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*

Jing Zhao, Marco M. Kessler, and Claire L. Moore‡
From the Department of Molecular Biology and Microbiology, Tufts University School of Medicine,
Boston, Massachusetts 02111-1800

Cleavage of pre-mRNA during 3′-end formation in yeast requires two protein factors, cleavage factor I (CF I) and cleavage factor (CF II). A 5300-fold purification of CF II indicates that four polypeptides of 150, 105, 100, and 90 kDa copurify with CF II activity. The 150-kDa protein is recognized by antibodies against Cft1, the yeast homologue of the 160-kDa subunit of the mammalian cleavage/polyadenylation specificity factor (CPSF). The 100-kDa subunit is identical to Brr5/Ysh1, a yeast protein with striking similarity to the 73-kDa subunit of CPSF. The 105-kDa protein, designated Cft2 (cleavage factor two) exhibits significant homology to the CPSF 100-kDa subunit. Cft2 is cross-linked to pre-mRNA substrate containing the poly(A) site and wild type upstream and downstream flanking sequences, but not to precleaved RNA lacking downstream sequences or to substrate in which the (UA)n processing signal has been deleted. The specific binding of Cft2 to the RNA substrate is ATP-dependent, in agreement with the requirement of ATP for cleavage. The sequence-specific binding of Cft2 and the similarities of CF II subunits to those of CPSF supports the hypothesis that CF II functions in the cleavage of yeast mRNA 3′-ends in a manner analogous to that of CPSF in the mammalian system. These results provide additional evidence that certain features of the molecular mechanism of mRNA 3′-end formation are conserved between yeast and mammals, but also highlight unexpected differences.

The formation of messenger RNA 3′-ends, an important step in maturation of eukaryotic mRNAs, requires two events, site-specific endonucleolytic cleavage of primary transcripts followed by poly(A) addition to the upstream fragment. While cleavage and polyadenylation are closely coupled in vivo, the two reactions can be experimentally uncoupled and assayed separately, allowing biochemical characterization of individual components. The 3′-end processing of mRNA depends on both cis-acting elements and trans-acting protein factors (for reviews, see Refs. 1–4). Six protein factors act in concert to recognize, cleave, and polyadenylate mammalian pre-mRNAs. Cleavage requires the cleavage and polyadenylation specificity factor (CPSF)3, cleavage stimulation factor (CstF), mammalian poly(A) polymerase (PAP), and two additional cleavage factors (CF Im and CF IIm). The poly(A) addition reaction needs PAP, CPSF, and poly(A)-binding protein II (PAB II). The recent purification of CF Im (5) leaves only CF IIm, uncharacterized, and cDNAs encoding many of the proteins have been cloned (for a recent review, see Ref. 4). The largest subunit (160 kDa) of CPSF binds specifically to the AAUAAA signal located upstream of the poly(A) site (6, 7), and the CstF binds to the GU- or U-rich downstream element via its 64-kDa subunit (8).

In the yeast Saccharomyces cerevisiae, three cis-acting elements are thought to be necessary and sufficient for mRNA 3′-end formation (Refs. 9 and 10 and references therein): the efficiency element formed by a UA repeat or related sequences, which functions by enhancing the efficiency of the positioning element; the positioning element, formed by AAUAAA or related sequences, which positions the cleavage site; and the actual cleavage site, a Py(A)n sequence. Both the efficiency and the positioning elements are found upstream of the poly(A) site.

Four functionally distinct factors are required for specific cleavage and polyadenylation of yeast pre-mRNA in vitro (11). Cleavage factors I and II (CF I and CF II) are sufficient for the cleavage reaction, while specific poly(A) addition needs CF I, PAP, and polyadenylation factor I (PF I). Yeast PAP has been purified to homogeneity (12), and the gene has been cloned (13, 14). Rapid progress is being made using genetics and biochemistry to identify components of the other factors (15). The RNA14 and RNA15 genes encode two subunits of CF I (16), and the gene for Pcf11, a 70-kDa CF I subunit, was obtained by virtue of its interaction with Rna14 and Rna15 in a two-hybrid screen (17). The latter approach also found a subunit of PF I, Fip1 (18), as a PAP-interacting protein. The genes for two additional proteins, Brr5/Ysh1 (19, 20) and Cft1 (21), were identified by their homology to genes encoding mammalian CPSF subunits. Cft1 is thought to be a subunit of the cleavage factor CF II (21), while Brr5/Ysh1 appears to have a role in both the cleavage and poly(A) addition steps (19, 20). For Rna14, Rna15, Fip1, Brr5/Ysh1, and Cft1, extracts deficient in these proteins because of conditional mutations or immunodepletion are defective in processing and can be rescued by the addition of crude fractions containing the appropriate activity, thus supporting the assignment as a component of CF I, CF II, or PF

© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

ATP-dependent Interaction with Precursor RNA*

(Received for publication, January 7, 1997, and in revised form, February 12, 1997)

* This work was supported by National Institutes of Health Grant R01 GM41752 (to C. L. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.
‡ To whom correspondence should be addressed: Dept. of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111-1800. Tel.: 617-636-7645; Fax: 617-636-0337, E-mail: cmoore@opal.tufts.edu.

* This work was supported by National Institutes of Health Grant R01 GM41752 (to C. L. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111-1800. Tel.: 617-636-7645; Fax: 617-636-0337, E-mail: cmoore@opal.tufts.edu.

The abbreviations used are: CPSF, cleavage and polyadenylation specificity factor; CstF, cleavage stimulation factor; CF Im and CF IIm, mammalian cleavage factor I and cleavage factor II; PAP, poly(A) polymerase; PAB II, poly(A)-binding protein II; CF I and CF II, cleavage factor I and cleavage factor II (yeast); PF I, polyadenylation factor I; BSA, bovine serum albumin; PBS, phosphate-buffered saline; FPLC, fast protein liquid chromatography.

This paper is available on line at http://www-jbc.stanford.edu/jbc/
Purification and Characterization of Yeast CF II

I. In the case of Pcf11, this subunit has been found in crude fractions containing Rna14 and Rna15 (17).

It is important to verify the composition of these factors with a reconstituted system using purified factors. We recently described the purification of the yeast CF I and its separation into two functional components, CF IA and CF IB (22). CF IA contains four polypeptides, and this biochemical analysis has confirmed the designation of Rna14, Rna15, and Pcf11 as subunits of CF IA (22). CF IB is a single 73-kDa protein. Here we report that four polypeptides (p150, p105, p100, and p90) copurify with CF IA activity. The 150-kDa subunit is recognized by antibodies against Cfl1, a yeast homologue of the 160-kDa subunit of mammalian CPSF (21). p100 is identical to the Brr5/Yah1 protein, which has similarity to the 73-kDa CPSF subunit (19, 20). p105, termed Cfl2, is a previously uncharacterized protein which binds pre-mRNA in a sequence-specific, ATP-dependent fashion and has significant homology to the 100-kDa subunit of mammalian CPSF. We discuss the implications of these findings to our growing understanding of the molecular mechanism of mRNA 3′-end formation in yeast.

EXPERIMENTAL PROCEDURES

In Vitro Synthesis of Pre-mRNA—Capped and uniformly 32P-labeled RNAs used as substrates for the cleavage and poly(A) addition reactions and ultraviolet cross-linking assays were prepared from the following plasmids by in vitro transcription of linearized DNA templates as described by Chen et al. (11). Full-length GAL7-1 RNA containing the GAL7 poly(A) site and flanking sequences was prepared from pJC-

GAL7-1 (11), and a mutated version of the full-length precursor, GAL7-2 RNA, which lacks the (UA)(U), repeat upstream of the poly(A) site, was prepared from the plasmid pJCGAL7-2 (21). Precleaved wild type GAL7-9 RNA, which lacks sequences downstream of the poly(A) site, was prepared from pJCGAL7-9 (23). All precursors were purified from 8.3 M urea, 5% polyacrylamide gels (24), precipitated with ethanol, and stored frozen at −20°C in 50 mM Tris-HCl, pH 7.0.

Proteins and Buffers—Highly purified CF I eluting from a heparin-HiTrap column (Pharmacia Biotech Inc.) was prepared as described (22). Broad range protein markers for SDS-polyacrylamide gel electrophoresis, prestarmed marker proteins, and purified bovine serum albumin (BSA) were from New England Biolabs Inc. Protein markers used for gel filtration were from Pharmacia, and ribonuclease A was from Sigma.

Protein concentrations were determined according to the Bradford assay (25) using the Bio-Rad kit and BSA as standard. The protein concentration in very dilute fractions was estimated by comparison with BSA standards run on slot blots stained with Coomassie Brilliant Blue or silver (26). The KCl concentration in column fractions was determined by measuring the conductivity of a 1/1000 dilution and comparing with standards made with the same composition and known amounts of KCl.

The pH of all buffering compounds was adjusted at room temperature and at a concentration of 1 M. Buffer A contained 20 mM potassium phosphate, pH 6.25, 0.02 mM EDTA, 0.5 mM dithiobiotrethol, and 10% glycerol. Buffer B was identical to buffer A except that Tris-HCl, pH 7.90, at 20 mM replaced the potassium phosphate. All buffers were supplemented with KCl as indicated. The concentration of protease inhibitors used in the chromatographic steps were as follows: Buffers used to apply the sample to the phosphocellulose column contained 0.25 mM phenylmethylsulfonyl fluoride, 0.03 mM leupeptin, and 0.1 mM pepstatin A. Buffers used to elute proteins from the phosphocellulose column and for application to the Bio-Rex column contained 0.1 mM phenylmethylsulfonyl fluoride, 0.06 mM leupeptin, and 0.2 μM pepstatin A. Buffers for elution of proteins from the Bio-Rex column and for application to Mono Q and heparin-Sepharose columns contained 0.05 mM phenylmethylsulfonyl fluoride, 0.03 mM leupeptin, and 0.1 mM pepstatin A. All further steps were carried out without protease inhibitors.

Purification of Yeast CF II—The purification protocol for CF II is shown schematically in Fig. 1. All procedures were performed at 0–4°C. After each step, samples were quickly frozen in liquid nitrogen and stored at −70°C. S. cerevisiae cells (bakers’ yeast) were purchased from Red Star (Randolph, MA). Whole cell extract preparation, ammonium sulfate precipitation, and Q-Sepharose chromatography were described as by Kessler et al. (22).

Fractions containing CF II, which eluted from a Q-Sepharose column with 300–500 mM KCl (~50 mg of protein), were pooled, dialyzed against 2 liters of buffer A supplemented with 100 mM KCl for 4 h, and applied to a 25-ml phosphocellulose P-11 column (Bio-Rad) equilibrated in the same buffer at 20 μl/h. After washing with two column volumes of buffer A + 100 mM KCl, proteins were eluted with a gradient of 100–500 mM KCl in three column volumes of buffer A.

Active fractions, which eluted at 150–300 mM KCl (~14 mg of protein), were pooled, dialyzed against 2 liters of buffer A + 50 mM KCl for 2 h, and applied to a 7-ml Mono Q column (Bio-Rad) equilibrated in the same buffer at 20 μl/h. The column was washed with 3 column volumes of buffer A + 50 mM KCl and was eluted with 2 column volumes of buffer A + 250 mM KCl and 3 column volumes of buffer A + 500 mM KCl. Most of the CF II activity (>80%) was eluted with buffer A + 250 mM KCl in a volume of 12 ml.

The CF II containing sample (~0.7 mg of protein) was dialyzed against 1 liter of buffer B + 50 mM KCl for 2 h and loaded onto a 1-ml Mono Q HR 5/5 FPLC column (Pharmacia) at 0.5 ml/min. The column was washed with 10 column volumes of the dialysis buffer and protein eluted with a gradient of 50–100 mM KCl (10 column volumes) and then with a gradient of 100–500 mM KCl (40 column volumes) at 0.5 ml/min. CF II activity came out between 280 and 330 mM KCl.

Fractions with CF II activity (~15 mg of protein) were pooled and dialyzed against 1 liter of buffer A + 100 mM KCl for 2 h and applied to a prepacked 1-ml heparin-HiTrap column (Pharmacia) at 10 ml/min. The column was washed with 5 column volumes of buffer A + 100 mM KCl and then with buffer A + 250 mM KCl, and CF II was eluted with buffer A + 500 mM KCl.

The CF II-containing sample (~0.15 mg of protein) was dialyzed against 1 liter of buffer B + 50 mM KCl for 2 h and loaded onto a 0.5-ml poly(A)-Sepharose column (Pharmacia) equilibrated in the same buffer at 2 ml/min. After washing the column with 3 column volumes of this buffer, the proteins were eluted with 3 × 0.5 ml steps of buffer B + 250 mM KCl and 3 × 0.5 ml buffer B + 500 mM KCl by gravity flow. CF II activity came out with 250 mM KCl.

The addition of 0.01% Nonidet P-40 to the pooled fractions containing CF II activity of the poly(A)-Sepharose column improved the recovery of CF II activity from the subsequent column about 10-fold. The pooled fractions (1 ml, 27 μg of protein) were concentrated to approximately 4-fold using Ultrafree-MC 30,000 nominal molecular weight limit concentrators (Millipore), and 0.2 ml of the concentrate was applied onto a 25-ml Superose 6 HR 10/30 FPLC column (Pharmacia) equilibrated in buffer A + 150 mM KCl at 0.01% Nonidet P-40 at 0.2 ml/min. The column was developed with the same buffer, and fractions of 0.5-ml volume were collected.

Cleavage Assays—Cleavage reactions were assembled on ice in a volume of 12 μl containing 1 mM magnesium acetate, 75 mM potassium acetate, 2% polyethylene glycol (polyethylene glycol 8000, Fisher), 2 mM ATP, 20 mM creatine phosphate, 10 mM radioactive full-length GAL7-1 RNA, 1.5 μM [α-32P]GAL7-1 RNA, 1.1 mM dithiobiotrethol, 0.4 units of RNasin (Promega) and 0.1 mg/ml purified BSA, 1.0 μl of partially purified CF II, and the amount of CF II indicated, usually 3.0–5.0 μl from gradient-eluted fractions and 0.5–1.0 μl from step-eluted fractions. The final concentration of KCl was kept below 50 mM. Creatine phosphate was eliminated from the reaction for the assay of CF II fractions obtained after the Bio-Rex 70 column. Reactions were incubated at 30°C for 20 min, stopped with proteinase K and SDS as described (11), diluted to 30 μl with 50 mM Tris-HCl, pH 7.0, extracted once with phenol:chloroform:isoamyl alcohol (25:24:1), and [1,10] of the reaction was separated on a 5% acrylamide-8.3 M urea gel and visualized by autoradiography or with a PhosphorImager (Storm 860, Molecular Dynamics).

For quantification of CF II activity, cleavage assay gels were exposed to a PhosphorImager screen for 1 h, the screen scanned using the PhosphorImager, and the ImageQuant program (version 1.0 for Macintosh, Molecular Dynamics) used to display the scanned image and determine the amount of precursor and cleavage products. All the values obtained were corrected by subtracting the background value determined from that region of an empty lane. Cleavage activity was calculated by dividing the total amount of upstream and downstream cleavage products by the sum of precursor and cleavage products. One activity unit is defined as the amount of CF II that can cleave 50% of the input precursor RNA under the conditions given above.

Aliquots of fractions from the Q-Sepharose, phosphocellulose, Bio-Rex, Mono Q, heparin, and poly(A)-Sepharose columns were dialyzed against 100–500 volumes of buffer A + 50 mM KCl for 1.5 h before use.
in processing assays. The fractions from Superose 6 gel filtration were concentrated 10-fold in Millipore microconcentrators before being assayed.

**Electrophoresis and Immunoblotting—**SDS-polyacrylamide gels were prepared and run according to Laemmli (27). Silver staining was according to Gottlieb and Chavko (28) using the silver staining plus kit from Bio-Rad. Immunoblots were carried out as described (29) using a 1:1000 dilution of polyclonal serum against Cr1, a generous gift of G. Stumpf and H. Domdey (21).

**UV Cross-linking—**Reactions (16 μl) were assembled on ice and contained 1 mM magnesium acetate, 75 mM potassium acetate, 1.5 μM tRNA, 1 mM ATP (added as indicated), 20 nM radioactive RNA substrate, and 10 μl of pure CF II from the Superose 6 column or 10 μl of highly purified CF II from the poly(A)-Sepharose column. The reactions were incubated at 30 °C for 5 min and then irradiated at room temperature with UV light for 10 × 30 s at 120 mJ/cm² (UV Stratalinker 1800, Stratagene) on paraffin lying on a glass plate at a distance of 15 cm. The samples were digested with ribonuclease A (Sigma, 0.9 μg/ml) for 30 min at 37 °C, and proteins were separated on a 7% polyacrylamide gel containing SDS. The gel was stained with silver, dried, and exposed to a PhosphorImager screen for 2 days.

**Peptide Microsequencing and Sequence Analysis—**Proteins were separated on an SDS-polyacrylamide gel and the gel stained with Coomassie Brilliant Blue. Bands were excised from the gel and peptides derived from these proteins subjected to microsequencing by William Lane at the Harvard University Microchemistry Facility (Cambridge, MA), using collisionally activated dissociation on a Finnigan TSQ 7000 triple quadrupole mass spectrometer. Homology searches were performed using the BLAST program (30) and sequence alignments made using DNASTAR programs (DNASTAR, Inc., Madison, WI), with the PAM250 Weight Table, in which amino acid residues with a score of greater than 0 are considered biochemically related and those with a score of 0 or higher as evolutionarily related.

**RESULTS**

**Purification of Yeast CF II—**Four functionally distinct factors (CF I, CF II, PF I, and PAP) must be combined to reconstitute cleavage and polyadenylation of yeast precursor mRNA (11). For this study, we undertook purification of the CF II factor, using seven chromatographic steps as shown schematically in Fig. 1. This purification is summarized in Table I. During purification, CF II activity tends to spread over many fractions when eluted with a salt gradient, and alternate gradient and stepwise elutions were used to keep CF II concentrated during chromatography. For example, the Bio-Rex 70 step did not result in increased specific activity (Table I), but it concentrated approximately 12-fold a sample that had become very dilute after two successive gradient-developed chromatographic steps (the Q-Sepharose column and the phosphocellulose column). Column fractions were assayed for reconstitution of endonucleolytic cleavage of pre-mRNA (GAL7-1 RNA) in a reaction containing CF I purified from a heparin-Sepharose column (22). After the final step, Superose 6 chromatography, an overall purification of about 5300-fold and yield of 2.7% was achieved.

The cleavage assay using fractions from the Superose 6 column is shown in Fig. 2A. CF II activity was spread over 10 fractions, but peaked in fractions 22–24, and only these fractions were used to determine yield. Proteins in the Superose 6 fractions were separated on a SDS-polyacrylamide gel and stained with silver. Four polypeptides with apparent molecular masses of 150, 105, 100, and 90 kDa copurified with CF II activity (Fig. 2A, lanes 5–7, and Fig. 2B, lanes 3–5). Additional polypeptides can be found in Fig. 24 (Fig. 2B, lane 5). Independent purifications using the same protocol, or using the same columns but in different sequence, showed the same four polypeptides copurifying with CF II activity.

From the silver-stained gel, the 150-, 105-, 100-, and 90-kDa polypeptides appear present in an approximately equimolar ratio. The combined mass of the CF II polypeptides is 445 kDa, assuming that each is represented only once in the complex. The elution position of CF II from the Superose 6 gel filtration column indicates that CF II is a large complex in the range of 600–700 kDa. Elution behavior upon gel filtration most accurately correlates with the Stokes radius of the particle rather than molecular size (31). The Stokes radius of CF II was determined to be near 100 Å by comparison with standard proteins (31). This is similar to that of mammalian CPSF, a complex of four polypeptides with combined molecular mass of 360 kDa and a Stokes radius of approximately 100 Å (32). The large Stokes radius of CF II relative to its molecular mass (~445 kDa) suggests that the shape of CF II is nonspherical.

**Interaction of Purified CF II with RNA Substrate—**Purified CF II when mixed with partially purified CF I can cleave wild type full-length RNA, but not mutant RNA, which lacks the (UA)₈ repeat (Fig. 3). This result indicates that substrate specificity is retained during the purification of CF II, if this specificity is rendered by CF II. Ultraviolet (UV) cross-linking analysis was then used to test whether CF II can bind precursor RNA. The CF II-containing fraction from the poly(A)-Sepharose column was incubated with different radioactive RNA substrates in the presence of tRNA as a nonspecific competitor. The mixture was then irradiated with UV light, treated with RNase A, and resolved on an SDS-polyacrylamide gel, and radioactively tagged proteins visualized by PhosphorImager analysis. The 105-kDa protein in CF II was cross-linked to the wild type full-length precursor RNA, but not to RNA lacking the (UA)₈ repeat (Fig. 4B, lanes 2–4). Silver staining of the same gel clearly shows the comigration of the 105-kDa protein band and the radioactively tagged species (Fig. 4A, A and B, lanes 3). A reaction using pure CF II from the Superose 6 column showed the exact same cross-linked band (Fig. 4B, lane 5), confirming the identity of the 105-kDa protein.
The reconstituted cleavage reaction requires ATP (11). Because of this property, we also tested the ability of CF II polypeptides to cross-link to precursor RNA in the presence and absence of ATP. This analysis showed that the cross-linking of the 105-kDa protein requires ATP, and none of the CF II proteins cross-linked in the absence of ATP (Fig. 4C).

Identification of CF II Subunits as CPSF Homologues—The 105- and 100-kDa protein bands were cut from SDS-polyacrylamide gels and submitted for microsequencing. The amino acid sequence obtained from these proteins was then used to search the yeast protein data base. This analysis of the 105 kDa protein identified its gene locus as GenBank™ YSCL9354, accession number U53878, on Chromosome XII, encoding a predicted polypeptide of 859 amino acids (Fig. 5) and estimated molecular mass of 94 kDa. We have called this gene CFT2 (Cleavage Factor Two 2). The sequence was examined for known RNA binding motifs such as the ribonucleoprotein-type RNA binding domain, a KH or RGG domain, or an arginine cluster (33), but none were found. Cft2 has numerous potential serine/threonine phosphorylation sites (34) scattered throughout the sequence and a possible tyrosine kinase site, DYL (34), at amino acid 32. There are two putative single cluster type nuclear localization signals (35). In a search of the data bases, the highest scoring match to the Cft2 sequence was the 100-kDa subunit of mammalian CPSF (7, 36), with 24% identity and 43% similarity, extending over the entire length of the protein (Fig. 6). This similarity increases to 73% if evolutionarily related residues are matched.

The peptide sequence EHHPDLSTTILR from p100 matched exactly to the sequence of the Brr5/Ysh1 protein (19, 20), a protein with strong homology to the 73-kDa protein of bovine CPSF, with 53% identity in the first 500 amino acids. The essential BRR5/YSH1 gene codes for a polypeptide of 859 amino acids and a predicted molecular mass of 94 kDa. We have called this gene CFT2 (Cleavage Factor Two 2). The sequence was examined for known RNA binding motifs such as the ribonucleoprotein-type RNA binding domain, a KH or RGG domain, or an arginine cluster (33), but none were found. Cft2 has numerous potential serine/threonine phosphorylation sites (34) scattered throughout the sequence and a possible tyrosine kinase site, DYL (34), at amino acid 32. There are two putative single cluster type nuclear localization signals (35). In a search of the data bases, the highest scoring match to the Cft2 sequence was the 100-kDa subunit of mammalian CPSF (7, 36), with 24% identity and 43% similarity, extending over the entire length of the protein (Fig. 6). This similarity increases to 73% if evolutionarily related residues are matched.

The peptide sequence EHHPDLSTTILR from p100 matched exactly to the sequence of the Brr5/Ysh1 protein (19, 20), a protein with strong homology to the 73-kDa protein of bovine CPSF, with 53% identity in the first 500 amino acids. The essential BRR5/YSH1 gene codes for a polypeptide of 859 amino acids and a predicted molecular mass of 94 kDa. We have called this gene CFT2 (Cleavage Factor Two 2). The sequence was examined for known RNA binding motifs such as the ribonucleoprotein-type RNA binding domain, a KH or RGG domain, or an arginine cluster (33), but none were found. Cft2 has numerous potential serine/threonine phosphorylation sites (34) scattered throughout the sequence and a possible tyrosine kinase site, DYL (34), at amino acid 32. There are two putative single cluster type nuclear localization signals (35). In a search of the data bases, the highest scoring match to the Cft2 sequence was the 100-kDa subunit of mammalian CPSF (7, 36), with 24% identity and 43% similarity, extending over the entire length of the protein (Fig. 6). This similarity increases to 73% if evolutionarily related residues are matched.

The 150-kDa protein of CF II was similar in size to Cft1, an essential protein of about 153 kDa with homology to the p160 of mammalian CPSF (21). To determine if Cft1 was a component of
CF II, immunoblots of CF II containing fractions from the last two chromatographic steps were treated with polyclonal antibodies against Cft1 (21). The 150-kDa protein in these fractions is recognized by the anti-Cft1 antibodies (Fig. 7, lanes 2 and 3).

DISCUSSION

We have described the purification of the yeast CF II, which in conjunction with CF I, catalyzes the cleavage of mRNA precursor at the poly(A) site. The four polypeptides of 150, 105, 100, and 90 kDa found in the purest CF II fractions co-eluted over seven different chromatographic steps, including the last step, a gel filtration column, and cofractionated during different purification procedures as well. This behavior argues that all four peptides are subunits of CF II and are associated with each other as a stable complex.

The Cft2 subunit of CF II can be cross-linked to GAL7 RNA substrate only if this RNA contains both the (UA)₆ efficiency element and sequences downstream of the poly(A) site. In vivo and in vitro studies have clearly shown that cis-acting elements are critical for yeast mRNA 3′-end formation. The sequence-specific binding of Cft2 is very interesting in that it is the first indication that one of the yeast processing factors is recognizing the precursor RNA in a sequence-dependent fashion. RNA with only seven nucleotides downstream of the GAL7 poly(A) site is an efficient cleavage substrate (11), and it is unlikely that CF II is recognizing a signal sequence present in this short stretch.
Instead, a stable complex between CF II and RNA may require interaction of CF II components with sequence at the cleavage site as well as with the efficiency element, and this may explain the failure to observe cross-linking to precleaved RNA. This conclusion is consistent with CF II's role in catalyzing the cleavage reaction. Interestingly, the three subunits of the mammalian CF I also cross-link to full-length precursor but not to precleaved RNA substrate. However, unlike Cft2, this binding is not dependent on an upstream polyadenylation signal (5).

Cft2 may be interacting directly with the (UA)₆ sequence, analogous to the binding of p160 of mammalian CPSF to the AAUAAA signal (6, 7). Alternatively, it may be the subunit in contact with the cleavage site, or it may interact with both sites. The first possibility would be reminiscent of early studies which showed that the binding of the p64 subunit of mammalian CstF was dependent on the AAUAAA sequence (37, 38). However, subsequent experiments demonstrated that p64 was in contact with the GU- or U-rich downstream element (8), an interaction stabilized by CPSF acting at the AAUAAA site (39, 40).

The specific RNA binding of Cft2 is ATP-dependent, in agreement with the requirement of ATP for cleavage (45). This is the first indication of how ATP might be utilized in this step of the reaction, though the exact molecular mechanism is not clear. Phosphorylation of CF II subunits by a kinase component could affect Cft2 binding. Extensive evidence exists for protein phosphorylation regulating spliceosome assembly (42) and the binding of sequence-specific transcription factors to DNA (42). Alternatively, a need for ATP might reflect the presence of an

Fig. 6. Comparison of the amino acid sequences of Cft2 of S. cerevisiae CF II and the 100-kDa subunit of bovine CPSF. The amino acids of each protein are numbered at the right. Dots in the sequence indicate gaps; vertical lines, identical residues; double dots, biochemically related residues; single plus double dots, evolutionarily related residues.
FIG. 7. Identification of 150-kDa subunit of CF II as Cft1 by immunoblot analysis. Aliquots of CF II-containing fractions were separated on an 8% polyacrylamide-SDS gel, blotted, and immunostained with antibodies directed against Cft1. Lane 1, molecular size markers stained with Coomassie Brilliant Blue; lane 2, 120 ng of CF II from the poly(A)-Sepharose column; lane 3, 20 ng of CF II from the Superose 6 column.

RNA-dependent helicase activity in CF II, which could remove RNA duplex structure blocking Cft2 binding. Several splicing factor proteins have the DEAD/DEAH motif indicative of RNA-dependent ATPases and ATP-dependent RNA helicases (43), but this has not been found in the mammalian or yeast polyadenylation factors for which sequence is available. The Brr5/Ysh1 protein does have near its carboxyl end the phosphate-binding loop (P-loop) motif GXXXGX(K/T)S commonly found in adenine and guanine nucleotide-binding proteins (44). Another possibility is that ATP provides energy for cleavage of yeast precursor and that ATP binding induces a conformational change in CF II in preparation for this event. All of these explanations would be in agreement with the finding that cleavage depends on hydrolysis of the β-γ bond of ATP (45).

The sequence-dependent binding suggested that Cft2 might be the yeast homologue of the p160 of CPSF. However, when a homology search was performed with the Cft2 sequence, we found that Cft2 had extensive homology to the p100 subunit of mammalian CPSF (Fig. 7). This was unexpected, since RNA binding activity has not been attributed to the CPSF p100 protein. Two of the other CF II subunits also show similarity to CPSF components. The p150 and p100 of CF II correspond to the Cft1 (21) and Brr5/Ysh1 proteins (19, 20), yeast homologues of the CPSF p160 and p73, respectively. The gene for the p90 protein of CF II has not yet been identified, and it remains to be seen whether any similarity will be found between it and other mammalian factors.

Our study establishes the function of CF II in cleavage of yeast mRNA precursor. The homologies with CPSF subunits and the specific interaction of CF II with a cis-acting efficiency element required for both steps of the reaction (analogous to the AAUAAA requirement in mammalian processing) raises the issue of whether CF II also has a role in the poly(A) addition step. However, CF I, CF II, and PAP were not sufficient to reconstitute poly(A) addition, and the presence of CF II did not stimulate this reaction (11). Chanfreau et al. (19) have shown recently that extracts from a brr5/ysh1 mutant strain were defective in poly(A) addition but not in cleavage, while wild type extracts immunodepleted of Brr5/Ysh1 (CF II p150) exhibited loss of cleavage activity as well. In a different study, extracts depleted of this protein by transcriptional repression of the gene were more severely impaired in poly(A) addition than cleavage (20). Depletion of extract with antibodies to Cft1 (CF II p150) abolished both cleavage and poly(A) addition (21).

Restoration of cleavage required only the addition of CF II containing fractions, while polyadenylation could be rescued by PF I and PAP. The PF I fraction used in this experiment contained Fip1 but not Cft1 (21). These results suggest that some components of CF II may be shared with PF I or that a stable association between these two factors exists in extracts. Resolution of these somewhat conflicting observations should come with further purification of the PF I activity. Previous studies have shown that immunodepletion of PAP from yeast processing extracts has an adverse effect on cleavage (16, 29), and it would not be surprising if all of the factors were assembled into a large complex via interactions that could be easily disrupted by chromatographic treatments.

The mammalian polyadenylation machinery utilizes six factors (CPSF, CstF, CF Im, CF II m, PAP, and PAB II) to generate cleaved mRNAs with poly(A) tails of the correct length. Cleavage requires the first five factors and poly(A) addition needs CPSF, PAP, and PAB II. In the yeast system, three factors (CF IA, CF IB, and CF II) are sufficient for cleavage in vitro. A conditional mutation in PAP can affect cleavage site choice in vivo (46), but this may be due to a degradation of Rna14 and Rna15, which occurs when the papI strain is shifted to growth at the nonpermissive temperature (47). CF IA contains proteins which are similar by sequence to CstF (15), while CF II contains homologues of CPSF subunits. It is not clear at the moment what mammalian factor might correspond to CF IB and why the cleavage step in yeast requires only three factors. It is possible that some proteins, which are essential to cleavage in mammals, are replaced in yeast with factors such as Ref2p (48), a nonessential RNA-binding protein that stimulates the processing at weak poly(A) sites.

In yeast, three factors (CF IA, CF IB, and PF I) are needed to recruit PAP to the appropriate substrate. The increased complexity of the yeast poly(A) addition reaction may reflect the fact that the major cis-acting elements in yeast are in general shifted to positions upstream of the cleavage site rather than flanking it as in higher eukaryotes, leading to the interesting possibility that the yeast positioning element might function not only in specifying the cleavage site but also in stabilizing the interaction of processing factors with the cleaved substrate.

The protein or combination of proteins that constitutes the actual endonuclease has not been identified for either yeast or mammals. Now that the cleavage factors of yeast have been purified, the identification of the genes encoding all of their components should soon follow, allowing more detailed studies of the molecular mechanism of mRNA 3′-end cleavage in a well defined, completely reconstituted system.

Acknowledgments—We thank Alexander Zelbelkovsky, Stefan Gross, Steffen Helming, and Neptune Mizrahi for helpful discussions. We are grateful to Ralph Isberg for the use of the FPLC system and to G. Stumpf and H. Domdey for antibodies against Cft1.

REFERENCES

1. Keller, W. (1995) Cell 81, 829–832.
2. Manley, J. L. (1995) Curr. Opin. Genet. Dev. 5, 222–228.
3. Wahle, E. (1995) Biochim. Biophys. Acta 1261, 183–194.
4. Wahle, E., and Keller, W. (1996) Trends Biochem. Sci. 21, 247–250.
5. Ruegsegger, U., Beyer, K., and Keller, W. (1996) J. Biol. Chem. 271, 6107–6113.
6. Keller, W., Bienroth, S., Lang, K. M., and Christofori, G. (1991) EMBO J. 10, 4241–4249.
7. Murthy, K. G. K., and Manley, J. L. (1995) Genes Dev. 9, 2672–2683.
8. MacDonald, C. C., Wilusz, J., and Shenk, T. (1994) Mol. Cell. Biol. 14, 6641–6654.
9. Guo, Z., and Sherman, F. (1995) Mol. Cell. Biol. 15, 5983–5990.
10. Guo, Z., and Sherman, F. (1996) Mol. Cell. Biol. 16, 2772–2776.
11. Chen, J., and Moore, C. L. (1992) Mol. Cell. Biol. 12, 3470–3483.
12. Lingner, J., Radtke, I., Wahle, E., and Keller, W. (1991) J. Biol. Chem. 266, 8741–8746.
13. Lingner, J., Kellermann, J., and Keller, W. (1991) Nature 354, 496–498.
14. Patel, D., and Butler, J. S. (1992) Mol. Cell. Biol. 12, 3297–3304.
Purification and Characterization of Yeast CF II

15. Manley, J. L., and Takagaki, Y. (1996) Science 274, 1481–1482
16. Minvielle-Sebastia, L., Preker, P. J., and Keller, W. (1994) Science 266, 1702–1705
17. Amrani, N., Minet, M., Wyers, F., Dufour, M., Aggerbeck, L., and Lacroute, F. (1997) Mol. Cell. Biol. 17, 1102–1109
18. Preker, P. J., Lingner, J., Minvielle-Sebastia, L., and Keller, W. (1995) Cell 81, 379–389
19. Chanfreau, G., Noble, S. M., and Guthrie, C. (1996) Science 274, 1511–1514
20. Jenny, A., Minvielle-Sebastia, L., Preker, P. J., and Keller, W. (1996) Science 274, 1514–1517
21. Stumpf, G., and Domdey, H. (1996) Science 274, 1517–1520
22. Kessler, M. M., Zhao, J., and Moore, C. L. (1996) J. Biol. Chem. 271, 27167–27175
23. Zhelkovsky, A. M., Kessler, M. M., and Moore, C. L. (1995) J. Biol. Chem. 270, 26715–26720
24. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, pp. 11.23–11.28, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
25. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
26. Dunbar, B. D. (1984) Protein Blotting: A Practical Approach, IRL Press, Oxford
27. Laemmli, U. K. (1970) Nature 277, 680–685
28. Gottlieb, M., and Chavko, M. (1987) Anal. Biochem. 165, 33–37
29. Kessler, M. M., Zhelkovsky, A. M., Skvorak, A., and Moore, C. L. (1995) Biochemistry 34, 1760–1769
30. Altschul, S., Gish, W., Miller, W., Meyers, E., and Lipman, D. (1990) J. Mol. Biol. 215, 403–410
31. Siegel, L. M., and Monty, K. J. (1966) Biochim. Biophys. Acta 112, 346–362
32. Bierroth, S., Wahle, E., Suter-Crazzolara, C., and Keller, W. (1991) J. Biol. Chem. 266, 19768–19776
33. Burd, C. G., and Dreyfuss, G. (1994) Science 265, 615621
34. Pearson, R. B., and Kemp, B. E. (1991) Methods Enzymol. 200, 62–81
35. Dingwall, C., and Laskey, R. A. (1991) Trends Biochem. Sci. 16, 478–481
36. Jenny, A., Hauri, H.-P., and Keller, W. (1994) Mol. Cell. Biol. 14, 8183–8190
37. Moore, C., Chen, J., and Whoriskey, J. (1988) EMBO J. 7, 3159–3169
38. Wilusz, J., and Shenk, T. (1988) Cell 52, 221–228
39. Murthy, K. G. K., and Manley, J. L. (1992) J. Biol. Chem. 267, 14804–14811
40. Weiss, E. A., Gilchrist, M. G., and Nevins, J. R. (1991) EMBO J. 10, 215–219
41. Takagaki, Y., and Manley, J. L. (1994) Nature 372, 471–474
42. Mermod, J. E., Cohen, P. T. W., and Lamond, A. I. (1994) EMBO J. 13, 5679–5688
43. Lamm, G. M., and Lamond, A. I. (1993) Biochem. Biophys. Acta 1173, 247–265
44. Saraste, M., Sihbald, P. R., and Wittinghofer, A. (1990) Trends Biochem. Sci. 15, 430–434
45. Butler, J. S., and Platt, T. (1988) Science 242, 1270–1274
46. Mandart, E., and Parker, R. (1995) Mol. Cell. Biol. 15, 6979–6986
47. Amrani, N., Dufour, M., Bonneaud, N., and Lacroute, F. (1996) Mol. & Gen. Genet. 252, 552–562
48. Russnak, R., Nehrke, K. W., and Platt, T. (1994) Mol. Cell. Biol. 15, 1689–1697
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

Jing Zhao, Marco M. Kessler and Claire L. Moore

*J. Biol. Chem.* 1997, 272:10831-10838.

doi: 10.1074/jbc.272.16.10831

Access the most updated version of this article at [http://www.jbc.org/content/272/16/10831](http://www.jbc.org/content/272/16/10831)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 22 of which can be accessed free at [http://www.jbc.org/content/272/16/10831.full.html#ref-list-1](http://www.jbc.org/content/272/16/10831.full.html#ref-list-1)