between Pap1p and CF IA, which binds the pre-mRNA through Rna15p, possibly onto the positioning element. The binding of CF I to the positioning element is hypothetical and has not been tested experimentally. However, as the positioning element determines the site of endonucleolytic cleavage, and as PF I is dispensable for the cleavage reaction, it is reasonable to assume that CF I binds to the positioning element. Likewise, PF I may bind to the efficiency element. Specific recognition of the correct 3' end processing site would, therefore, be imparted by the assembly of a complex network of RNA–protein and protein–protein interactions.

In mammals, the requirement for CPSF is reflected by the specific interaction of its 30 and 160 kDa subunits with the almost invariant 3' end processing signal AAUAAA (Keller *et al.*, 1991; Murthy and Manley, 1995). Gel retardation assays indicate that yeast PF I binds to pre-

Fig. 5. Schematic representation of the yeast polyadenylation complex. The binding of CF I and PF I–Pap1p to the positioning and efficiency elements, respectively, is hypothetical. CF I and PF I–Pap1p subunits are shaded in light and dark grey, respectively. Pab1p is associated loosely with CF IA and is thus not considered as an actual subunit of any of these factors (Minvielle-Sebastia *et al.*, 1997). Protein–protein interaction between polypeptides of the two factors is indicated by a double-headed arrow.

mRNAs *in vitro* (data not shown). Strikingly, both the recombinant 30 kDa subunit of CPSF and Yth1p do interact with homopolymeric RNA *in vitro* (Barabino *et al.*, 1997). As mentioned above, direct binding of the PF I–Pap1p holoenzyme to RNA would probably increase the processivity of polyadenylation and might thus explain the more rapid polyadenylation exerted by the complex as compared with Pap1p alone (Figure 4). In mammals, PAP is activated by two factors, PAB II and CPSF (Wahle, 1991; Bienroth *et al.*, 1993). These factors increase the processivity of mammalian PAP to various extents. The sequence similarities between yeast PF I and CPSF suggest that some of the PF I subunits may act similarly on Pap1p. This issue should be addressed by a careful analysis of the kinetic parameters of the PF I–Pap1p complex. The PF I–Pap1p complex extends poly(A) tails far beyond the length observed *in vivo*. Normal polyadenylation can be reproduced by the addition of purified CF IA and partially purified CF II/CF IB, implying that these fractions contain an activity that controls the length of the poly(A) tail synthesized by PF I–Pap1p (Minvielle-Sebastia *et al.*, 1997). We characterized the different polypeptides co-purifying with CF IA activity as Rna14p, Rna15p, Pcf11p (Amrani *et al.*, 1997), a new protein called Clp1p (L.Minvielle-Sebastia *et al.*, unpublished data) and, interestingly, Pab1p (Minvielle-Sebastia *et al.*, 1997). We showed that Pab1p is required for the synthesis of normal poly(A) tails, and might thus represent the activity that controls the length of the poly(A) tail synthesized by PF I–Pap1p. Possibly, Pab1p acts directly on Pap1p by inhibiting its activity once the poly(A) tails have reached the normal length. The tentative arrangement of the factors involved in the polyadenylation reaction is depicted in the model shown in Figure 5.

Until now, no mammalian protein with highly significant similarity to Pta1p or Fip1p was found in the databanks. However, putative Pta1p and Fip1p homologues do exist in the fission yeast *Schizosaccharomyces pombe*. It thus remains an open question whether the PF I subunits other than Pap1p and those related to CPSF are unique to lower eukaryotes, or whether their mammalian relatives have merely not yet been discovered.

Materials and methods

Buffers

All buffers used for PF I purification contained 0.5 mM dithiothreitol (DTT), unless otherwise noted, and a cocktail of protease inhibitors (0.5 mM phenylmethylsulfonyl fluoride, 0.7 µg/ml pepstatin and 0.4 µg/ml leupeptin hemisulfate). Buffer B: 20 mM HEPES–KOH pH 7.0, 1.5 mM Mg acetate and 10 mM K acetate; buffer D: 20 mM Tris–HCl pH 8.0, 0.5 mM EDTA, 20% glycerol and KCl at the millimolar concentration indicated following a hyphen; buffer F: as buffer D, but with HEPES–KOH pH 7.4 instead of Tris–HCl and 1 mM β-mercaptoethanol instead of DTT; buffer G: as buffer F, but with K acetate instead of KCl and 0.02% NP-40; buffer N: 20 mM HEPES–KOH pH 7.4, 20% glycerol, 50 mM K acetate, 0.02% NP-40 and 1 mM β-mercaptoethanol; buffer I: 20 mM Tris–HCl pH 8.3, 0.5 mM EDTA, 0.05% NP-40 and K acetate concentrations as indicated.

Plasmids, *S.cerevisiae* strains and disruption of *YDH1* and *YHH1/CFT1*

Standard cloning procedures were performed as described (Sambrook *et al.*, 1989). Polymerase chain reaction (PCR) amplifications were done with Ampli-taq DNA polymerase (Pharmacia) in 50 µl of buffer (provided by the manufacturer) supplemented with 50 pmol of each primer and 100 ng of plasmid DNA or 200 ng of genomic yeast DNA. Yeast media and standard genetic manipulations were as described elsewhere (Guthrie and Fink, 1991).

A plasmid for expression of double-tagged Fip1p in yeast was constructed as follows: the coding sequence of *FIP1* was amplified from a genomic clone by PCR utilizing a 5' primer (5'-GGGGCGGCCGCA-GCTCCAGTGAAGACG-3') and a 3' primer (5'-GGGCTTAGGGT-CATTTCGAATTTTG-3'). The resulting fragment was restricted with *NotI* (underlined) and *HindIII* and cloned in the corresponding sites of pHH1, a vector designed for the expression of HA1-His₆-tagged proteins in yeast (L.Minvielle-Sebastia, unpublished), yielding plasmid pIA96. The 0.3 kb *XhoI* (blunt ended)–*SacI* fragment, carrying the *CYC1* promoter, the double tag and the first codon of *FIP1* following the initiation codon, was isolated from this plasmid and inserted into pIA22 (*CEN4-TRP1-FIP1*; Preker *et al.*, 1995) restricted with *PstI* (blunt ended) and *SacI*. Fip1p encoded by the resulting plasmid (pIA97) contains an additional 19 amino acids (MYPYDVPDYAHHHHHHAAA) at its amino-terminus (the HA1 epitope peptide is underlined). This vector was transformed into strain PJP24 [relevant genotype: *fip1::LEU2 trp1 ura3-52* pIA24 (*CEN4-URA3-FIP1*)] and the residual *URA3*-marked plasmid was eliminated on synthetic complete medium containing 5-fluoroorotic acid (5-FOA), yielding PJP14.

For cloning of *YDH1*, an ~330 bp fragment of the amino-terminal part of the coding region was amplified by PCR on yeast genomic DNA and used as a probe to screen an *S.cerevisiae* genomic plasmid bank (Pick, 1995). Three independent clones were obtained that contained the entire *YDH1* open reading frame and its 5' and 3'-regulatory elements. A 3.3 kb *PmlI*–*ApaI* fragment containing *YDH1* was subcloned into *SmaI*–*ApaI*-restricted pBluescriptKS– (Stratagene) to yield plasmid pIA111. The shuttle vector pIA115 was generated by insertion of a *Bam*HI–*KpnI* fragment from pIA111 into the same sites of plasmid pFL38 (*CEN4-URA3*; Bonneaud *et al.*, 1991).

Strains heterozygous for a deletion of *YDH1* and *YHH1/CFT1* were generated by PCR-mediated, single step disruption of the respective open reading frames (Baudin *et al.*, 1993). For that, a *Bgl*III fragment of plasmid pFL38 containing the *TRP1* marker was amplified with two primers that contained 40–45 bp homologous to sequences flanking the gene of interest at their ends. PCR products were transformed into the diploid strain BMA41-2N (*MATa/MATα ura3-1/ura3-1 Δtrp1/Δtrp1 ade2-1/ade2-1 leu2-3,112/leu2-3,112 his3-11,15/his3-11,15*; Baudin-Baillieu *et al.*, 1997). Transformants were selected on medium lacking tryptophan, and integration of the marker at the correct loci was verified by Southern blotting. Upon sporulation and tetrad dissection, no more than two spores gave rise to growing colonies at 22°C, and all viable clones were tryptophan auxotrophs. The deletion of *YDH1* could be rescued by transformation of the diploid strain with plasmid pIA115 (*CEN4-URA3-YDH1*) prior to sporulation. In this case, all tryptophan heterotrophs were sensitive to 5-FOA.

A *TRP1*-marked plasmid (pIA114) for the expression of wild-type *PTA1* in yeast was generated by subcloning a ~5.3 kb *Hind*III fragment from YCpPTA1 (O'Connor and Peebles, 1992) into the *Hind*III site of YCplac22ΔBglIII/Δ*SacI* (Preker *et al.*, 1995).

LM113 is a *pta1-1* mutant segregant of a backcross of POC8-23d (*pta1-1 MATa ade2-1 leu2-Δ1 lys2-801 trp1-Δ101 ura3-52*; O'Connor

and Peebles, 1992) to W303 (*MAT α ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1*; R.Rothstein, Columbia University, New York). The temperature-sensitivity of LM113 can be complemented by transformation with pIA114 (*CEN4-TRP1-PTA1*). Other strains and their relevant genotypes are LM88 (*rna14-1 ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1*; Minvielle-Sebastia *et al.*, 1994) and LM96 (*fip1-1 ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1*). LM96 was obtained by integration of the temperature-formamide-sensitive *fip1-1* mutant allele (Preker *et al.*, 1995) in place of the wild-type *FIP1* allele.

Cell culturing and extract preparation

Saccharomyces cerevisiae strain PJP14 was cultured in YPD medium (1% yeast extract, 2% peptone and 2% dextrose) supplemented with adenine hemisulfate (20 mg/l) and ampicillin (50 mg/l). Five times 10 l of medium were inoculated with an exponentially growing culture and vigorously aerated. The temperature of the culture was ~28°C and the generation time was ~3 h. At an OD₆₀₀ of 3–4, 200 g of dry dextrose were added. At an OD₆₀₀ of ~6, cells were harvested by centrifugation (3000 r.p.m., 10 min) in a Cryofuge 6000 (Heraeus Sepatech) at room temperature and converted into spheroplasts by treatment with Zymolyase-100T (Seikagaku, Tokyo, Japan) according to Butler *et al.* (1990). All subsequent steps were done on ice or at 4°C. The spheroplasts were resuspended in 200 ml of buffer B per 10 l of initial culture volume and lysed by 20 strokes in a 60 ml Dounce homogenizer with a tight-fitting plunger (type S; B.Braun Biotech). The lysate was clarified by centrifugation (10 000 r.p.m., 15 min) in a Sorvall GSA rotor. The supernatant was brought to 200 mM K acetate and stirred gently for 30 min. The cell debris was removed by centrifugation at 290 000 *g*_{max} for 90 min in a Kontron TFT 45.94 rotor. The supernatant was removed with a pipette, taking care to avoid the flaky white layer at the top of the tubes, and adjusted to 40% saturation by the addition of solid ammonium sulfate. After incubation overnight, precipitated proteins were pelleted by centrifugation (10 000 r.p.m., 20 min) in a Sorvall GSA rotor and the pellet was resuspended in buffer D-50 to a final protein concentration of ~15 mg/ml. Typically, the precipitate contained 20–25% of the total protein present in the crude extract. The residual ammonium sulfate was removed by extensive dialysis against two changes of buffer D-50. The dialysate was diluted with buffer D-0 to adjust the conductivity to that of the dialysis buffer and clarified by centrifugation (10 000 r.p.m., 20 min) in a Sorvall GSA rotor.

Small-scale extracts were prepared from strains grown in 0.5 l of YPD at 30°C on a shaker and processed as described by Butler *et al.* (1990) except that buffers were as for the large-scale extracts (see above).

Purification of PF I

All manipulations were done at 0–4°C. Column material was purchased from Pharmacia, Bio-Rad and Qiagen. Columns were run at one column volume per hour, except for fast protein liquid chromatography (FPLC) columns. Fractions were frozen in liquid nitrogen and stored at –70°C.

Ammonium sulfate-fractionated extract was applied to a 430 ml Macro-Prep high Q anion exchange column equilibrated in buffer D-50. The column was washed with one column volume of the above buffer and developed with a 2.4 l gradient of buffers D-50 to D-500. PF I activity desorbed between 200 and 330 mM KCl. Peak fractions were pooled and dialysed for 3 h against 4 l of buffer D-0 and for 2 h against 2 l of buffer D-50. The dialysate was loaded on a 100 ml Blue-Sepharose column that had been equilibrated in buffer D-80. The column was washed with 80 ml of buffer D-80 and eluted with an eight column volume gradient from 80 to 500 mM KCl in buffer D. Fractions containing PF I activity were dialysed for 4 h against two changes of buffer F-0 and applied to a 100 ml heparin-Sepharose CL-6B column equilibrated with buffer F-100. After washing with one column volume of buffer F-100, PF I activity was recovered at ~420 mM KCl in a 400 ml gradient of buffers F-100 to F-800. Mg acetate and NP-40 were added to the eluate to final concentrations of 2 mM and 0.02% (v/v), respectively, before loading on a 4 ml Ni²⁺-NTA-agarose column equilibrated in buffer N-50. The column was washed with 4 vols of buffer N-50 and developed with a gradient (12 column volumes) from 0 to 250 mM imidazole in buffer N-50. PF I activity eluted in a broad peak between 30 and 170 mM imidazole. The pool of activity was loaded directly on a 1 ml Mono S HR 5/5 column equilibrated with buffer G-100 at a flow rate of 0.3 ml/min. The column was washed with 2 ml of the above buffer and developed with a gradient from 100 to 500 mM K acetate in buffer G over 20 column volumes. PF I was recovered between 160 and 400 mM K acetate.

An immunoaffinity column was prepared by covalently coupling 200 µg of monoclonal antibody 12CA5 (Boehringer Mannheim) to

600 µl of protein A-Sepharose CL-4B with dimethyl pimelimidate as cross-linking reagent according to Harlow and Lane (1988). The slurry was loaded in a column (1 cm diameter) and equilibrated in buffer G-300. Of the Mono S pool, 20% (2 ml, 1.2 mg of protein) was injected at a rate of 0.02 ml/min. The flow trough was reloaded at a rate of 0.04 ml/min, and the column was washed with buffer G-300 until the OD₂₈₀ of the eluate was stable. The affinity beads were transferred to a 5 ml centrifuge tube with 3 ml of the above buffer, recovered by centrifugation and suspended in two volumes of buffer G-300 containing HA1 epitope peptide (Berkeley Antibody Co., Richmond, CA) at a concentration of 100 µM. Elution was done for 10–20 min at 24–30°C with occasional mixing. The supernatant was recovered after centrifugation, and the elution step was repeated two or three times. The resin was regenerated by washing with 2 ml of glycine-HCl pH 2.8 and 4 ml of buffer G-300, and re-used twice without detectable loss in protein yield.

The eluate from one of these columns was dialysed against buffer D-50 and loaded on a Mini Q PC 3.2/3 column (bed volume 240 µl) equilibrated in loading buffer at a rate of 0.1 ml/min. The column was washed with 3 ml of buffer D-50 and connected to a 4.8 ml gradient from 50 to 500 mM KCl in buffer D. PF I activity was recovered at ~290 mM KCl.

3' end processing assays

Substrate RNAs were produced as run-off transcripts with pG4-CYC1 or pG4-CYC1pre as templates that had been linearized with *EcoRI* or *NdeI*, respectively (Preker *et al.*, 1995).

During PF I purification, 100 µl samples from every other column fraction were dialysed against buffer G-50, and PF I-containing fractions were detected by their ability to restore polyadenylation activity of *fip1-1* mutant extracts prepared from LM96 cells as described (Preker *et al.*, 1995).

Antibodies, Western blotting and immunoprecipitations

Polyclonal antibodies directed against purified recombinant Pap1p and Fip1p were affinity purified on immobilized proteins (Preker *et al.*, 1995). Rabbit antisera to Cft1p (Stumpf and Domdey, 1996), Ysh1p (Jenny *et al.*, 1994) and Yth1p (Barabino *et al.*, 1997) were gifts from the respective authors.

Immunoblotting assays were done by standard procedures (Sambrook *et al.*, 1989). Monoclonal antibody 12CA5 specific for the HA1 epitope (Boehringer Mannheim) was used at a dilution of 2 µg/ml in buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.05% Tween-20). Polyclonal antibodies were diluted in the above buffer containing 5% (w/v) powdered milk. Peroxidase-conjugated secondary antibodies (swine anti-rabbit and rabbit anti-mouse immunoglobulins; DAKO, Denmark) were diluted 1:2000 and 1:1000, respectively. Detection was done with the enhanced chemiluminescence kit according to the protocol provided by the manufacturer (Amersham).

For co-immunoprecipitation experiments, 80 µl from a Mono S peak fraction were first pre-adsorbed to 400 µl of a 15% protein A-Sepharose slurry in buffer I-300, to remove proteins that bind to the affinity resin unspecifically. Following incubation on a roller at 4°C for 30 min, the resin was pelleted and the supernatant was recovered. Monoclonal antibody (10 µg) or 20–40 µl of polyclonal antiserum were coupled to protein A-Sepharose in buffer I-300. After extensive washing with the above buffer, the antibody beads were combined with 300 µl of the pre-adsorbed PF I fraction and binding was performed on a roller for 2 h at 4°C. The resin was pelleted by centrifugation in a microfuge (5000 r.p.m., 30 s) and washed four times with 1 ml of buffer I-300 at room temperature. Alternatively, the antibody beads were washed three times with buffer I-100, twice with the same buffer containing 1.2 M K acetate, and once with buffer I-100. After the final wash, bound proteins were eluted in 40 µl of sample buffer for 5 min at 90°C and resolved by electrophoresis on SDS-polyacrylamide gels.

Amino acid sequencing of PF I components

Immunopurified PF I from three runs of the anti-HA1 affinity column (corresponding to 30 l of starting cell culture) was pooled and proteins were precipitated by the addition of trichloroacetic acid to a final concentration of 17% (w/v). The PF I pool contained ~100 µg/ml of the HA1 nonapeptide and we believe that this served as a carrier for precipitation of the polypeptides. After centrifugation at 4°C in a microfuge at maximal speed for 1 h, pellets were washed subsequently with ice-cold 80 and 100% acetone, air dried and resuspended in 250 µl of protein sample buffer. The amount of PF I was estimated by comparison of a sample run on a polyacrylamide gel with known

amounts of Pap1p run on the same gel. The following steps were done by the TopLab company (München, Germany). Of the 150, 105 and 85 kDa subunits of PF I, 40–60 pmol each were excised from a polyacrylamide gel and digested *in situ* with endoproteinase Lys-C. Peptides were eluted from the gel slices, separated by reverse phase chromatography, and microsequenced with an automated sequencer (Porton 3600; Beckman Instruments, Fullerton). The sequences obtained from the 150 kDa subunit were NIIDIQLK, FHGLITDIGLIPQK, SNIYYIQMEAEGRLLI, LVXAGNTTISK and VIGYDENVPXAEG-FQSGILLINP. The sequences from the 105 kDa subunit were SYG-TVVDFTMFLPDDSS, NLN(S/N)QYSGFSGT(G/E)EAENFDNLD and GALSIGDVRLAQLK. From the 85 kDa subunit the sequence obtained was FISEVVLSTQ.

Amino acid sequences were compared with the yeast protein database at MIPS, Martinsried. For further sequence analysis, the BLASTp program (Altschul *et al.*, 1994) at the Baylor College of Medicine (Houston, TX) server was used.

Acknowledgements

We thank S.Barabino, A.Jenny, G.Stumpf and H.Domdey for antibodies; A.Baudin, F.Lacroute and P.O'Connor for strains and plasmids; H.Pick and P.Philippsen for the *S.cerevisiae* genomic plasmid bank; and S.Barabino, U.Rüegsegger and E.Wahle for comments on the manuscript and helpful discussions. This work was supported by the Kantons of Basel, and by grants from the Schweizerischer Nationalfonds and the European Union (provided via the Bundesanstalt für Bildung und Wissenschaft, Bern). P.J.P. was supported by a long-term fellowship of the Boehringer Ingelheim Fonds.

References

- Altschul,S.F., Boguski,M.S., Gish,W. and Wootton,J.C. (1994) Issues in searching molecular sequence databases. *Nature Genet.*, **6**, 119–129.
- Amrani,N., Minet,M., Wyers,F., Dufour,M.-E., Aggerbeck,L.P. and Lacroute,F. (1997) *PCF11* encodes a third protein component of the yeast cleavage and polyadenylation factor I. *Mol. Cell. Biol.*, **17**, 1102–1109.
- Barabino,S.M.L., Hübner,W., Jenny,A., Minvielle-Sebastia,L. and Keller,W. (1997) The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev.*, **11**, in press.
- Baudin,A., Ozier-Kalogeropoulos,O., Denouel,A., Lacroute,F. and Cullin,C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 3329–3330.
- Baudin-Baillieu,A., Guillemet,E., Cullin,C. and Lacroute,F. (1997) Construction of a yeast strain deleted for the *TRP1* promoter and coding region that enhances the efficiency of the polymerase chain reaction-disruption method. *Yeast*, **13**, 353–356.
- Bienroth,S., Wahle,E., Suter-Crazzolara,C. and Keller,W. (1991) Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. *J. Biol. Chem.*, **266**, 19768–19776.
- Bienroth,S., Keller,W. and Wahle,E. (1993) Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.*, **12**, 585–594.
- Bonneaud,N., Ozier-Kalogeropoulos,O., Li,G., Labouesse,M., Minvielle-Sebastia,L. and Lacroute,F. (1991) A family of low and high copy replicative, integrative and single-stranded *S.cerevisiae*/*E.coli* shuttle vectors. *Yeast*, **7**, 609–615.
- Butler,J.S., Sadhale,P.P. and Platt,T. (1990) RNA processing in vitro produces mature 3' ends of a variety of *Saccharomyces cerevisiae* mRNAs. *Mol. Cell. Biol.*, **10**, 2599–2605.
- Chanfreau,G., Noble,S.N. and Guthrie,C. (1996) Essential yeast protein with unexpected similarity to subunits of mammalian cleavage and polyadenylation specificity factor (CPSF). *Science*, **274**, 1511–1514.
- Chen,J. and Moore,C. (1992) Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. *Mol. Cell. Biol.*, **12**, 3470–3481.
- Field,J., Nikawa,J.-I., Broek,D., MacDonald,B., Rodgers,L., Wilson,I.A., Lerner,R.A. and Wigler,M. (1988) Purification of a RAS-responsive adenyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method. *Mol. Cell. Biol.*, **8**, 2159–2165.
- Guo,Z. and Sherman,F. (1996) 3' End forming signals of yeast mRNA. *Trends Biochem. Sci.*, **21**, 477–481.
- Guthrie,C. and Fink,G.R. (1991) *Guide to Yeast Genetics and Molecular Biology*. Academic Press San Diego, CA.
- Haff,L.A. and Keller,E.B. (1975) The polyadenylate polymerases from yeast. *J. Biol. Chem.*, **250**, 1838–1846.
- Harlow,E. and Lane,D. (1988) Immunoaffinity purification In Harlow,E. and Lane,D. (eds), *Antibodies. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 511–552.
- Hopper,A.K., Banks,F. and Evangelidis,V. (1978) A yeast mutant which accumulates precursor RNAs. *Cell*, **14**, 211–219.
- Jenny,A. and Keller,W. (1995) Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyadenylation specificity factor. *Nucleic Acids Res.*, **23**, 2629–2635.
- Jenny,A., Hauri,H.-P. and Keller,W. (1994) Characterization of cleavage and polyadenylation specificity factor and cloning of its 100-kilodalton subunit. *Mol. Cell. Biol.*, **14**, 8183–8190.
- Jenny,A., Minvielle-Sebastia,L., Preker,P.J. and Keller,W. (1996) Sequence similarity between the 73-kilodalton protein of mammalian CPSF and a subunit of yeast polyadenylation factor I. *Science*, **274**, 1514–1517.
- Keller,W., Bienroth,S., Lang,K.M. and Christofori,G. (1991) Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3'-processing signal AAUAAA. *EMBO J.*, **10**, 4241–4249.
- Kessler,M.M., Zhao,J. and Moore,C.L. (1996) Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. *J. Biol. Chem.*, **271**, 27167–27175.
- Keys,D.A., Vu,L., Steffan,J.S., Dodd,J.A., Yamamoto,R.T., Nogi,Y. and Nomura,M. (1994) *RRN6* and *RRN7* encode subunits of a multiprotein complex essential for the initiation of rDNA transcription by RNA polymerase I in *Saccharomyces cerevisiae*. *Genes Dev.*, **8**, 2349–2362.
- Lingner,J., Kellermann,J. and Keller,W. (1991a) Cloning and expression of the essential gene for poly(A) polymerase from *S.cerevisiae*. *Nature*, **354**, 496–498.
- Lingner,J., Radtke,I., Wahle,E. and Keller,W. (1991b) Purification and characterization of poly(A) polymerase from *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **266**, 8741–8746.
- MacDonald,C.C., Wilusz,J. and Shenk,T. (1994) The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol. Cell. Biol.*, **14**, 6647–6654.
- Minvielle-Sebastia,L., Preker,P.J. and Keller,W. (1994) RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. *Science*, **266**, 1702–1705.
- Minvielle-Sebastia,L., Preker,P.J., Wiederkehr,T., Strahm,Y. and Keller,W. (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. *Proc. Natl Acad. Sci. USA*, **94**, in press.
- Murthy,K.G.K. and Manley,J.L. (1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev.*, **9**, 2672–2683.
- O'Connor,J.P. and Peebles,C.L. (1992) *PTA1*, an essential gene of *Saccharomyces cerevisiae* affecting pre-tRNA processing. *Mol. Cell. Biol.*, **12**, 3843–3856.
- Pick,H. (1995) Expression studies and cellular analysis of Npk1p and Cdc15p, two protein kinases of *Saccharomyces cerevisiae*. Ph.D. thesis, Universität Basel, Switzerland.
- Preker,P.J., Lingner,J., Minvielle-Sebastia,L. and Keller,W. (1995) The *FIP1* gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell*, **81**, 379–389.
- Rüegsegger,U., Beyer,K. and Keller,W. (1996) Purification and characterization of human cleavage factor I_m involved in the 3' end processing of messenger RNA precursors. *J. Biol. Chem.*, **271**, 6107–6113.
- Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Simos,G., Tekotte,H., Grosjean,H., Segref,A., Sharma,K., Tollervey,D. and Hurt,E.C. (1996) Nuclear pore proteins are involved in the biogenesis of functional tRNA. *EMBO J.*, **15**, 2270–2284.
- Stumpf,G. and Domdey,H. (1996) Dependence of yeast pre-mRNA 3'-end processing on CFT1: a sequence homolog of the mammalian AAUAAA binding factor. *Science*, **274**, 1517–1520.
- Takagaki,Y. and Manley,J.L. (1994) A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein. *Nature*, **372**, 471–474.
- Takagaki,Y., Ryner,L.C. and Manley,J.L. (1989) Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes Dev.*, **3**, 1711–1724.

- Takagaki, Y., MacDonald, C.C., Shenk, T. and Manley, J.L. (1992) The human 64-kDa polyadenylation factor contains a ribonucleoprotein-type RNA binding domain and unusual auxiliary motifs. *Proc. Natl Acad. Sci. USA*, **89**, 1403–1407.
- Wahle, E. (1991) A novel poly(A)-binding protein acts as a specificity factor in the second phase of messenger RNA polyadenylation. *Cell*, **66**, 759–768.
- Wahle, E. and Keller, W. (1996) The biochemistry of polyadenylation. *Trends Biochem. Sci.*, **21**, 247–250.

Received on March 18, 1997; revised on May 5, 1997

Note added in proof

While this manuscript was under review, it was reported that the homologues of the three large subunits of CPSF are present in the purified yeast CF II factor [Zhao, J., Kessler, M.M. and Moore, C.L. (1997) Cleavage factor II of *Saccharomyces cerevisiae* contains homologues to subunits of the mammalian cleavage/polyadenylation specificity factor and exhibits sequence-specific, ATP-dependent interaction with precursor RNA. *J. Biol. Chem.*, **272**, 10831–10838]. This is in contrast to our findings reported here. The discrepancy could perhaps be explained by assuming that a complex consisting of PF I/PAPp and CF II exists in the cell which upon fractionation can separate into different subcomplexes, one of which has CF II activity and the other PF I/PAP activity. The CPSF homologues might form a core of polypeptides contained in both subcomplexes. The fact that we did not detect these proteins in our partially purified CF II fractions (Figure 3B, lanes 1 and 4) may be because the assay for CF II activity is more sensitive than the detection of proteins in the Western blot.