BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Reconstitution of CF IA from Overexpressed Subunits Reveals Stoichiometry and Provides Insights into Molecular Topology

James M. B. Gordon,[†,⊥] Sergei Shikov,[†,⊥] Jason N. Kuehner,[‡] Melissa Liriano,[†] Eunhee Lee,[§] Walter Stafford,[§] Mathias Bach Poulsen,[‡] Celia Harrison,[†] Claire Moore,[‡] and Andrew Bohm*[,†]

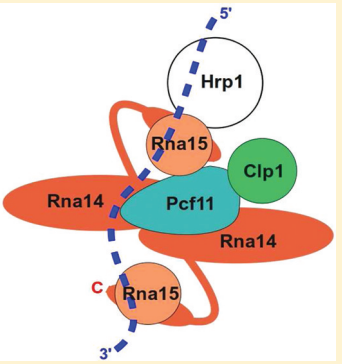
[†]Department of Biochemistry, Tufts University School of Medicine, Boston, Massachusetts 02111, United States

[‡]Department of Molecular and Microbiology, Tufts University School of Medicine, Boston, Massachusetts 02111, United States

[§]Boston Biomedical Research Institute, 64 Grove St., Watertown, Massachusetts 02472, United States

Supporting Information

ABSTRACT: Yeast cleavage factor I (CF I) is an essential complex of five proteins that binds signal sequences at the 3′ end of yeast mRNA. CF I is required for correct positioning of a larger protein complex, CPF, which contains the catalytic subunits executing mRNA cleavage and polyadenylation. CF I is composed of two parts, CF IA and Hrp1. The CF IA has only four subunits, Rna14, Rna15, Pcf11, and Clp1, but the structural organization has not been fully established. Using biochemical and biophysical methods, we demonstrate that CF IA can be reconstituted from bacterially expressed proteins and that it has 2:2:1:1 stoichiometry of its four proteins, respectively. We also describe mutations that disrupt the dimer interface of Rna14 while preserving the other subunit interactions. On the basis of our results and existing interaction data, we present a topological model for heterohexameric CF IA and its association with RNA and Hrp1.



All eukaryotic mRNA precursors undergo 3′ end processing. This processing influences transcriptional termination, it facilitates mRNA export from the nucleus, it influences the efficiency with which mRNA molecules are utilized by ribosomes, and it determines the longevity of individual mRNA molecules in the cytoplasm.[1] Processing of the 3′ end of the nascent mRNA proceeds in two biochemical steps: endonucleolytic cleavage and addition of a poly(A) tail.[2] Defects and alterations in 3′ end processing are manifested in a variety of human diseases including human cancers, thalassemias, influenza, and HIV.[3] Despite the universality and essential nature of this processing, our understanding of the complex of proteins responsible for this aspect of mRNA maturation is highly incomplete. Three-dimensional structures are available for many fragments of the processing complex, but there is no consensus as to how these pieces function together. Even the stoichiometry of the individual subunits within the complex is not well understood.

In yeast, the 3′ end processing complex is composed of at least 21 different proteins.[1,2] It recognizes *cis*-acting processing signals on pre-mRNA, cleaves at a site determined by these sequences, elongates the RNA molecule with a string of adenosines, and then terminates elongation once the poly(A) tails have reached their appropriate length. Early fractionation of this complex yielded a series of smaller, stable protein complexes, most notably cleavage factor I (CF I) and cleavage polyadenylation factor (CPF);[4−6] CPF contains the catalytically active subunits including the mRNA-cleaving nuclease, Ysh1, and the template-independent poly(A) polymerase, Pap1. The activity of these proteins is not, however, sequence specific.[7,8] They require the remaining proteins of CPF and CF I to direct them to authentic sites of mRNA processing. CF I also tethers the entire 3′ end processing complex to RNA polymerase II[9] and it is required for efficient export of mature mRNAs.[10,11]

In recent years it has become apparent that many genes contain alternative 3′ end processing sites. Alternative processing is common within different tissues, and it has been shown to be particularly prevalent during development and cell reprogramming.[12−14] It is estimated that roughly half of human genes can undergo alternative 3′ processing.[15,16] This alternative processing sometimes changes the length of the mature protein and often affects gene expression by altering the 3′ untranslated region (UTR) of the mature transcript. The 3′ UTRs are targeted by microRNAs and other factors that alter translation and mRNA longevity in eukaryotes.[17−19] During germline development in *C. elegans*, for example, gene regulation occurs primarily through the 3′ UTRs and not promoters.[20] Proliferating cells of the immune system generally express mRNAs with longer 3′ untranslated regions (UTRs) than those that are nonproliferating or slowly proliferating.[21] The 3′ end of mRNAs also differ in cancer cells, where

Received: June 23, 2011
Revised: September 29, 2011
Published: October 25, 2011

 © 2011 American Chemical Society

alternative 3′ end processing that removes part of the UTR may provide a mechanism for oncogenic activation without genetic sequence alterations of the proto-oncogene.[22,23]

To better understand 3′ end site selection, we have undertaken studies of the yeast CF IA complex. This complex contains four of the five proteins in CF I, and it is stable in the absence of RNA.[24] CF IA is composed of two tightly associated subcomplexes: Rna14 (80 kDa) associates tightly with Rna15 (33 kDa), and Pcf11 (72 kDa) binds tightly to Clp1 (50 kDa). Rna15 binds pre-mRNA with low degree of specificity, but its interaction with the nucleotide substrate of the processing reaction is strengthened by Rna14.[24,25] Thus, Rna14 appears to act as a central scaffold around which CF I is assembled.

High-resolution structures are available for fragments of all five CF I subunits.[26−35] While the structures have revealed numerous molecular details, they have not addressed the overall architecture or stoichiometry. Herein we show that CF IA is a heterohexameric complex with 2:2:1:1 stoichiometry of Rna14, Rna15, Pcf11, and Clp1, respectively. We also present a topological model of the complex based on our biochemical results and the body of existing data regarding interactions among the proteins and with RNA.

METHODS

Protein Expression and Purification. *Saccharomyces cerevisiae* Rna14 and Rna15 were coexpressed from a pETDuet vector as described previously.[25] Clp1 and Pcf11 were cloned into and then coexpressed from a pRSFDuet vector (Novagen). Pcf11 was cloned into the site 1 position within the vector using *Bam*HI and NotI restriction enzymes. Clp1 was cloned into the second position using NdeI and *Kpn*I restriction sites. To generate the version of Clp1 with MBP on the C-terminal end, the MalE gene from *E. coli* was cloned into the vector expressing Pcf11 and Clp1 using the *Kpn*I and AvrII restriction sites downstream of the *CLP1* gene. To generate the pETDuet plasmid that allowed expression of and purification of Rna14 in the absence of Rna15, a MBP tag was placed on the N-terminus of Rna15. The MalE gene from *E. coli* was cloned into the NcoI site upstream of the *RNA15* gene in the pETDuet vector. Isolated Rna15 was expressed as described previously.[24] All proteins were expressed in *E. coli* strain BL21 (DE3) pLysS. Protein expression was induced with 0.2 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) overnight at 16 °C. Harvested cells were lysed in Tris-HCl, pH 8.0, 250 mM NaCl, 5% glycerol, 0.01% βME, 5 mM PMSF using an Avestin C5 homogenizer, and debris were removed by centrifugation at 30000 rpm. Overexpressed proteins in the cleared lysates were purified using Ni-NTA. In some cases, the material was further purified through ion exchange and/or size exclusion chromatography.

Analytical Size Exclusion Chromatography (SEC). The proteins were applied onto a Superdex-200 HR10/30 (GE Healthcare) column in a buffer containing 20 mM, Tris-HCl, pH 8.0, 250 mM NaCl, 5 mM βME. The eluted protein fractions were run on SDS-PAGE followed by staining with either Coomassie blue or Imperial protein stain (Pierce).

Site Directed Mutagenesis. The QuikChange kit (Agilent) was used to introduce a double mutation of R562E and Y563S in the Rna14/Rna15 pETDuet plasmid according to the manufacturer's protocol. The following primers were used: Rna14 ES for 5′: GAA GTT TTC ACA AGT CGT AGT CAA ATT CAA AAC TCC AAC 3′ and Rna14 ES Rev for 5′: GTT GGA GTT TTG AAT TTG ACT ACG ACT TGT GAA AAC TTC 3′.

Yeast Strains and Plasmids. The *RNA14* gene (positions −239 to +2299 relative to ATG start codon) was amplified from yeast genomic DNA and cloned into *Bam*HI-*Pst*I-digested plasmid pRS315 (*LEU2*). Mutations were introduced via the QuikChange method (Agilent). Mutant yeast growth phenotypes were assayed using a plasmid-shuffle complementation assay.[36] Briefly, the haploid yeast strain LM21 (*MAT**a**, rna14::TRP1 ade2 ura3 his3 trp1 leu2* [pLM28-*RNA14, URA3*])[37] was transformed with either pRS315-*RNA14* or pRS315-*rna14-ES*, and loss of pLM28-*RNA14* (*URA3*) was selected for on medium containing 1 mg/mL of 5-fluoroorotic acid.

***In Vitro* RNA Analyses.** Preparation of yeast cell extracts, transcription of [α^{32}-P]UTP-labeled full-length GAL7-1 RNA or precleaved GAL7-9 RNA, and 3′ end processing assays were performed as described previously.[38] In the rescue reactions 1 μL of a 1 mg/mL solution of CF IA was added to the standard assay.

Western Blot Analysis. 50 μg of the yeast whole-cell extract used for *in vitro* processing reactions was analyzed by immunoblotting using polyclonal antibodies against Rna14, Rna15, and actin (mouse monoclonal; mAbcam 8224).

Electrophoretic Mobility Shift Assay (EMSA). Binding reactions were carried out for 20 min at 4 °C in 30 μL of binding buffer [25 mM Hepes·KOH (pH 7.8), 150 mM KCl, 5 mM βME, 5% glycerol]. Increasing amounts of protein [0.2−2 uM] were incubated with a fixed concentration of ^{32}P labeled *GAL7-1* RNA (∼1 μM). Electrophoresis was performed in nondenaturing 4.5% (29:1) polyacrylamide gels. The gels were run for 1 h in 1× TBE buffer (90 mM Tris, 64.6 mM boric acid, 2.5 mM EDTA, pH 8.4) at 100 V at 4 °C. The gels were dried and exposed to a PhosphorImager screen overnight. The screen was scanned on a Molecular Dynamics Storm 840 PhosphorImager.

Circular Dichroism Spectroscopy. Prior to data collection, proteins were exhaustively dialyzed against 20 mM Tris phosphate, pH 7.8, 150 mM NaF, 150 mM K_2S0_4, 5 mM βME. CD measurements were performed using a Jasco J-810 spectropolarimeter at 4 °C.

Analytical Ultracentrifugation. Prior to centrifugation, the samples were exhaustively dialyzed against a buffer containing 20 mM Tris-HCl, pH 7.5, 250 mM NaCl, 5 mM βME. The dialysate used for all dilutions and as a reference. Experiments were carried out on a Beckman Instruments Optima XL-I Analytical Ultracentrifuge with Rayleigh optics. Sedimentation equilibrium analyses were carried out at three speeds and three loading concentrations for the wild-type complex and two speeds and three concentrations for the mutant complex. Interference optics were used, and optical blank runs were carried out at the same speeds as the run and subtracted from the run scans to correct for optical background systematic noise. The data were fit with SEDANAL for the analysis of interacting and noninteracting systems.[39]

Small-Angle X-ray Scattering Analysis (SAXS). Protein samples were dialyzed against a buffer containing 20 mM Tris-HCl, pH 7.5, 250 mM NaCl, 5% glycerol, and 5 mM βME. Scattering intensity was measured using X-rays of wavelength λ = 1.0 Å at the beamline 12.3.1 at Lawrence Berkeley National Laboratory. Two-dimensional data reduction consisted of normalization for detector response, exposure time and sample transmission, absolute intensity calibration, azimuthal integra-

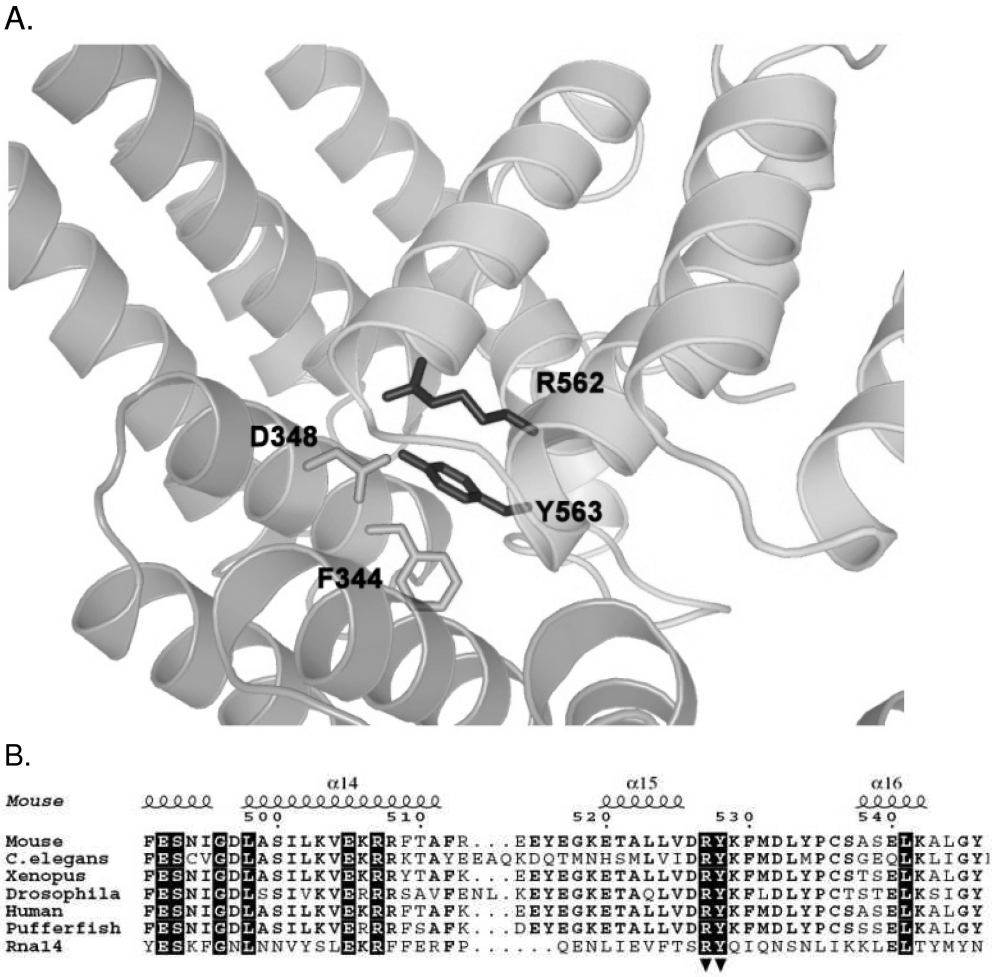


Figure 1. Molecular model of the Rna14 homodimer interface. (A) Ribbon diagram of the conserved Rna14 interface is based on the crystal structure of the murine CstF-77. Residues mentioned in the text are labeled. (B) Partial sequence alignment of Rna14 proteins from various species. Secondary structure elements are as in the murine Rna14 X-ray crystal structure (PDB code 2OND). Residues with 100% conservation are highlighted in black. Residues mutated in this work are marked with arrowheads beneath the sequence.

tion, and background subtraction from buffer alone to obtain the normalized scattered intensity I as a function of q, where $q = 2\pi s = 4\pi \sin(\theta)/\lambda$, and 2θ is the scattering angle. SAXS data were analyzed using Guinier plots to determine the radius of gyration R_g. The program GNOM[40] was used to determine the distance distribution function $p(r)$.

RESULTS

Disruptions of the Rna14 Dimer Interface Do Not Perturb the Rna14/Rna15 Interaction. Earlier biophysical studies of the yeast Rna14/Rna15 complex revealed that this complex has 2:2 stoichiometry.[25] The crystal structures of the Rna14 homologue in higher eukaryotes, CstF-77, showed a highly elongated dimer with an extensive protein−protein interface composed of both hydrophobic and ionic interactions.[26,27] On the basis of the structure of murine CstF-77, we constructed a molecular model of yeast Rna14 using MODELER[41] (Figure 1A). Using this model as a guide, we selected two conserved residues, R562 and Y563, for mutagenesis with the aim of disrupting the Rna14 dimer. Arginine 562, which normally interacts with Asp348 on the other monomer, was mutated to glutamic acid to cause charge−charge repulsion. Tyrosine 563, which makes a hydrogen bond via its side chain hydroxyl and participates in a hydrophobic cluster including Phe344 and Gln380 on the other monomer, was mutated to serine (Figure 1B). This Rna14 double mutation will be referred to as Rna14-ES.

Wild-type Rna14 is mostly insoluble when bacterially expressed on its own, but it forms a soluble complex with Rna15 when both proteins are expressed from the same plasmid.[25] We coexpressed both the wild-type Rna14/Rna15 complex and the complex between Rna14-ES and Rna15 and then purified each complex using an N-terminal hexahistidine tag on Rna14. To determine whether the ES mutation in Rna14 altered the secondary structure of the complex, circular dichroism (CD) spectroscopy was performed on the wild-type and mutant assemblies. The UV wavelength scan of wild-type revealed a spectrum with two distinct minima at 208 and 222 nm (Figure 2A), suggesting that a significant fraction of the amino acids are in an α-helical conformation.[42] The ES mutant and WT spectra are essentially identical, indicating that the mutation did not grossly alter the secondary structure of Rna14.

We then subjected both the wild-type and mutant complexes to size exclusion chromatography (SEC) to determine whether the mutation disrupts the Rna14 dimer interface. While the WT Rna14/Rna15 complex migrates with an apparent molecular weight of 380 kDa, the apparent molecular weight of the Rna14/Rna15 ES mutant is 205 kDa (Figure 2B). The

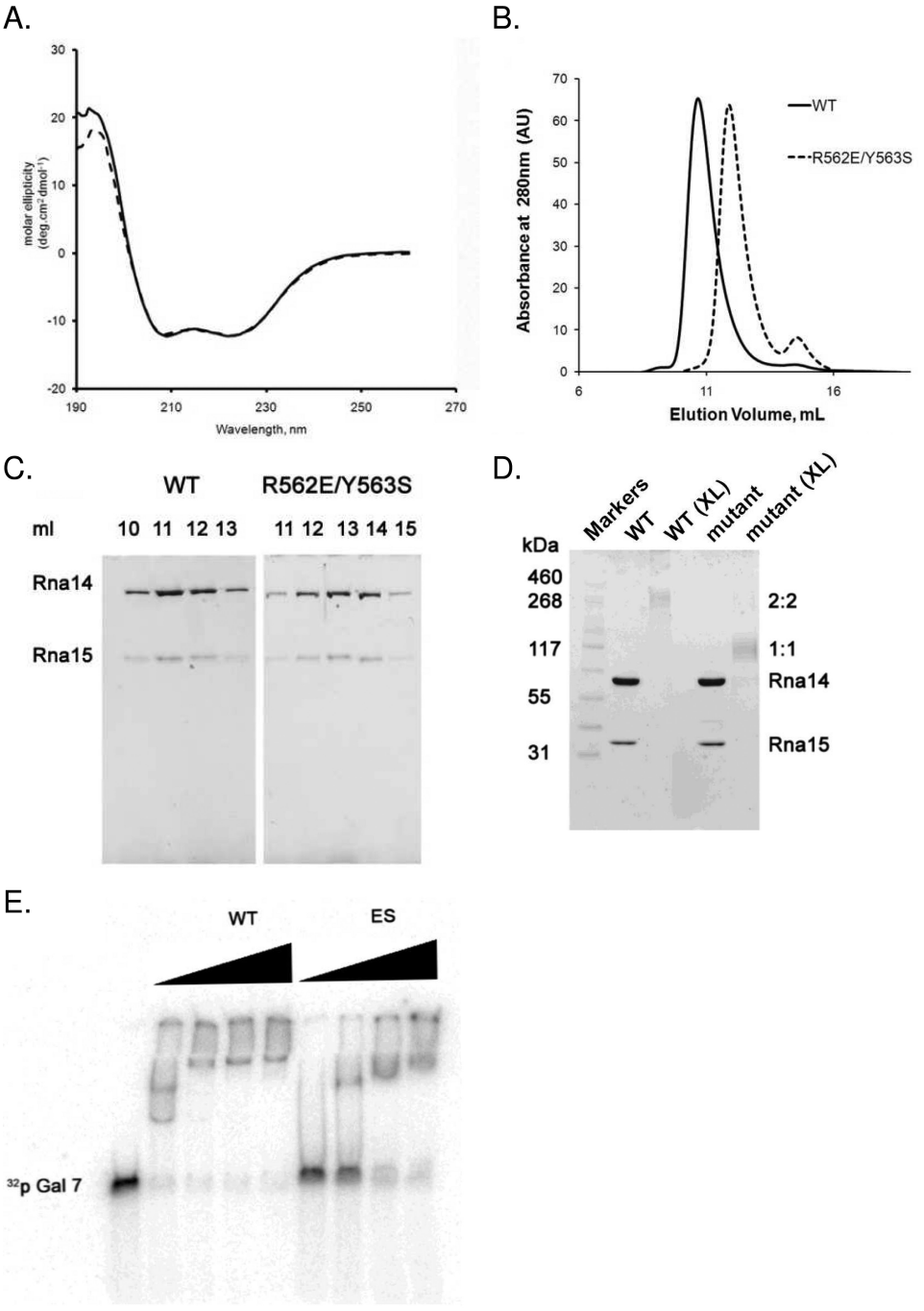


Figure 2. Disruption of the Rna14 homodimer interface. (A) CD spectra of WT (solid line) and ES mutant (dashed line) Rna14/Rna15 complex. (B) Size exclusion chromatograms of WT and ES mutant Rna14/Rna15 complexes applied onto a Superdex-200 HR10/30 column. The WT Rna14/15 complex elutes as a single peak at 11 mL, corresponding to the molecular weight of about 380 kDa. ES mutant complex elutes at 13 mL, corresponding to a molecular weight of 205 kDa. (C) SDS-PAGE analysis of the WT and ES mutant Rna14/Rna15 complex peak fractions showing that Rna15 is retained in both complexes. (D) Formaldehyde cross-linking of WT and ES mutant Rna14/Rna15 complex. Samples denoted (XL) have been cross-linked with 0.5% formaldehyde. (E) EMSA analysis of *GAL7* RNA binding. ES Rna14/Rna15 mutant still binds the *GAL7* RNA substrate, however with weaker affinity.

predicted molecular mass of the wild-type Rna14/Rna15 heterotetramer is 226 kDa, and the abnormally large experimental value for the wild-type is most likely due to the elongated shape of the Rna14 homodimer. Both the wild type and mutant are well separated from the void volume peak of the column, indicating that the peaks are from soluble proteins rather than molecular aggregates. Thus, consistent with dimer disruption, the wild-type Rna14/Rna15 complex migrates with approximately twice the molecular weight of the mutant. Moreover, SDS-PAGE of the peak fractions from the gel filtration column revealed that both the wild type and mutant yielded peaks that contain Rna15 (Figure 2C). The ratio of Rna14 to Rna15, measured by integrating the intensity of the stained bands from the gel filtration peak fractions, is 2.7 and

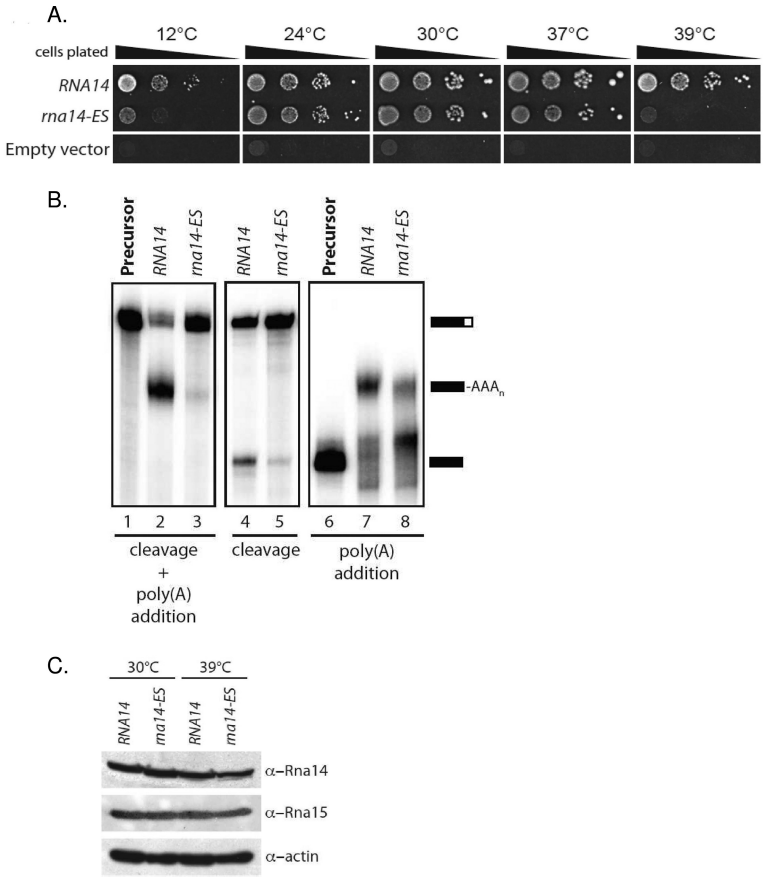


Figure 3. Effect of the R562E/Y563S mutation in Rna14 on yeast growth and mRNA processing. (A) Viability of strains at various temperatures when the *RNA14* shuffle strain is plated on 5FOA-containing media to force loss of the wild-type covering plasmid and leave only plasmid containing no insert (Empty vector) or the *RNA14* or *rna14- R562E/Y563S* genes. (B) *In vitro* processing assays using the indicated yeast extracts. Coupled cleavage and poly(A) addition assays were performed with precursor containing sequences upstream and downstream of the *GAL7* poly(A) site (left panel) and a cleavage only assay (middle panel). Poly(A) addition in the absence of cleavage was assayed using substrate that ends at the *GAL7* poly(A) site (right panel). The positions of precursors and products are shown on the right. (C) Western blot to test for expression levels of Rna14, rna14-ES, and Rna15 in yeast grown at 30 and 39 °C (shifted for 1 hour).

2.5 for wild type and ES mutant, respectively. The similarity in the Rna14/Rna15 ratios indicates that though the ES mutation disrupts the Rna14 interface, the absence of the Rna14 dimer does not cause an obvious defect in Rna15 binding. This suggests that the most important determinants of the Rna15 interface on Rna14 are formed by only one of the two Rna14 subunits. Thus, Rna15 most likely does not bind across the Rna14 dimer interface. These data are completely consistent with earlier studies that have shown that the C-terminal region of Rna14 is necessary and sufficient for Rna15 binding.[27,35]

To better address the stoichiometry of the ES mutant, we performed cross-linking experiments with formaldehyde. As shown in Figure 2D, the Rna14/Rna15 complex runs with molecular mass of about 260 kDa, consistent with the 2:2 stoichiometry that has been previously established for the wild-type complex.[25] The ES complex runs with a molecular mass of about 120 kDa, suggesting 1:1 stoichiometry. The predicted molecular mass of the 1:1 and 2:2 complexes are 113 and 226 kDa, respectively.

We also examined the stoichiometry using analytical ultracentrifugation (AUC). Consistent with an earlier report,[25] the sedimentation equilibrium AUC data of the wild-type Rna14/Rna15 complex fit best to a model wherein a stable 2:2 complex of Rna14 and Rna15 subunits weakly associates to form a 4:4 complex (Figure S1). We do not believe the heterooctamer species is of physiological significance because the interaction among heterotetramers is relatively weak (45uM). Moreover, the heterooctamer is not observed in later experiments wherein the Rna14/Rna15 complex is mixed with Clp1 and Pcf11. AUC data from the ES mutant complex proved difficult to fit to a model with simple, reversible equilibrium. The data were successfully fit to a model wherein noninteracting species of MW 114 ± 6 and 283 ± 40 kDa are present. The amount of the higher molecular weight species ranged from 13 to 20% of the total depending on the experiment. We believe this is most likely due to protein aggregation, though weak 2:2 complex formation in the context of the mutant cannot be ruled out (Figure S2). In summary, the AUC results are consistent with a model wherein the mutant largely disrupts the Rna14 dimer interface (yielding a species with a predicted molecular weight of 113 kDa). Consistent with the gel densitometry and cross-linking results, AUC also supports that the Rna14 mutant does not interfere significantly with the interaction between Rna14 and Rna15.

Disruption of the Rna14 Dimer Interface Impairs the Rna14/Rna15 Interaction with RNA and Its Function in mRNA 3′ End Processing. Rna15 binds the pre-mRNA via its N-terminal RNA recognition motif (RRM) domain.[30,34] Since there are two copies of Rna15 in the native complex, we tested the ability of the complex to bind RNA in the presence

of the ES mutation. *GAL7* 3′UTR RNA radiolabeled with ^{32}P was used in electromobility gel shift assay (EMSA). As shown in Figure 2E, the RNA bound less tightly to the mutant complex than to the wild type. A similar pattern is observed with *CYC1* 3′ UTR, a shorter RNA substrate (not shown). Interestingly, at higher ratios of protein to RNA, the electrophoretic mobility of the mutant and wild-type complexes are nearly identical. The fact that the 2:2 complex binds RNA more tightly than the 1:1 complex suggests that both copies of Rna15 may engage the RNA in the wild-type complex. The observation that the 1:1 complex exhibits similar mobility to the wild type at high protein concentrations reinforces this interpetation, as it appears that the RNA can act as a bridge, bringing together two 1:1 complexes in the context of the Rna14-ES mutant.

To test the functional consequences of this mutation *in vivo*, we introduced the *rna14-ES* mutant allele into yeast as the sole copy of the gene and characterized growth across a range of temperatures. Compared with the isogenic wild-type strain (*RNA14*), the *rna14-ES* cells were temperature-sensitive at 12 and 39 °C (Figure 3A). Next, we prepared whole-cell extracts from wild-type and *rna14-ES* strains grown at 30 °C and tested for their ability to cleave and polyadenylate synthetic radiolabeled *GAL7* precursor RNA *in vitro*. The wild-type extract results in efficient cleavage and polyadenylation of the substrate, but the processing activity of *rna14-ES* extract is strongly impaired (Figure 3B, lane 3). To discriminate between effects of the *rna14-ES* mutation on cleavage and poly(A) addition, cleavage only and polyadenylation only assays were carried out. The latter used a precleaved RNA substrate containing sequences upstream of the *GAL7* poly(A) site. Compared to wild-type extract, both the cleavage and polyadenylation activity of *rna14-ES* extract are impaired (Figure 3B, lanes 5 and 8). This demonstrates that the Rna14 dimerization mutant is defective for both steps of processing. To rule out protein instability as the cause of the observed defects, Western blot analysis was used to verify stable expression of Rna14 and Rna15 at both 30 and 39 °C (Figure 3C).

SAXS Analyis Suggests That the Two Rna15 Molecules Are near the Center of the Rna14 Dimer. To estimate the dimensions of the Rna14/Rna15 complex in solution, we employed small-angle X-ray scattering. The human Rna14 homologue, CstF-77, was used as a control for our methodologies since the dimensions of the very closely related murine CstF-77 are known from the X-ray structure. Moreover, unlike isolated Rna14 (which is prone to aggregation) full-length CstF-77 is very well behaved in the absence of other proteins. Neither yeast Rna14/Rna15 nor full-length human CstF-77 exhibited obvious concentration dependence to the shape of their scattering curve (not shown), confirming that the samples are well behaved in solution. As shown in Figure 4A, the scattering profiles of CstF-77 and Rna14/Rna15 are quite different, with the CstF-77 sample yielding a pronounced dip relative to the yeast Rna14/Rna15. The complex of Rna14 (677 amino acids) and Rna15 (296 residues) is somewhat larger than CstF-77 (717 residues). Thus, the difference in the scattering curves is likely due to both the added electron density from Rna15 and the fact that Rna14 is 40 residues shorter than its human homologue. We applied an indirect Fourier transformation procedure using the program GNOM to obtain the radii of gyration (R_g) and the pair distribution function $p(r)$. The calculated R_g values for Rna14/Rna15 and for CstF-77

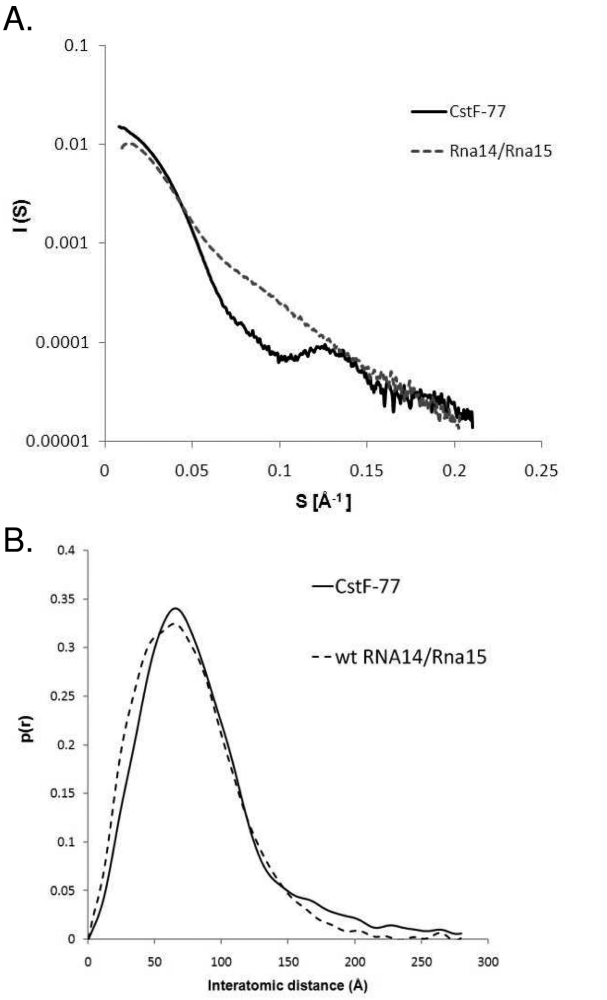


Figure 4. Small-angle X-ray scattering analysis. (A) The experimental SAXS curves for human CstF-77 (continuous line) and yeast Rna14/Rna15 complex (broken line). (B) $p(r)$ plots calculated from the SAXS data.

were 67.0 and 70.9, respectively. As shown in Figure 4B, the $p(r)$ function indicates that the Rna14/Rna15 complex has more interatomic vectors in the 0−50 Å range than CstF-77. This suggests that the extra amino acids are likely to be near the center of the Rna14 dimer rather than at the distal ends. The structure of murine CstF-77 showed it to be an elongated dimer, approximately 55 Å wide and 165 Å long. Consistent with this, the $p(r)$ curve indicates a significant number of interatomic vectors with lengths greater than 150 Å. Thus, while it is not possible to draw completely firm structural conclusions, the SAXS data suggest that the two Rna15 molecules are near the center of the Rna14 dimer. This is consistent with the crystal structures of truncated CstF-77 from mouse and from *E. cuniculi*. These structures lack the C-terminal domains necessary for Rna15 binding, but in both cases the truncated C-termini are also near the center of the complex.[26,27]

CF IA Complex Has Two Copies of the Rna14/Rna15 Dimer and One Copy of Clp1/Pcf11. To better understand the architecture of the larger complex, we overexpressed two additional subunits Clp1 and Pcf11 in bacteria. Along with Rna14 and Rna15, these subunits comprise the yeast complex known as CF IA. This complex is known to be stable in the

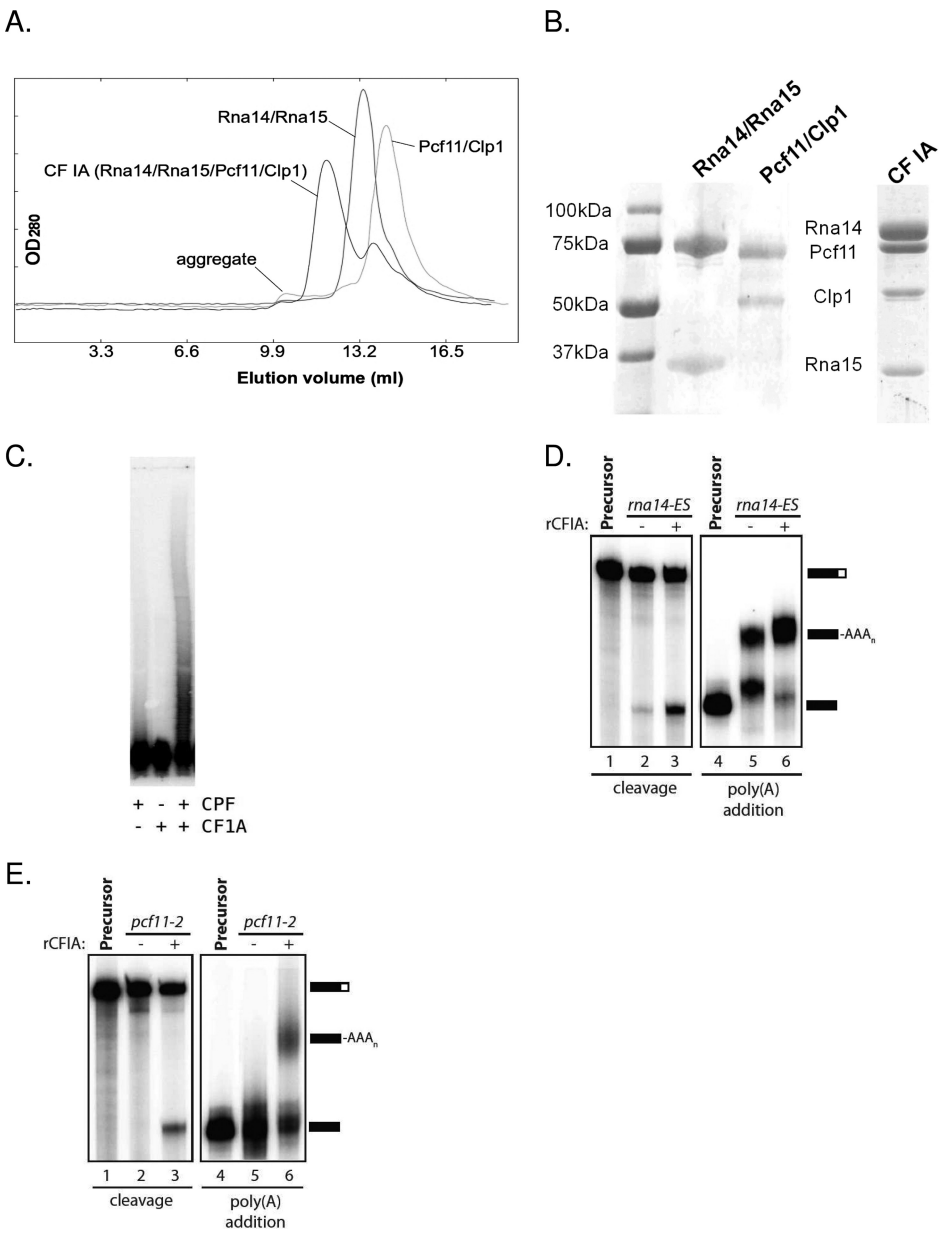


Figure 5. Reconstitution of the CF IA complex. (A) Size exclusion chromatography of Pcf11/Clp1, Rna14/Rna15, and the assembled CF IA. (B) SDS PAGE of the individual subcomplexes and the peak fractions of the wt CF IA complex. (C) Addition of recombinant CF IA to TAP-purified CPF enhances polyadenylation activity. (D) Addition of recombinant CF IA to a yeast extract with the *rna14-ES* mutation rescues both cleavage and polyadenylation activity. (E) Recombinant CF IA rescues activity when combined with yeast extract from the *pcf11-2* strain.

absence of RNA and can be purified, albeit with low yield, by placing a tandem affinity purification tag on Rna15.[43] As noted above, Clp1 and Pcf11 are known to form a tight binary complex.[32] Recombinant Pcf11 is largely insoluble on its own. To facilitate its expression, the *PCF11* and *CLP1* genes were cloned into a polycistronic vector and coexpressed. A hexahistidine tag on the N-terminus of Pcf11 allows the Clp1/Pcf11 complex to be purified by nickel affinity chromatography. Consistent with its predicted 1:1 stoichiometry, the complex migrates with an apparent molecular mass of 120 kDa by gel filtration (predicted MW = 122 kDa) (Figure 5A). Upon mixing Pcf11/Clp1 with recombinant wt Rna14/Rna15, the resulting SEC peak appears earlier than either the 2:2 Rna14/Rna15 complex or the Clp1/Pcf11 complex. The complex elutes well after dextran blue (MW 2 MDa), indicating that it is not simply a molecular aggregate. SDS-PAGE analysis confirmed the presence of all four protein subunits in the peak fraction (Figure 5B).

CF IA has no catalytic activity on its own. Its role is to tether the pre-mRNA to the larger complex, CPF, that contains the nonspecific RNA endonuclease, Ysh1, and the poly(A) polymerase, Pap1. We performed three tests to demonstrate that our reconstituted CF IA complex is functionally active. First, we added it to CPF that had been purified using a tandem affinity tag on the N-terminal end of Pta1 as described previously.[43] The recombinant complex significantly enhances polyadenylation activity (Figure 5C), but the poly(A) tails are not of a defined length. The indeterminate length is consistent with earlier studies which have shown that poly(A) binding proteins and Hrp1 (also known as CF IB) are necessary for tail length control.[43]

dx.doi.org/10.1021/bi200964p | *Biochemistry* 2011, 50, 10203–10214

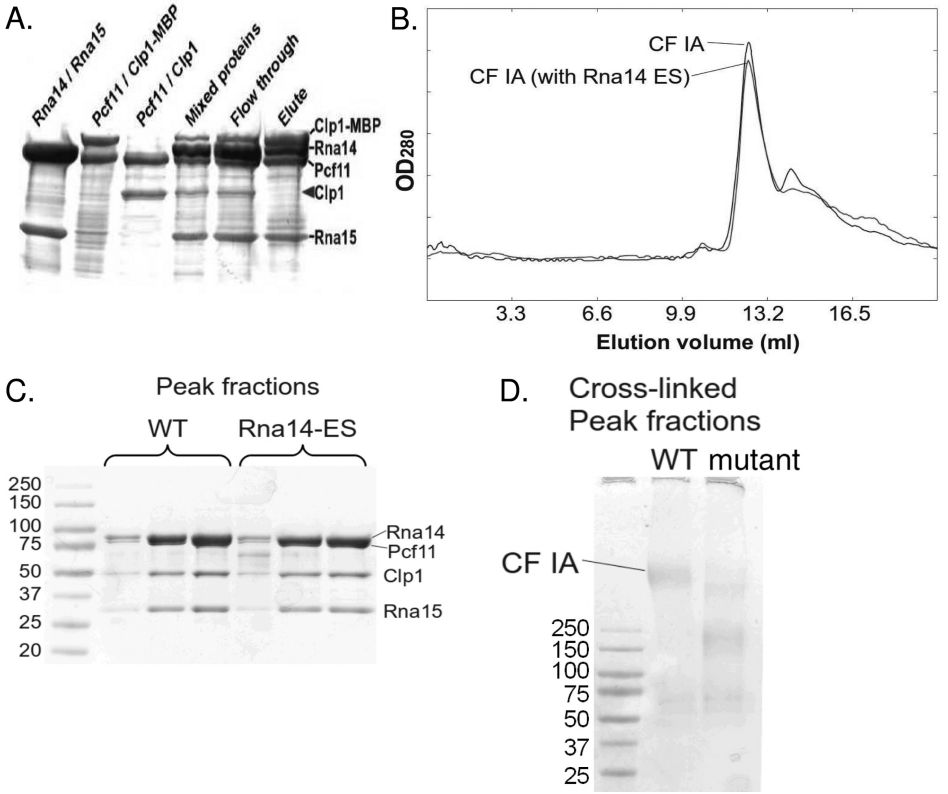


Figure 6. Stoichiometry of CF IA subunits. (A) The MBP-Clp1 fusion purifies along with Pcf11, and when added to the Rna14/Rna15 complex, CF IA can be purified using the MBP tag. MBP-tagged Clp1/Pcf11 was mixed with the untagged Clp1/Pcf11 and Rna14/15, and the resulting complex was purified using amylose resin. The untagged Clp1 did not copurify with the tagged Clp1. (B) Size exclusion chromatography of CF IA complex assembled from Clp1/Pcf11 and wt Rna14/Rna15 subcomplexes (solid line) and assembled with ES Rna14/Rna15 mutant complex (dashed line). (C) SDS gel showing peak fractions from the SEC run in 6B. (D) SDS gel showing cross-linked peak fractions using glutaraldehyde. A fraction of the mutant complex migrates at the same position on the gel as the wild type.

Next we tested the activity of recombinant CF IA by adding it to two functionally defective yeast extracts. The first extract was from the Rna14-ES mutant described earlier. The second contains the previously characterized Pcf11-2 mutation. Pcf11-2 causes temperature-sensitive yeast growth and a defect in in both cleavage and polyadenylation.[44] Addition of recombinant, reconstituted CF IA to either the extract with defective Rna14 or the extract with defective Pcf11 resulted in the rescue of both cleavage and polyadenylation activity (Figure 5D,E).

The relative intensity of the Coomassie stained bands of the peak fractions from gel filtration (Figure 5B) suggests a stoichiometry of 2:2:1:1 for Rna14, Rna15, Pcf11, and Clp1, respectively. The intensity of the four bands in the CF IA peak fractions was insensitive to changes in the ratio of the two binary subcomplexes loaded onto the column, and when mixed at a 2:2:1:1 ratio, we found that almost all of both proteins appeared in the peak fractions. When mixed at different ratios, we observed additional peaks corresponding to the excess complex (not shown). To further verify the stoichiometry, we introduced a maltose binding protein (MBP) tag (46 kDa) at the N-terminal end of Clp1. This tag allows us to purify the CF IA complex via an amylose column. We mixed the MBP-tagged Pcf11/Clp1 complex with the untagged Pcf11/Clp1 complex. This mixture was added to an equal quantity of the WT Rna14/Rna15 complex, and CF IA was purified via the MBP tag. Input and eluted fractions were analyzed by SDS-PAGE (Figure 6A). If the four subunits were present with the same stoichiometry, the untagged Clp1 would be detected in the eluate with the MBP-tagged complex. As this was not observed, we conclude that there is only one copy of the Clp1/Pcf11 subcomplex in CF IA.

Bridging by the Pcf11/Clp1 Complex Can Assemble Dimerization-Defective Rna14/Rna15 into CF IA. We were interested to know whether the Rna14-ES/Rna15 mutant complex (which exists as a 1:1 heterodimer) would retain its ability to bind Clp1/Pcf11. When the ES mutant Rna14/Rna15 complex was mixed with Pcf11/Clp1 and run on a gel filtration column, it formed a complex that exhibited the same chromatographic mobility as wild-type CF IA (Figure 6B), and the ratio of Rna14/Rna15 to Pcf11/Clp1 is similar to wild type (Figure 6C and Figure S3). Interestingly, cross-linking with guteraldehyde reveals that the mutant and wild-type complexes are probably not completely identical in structure. Cross-linking of the wild-type CF IA complex reveals a prominent band of MW greater than 300 kDa. Cross-linking of the mutant under the same conditions reveals only a bit of this high molecular weight species. Much of cross-linked product runs at ~200 kDa (Figure 6D). These data suggests that the Pcf11/Clp1 complex bridges the two copies of defective Rna14/Rna15 heterodimer, resulting in wild-type stoichiometry, albeit perhaps with a somewhat different structure. This structural alteration may be the reason the Rna14-ES mutant exhibits a temperature-sensitive phenoype when expressed in yeast and exhibits significantly reduced activity in our *in vitro* RNA processing assays.

Rna15 Is Attached to Rna14 through a Proteolytically Sensitive Linker. To further probe the molecular architecture, the Rna14/Rna15 complex was subjected to limited digestion with trypsin. Using a protease to substrate ratio of 1:7000, most of a ~10 kDa region in Rna14 was removed within 3 h. Rna15 was largely unaffected by the protease treatment under these conditions (Figure 7A). As in the earlier

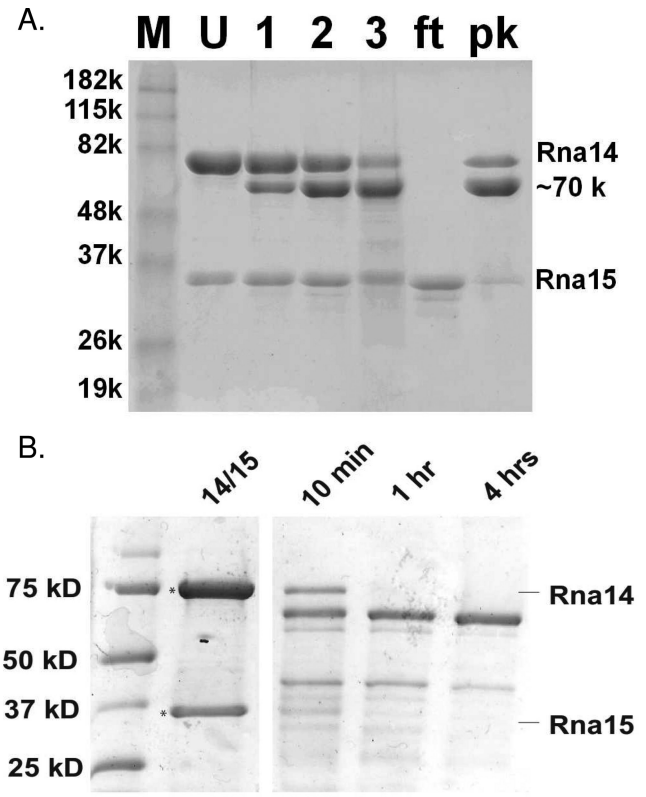


Figure 7. Limited proteolysis. (A) Trypsin cleavage of Rna14/Rna15. Lanes: M, molecular weight markers; U, undigested Rna14/Rna15; 1, Trypsin digest after 1 h; 2, after 2 h; 3, after 3 h; ft; flow through of nickel affinity column; pk, peak fraction from the nickel affinity column. The digested Rna14 product has an apparent molecular weight of ~70 kDa. (B) Chymotrypsin cleavage of Rna14/Rna15.

reconstitution experiments, the Rna14/Rna15 complex was purified via an N-terminal His-tag on Rna14, and Rna15 was untagged. After proteolysis, the truncated Rna14 fragment rebound to the Ni-NTA resin (labeled pk), indicating that the N-terminus of Rna14 was unaffected by proteolysis. The untagged Rna15, however, was in the flow-through fraction, indicating, as previously reported, that it binds the C-terminal end of Rna14.[27] In the crystal structures of murine and *E. cuniculi* Rna14 homologues, the C-termini were removed through genetic manipulation to facilitate crystallization.[26,27] Limited digestion of the Rna14/Rna15 complex with chymotrypsin produces an Rna14 fragment of similar size as with trypsin; however, in this case Rna15 is also degraded (Figure 7B). The susceptibility of the Rna14 C-terminal region to cleavage by proteases with very different sequence specificities indicates that it is likely to be flexible. As discussed below, such flexibility may be functionally important. It suggests that the relative orientation of the RNA-binding domains within the two copies of Rna15 may not be strictly defined by the other proteins in the complex.

Interaction of Rna15 Strengthens the Interaction of Rna14 with Pcf11/Clp1. To further investigate potential interactions among CF IA subunits, we expressed Rna15 in the absence of Rna14 and purified Rna15 via an N-terminal hexahistidine tag. Isolated Rna15 ran a monomer on SEC, and when run together with Pcf11/Clp1, no interaction was observed (Figure 8A). Pull-down experiments have shown a

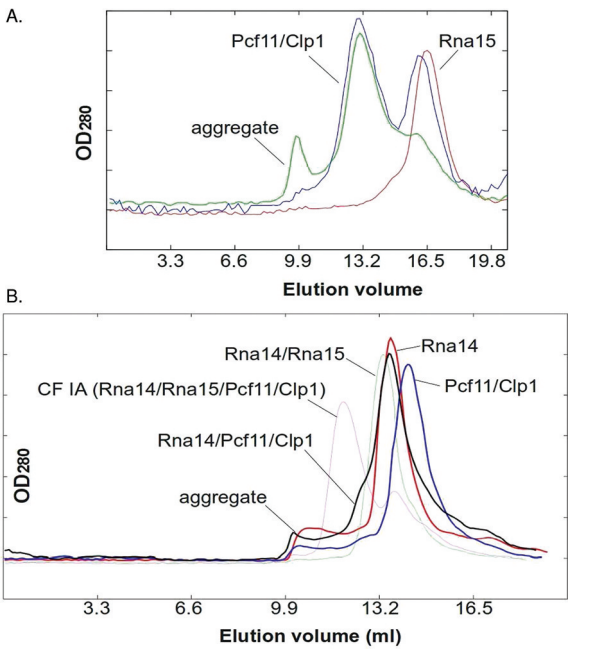


Figure 8. Gel filtration with isolated Rna15 and Rna14. (A) Size exclusion chromatography of isolated Rna15 (red), Pcf11/Clp1 (green), and Rna15/Pcf11/Clp1 (blue). (B) Size exclusion chromatography of isolated Rna14 (red), Pcf11/Clp1 (blue), and Rna14/Pcf11/Clp1 (black). The Rna14/Rna15 complex (green) and CF IA (magenta) traces are also shown for reference.

weak interaction between Rna15 and Pcf11 that requires the C-terminal 53 amino acids of Rna15.[28] The absence of an interaction as assayed by gel filtration suggests that Rna15 alone interacts relatively weakly with Pcf11/Clp1. In contrast, and as shown in Figure 5, the Rna14/Rna15 complex forms a tight complex with Pcf11/Clp1 in the same assay.

We next sought to examine the interaction of isolated Rna14 with Pcf11/Clp1. Rna14 is normally insoluble when expressed on its own in bacteria, but we were able to obtain full-length, soluble His6-Rna14 in the absence of Rna15 by coexpressing Rna14 with a version of Rna15 that contained maltose binding protein at its N-terminus. The MBP-Rna15 was soluble, but it did not interact with Rna14 (not shown). This suggests that the N-terminus of Rna15 is physically close to Rna14 within the complex. The isolated Rna14 appeared to run as a dimer by SEC gel filtration, and we repeatedly observed a shoulder on the main gel filtration peak, indicating formation of a higher molecular weight complex in the presence of Pcf11/Clp1 (Figure 8B). This binding is clearly not as strong as when the Rna14/Rna15 complex is mixed with Pcf11/Clp1. Thus, the gel filtration results indicate that that the Rna14 interaction with Pcf11/Clp1 is tighter than that of Rna15, but less tight than that of the Rna14/Rna15 complex.

DISCUSSION

Though we do not yet have sufficient information to build a molecular model of yeast CF I, our results, along with existing biochemical and genetic findings, have allowed us to propose a topological model of this complex (Figure 9). Support for this model comes from analysis of functional, recombinant CF IA and in particular from the Rna14-ES mutant which disrupts the Rna14 dimer interface. This mutant has no obvious effect on the interaction of Rna14 with Rna15, and the mutant complex retains its ability to bind RNA *in vitro*. SEC, AUC, and cross-linking experiments indicate that whereas the wild-type Rna14/Rna15 complex has 2:2 stoichiometry, the mutant forms a tight 1:1 complex. This suggests that the primary Rna15 binding site on Rna14 is formed by a single Rna14 subunit. Consistent with recent results showing that that the C-terminal "monkey tail" of Rna14 is responsible for Rna15 binding,[27,35] we found that digestion of the Rna14 C-terminus with trypsin caused Rna15 to be lost from the complex. In addition, SAXS results reported here indicate that the two Rna15 subunits are likely to be near the center of the elongated Rna14 dimer and that the maximum dimension of the Rna14/Rna15 complex is similar to that of a human CstF-77 sample that lacked the Rna15 homologue. Yeast expression of the mutant Rna14 in place of the wild-type protein resulted in temperature sensitive growth and extracts from these yeast were defective in 3′ end processing.

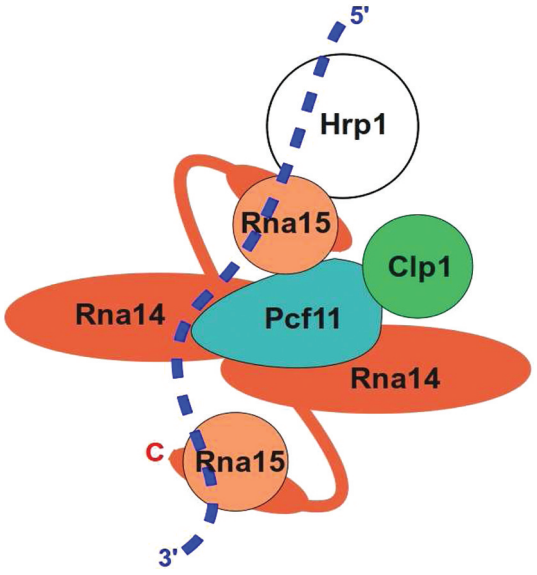


Figure 9. Topological model of the CF I complex. In this model, the Rna14 dimer interacts in an asymmetric manner with Pcf11, and both copies of Rna15 interact with the pre-mRNA (dotted line). Both copies of Rna15 are tethered to the complex via the flexible region of Rna14 that separates the structured HAT repeats (large red ovals) from the monkey tail domain at the C-terminus that interacts with Rna15 (smaller red ovals behind Rna15). One of the Rna15 subunits makes additional interactions with Pcf11 and Hrp1. We do not know which of the two Rna15 subunits interacts with Pcf11. Clp1 is shown binding Pcf11, but not other subunits.

We also report a method for soluble, bacterial expression of the Pcf11/Clp1 complex. Using cross-linking, SEC, and a mixing experiment wherein Clp1 is tagged with MBP, we show that the Rna14/Rna15 2:2 heterotetramer binds only one copy of the Pcf11/Clp1heterodimer. The resulting 2:2:1:1 complex can be purified by gel filtration, and it is functionally active in assays where the recombinant, reconstituted complex was added to either functionally deficient yeast extracts or TAP-purified CPF lacking CF I. The mutant which disrupts the Rna14 dimer exhibits the same stoichiometry as the wild type when combined with the Pcf11/Clp1 complex, suggesting that the Pcf11/Clp1 complex binds in such a way as reassemble the Rna14/Rna15 heterotetramer.

We have shown that isolated Rna14 forms a complex with Pcf11/Clp1 that is less tight than the complex formed when Rna14/Rna15 are mixed with Pcf11/Clp1.

Earlier work demonstrated that Clp1 interacts tightly with Pcf11, but not with Rna14 or Rna15, and that Pcf11 interacts most tightly with Clp1 and Rna14.[24] Our results with the full-length bacterially overexpressed proteins are completely consistent with these earlier data. In our experiments, isolated Rna15 did not appear to form a complex with Pcf11/Clp1 on gel filtration, but since Rna15 clearly strengthens the connection between Rna14 and Pcf11/Clp1, it is not unlikely that such an interaction exists. Indeed, interaction of these two proteins was seen in earlier pull-down experiments and the Pcf11 interaction was shown to involve the structured, helical domain at the Rna15 C-terminus.[28] We suspect the gel filtration assay used in this study may be too stringent to allow the Rna15−Pcf11 interaction to be seen.

The stoichiometry of CF IA requires that the single copy of Pcf11 within the complex form an asymmetric interaction with the 2:2 assembly of Rna14 and Rna15. CstF-77, and most likely Rna14, forms a symmetric dimer in the absence of other proteins. Thus, in principle, there should be two equivalent binding sites for Pcf11. That only one of these sites is occupied indicates that the other is either sterically blocked (i.e., by Pcf11 when it binds the first site) or that the Rna14 dimer is distorted by Pcf11 binding such as to disrupt the second binding site. In either case, Pcf11 most likely binds near the center of the Rna14 dimer, for this is where the dimer interface is. Our SAXS results also place Rna15 relatively close to the center of the complex. The stoichiometry of the entire cleavage/polyadenylation complex is not yet known, but it makes sense that the symmetry of the Rna14/Rna15 heterotetramer should be broken. Were it not, CF IA would presumably recruit two copies of the larger complex, CPF, to the pre-mRNA. This would entail two copies of the RNA-processing nuclease, Ysh1, two copies of the poly(A) polymerase, Pap1, and presumably two copies of the 13 other CPF proteins.

Higher eukaryotes contain CstF-50, a dimeric protein not found in yeast that associates tightly with the Rna14 and Rna15 homologues (CstF-77 and CstF-64, respectively).[45] We anticipate that mammalian Pcf11 will break the symmetry of the CstF complex in the same way as the yeast protein breaks the symmetry of the Rna14/Rna15 complex. Mammals also contain two copies of the proteins that constitute the CF Im complex, and RNA is bound by both copies.[46,47] Although the topology of the CFIm25/CFIm 68 complex appears similar to that of Rna14/Rna15, CF Im binds upstream of the CstF and CPSF complexes and has no obvious homology with CstF or yeast CF I.

The two Rna15 molecules of CF IA are attached to the complex primarily through the proteolytically sensitive C-terminal ends of Rna14. Limited trypsin and chymotrypsin digests of the Rna14/Rna15 complex result in Rna14 fragments of very similar size (~70 kDa). The ~10 kDa C-terminal fragment of Rna14 that is removed includes the monkey tail region that forms a tight complex with the central hinge region

of Rna15.[35] That proteases with different specificities cleave Rna14 in the same region suggests that this part of the molecule, between the structured N-terminus (residues 30–584, based on our homology modeling) and the monkey tail domain (residues 626–677), is at least partially disordered. This suggests that the two copies of Rna15 may be capable of movement relative to the bulk of Rna14 which is much less sensitive to proteolysis. Importantly, SEC of Rna15 indicates that the isolated protein does not self-associate. Thus, the two copies of Rna15 may be physically separated. Moreover, since there is only one copy of Pcf11, we expect that the interaction between Rna15 and Pcf11 involves one, and not necessarily both, copies of Rna15. The unbound copy of Rna15 is probably attached only via the Rna14 monkey tail domain. RNA binding involves the RRM domain within Rna15. Having two copies of Rna15 may facilitate binding to a range of RNA sequences. The flexibility of the C-terminal region of Rna14 may also help break the structural symmetry of the Rna14/Rna15 hetero-tetramer when other proteins bind.

Rna15 and Hrp1 have both been shown to interact with RNA via their respective RRM domains,[29,33] and recently the structures of the RRM domains of both of these proteins have been been solved bound to a single RNA molecule that includes the efficiency element (UAUAUA) and an adjacent positioning element (AAUAAU) sequence.[34] This structure and related biochemical work allows us to extend our topological model of CF IA to include Hrp1 (aka CF IB). The model shown in Figure 9 has Pcf11 interacting with the copy of Rna15 that also interacts with Hrp1. However, the available data cannot distinguish between this arrangement of subunits and one in which Pcf11 interacts with the other copy of Rna15. We also cannot distinguish between these models and one in which Pcf11 interacts with both Rna15 subunits, but this seems unlikely because there is only one copy of Pcf11 and the Pcf11 sequence does not exhibit significant internal repeats. Regardless of which model is correct, the second Rna15 subunit is likely to play a role in RNA binding. Isolated Rna15 binds U-rich and GU rich sequences with ~10 uM affinity,[29] and such sequences are often present on the 5′ and 3′ sides of the cleavage site.

Clearly, a complete validation of the CF I model and its implications for RNA binding and specificity will require additional work. We believe this work will be facilitated by the recombinant CF IA complex described here and that the topological considerations discussed will help shape thinking regarding the structure and function of this important macromolecular complex. Together with work on the other 3′ end processing subcomplexes, these studies will provide a structural framework for understanding how specific RNA sequences direct 3′ end processing and why this processing often occurs at alternative sites on the pre-mRNA.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Sedimentation equilibrium data for the Rna14-ES mutant complex as well as additional cross-linking data and gel filtration runs. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 617-636-2994. Fax: 617-636-2409. E-mail: Andrew.Bohm@tufts.edu.

Author Contributions

⊥These two authors contributed equally to this work.

Funding

This work was supported by NSF grant MCB-112768 (A.B.) and NIH grants GM65972 (A.B.), GM41752 (C.M. and J.K.), and K12GM074869 (J.K.).

■ ABBREVIATIONS

CF I, cleavage factor I; CF IA, cleavage factor IA; CPF, cleavage polyadenylation factor; AUC, analytical ultracentrifugation; SEC, size exclusion chromatography; SAXS, small-angle X-ray scattering; CstF, cleavage stimulation factor; RRM, RNA recognition motif.

■ REFERENCES

(1) Millevoi, S., and Vagner, S. (2010) Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. *Nucleic Acids Res. 38*, 2757–2774.

(2) Mandel, C. R., Bai, Y., and Tong, L. (2008) Protein factors in pre-mRNA 3′-end processing. *Cell. Mol. Life Sci. 65*, 1099–1122.

(3) Danckwardt, S., Hentze, M. W., and Kulozik, A. E. (2008) 3′ end mRNA processing: molecular mechanisms and implications for health and disease. *EMBO J. 27*, 482–498.

(4) Ohnacker, M., Barabino, S. M., Preker, P. J., and Keller, W. (2000) The WD-repeat protein pfs2p bridges two essential factors within the yeast pre-mRNA 3′-end-processing complex. *EMBO J. 19*, 37–47.

(5) Chen, J., and Moore, C. (1992) Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. *Mol. Cell. Biol. 12*, 3470–3481.

(6) Kessler, M. M., Zhao, J., and Moore, C. L. (1996) Purification of the Saccharomyces cerevisiae cleavage/polyadenylation factor I. Separation into two components that are required for both cleavage and polyadenylation of mRNA 3′ ends. *J. Biol. Chem. 271*, 27167–27175.

(7) Mandel, C. R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J. L., and Tong, L. (2006) Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. *Nature 444*, 953–956.

(8) Balbo, P. B., and Bohm, A. (2007) Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure 15*, 1117–1131.

(9) Barillà, D., Lee, B. A., and Proudfoot, N. J. (2001) Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A. 98*, 445–450.

(10) Hammell, C. M., Gross, S., Zenklusen, D., Heath, C. V., Stutz, F., Moore, C., and Cole, C. N. (2002) Coupling of termination, 3′ processing, and mRNA export. *Mol. Cell. Biol. 22*, 6441–6457.

(11) Brodsky, A. S., and Silver, P. A. (2000) Pre-mRNA processing factors are required for nuclear export. *RNA 6*, 1737–1749.

(12) Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., and Schroth, G. P. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature 456*, 470–476.

(13) Ji, Z., and Tian, B. (2009) Reprogramming of 3′ untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS One 4*, e8419.

(14) Shepard, P. J., Choi, E.-A., Lu, J., Flanagan, L. A., Hertel, K. J., and Shi, Y. (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA 17*, 761–772.

(15) Yan, J., and Marr, T. G. (2005) Computational analysis of 3′-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. *Genome Res. 15*, 369–375.

(16) Tian, B., Hu, J., Zhang, H., and Lutz, C. S. (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res. 33*, 201–212.

(17) Jan, C. H., Friedman, R. C., Ruby, J. G., Bartel, D. P. (201AD) Formation, regulation and evolution of Caenorhabditis elegans 3′UTRs, *Nature 469*, 97−101.
(18) Chan, C. S., Elemento, O., and Tavazoie, S. (2005) Revealing posttranscriptional regulatory elements through network-level conservation. *PLoS Comput. Biol 1*, e69.
(19) Friedman, R. C., Farh, K. K.-H., Burge, C. B., and Bartel, D. P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res. 19*, 92−105.
(20) Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008) 3′ UTRs are the primary regulators of gene expression in the C. elegans germline. *Curr. Biol. 18*, 1476−1482.
(21) Sandberg, R., Neilson, J. R., Sarma, A., and Sharp, P. A. (2008) Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. *Science 320*, 1643−1647.
(22) Mayr, C., and Bartel, D. P. (2009) Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell 138*, 673−684.
(23) Fu, Y., Sun, Y., Li, Y., Li, J., Rao, X., Chen, C., and Xu, A. (2011) Differential genome-wide profiling of tandem 3′ UTRs among human breast cancer and normal cells by high-throughput sequencing. *Genome Res. 21*, 741−747.
(24) Gross, S., and Moore, C. (2001) Five subunits are required for reconstitution of the cleavage and polyadenylation activities of Saccharomyces cerevisiae cleavage factor I. *Proc. Natl. Acad. Sci. U. S. A. 98*, 6080−6085.
(25) Noble, C. G., Walker, P. A., Calder, L. J., and Taylor, I. A. (2004) Rna14-Rna15 assembly mediates the RNA-binding capability of Saccharomyces cerevisiae cleavage factor IA. *Nucleic Acids Res. 32*, 3364−3375.
(26) Bai, Y., Auperin, T. C., Chou, C.-Y., Chang, G.-G., Manley, J. L., and Tong, L. (2007) Crystal structure of murine CstF-77: dimeric association and implications for polyadenylation of mRNA precursors. *Mol. Cell 25*, 863−875.
(27) Legrand, P., Pinaud, N., Minvielle-Sébastia, L., and Fribourg, S. (2007) The structure of the CstF-77 homodimer provides insights into CstF assembly. *Nucleic Acids Res. 35*, 4515−4522.
(28) Qu, X., Perez-Canadillas, J.-M., Agrawal, S., De Baecke, J., Cheng, H., Varani, G., and Moore, C. (2007) The C-terminal domains of vertebrate CstF-64 and its yeast orthologue Rna15 form a new structure critical for mRNA 3′-end processing. *J. Biol. Chem. 282*, 2101−2115.
(29) Pancevac, C., Goldstone, D. C., Ramos, A., and Taylor, I. A. (2010) Structure of the Rna15 RRM-RNA complex reveals the molecular basis of GU specificity in transcriptional 3′-end processing factors. *Nucleic Acids Res. 38*, 3119−3132.
(30) Pérez Cañadillas, J. M., and Varani, G. (2003) Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *EMBO J. 22*, 2821−2830.
(31) Meinhart, A., and Cramer, P. (2004) Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. *Nature 430*, 223−226.
(32) Noble, C. G., Beuth, B., and Taylor, I. A. (2007) Structure of a nucleotide-bound Clp1-Pcf11 polyadenylation factor. *Nucleic Acids Res. 35*, 87−99.
(33) Pérez-Cañadillas, J. M. (2006) Grabbing the message: structural basis of mRNA 3′UTR recognition by Hrp1. *EMBO J. 25*, 3167−3178.
(34) Leeper, T. C., Qu, X., Lu, C., Moore, C., and Varani, G. (2010) Novel Protein-Protein Contacts Facilitate mRNA 3′-Processing Signal Recognition by Rna15 and Hrp1. *J. Mol. Biol. 401*, 334−349.
(35) Moreno-Morcillo, M., Minvielle-Sébastia, L., Fribourg, S., and Mackereth, C. D. (2011) Locked Tether Formation by Cooperative Folding of Rna14p Monkeytail and Rna15p Hinge Domains in the Yeast CF IA Complex. *Structure 19*, 534−545.
(36) Boeke, J. D., Trueheart, J., Natsoulis, G., and Fink, G. R. (1987) 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol. 154*, 164−175.
(37) Bonneaud, N., Minvielle-Sebastia, L., Cullin, C., and Lacroute, F. (1994) Cellular localization of RNA14p and RNA15p, two yeast proteins involved in mRNA stability. *J. Cell. Sci 107* (Pt 4), 913−921.
(38) Zhao, J., Kessler, M. M., Helmling, S., O'Connor, J. P., and Moore, C. L. (1999) Pta1,a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol. Cell. Biol. 19* (11), 7733−40.
(39) Stafford, W. F., and Sherwood, P. J. (2004) Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. *Biophys. Chem. 108*, 231−243.
(40) Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr. 25*, 495−503.
(41) Eswar, N., Eramian, D., Webb, B., Shen, M.-Y., and Sali, A. (2008) Protein structure modeling with MODELLER. *Methods Mol. Biol. 426*, 145−159.
(42) Sreerama, N., and Woody, R. W. (2004) Computation and analysis of protein circular dichroism spectra. *Methods Enzymol. 383*, 318−351.
(43) Dheur, S., Nykamp, K. R., Viphakone, N., Swanson, M. S., and Minvielle-Sebastia, L. (2005) Yeast mRNA Poly(A) tail length control can be reconstituted in vitro in the absence of Pab1p-dependent Poly(A) nuclease activity. *J. Biol. Chem. 280*, 24532−24538.
(44) Amrani, N., Minet, M., Wyers, F., Dufour, M. E., Aggerbeck, L. P., and Lacroute, F. (1997) PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Mol. Cell. Biol. 17*, 1102−1109.
(45) Moreno-Morcillo, M., Minvielle-Sébastia, L., Mackereth, C., and Fribourg, S. (2011) Hexameric architecture of CstF supported by CstF-50 homodimerization domain structure. *RNA 17*, 412−418.
(46) Yang, Q., Coseno, M., Gilmartin, G. M., and Doublié, S. (2011) Crystal Structure of a Human Cleavage Factor CFI(m)25/CFI(m)68/RNA Complex Provides an Insight into Poly(A) Site Recognition and RNA Looping. *Structure 19*, 368−377.
(47) Li, H., Tong, S., Li, X., Shi, H., Ying, Z., Gao, Y., Ge, H., Niu, L., and Teng, M. (2011) Structural basis of pre-mRNA recognition by the human cleavage factor I(m) complex. *Cell Res. 21*, 1039−1051.