REPORT

Evidence that polyadenylation factor CPSF-73 is the mRNA 3′ processing endonuclease

KEVIN RYAN,[1,2] OLGA CALVO,[1,3] and JAMES L. MANLEY[1]
[1]Department of Biological Sciences, Columbia University, New York, NY 10027, USA

ABSTRACT

Generation of the polyadenylated 3′ end of an mRNA requires an endonucleolytic cleavage followed by synthesis of the poly(A) tail. Despite the seeming simplicity of the reaction, more than a dozen polypeptides are required, and nearly all appear to be necessary for the cleavage reaction. Because of this complexity, the identity of the endonuclease has remained a mystery. Here we present evidence that a component of the cleavage-polyadenylation specificity factor CPSF-73 is the long-sought endonuclease. We first show, using site-specific labeling and UV-cross-linking, that a protein with properties of CPSF-73 is one of only two polypeptides in HeLa nuclear extract to contact the cleavage site in an AAUAAA-dependent manner. The recent identification of CPSF-73 as a possible member of the metallo-β-lactamase family of Zn^{2+}-dependent hydrolytic enzymes suggests that this contact may identify CPSF-73 as the nuclease. Supporting the significance of the putative hydrolytic lactamase domain in CPSF-73, we show that mutation of key residues predicted to be required for activity in the yeast CPSF-73 homolog result in lethality. Furthermore, in contrast to long held belief, but consistent with properties of metallo-β-lactamases, we show that 3′ cleavage is metal-dependent, likely reflecting a requirement for tightly protein-bound Zn^{2+}. Taken together, the available data provide strong evidence that CPSF-73 is the 3′ processing endonuclease.

Keywords: polyadenylation; mRNA; endonuclease; Zn^{2+}-dependent

INTRODUCTION

All mRNA precursors in eukaryotic cells must undergo 3′ processing in the nucleus before their transport to the cytoplasm. This almost always entails a two-step reaction in which the pre-mRNA first undergoes site-specific endonucleolytic cleavage to generate the mRNA 3′ end, which then serves as a primer for synthesis of a poly(A) tail. Although both reactions, which are coupled in vivo but can be separated in vitro, are in principle quite simple, in fact, they require a complex set of protein factors (for review, see Colgan and Manley 1997; Zhao et al. 1999a; Edmonds 2002). These include, in mammals, cleavage-polyadenylation specificity factor (CPSF), which recognizes the nearly ubiquitous AAUAAA signal; cleavage stimulatory factor (CstF), which interacts with a less-conserved G/U-rich sequence situated downstream of the cleavage site; cleavage factors I and II (CFI and CFII), and finally poly(A) polymerase (PAP). Remarkably, all these factors, including in most cases PAP, are necessary for the cleavage reaction. In addition, the carboxy-terminal domain of the RNA polymerase II (RNAP II) largest subunit (CTD) was also found to be required for efficient pre-mRNA cleavage (for review, see Hirose and Manley 2000; Bentley 2002; Proudfoot and O'Sullivan 2002). As all of these factors, except PAP, are multisubunit, more than a dozen polypeptides (not including RNAP II) are required for the 3′ cleavage reaction. Although the mRNA signals for 3′ end formation in yeast are distinct from these in mammals, many of the mammalian factors have yeast counterparts, and the polyadenylation machinery in yeast is at least as complex as its mammalian counterpart (Zhao et al. 1999a).

The current compilation of cleavage/polyadenylation factors is the result of years of biochemical fractionation and analysis, and more recently genetic studies in yeast. Although the list is believed to be essentially complete, a more confident assessment of this awaits total reconstitution in vitro. Although specific functions for many of the polypeptides are known, for example, the key RNA-binding subunits appear to be the CPSF subunits CPSF-160 and CPSF-30 and the CstF-64 subunit of CstF, the precise roles of

Reprint requests to: James L. Manley, Department of Biological Sciences, Columbia University, 1117 Fairchild Center, New York, NY 10027, USA; e-mail: Jlm2@columbia.edu; fax: (212) 865-8246.
Present addresses: [2]Department of Chemistry, City College of New York–CUNY, 138th Street at Convent Avenue, New York, NY 10031, USA; [3]Centro de Investigación del Cáncer, (Universidad de Salamanca-CSIC), Campus Miguel de Unamuno, E-37007-Salamanca, Spain.
Article and publication are at http://www.rnajournal.org/cgi/doi/10.1261/rna.5214404.

many of the proteins, such as CPSF-73, CPSF-100, and the various subunits of CFI and CFII, remain largely undefined. Most important, whereas 3′ processing has been known to entail an endonucleolytic cleavage for nearly a quarter century, the identity of the actual endoribonuclease remains unknown (for discussion, see Shatkin and Manley 2000).

After so many years of study one might ask what, in a process whose activity is divided among so many required subunits, will proof of an endonucleolytic role actually entail? Might one expect to find an individual subunit, or small group of subunits, that will cleave a pre-mRNA substrate in vitro? If so, would such cleavage be sequence-specific, or result in more random degradation of a test RNA substrate? To our knowledge, no experiments, using either fractionated factors or recombinant subunits, have revealed a specific factor, let alone single polypeptide, with nuclease activity unequivocally ascribable to it. We describe here our efforts to solve this problem. This involved first identification of components of the precleavage complex that contact directly the actual cleavage site of a pre-mRNA substrate. Accordingly, we prepared a substrate RNA whose scissile phosphate carried the only radioactive label in the molecule, and characterized proteins that contact the cleavage site by UV cross-linking. We found that this substrate cross-links in an AAUAAA-dependent manner to only two proteins in the HeLa nuclear extract, CstF-64, and another protein of ~75–80 kD. Additional genetic and biochemical experiments, together with recent bioinformatic studies, lead us to conclude that this protein corresponds to CPSF-73, and that CPSF-73 is likely to be the long-sought 3′ processing endoribonuclease.

RESULTS AND DISCUSSION

In an effort to identify the 3′ pre-mRNA processing endonuclease, we first set out to identify proteins that contact the scissile phosphate at the cleavage site in a poly(A) signal (i.e., AAUAAA)-dependent manner. For this, we prepared wild-type (AAUAAA) and mutant (AAAAAA) derivatives of the SV40 late polyadenylation substrate containing a single, site-specific ^{32}P label at this phosphate for utilization in UV-cross-linking assays with HeLa cell nuclear extract (NE; see Fig. 1A and Materials and Methods). The site-specifically labeled SV40 late poly(A) substrate (SSL) extended 30-nucleotides on the 3′ side of the cleavage site so as to include the G/U-rich (CstF-binding) downstream element (e.g., Beyer et al. 1997; Takagaki and Manley 1997). A uniformly labeled version of this substrate, prepared by in vitro

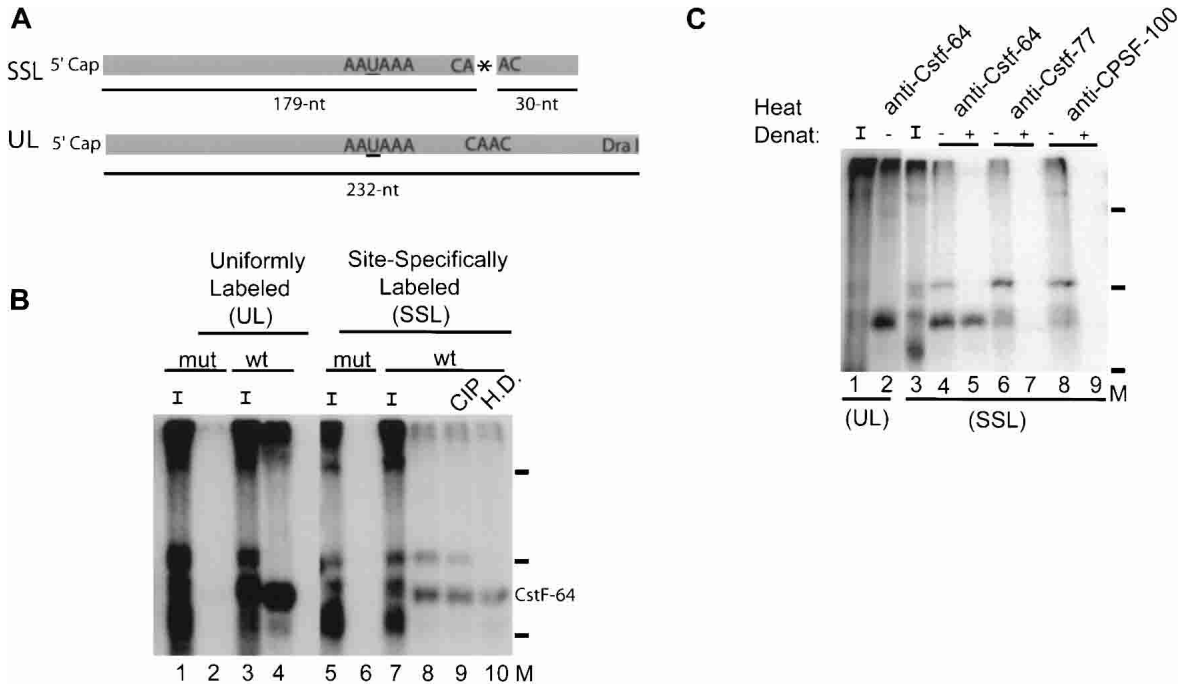

FIGURE 1. Two Hela cell proteins photo cross-link to the SV40 L poly(A) site in vitro. (*A*) Diagram of the two SV40L pre-mRNA substrates used in this study. The asterisk in the site-specifically labeled (SSL) substrate denotes the natural SV40L cleavage site, the site-specific ^{32}P phosphate label, and the site of ligation between the 179-nucleotide upstream in vitro transcribed piece and the 30-nt downstream chemically synthesized piece. The uniformly labeled substrate (UL) was made by in vitro transcription of the DraI-digested pG3SVL-A plasmid using either [α-^{32}P]GTP or UTP. The underlined U was changed to A in the mutated poly(A) substrate (mutant). (*B*) UV cross-linking of both proteins is AAUAAA-dependent. Either the mutant or wild-type substrates, UL or SSL, were irradiated in nuclear extract, digested with RNAse A then immunoprecipitated with the anti-Cstf-64 monoclonal antibody 3A7. (CIP) calf intestinal alkaline phosphatase treatment just before IP; (H.D.) heat-denaturation just before IP. Input, I, either 7.5% or 10% of one 10-μL reaction. Prestained protein markers (M) are indicated with dashes. (*C*) As in *B* except: wild-type SV40L substrates and immunoprecipitated with the indicated antibody.

transcription, oligonucleotide-directed RNase H digestion, and gel purification, was processed in NE as efficiently as the standard, uniformly labeled cleavage substrate (UL), also diagrammed in Figure 1A (data not shown). To test the poly(A) signal dependence of cross-linking we made mutant versions of both substrates in which the U (underlined here and in Fig. 1A) of the CPSF recognition site AAUAAA was changed to A. In addition, all four substrates, SSL and UL, each either wild type or mutated, contain only natural RNA nucleotides. Although we considered use of non-natural, photo-cross-linkable chemical functional groups to increase the cross-linking efficiency, we instead used the natural RNA to avoid any perturbation of the precleavage complex. Thus, our SSL substrates were composed of only natural RNA nucleotides, and included the appropriate *cis*-acting elements needed to elicit efficient in vitro processing.

The results of a representative UV cross-linking experiment are shown in Figure 1B. The first four lanes characterize the interaction between the uniformly labeled wild-type SV40 cleavage substrate (UL) and HeLa nuclear proteins. As observed previously (e.g., Takagaki et al. 1990), this identified an AAUAAA-dependent ~65-kD protein that was confirmed to be the CstF-64 component of CstF by immunoprecipitation with the anti-CstF-64 monoclonal antibody (mAb) 3A7 (Fig. 1B, lane 4). The cross-linking reactions using the site-specifically labeled substrates showed a group of poorly resolved bands and very little AAUAAA dependence (Fig. 1B, lanes 5,7). However, when the crude reaction was immunoprecipitated using the 3A7 mAb, the difference was striking: The wild-type AAUAAA-containing RNA but not the mt RNA was observed to cross-link to two proteins that were precipitated by the anti-CstF-64 mAb (Fig. 1B, lane 8). Somewhat surprisingly, one comigrated with the CstF-64 cross-linked to the uniformly labeled RNA substrate. Treatment of the reaction mixture with phosphatase before immunoprecipitation (Fig. 1B, lane 9) demonstrated that the more slowly migrating band was not a phosphorylated form of CstF-64, but rather a protein that was coimmunoprecipitated by the 3A7 antibody. This was confirmed by heating (4 min, 100°C) the reaction mixture before immunoprecipitation to disrupt noncovalently bound complexes, and this eliminated precipitation of the low mobility species but not CstF-64 (Fig. 1B, lane 10). In this and other experiments (see below) we found that the 3A7 antibody was unusual in that it recognized its epitope even after vigorous heat denaturation of the extract. Thus, the SSL substrate transferred its labeled phosphate to two proteins in an AAUAAA-dependent manner, and one of these proteins was identified as CstF-64. This was unexpected because CstF-64 is known to bind directly to the downstream G/U-rich element. The interaction with the cleavage site is intriguing, but additional studies will be required to determine its significance. On the basis of the known properties of CstF-64 and experiments described below, CstF-64 is unlikely to be the endonuclease. The more slowly migrating protein had an apparent molecular mass in the range of roughly 75–80 kD, and seemed to be a better candidate.

The data in Figure 1B showed that two proteins labeled in the UV cross-linking were precipitated by the anti-CstF-64 antibody. But this experiment could not determine either the identity of the higher molecular weight (MW) protein or whether it bound CstF-64 directly or through the mediation of other, unlabeled proteins also contained in a post-cross-linking complex. To investigate this further, we repeated the experiment with antibodies to other known cleavage factors, including CstF-77, CPSF-100, PAP, and the RNAP II CTD. The CPSF-100 antibody used is known to cross-react with CPSF-73 by virtue of the sequence similarity between these two CPSF subunits (e.g., Takagaki and Manley 2000). The results using some of these antibodies are shown in Figure 1C. Both the anti-CstF-77 (Fig. 1C, lane 6) and the anti-CPSF-100 (Fig. 1C, lane 8) antibodies precipitated the same two labeled proteins that were precipitated by the anti-CstF-64 antibody (Fig. 1C, lane 4). This indicated that CstF-77, CPSF-73, or CPSF-100 was also present in the post-cross-linking 3′-processing complex. However, neither the anti-CstF-77 (Fig. 1C, lane 7) nor the anti-CPSF-100 (Fig. 1C, lane 9) antibody precipitated the labeled proteins after heat denaturation, leading to the possibility that neither labeled band corresponds to the antigens of these antibodies, or that the epitope was destroyed by boiling. We tested this in a separate experiment using Western blotting of heat denatured, immunoprecipitated NE. These two antibodies, in contrast to the 3A7 mAb, were both unable to immunoprecipitate their respective antigens after the extract had been boiled (not shown). Therefore, we were unable to determine whether the antibodies to CPSF-100/73 and CstF-77 precipitated the slower migrating species directly or indirectly, and hence could not identify that protein conclusively by these experiments. The antibodies to RNAP II CTD and PAP did not precipitate any proteins labeled in the cross-linking reaction (not shown), and these proteins were therefore considered unlikely to be part of the post-cross-linking processing complex.

Among the known 3′ cleavage factors, only two have molecular weights that make them candidates for the unidentified protein detected in the cross-linking reactions. Both CPSF-73 (MW 77,486) and CstF-77 (MW 82,921) fall within this range. Unfortunately, these experiments, as well as related ones with several additional antibodies (not shown), were unable to conclusively distinguish between these two polypeptides. However, we believe that CPSF-73 is more likely to be the unidentified protein, and in fact the 3′ endonuclease, for the following three reasons: First, the CPSF-73 sequence is the most highly conserved among all the eukaryotic 3′ processing factors (see below), a trait that would be expected for the protein directly responsible for catalysis. Second, cytoplasmic CPSF specifically lacks CPSF-73 (Dickson et al. 1999; Mendez et al. 2000). Although

cytoplasmic CPSF retains the ability of the nuclear CPSF to direct polyadenylation by interaction with PAP and the mRNA, it has not been implicated in any RNA cleavage activity. Indeed, the presence of the endonuclease could be deleterious, as it could conceivably cleave the short poly(A) tail from cytoplasmic mRNAs that are in fact substrates for poly(A) extension. The third and perhaps most persuasive line of reasoning implicating CPSF-73 as the 3′ nuclease is the identification of CPSF-73 as a member of the metallo-β-lactamase fold superfamily; a diverse group of proteins many of which use zinc as a cofactor in their role as hydrolytic enzymes such as esterases and lactamases (Aravind 1999). More recently, CPSF-73 has been classified as a founding member of the β-CASP subfamily of metallo-β-lactamases (Callebaut et al. 2002). Proteins of this subgroup have been suggested to participate in nucleic acid metabolism, and significantly include the DNA endonuclease Artemis (Ma et al. 2002). CPSF-100, being related to CPSF-73, has also been classified as a β-CASP protein but, based on alterations in the conserved β-CASP sequence motifs, is assumed not to be capable of any hydrolytic catalysis (Callebaut et al. 2002). This is consistent with our inability to detect CPSF-100 in contact with the cleavage site. These findings not only point to CPSF-73 as a potential hydrolytic enzyme, and thus a candidate for the 3′ endonuclease, but also lead to the prediction that a group of specific residues in CPSF-73 are zinc- or other divalent metal-chelating residues, and that these residues may play an essential, catalytic role in hydrolysis of the pre-mRNA.

To test this prediction, we studied the yeast homolog of CPSF-73, Ysh1p. Ysh1/Brr5 was independently cloned and characterized by Chanfreau et al. (1996) and Jenny et al. (1996) and shown to be essential for cell viability and for 3′ end processing both in vivo and in vitro. The Ysh1p sequence shows similarity to both the 100-kD and 73-kD subunits of mammalian CPSF, being 23% identical over the entire length of CPSF-100 and a striking 53% identical in the first 500 amino acids to CPSF-73 (Chanfreau et al. 1996; Jenny et al. 1996).

Typical metallo-β-lactamase fold proteins contain five highly conserved sequence motifs (1–5) that participate in zinc coordination and catalysis. The β-CASP subfamily, however, is distinct, containing in place of motif 5 the β-CASP domain, which in turn contains three conserved motifs, A–C (Fig. 2A; Callebaut et al. 2002). Thus, β-CASP proteins are defined by the presence of seven sequence motifs, all of which include an acidic or histidine residue known or predicted to par-

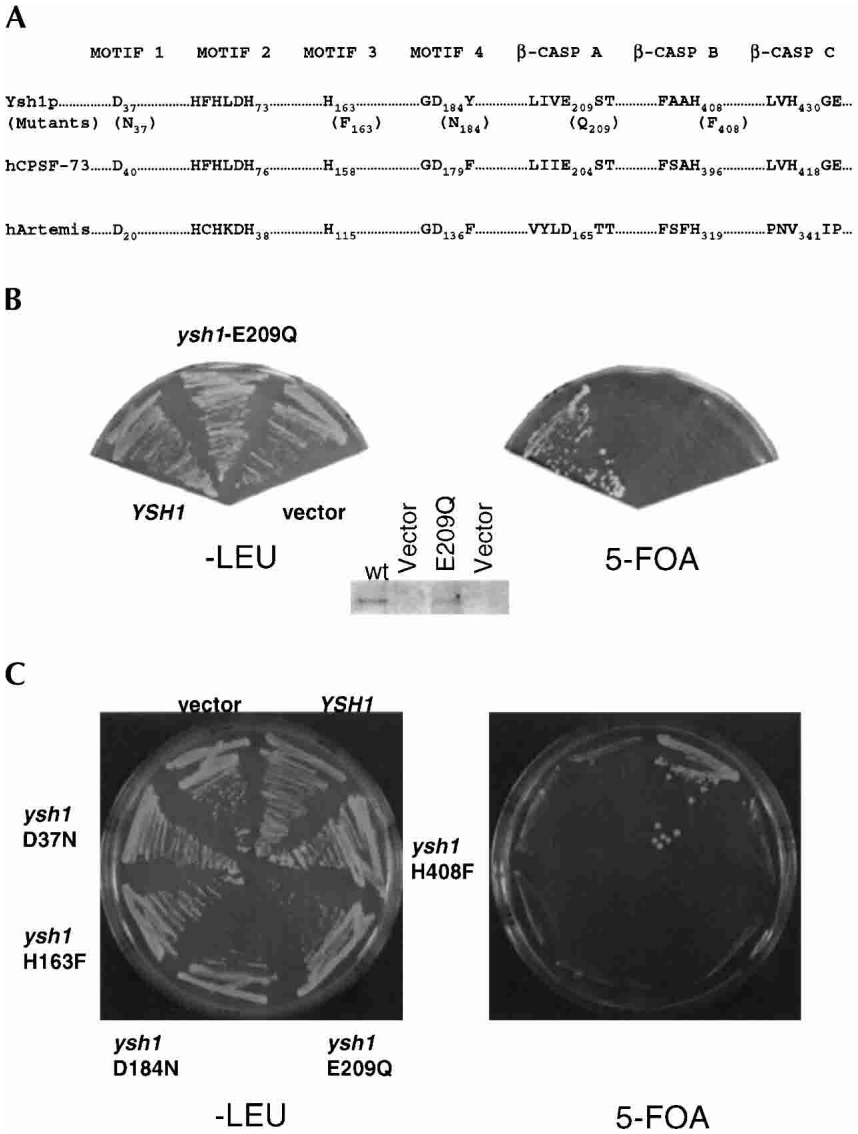

FIGURE 2. Mutation of conserved residues in the metallo-β-lactamase and β-CASP consensus motifs of the yeast CPSF-73 homolog Ysh1p causes lethality. (*A*) Abbreviated sequence alignment between scYsh1p, hCPSF-73, and hArtemis showing the β-lactamase motifs 1–4 and β-CASP motifs A, B, and C (adapted from Callebaut et al. 2002). Residues substituted individually in the plasmid shuffle assay are shown directly *below* the native Ysh1p residue. (*B*) Plasmid shuffling was used to test whether E209 of the β-CASP consensus in Ysh1p is essential for cell viability. Chromosomal *YSH1* was disrupted with *HIS3* and replaced by *YSH1* on a *URA3* plasmid. Cells were transformed with a *LEU* plasmid bearing wild-type *YSH1*, *ysh1-209*, or no insert (vector). Individual colonies were isolated and streaked on plates with or without 5-FOA. (*Inset*) Western blot of cell extracts comparing flu epitope-tagged wild-type and mutant ysh1p. (*C*) Plasmid shuffling was done as in *B* with four additional *ysh1* mutants that alter the residues indicated (see text).

ticipate in metal binding and hydrolysis. For example, motif A contains an acidic residue that has been shown to be required for activity of the DNA endonuclease Artemis (D165; Ma et al. 2002). To investigate whether the corresponding residue, E209, is essential for Ysh1p function, we used site-directed mutagenesis and a plasmid-shuffling procedure. We created a strain with a chromosomal deletion of *YSH1* (*ysh1*::*HIS3*) that contained wild-type *YSH1* on a plasmid with a *URA3*–selective marker and, in addition, either wild-type *YSH1*, the *ysh1*–209 mutant (E209Q) or no insert on a plasmid containing a *LEU2*-selective marker. The plasmid with the *URA3* marker was then counterselected on medium containing 5-fluoro-orotic acid (5-FOA). Wild-type *YSH1*, but not *ysh1*–209 or vector alone, complemented the chromosomal *YSH1* deletion (Fig. 2B). Western blots revealed that flu epitope-tagged wild-type and mutant proteins accumulated to comparable levels (Fig. 2B). To extend this result, we analyzed in an identical way four additional mutations that each alter a residue predicted to be essential for activity, one in β-CASP motif B and the others in motifs 1, 2, and 4 (see Fig. 2A). Strikingly, all four resulted in lethality upon loss of *YSH1* (Fig. 2C). These results strongly suggest that the putative metallo-β-lactamase/β-CASP domain exists in Ysh1p and is essential for activity. This in turn supports the hypothesis that CPSF-73/Ysh1p is the 3′ processing endonuclease.

It has long been known that dialysis of NE leads to partial loss of 3′ processing activity (e.g., Wahle and Keller 1994). It has likewise long been assumed that the 3′ endonuclease is metal independent because the reaction proceeds in vitro in the presence of high concentrations of EDTA (e.g., Hirose and Manley 1997 and references therein). But this seemed inconsistent with the presence of a functional metallo-β-lactamase fold in CPSF-73, which is supported by the genetic analysis described above. One possible explanation for the dialysis-dependent reduction of NE cleavage activity is that a low molecular weight cofactor, such as zinc, is partially lost during dialysis. Although physiological zinc concentrations are very low (Outten and O'Halloran 2001) and protein-bound zinc is tightly held, perhaps explaining the resistance of cleavage to EDTA, it is conceivable that during dialysis in the presence of low levels of EDTA CPSF-73 loses some of its zinc, thereby inactivating a fraction of the protein. In fact, some metallo-β-lactamases bind two zinc atoms, one tightly and one loosely, with both needed for efficient catalysis but one sufficient for minimal activity (Wang et al. 1999), and CPSF-73 could fall into this category.

We first tested whether $ZnCl_2$ is able to restore cleavage activity lost during dialysis of NE. Figure 3A, lane 1, shows a cleavage reaction using the SV40 late substrate, without dialysis of the extract, and in the presence of a typically high concentration (2 mM) of EDTA. This EDTA concentration also prevents subsequent polyadenylation of the 5′ cleavage product. Removal of all the EDTA but that contributed by the nuclear extraction buffer led to cleavage as well as poly-

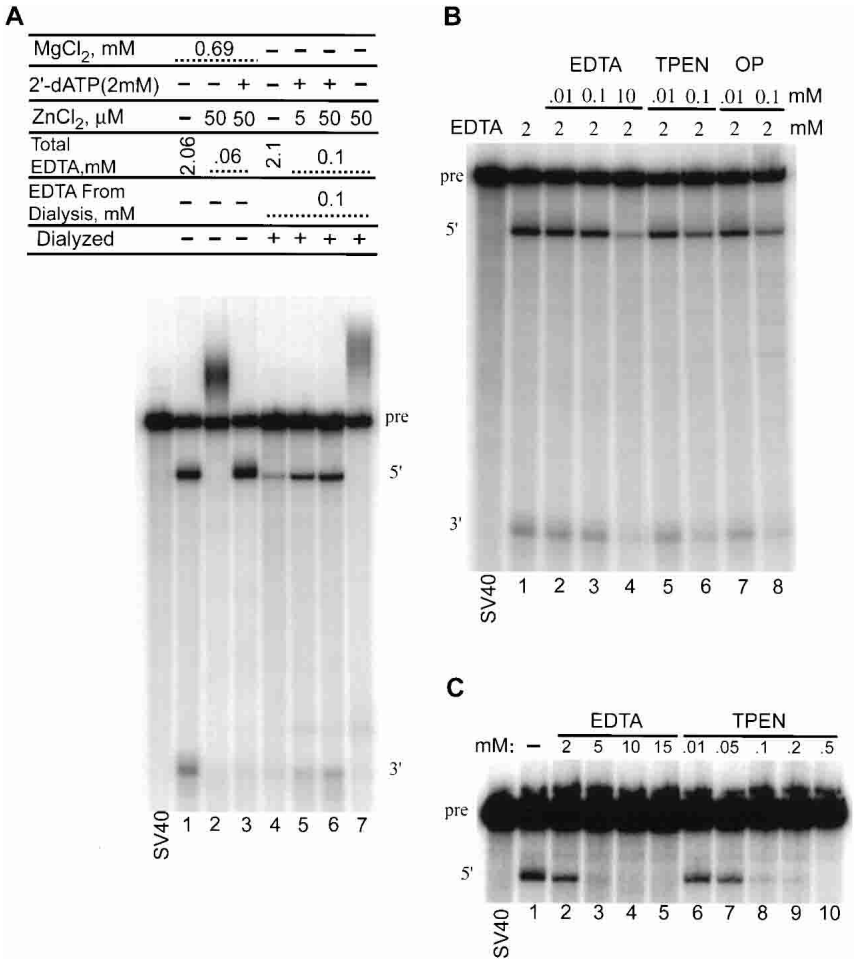

FIGURE 3. 3′ Cleavage is divalent cation-dependent, and zinc supports both cleavage and polyadenylation in vitro. (*A*) $ZnCl_2$ partially restores cleavage activity lost during nuclear extract dialysis. (Lane *1*) Cleavage of uniformly labeled SV40L pre-mRNA in the presence of 2 mM EDTA. Cleavage (lane *3*) or cleavage and polyadenylation (lane *2*) in the presence of added $ZnCl_2$. (Lane *4*) Loss of cleavage activity upon dialysis (cf. lane *1*). (Lanes *5*,*6*) Partial restoration of cleavage activity by addition of increasing zinc concentration. (Lane *7*) $ZnCl_2$ supports poly(A) addition in the absence of exogenous $MgCl_2$. (*B*) EDTA at high concentrations and the zinc-specific chelators TPEN and OP at low concentrations inhibit cleavage. In addition to the chelators listed above the gel, each sample contained 2 mM EDTA. (*C*) Inhibition of cleavage by EDTA and TPEN is concentration dependent.

adenylation, the latter likely supported by the $MgCl_2$ contributed by the buffer (Fig. 3A, lane 2). Including 2 mM 2′-dATP in the reaction (Fig. 3A, lane 3) terminated this polyadenylation activity. Figure 3A, lane 4, shows the extent to which cleavage activity was reduced by dialysis of the extract against a buffer containing 0.2 mM EDTA. Most significantly, Figure 3A, lanes 5 and 6, show that partial restoration of activity was achieved when increasing amounts of $ZnCl_2$ were added. Withholding 2′-dATP from the reaction (Fig. 3A, lane 7) revealed that $ZnCl_2$ could also support polyadenylation, suggesting that the poly(A) addition observed in Figure 3A, lane 2, may have been supported by either $ZnCl_2$ or the residual $MgCl_2$ from the extraction buffer. In any event, this experiment demonstrated that $ZnCl_2$ stimulated 3′ cleavage in a concentration-dependent manner in a divalent cation-deficient extract.

We also approached the question of possible zinc-dependent 3′ cleavage by adding either a higher concentration of EDTA or relatively low concentrations of two different zinc-specific chelators to in vitro cleavage reactions using the nondialyzed NE (Fig. 3B). Increasing the EDTA concentration from 2 to 12 mM significantly inhibited in vitro cleavage (Fig. 3B, lanes 3,4), consistent with our previous results showing that cleavage was resistant to 4 mM EDTA and modestly inhibited by 8 mM (Hirose and Manley 1997). Strikingly, the zinc-specific chelators TPEN (N,N,N′,N′-tetrakis(2-pyridylmethyl)ethylenediamine) and OP (orthophenanthroline; Siemann et al. 2002) each partially inhibited cleavage at 100 μM, far below the EDTA concentrations required for inhibition. Figure 3C extends these results, showing that inhibition by TPEN was concentration dependent and ~50-fold more effective than EDTA. (Note that the NE used in Fig. 3, B and C, was the same, although both cleavage efficiency and resistance to EDTA and TPEN were reduced in the experiment shown in Fig. 3C. This reflects time-dependent loss of activity that occurs with nondialyzed extracts, even when stored at −80°C [K. Ryan and F. Kleiman, unpubl. data]. Although the basis for this is unknown, it is intriguing that decreased cleavage efficiency and increased sensitivity to chelators are correlated.)

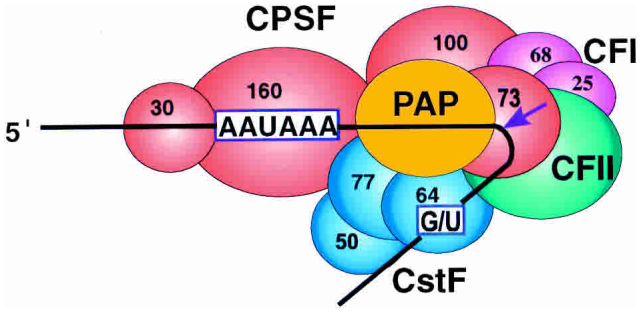

FIGURE 4. CPSF-73 as the 3′ processing nuclease. The core polyadenylation machinery (e.g., RNAP II is not included) is illustrated as it would assemble in a precleavage complex on an mRNA precursor, with CPSF-73 at the site of cleavage. The AAAUAA and GU-rich signal sequences are indicated, and the arrow denotes the cleavage site. The protein–protein and protein–RNA interactions depicted are consistent with available data. See text for details of protein factors.

The data in Figure 3 show that the 3′ pre-mRNA cleavage, in contrast to long-standing belief, is in fact a metal-dependent process. Because cleavage activity could be partially restored to a dialyzed extract by adding back $ZnCl_2$, and because cleavage in the nondialyzed extract was inhibited by low concentrations of zinc-specific chelators, our data also suggest that zinc is the metal cofactor needed for cleavage. This is consistent with the properties of the metallo-β-lactamase domain in CPSF-73. However, it is also possible that the zinc dependence reflects the need for zinc by CPSF-30, a CPSF subunit known to bind RNA in vitro and to have five zinc fingers and one zinc knuckle (Barabino et al. 1997). Indeed, it has been suggested that the *Drosophila* homolog of CPSF-30 possesses endonuclease activity (Bai and Tolias 1996). However, several observations suggest that CPSF-30 is unlikely to be the 3′ endonuclease. First, it has no motifs or domains consistent with nuclease activity. Second, our cross-linking experiments failed to provide evidence that CPSF-30 contacts the pre-mRNA in the vicinity of the cleavage site. Finally, reconstitution experiments with both yeast (Zhao et al. 1999b) and mammalian (Murthy and Manley 1992) factors provided evidence that CPSF-30 can be dispensable for cleavage in vitro.

Our data together with other findings strongly suggest that CPSF-73 is the long sought 3′ endonuclease. Although formal proof of this would require the demonstration that purified CPSF-73 display endonuclease activity in vitro, this may well be impossible. Technically, recombinant bacterially expressed CPSF-73 is exceptionally insoluble (data not shown) and therefore difficult to work with. Perhaps more significant, both biochemical and genetic experiments suggest that a dozen or more polypeptides are required for authentic 3′ cleavage, in vitro as well as in vivo. It thus may be that CPSF-73 will display activity only in a multiprotein complex, which would in turn make it impossible from such experiments to state conclusively that CPSF-73 is the actual nuclease. Nevertheless, the available evidence strongly suggests that CPSF-73 is the pre-mRNA 3′ endonuclease. Figure 4 depicts a revised view of the core polyadenylation machinery that incorporates the positioning of CPSF-73 at the cleavage site.

MATERIALS AND METHODS

Extracts and antibodies

Nuclear extract was prepared from HeLa cells as described by Dignam et al. (1983) with minor modifications (Takagaki et al. 1988), and either dialysed in buffer D (4°C, 10 h in 20% glycerol, 20 mM Na-HEPES, pH 7.9, 0.5 mM DTT, 50 mM ammonium sulfate, 0.2 mM EDTA, and 0.2 mM PMSF) or left in the extraction buffer without dialysis (15% glycerol, 14 mM Na-HEPES, pH

7.9, 0.5 mM DTT, 0.25 M NaCl, 2.5 mM KCl, 1.38 mM $MgCl_2$, 0.12 mM EDTA). Total nuclear extract protein concentration as measured by the Bio-Rad Protein assay (Bio-Rad) ranged from 4.4 to 4.9 mg/mL. Immunoprecipitation (IP) buffer IP2 was 50 mM Tris-HCl (pH 7.6), 50 mM NaCl, and included 0.1 mg/mL BSA and 0.1% NP-40 unless otherwise indicated. CstF-77 and CPSF-73 rabbit polyclonal antibodies were kindly provided by D. Bentley (Univ. of Colorado, Denver). The PAP rabbit polyclonal antibody was a gift of V. Vethantham (Columbia Univ., New York). The CTD antibody 8WG16 was purchased from Covance Research Products.

Preparation of cross-linking/processing substrates

The site specifically scissile phosphate labeled SV40L substrate (SSL) was made from two RNA fragments that were ligated using the standard RNA ligation procedure (Moore and Sharp 1993). To produce the capped upstream piece (uppercase in Fig. 1A) the plasmid pG3SVL-A (Takagaki et al. 1988) was digested with MfeI and HpaI, treated with calf intestinal phosphatase, then ligated to the preformed oligonucleotide duplex 5′pAACAA CATATG/5′pAATTCATATGTTGTT. After sequencing and large-scale preparation, the resulting plasmid pSVUP was sequentially digested with NdeI and mung bean nuclease (NEB), then gel purified on 0.9% agarose to yield a blunt-ended SP6 RNA polymerase transcription template whose transcript terminates exactly at the wild-type poly(A) site. The AAUAAA poly(A) site mutant version, pmutSVUP, was made analogously starting from a site-specific mutant version of pG3SVL-A containing the U→A mutation. Trace labeling with [α-^{32}P]UTP during a large-scale transcription reaction including the cap analog (NEB) allowed for quantitation of the gel-purified product. The downstream piece was synthesized by Dharmacon and deprotected according to the manufacturer's instructions. After purification on a 20% denaturing polyacrylamide gel and phosphorylation using T4 polynucleotide kinase (NEB) and [γ-^{32}P]ATP (6000 Ci/mmole), the upstream and downstream pieces were mixed in a 1:1:1 molar ratio with the gel-purified oligodeoxynucleotide ligation splint 5′-GAATGCA ATTGTTGTTGTTAACTTG using T4 DNA ligase (Promega, 20 U/μL). Typically, 25 pmoles of each component were ligated in a 16 μL ligation at 30°C for 4.5 h. The final product was gel purified on 6% DPAGE and resuspended in DEPC-treated water. The uniformly labeled (UL), 5′-capped substrate was prepared by standard run-off transcription using DraI-digested pG3SVL-A plasmid (or U→A mutant version) as the template and either [α-^{32}P]UTP orGTP as the internal label. This substrate is 24 nucleotides longer at the 3′end than the SSL substrate but underwent 3′ cleavage in nuclear extract no more efficiently than did a uniformly labeled version of the SSL substrate (not shown).

UV cross-linking

Cross-linking reactions (10 μL) were assembled on ice using 5 μL dialyzed Hela cell nuclear extract and the following components at these final concentrations: tRNA 0.1 mg/mL; DTT 2 mM; RNAseIn (Promega) 0.05 U/μL; EDTA 2 mM; creatine phosphate 40 mM; RNA substrate ~1–5 nM. The mixture was incubated at 30°C for 30 min then transferred to a microwell mini-tray (Nalge Nunc) that had been frozen into a tray of ice. The tray was placed 3.5 inches from the 254-nm light source of a Stratalinker UV irradiator (Stratagene) and irradiated for 8 min. The contents of the wells were transferred to a clean microfuge tube, mixed with 1 μL of RNAse A (10 mg/mL stock solution) and incubated at 37°C for 2–4 h. For Figure 1B, lane 9, 1 μL of calf intestinal alkaline phosphatase (NEB, 10 U/μL) was added and the incubation continued for 20 min. Cross-linking with the UL substrate was done identically. Due to the greater number of radioactive phosphates in the UL substrate, the gels containing reactions with the two different substrates were exposed for different times and the images aligned in the Figure to reproduce the original gel.

Immunoprecipitations

The crude cross-linked reaction was diluted with 140 μL of IP2, precleared with rocking at 4°C from 4–14 h, and then transferred to 12 μL (washed and drained) of IP beads for the same amount of time. For those samples where the reaction was heat denatured before the immunoprecipitation, the crude reaction was diluted with 140 μL of IP2 and boiled for 4 min. The mixture was allowed to cool to room temperature and spun in a microfuge to clarify. The IP beads consisted of protein A–sepharose (CL-4B, Pharmacia) with (for α-Cstf-64 3A7 and α-CTD 8WG16 mouse mAbs) or without (for polyclonal antibodies) prebinding a rabbit anti-mouse IgG Fcγ fragment-specific antibody (Jackson ImmunoResearch). After the immunoprecipitation, the beads were washed with IP2 (3 × 500 μL), 1 × PBS (1 × 500 μL) and finally with IP2 without BSA or NP-40 (1 × 500 μL). The washed beads were boiled for 4 min in 20 μL of 3× SDS loading buffer and the entire mixture was loaded onto an 8% SDS polyacrylamide gel. The gel was fixed and stained with coomassie before drying on paper and exposure to either Kodak X-ray film or a phosphorImager screen.

In vitro pre-mRNA 3′ processing reactions

In vitro cleavage reactions were assembled on ice and contained 6.25 μL of the Hela cell nuclear extract described above per 12.5 μL of reaction volume. In addition to the extract buffer components the following were included at these final concentrations unless otherwise indicated. For the experiment shown in Figure 3A, tRNA 0.1 mg/mL; DTT 2 mM; RNAseIn (Promega) 0.056 U; creatine phosphate 40 mM; polyvinyl alcohol (PVA) 2.5%; SV40 L/ Dra I 5′capped, uniformly labeled with [α-^{32}P]GTP or UTP RNA substrate 1–5 nM. Zinc chloride, EDTA, and 2′-dATP (but not ATP) were also added as indicated. In the reactions shown in Figure 3B,C, the metal chelators were incubated with the nuclear extract (nondialyzed) for 5 min at 30°C before replacing on ice and adding the substrate and other components. All cleavage reactions proceeded for 2 h at 30°C and were then stopped by treatment with proteinase K, phenol:chloroform extraction, and ethanol precipitation. Products were resolved on a 6% denaturing sequencing gel, dried, and exposed to either Kodak film or a phosphorImager screen. The zinc-specific chelators TPEN and OP were purchased from Fluka.

Yeast methods

We constructed a *Saccharomyces cerevisiae* strain (OCY1) with a chromosomal deletion of the *YSH1* gene (*ysh1*::*HIS3*), containing

wild-type *YSH1* on the *URA3* plasmid pMGL4. Plasmid shuffling was performed by transforming OCY1 strain with a *LEU2* plasmid, pMGL3, containing either wild-type *YSH1*, mutant ysh1-E209Q or any gene, and subsequently counterselecting the *URA3* marker on SC plates lacking leucine and histidine and containing 5-FOA. Mutant alleles of *YSH1* were created by oligonucleotide-directed mutagenesis using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. The yeast strains were transformed by the lithium acetate technique (Gietz and Schiestl 1991), and plasmids were maintained by growth in selective medium. Media preparation, tetrad dissection, genomic DNA isolation, plasmid shuffling, and all other yeast manipulations were performed by standard methods (Ausubel et al. 1991; Guthrie and Fink 1991). All the *S. cerevisiae* yeast strains used were isogenic derivatives of W303. OCY1 (*MATa/MATα ura3-52/ ura3-52 trp1-1/ trp1-1 ade2-1/ ade2-1 leu2-3,112/ leu2-3,112, his3-11,15/ his3-11,15 ade2-1/ ade2-1 YSH1/ysh1::HIS* [pMGL4-YSH1]); OCY2 (*MATa ura3-52 trp1-1 ade2-1 leu2-3,112, his3-11,15 ade2-1 YSH1/ysh1::HIS3* [pMGL4-YSH1]); OCY3 (*MATa ura3-52 trp1-1 ade2-1 leu2-3,112, his3-11,15 ade2-1 YSH1/ysh1::HIS3;* [pMGL4-YSH1, pMGL3-YSH1]); OCY4 (*MATa ura3-52 trp1-1 ade2-1 leu2-3,112, his3-11,15 ade2-1 YSH1/ysh1::HIS3* [pMGL4-YSH1, pMGL3-ysh1-209]); OCY5 (*MATa ura3-52 trp1-1 ade2-1 leu2-3,112, his3-11,15 ade2-1 YSH1/ysh1::HIS3* [pMGL4-YSH1, pMGL3]).

ACKNOWLEDGMENTS

We are grateful to K. Murthy and S. Kaneko for advice, N. Rao for technical assistance, and I. Boluk for help preparing the manuscript. This work was supported by National Institutes of Health grant GM-28983.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

Received October 22, 2003; accepted December 29, 2003.

REFERENCES

Aravind, L. 1999. An evolutionary classification of the metallo-β-lactamase fold proteins. *In Silico. Biol.* **1:** 69–91.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., eds. 1991. *Current protocols in molecular biology*. Greene Publishing Associates/Wiley-Interscience, New York, NY.

Bai, C. and Tolias, P.P. 1996. Cleavage of RNA hairpins mediated by a developmentally regulated CCCH zinc finger protein. *Mol. Cell. Biol.* **16:** 6661–6667.

Barabino, S.M., Hubner, 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:** 1703–1716.

Bentley, D. 2002. The mRNA assembly line: Transcription and processing machines in the same factory. *Curr. Opin. Cell. Biol.* **14:** 336–342.

Beyer, K., Dandekar, T., and Keller, W. 1997. RNA ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3′-end processing of pre-mRNA. *J. Biol. Chem.* **272:** 26769–26779.

Callebaut, I., Moshous, D., Mornon, J.P., and de Villartay, J.P. 2002. Metallo-β-lactamase fold within nucleic acids processing enzymes: The β-CASP family. *Nucleic Acids Res.* **30:** 3592–3601.

Chanfreau, G., Noble, S.M., and Guthrie, C. 1996. Essential yeast protein with unexpected similarity to subunits of mammalian cleavage and polyadenylation specificity factor (CPSF). *Science* **274:** 1511–1514.

Colgan, D.F. and Manley, J.L. 1997. Mechanism and regulation of mRNA polyadenylation. *Genes & Dev.* **11:** 2755–2766.

Dickson, K.S., Bilger, A., Ballantyne, S., and Wickens, M.P. 1999. The cleavage and polyadenylation specificity factor in *Xenopus laevis* oocytes is a cytoplasmic factor involved in regulated polyadenylation. *Mol. Cell. Biol.* **19:** 5707–5717.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11:** 1475–1489.

Edmonds, M. 2002. A history of poly A sequences: From formation to factors to function. *Prog. Nucleic Acid Res. Mol. Biol.* **71:** 285–389.

Gietz, R.D. and Schiestl, R.H. 1991. Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier. *Yeast* **7:** 253–263.

Guthrie, C. and Fink, G.R. 1991. Guide to yeast genetics and molecular biology. *Methods Enzymol.* **194:** 3–21.

Hirose, Y. and Manley, J.L. 1997. Creatine phosphate, not ATP, is required for 3′ end cleavage of mammalian pre-mRNA in vitro. *J. Biol. Chem.* **272:** 29636–29642.

———. 2000. RNA polymerase II and the integration of nuclear events. *Genes & Dev.* **14:** 1415–1429.

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.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M.R. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108:** 781–794.

Mendez, R., Murthy, K.G., Ryan, K., Manley, J.L., and Richter, J.D. 2000. Phosphorylation of CPEB by Eg2 mediates the recruitment of CPSF into an active cytoplasmic polyadenylation complex. *Mol. Cell* **6:** 1253–1259.

Moore, M.J. and Sharp, P.A. 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **365:** 364–368.

Murthy, K.G. and Manley, J.L. 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.* **267:** 14804–14811.

Outten, C.E. and O'Halloran, T.V. 2001. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* **292:** 2488–2492.

Proudfoot, N. and O'Sullivan, J. 2002. Polyadenylation: A tail of two complexes. *Curr. Biol.* **12:** R855–R857.

Shatkin, A.J. and Manley, J.L. 2000. The ends of the affair: Capping and polyadenylation. *Nat. Struct. Biol.* **7:** 838–842.

Siemann, S., Brewer, D., Clarke, A.J., Dmitrienko, G.I., Lajoie, G., and Viswanatha, T. 2002. IMP-1 metallo-β-lactamase: Effect of chelators and assessment of metal requirement by electrospray mass spectrometry. *Biochim. Biophys. Acta.* **1571:** 190–200.

Takagaki, Y. and Manley, J.L. 1997. RNA recognition by the human polyadenylation factor CstF. *Mol. Cell. Biol.* **17:** 3907–3914.

———. 2000. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* **20:** 1515–1525.

Takagaki, Y., Ryner, L.C., and Manley, J.L. 1988. Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* **52:** 731–742.

Takagaki, Y., Manley, J.L., MacDonald, C.C., Wilusz, J., and Shenk, T.

1990. A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. *Genes & Dev.* **4:** 2112–2120.

Wahle, E. and Keller, W. 1994. 3′-End processing of mRNA. In: *RNA Processing: A practical approach* (eds. S.J. Higgins and B.D. Hames), pp. 1–33. Oxford University Press, Oxford, UK.

Wang, Z., Fast, W., Valentine, A.M., and Benkovic, S.J. 1999. Metallo-β-lactamase: Structure and mechanism. *Curr. Opin. Chem. Biol.* **3:** 614–622.

Zhao, J., Hyman, L., and Moore, C. 1999a. Formation of mRNA 3′ ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63:** 405–445.

Zhao, J., Kessler, M., Helmling, S., O'Connor, J.P., and Moore, C. 1999b. Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol.* **19:** 7733–7740.