The C-terminal Domains of Vertebrate CstF-64 and Its Yeast Orthologue Rna15 Form a New Structure Critical for mRNA 3’-End Processing*§

Xiangping Qu†1, Jose-Manuel Perez-Canadillas§1,2, Shipra Agrawal‡3, Julia De Baecke§4, Hailing Cheng§4, Gabriele Varani§5, and Claire Moore‡6

From the †Department of Molecular Microbiology, Tufts University School of Medicine and the Sackler Graduate School of Biomedical Sciences, Boston, Massachusetts 02111 and the §Department of Biochemistry and Department of Chemistry, University of Washington, Seattle, Washington 98195

Yeast Rna15 and its vertebrate orthologue CstF-64 play critical roles in mRNA 3’-end processing and in transcription termination downstream of poly(A) sites. These proteins contain N-terminal domains that recognize the poly(A) site, but little is known about their highly conserved C-terminal regions. Here we show by NMR that the C-terminal domains of CstF-64 and Rna15 fold into a three-helix bundle with an uncommon topological arrangement. The structure defines a cluster of evolutionary conserved yet exposed residues we show to be essential for the interaction between Pcf11 and Rna15. Furthermore, we demonstrate that this interaction is critical for the function of Rna15 in 3’-end processing but dispensable for transcription termination. The C-terminal domain of the Rna15 homologue Pti1 contains critical sequence alterations within this region that are predicted to prevent Pcf11 interaction, providing an explanation for the distinct functions of these two closely related proteins in the 3’-end formation of RNA polymerase II transcripts. These results define the role of the C-terminal half of Rna15 and provide insight into the network of protein/protein interactions responsible for assembly of the 3’-end processing apparatus.

The formation of the 3’-ends of eukaryotic mRNA by cleavage and polyadenylation occurs co-transcriptionally and is intimately linked to splicing, assembly of an export-competent messenger ribonucleoprotein complex (mRNP), and nuclear surveillance mechanisms that monitor the quality of the exported mRNP (1–3). Transcription termination requires the 3’-end processing machinery as well as a 5’-3’-exonuclease (4–6). The precise nature of the interactions between the polyadenylation and transcription machineries and of their changes during transcription termination remains to be elucidated. To fully understand these connections and their regulation, it is important to develop a clear picture of the organization and function of the polyadenylation machinery at a molecular level.

Among the proteins that ensure a communication between transcription and 3’-end processing are human CstF-64 and its Saccharomyces cerevisiae orthologue Rna15, two of the earliest identified components of the polyadenylation machinery (7, 8). These proteins play a key role in mRNA 3’-end processing by recognizing one of the signal sequences that specify poly(A) sites (9, 10). Altering the levels of CstF-64 or mutating Rna15 can shift the pattern of poly(A) site utilization when multiple processing sites are present on the same transcript (11–15). Defects in Rna15 also have consequences on mRNP biogenesis that transcend 3’-end processing, such as hindering transcription elongation and preventing transcription termination downstream of poly(A) sites (16–18). The G/U-rich sequences that represent binding sites for CstF-64 are essential for proper termination of transcription (19), indicating that the functions of Rna15 and CstF-64 in transcription are conserved as well. Mutation of RNAS15 also causes poor recruitment of the Yra1 export factor to actively transcribed genes (20), retention of mRNA in the nucleus (18, 21), and poor release of mRNA from the site of transcription (22). The defective mRNPs produced by Rna15 mutant cells are recognized and degraded by the nuclear exosome (23).

* This work was supported by National Institutes of Health Grants GM041752 (to C. M.) and GM64440 (to G. V.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† The on-line version of this article (available at http://www.jbc.org/) contains supplemental Tables I and II and Figs. 1 and 2.

‡ Both authors contributed equally to this work.

§ Present address: Instituto de Química Física Rocosolano. CSIC, Serrano 119, Madrid 28006, Spain.

§ Present address: Comprehensive Cancer Center, Ohio State University, Columbus, Ohio 43221.

© Present address: Dept. of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115.

† To whom correspondence may be addressed: Dept. of Chemistry, University of Washington, Bagley Hall, Seattle, WA 98195-1700. Tel.: 206-543-7113; Fax: 206-685-8663; E-mail: varanil@chem.washington.edu.

‡ To whom correspondence may be addressed: Dept. of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Ave., Boston, MA 02111. Tel.: 617-636-6935; Fax: 617-636-0337; E-mail: Claire.moore@tufts.edu.

The abbreviations used are: mRNP, messenger ribonucleoprotein; RNA, RNA polymerase; HSQC, heteronuclear single quantum coherence; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; ChIP, chromatin immunoprecipitation; PMSF, phenylmethylsulfonyl fluoride; DTT, dithiothreitol; MBP, myelin basic protein; snoRNA, small nucleolar RNA; CTD, C-terminal domain; RRM, RNA recognition motif; CF IA, cleavage factor Iα; CPF, cleavage and polyadenylation factor.
Function of the Rna15 C Terminus in mRNA 3′-End Processing

The Rna15 and CstF-64 proteins are closely related in sequence and structure at their N and C termini (Fig. 1A). These two proteins are also homologous to Pti1 (24–26), a component of the APT complex involved in S. cerevisiae snRNA 3′-end formation (26). These three proteins contain RRM-type RNA binding domains at their N termini that are both necessary and sufficient for RNA binding (27–29). Mutation of conserved amino acids in this motif within Rna15 impairs the ability of Rna15 to interact with RNA and to function in mRNA 3′-end processing (10). Metazoan CstF-64 proteins have a conserved 100 amino acid region called the “hinge” immediately C-terminal to the RRM that is also found in Pti1 and Rna15 (24). This region of CstF-64 interacts with CstF-77 and with Symplekin, a protein implicated in the assembly of several RNA 3′-end processing machineries (30–34). The hinge domain of Pti1 interacts with Rna14 and Pta1, the yeast orthologues of CstF-77 and Symplekin, respectively (24). The central region of CstF-64 is variable in length and composition, and its function is unknown. In most metazoans, it is rich in proline and glycine and contains up to a dozen repeats of the MEAR(A/G) motif, which are missing altogether in worms and fungi (32).

The 50 amino acids at the very C terminus of the Rna15/CstF-64 family are even more highly conserved than the N-terminal RRM. A mutant of the Schizosaccharomyces pombe homologue, Ctf1, lacking the last 70 amino acids (including this C-terminal domain) caused transcription to continue downstream of the poly(A) site, leading to the proposal that this region participates in termination of transcription (35). Consistent with this idea, the last 100 or so amino acids of Rna15/hCstF-64, and of Ctf1, were shown genetically and biochemically to interact directly with the transcription factors Sub1/PC4 and Res2 (35, 36). This region in Ctf1 is thought to function only in transcription termination and not in 3′-end processing, because its loss does not affect accumulation of properly processed mRNA in vivo (35). However, the precise role of the Rna15/CstF64 C terminus in cleavage and polyadenylation has not been addressed.

Rna15 is part of CF IA (37, 38) and interacts directly with the Rna14 and Pcf11 subunits of this factor (39), and with the poly(A)-binding protein Pab1 (40). The interaction between Rna15 and Pcf11 is particularly interesting in light of recent reports that provide a role for the N terminus of Pcf11 in the promotion of transcription termination (41, 42). However, the domains of Rna15 important for the contacts with Rna14 and Pcf11, and the role of these interactions in 3′-end processing and termination, have not been delineated. The three-dimensional organization of the regions of the protein beyond the RRM also remains to be determined.

In this study, we have investigated the interaction of the C-terminal half of Rna15 with Rna14, Sub1, and Pcf11. We report that the C-terminal domain of the Rna15/CstF-64 protein family forms an unusual structure of three helices that positions highly conserved residues on the protein surface to mediate protein/protein interactions. We demonstrate that this highly conserved domain functions in the cleavage and polyadenylation of mRNA precursors by providing a contact site for Pcf11, whereas the region of the protein just N-terminal to it interacts with Rna14 and Sub1. These studies provide structural and biochemical insight into highly conserved protein/protein interactions responsible for assembly of the 3′-end processing apparatus in all eukaryotes.

Experimental Procedures

Yeast Strains, Media, and Plasmids—The yeast strains LM31 (MATa RNA15::TRP1 ura3-1 trp1-1 ade2-1 leu2-3,112 his3-11, plus pLM13 containing RNA15 and URA3) was a gift from F. Lacroute (43), and the pcf11-9 and isogenic wild-type strains were a gift from W. Keller (44). Yeasts were grown and maintained on YPD (1% yeast extract, 2% peptone, 2% glucose) or on selective media as needed. For temperature sensitivity assays, yeasts were grown at 16, 25, 30, and 37 °C. Loss of URA3 plasmids from yeast cells was accomplished by plating on solid medium containing 5-fluoroorotic acid.

The RNA15-His6 coding sequence from plasmid pET21b/Rna15 (37) was subcloned into the LEU2 vector YeaPlac181 to generate YeaPlac-RNA15H6. Truncations of the RNA15 open reading frame from the 3′-end were amplified using the primers listed in Table 1 and cloned into pET21b (Novagen) between the Ndel and Xhol sites. The mutants rna15-CTD1 and rna15-CTD2 were constructed using a forward PCR primer (supplemental Table 1) containing the desired mutation and a reverse primer corresponding to the end of the RNA15 coding sequence. The wild-type BamHI/Ndel fragment on pET21b/Rna15 was then replaced with the mutated PCR fragment digested with BamHI and Ndel. Mutations were confirmed by sequencing. BclI-Bpl fragments from these mutant rna15 constructs were subcloned into BclI-Bpl-treated YeaPlac-RNA15H6. This was accomplished by first growing the plasmids in the dam− Escherichia coli strain UF253. All of the YeaPlac-RNA15H6 variants were transformed into LM31 for expression in yeast and phenotypic studies. Additional truncations of Rna15 (t1, t2, t3, t4, and t5, see supplemental Table 1) were also made. These behaved identically to rna15-t6 in terms of cell viability and interactions with RNA, Pcf11, Rna14, and Sub1 (data not shown). The plasmids pET21b/RNA14 (37) and pRSSETB-SUB1 (36) were used for in vitro translation of Rna14 and Sub1. The Rna15/His6-Rna14 complex was expressed using the pET-Duet-1 plasmid (Novagen), a gift from I. Taylor (45). To express Pcf11, the PCF11 coding sequence was cloned into the pMAL vector (New England Biolabs) to give pMAL-Pcf11. For purification of Sub1 from yeast, the SUB1 open reading frame was cloned into the pYES2CT vector (Invitrogen), which introduced the V5/His6 tag onto the C terminus and allowed galactose-induced expression from a high copy plasmid.

Protein Expression and Purification—His6-tagged recombinant Rna15 and its truncations were expressed in E. coli BL21(DE3). A 1-liter culture was grown to an OD of 0.6, followed by induction with 1 mM isopropyl-β-D-thiogalactoside for 3–4 h at 37 °C before harvesting. Protein purification was done as described (39) with the following modifications. Cells were disrupted in binding buffer containing 20 mM Tris (pH 7), 500 mM KCl, 5 mM β-mercaptoethanol, 10% glycerol, 0.1% Nonidet P-40, 1 mM phenylmethylsulfonyl fluoride (PMSF), 2 mM pepstatin, and 0.6 mM leupeptin. The cell lysate was allowed to bind...
to 1 ml of Talon resin (Clontech) or nickel-nitrilotriacetic acid-agarose (Qiagen) for 1 h at 4 °C with gentle agitation. Washings were performed for 10 min each for two times in wash buffer 1 (binding buffer plus 5 mM imidazole (pH 7.4)), 10 min each for two times in wash buffer 2 (wash buffer 1 lacking 0.1% Nonidet P-40), and 10 min each for two times in wash buffer 3 (same as wash buffer 2 but with 125 mM KCl). Proteins were eluted with 100 mM imidazole in wash buffer 3. The eluate was dialyzed against buffer containing 20 mM Tris (pH 8.0), 50 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 20% glycerol, 1 mM PMSE, 2 mM pepstatin A, and 0.6 mM leupeptin with two buffer changes for 1 h each (3 h total). Pcf11 was expressed as an MBP fusion from pMAL-PCF11 in *E. coli* and purified on amylose resin following the protocol provided by New England Biolabs. The His6-Rna14/Rna15 heterodimer was described by Noble et al. (45) and purified by nickel affinity chromatography. Sub1-V5/His6 was purified from the yeast strain INVSc1 (Invitrogen) transformed with pYES-Sub1, in which expression is induced by growth in selective medium containing 2% galactose. Cells were lysed with the yeastBuster protein extraction reagent (Novagen) according to the manufacturer’s protocol, and Sub1-V5/His6 was purified by nickel-affinity purification (Qiagen) using a wash buffer containing 500 mM NaCl. 

35S-Labeled Rna14 and Sub1 proteins were made by *in vitro* coupled transcription and translation (TNT) reactions using the rabbit reticulocyte lysate system (Promega) in the presence of [35S]methionine (Amersham Biosciences) as described by the manufacturer. One microgram of plasmid was used in each TNT reaction.

For structural investigations, different CstF-64, Rna15, and PC4 protein constructs were subcloned into pET-28 (Novagen). Proteins were overexpressed in BL21 (DE3) cells or Rosetta (DE3) pLysS strains (Novagen) by induction with 1 mM isopropylthiogalactoside at a cell density of 0.5 OD. Expression was induced by growth in selective medium containing 2% galactose. Cells were lysed with the yeastBuster protein extraction reagent (Novagen) according to the manufacturer’s protocol, and Sub1-V5/His6 was purified by nickel-affinity purification (Qiagen) using a wash buffer containing 500 mM NaCl. 35S-Labeled Rna14 and Sub1 proteins were made by *in vitro* coupled transcription and translation (TNT) reactions using the rabbit reticulocyte lysate system (Promega) in the presence of [35S]methionine (Amersham Biosciences) as described by the manufacturer. One microgram of plasmid was used in each TNT reaction.

For structural investigations, different CstF-64, Rna15, and PC4 protein constructs were subcloned into pET-28 (Novagen). Proteins were overexpressed in BL21 (DE3) cells or Rosetta (DE3) pLysS strains (Novagen) by induction with 1 mM isopropylthiogalactoside at a cell density of 0.5 OD. Expression was induced by growth in selective medium containing 2% galactose. Cells were lysed with the yeastBuster protein extraction reagent (Novagen) according to the manufacturer’s protocol, and Sub1-V5/His6 was purified by nickel-affinity purification (Qiagen) using a wash buffer containing 500 mM NaCl. 35S-Labeled Rna14 and Sub1 proteins were made by *in vitro* coupled transcription and translation (TNT) reactions using the rabbit reticulocyte lysate system (Promega) in the presence of [35S]methionine (Amersham Biosciences) as described by the manufacturer.

For structural investigations, different CstF-64, Rna15, and PC4 protein constructs were subcloned into pET-28 (Novagen). Proteins were overexpressed in BL21 (DE3) cells or Rosetta (DE3) pLysS strains (Novagen) by induction with 1 mM isopropylthiogalactoside at a cell density of 0.5 OD. Expression was induced by growth in selective medium containing 2% galactose. Cells were lysed with the yeastBuster protein extraction reagent (Novagen) according to the manufacturer’s protocol, and Sub1-V5/His6 was purified by nickel-affinity purification (Qiagen) using a wash buffer containing 500 mM NaCl. 35S-Labeled Rna14 and Sub1 proteins were made by *in vitro* coupled transcription and translation (TNT) reactions using the rabbit reticulocyte lysate system (Promega) in the presence of [35S]methionine (Amersham Biosciences) as described by the manufacturer. One microgram of plasmid was used in each TNT reaction.

For structural investigations, different CstF-64, Rna15, and PC4 protein constructs were subcloned into pET-28 (Novagen). Proteins were overexpressed in BL21 (DE3) cells or Rosetta (DE3) pLysS strains (Novagen) by induction with 1 mM isopropylthiogalactoside at a cell density of 0.5 OD. Expression was induced by growth in selective medium containing 2% galactose. Cells were lysed with the yeastBuster protein extraction reagent (Novagen) according to the manufacturer’s protocol, and Sub1-V5/His6 was purified by nickel-affinity purification (Qiagen) using a wash buffer containing 500 mM NaCl. 35S-Labeled Rna14 and Sub1 proteins were made by *in vitro* coupled transcription and translation (TNT) reactions using the rabbit reticulocyte lysate system (Promega) in the presence of [35S]methionine (Amersham Biosciences) as described by the manufacturer. One microgram of plasmid was used in each TNT reaction.

For structural investigations, different CstF-64, Rna15, and PC4 protein constructs were subcloned into pET-28 (Novagen). Proteins were overexpressed in BL21 (DE3) cells or Rosetta (DE3) pLysS strains (Novagen) by induction with 1 mM isopropylthiogalactoside at a cell density of 0.5 OD. Expression was induced by growth in selective medium containing 2% galactose. Cells were lysed with the yeastBuster protein extraction reagent (Novagen) according to the manufacturer’s protocol, and Sub1-V5/His6 was purified by nickel-affinity purification (Qiagen) using a wash buffer containing 500 mM NaCl. 35S-Labeled Rna14 and Sub1 proteins were made by *in vitro* coupled transcription and translation (TNT) reactions using the rabbit reticulocyte lysate system (Promega) in the presence of [35S]methionine (Amersham Biosciences) as described by the manufacturer.
Function of the Rna15 C Terminus in mRNA 3'-End Processing

A

H.s CstF-64

S.c Rna15

Gln261/269Ala

t10

t9

t8

t7

t6

JUFO
PSIPRED
Jpred

ΔCα
tions were prepared on ice and incubated first at 4 °C for 10 min and then at 30 °C for 20 min. Reactions were stopped by the addition of protease inhibitor K and SDS, brought to a volume of 30 μl with Tris-EDTA (pH 7.5), and extracted once with phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v). One-tenth of the reaction was resolved on a 5% acrylamide, 8.3 M urea gel and then visualized by using a PhosphorImager. For ChIP analysis, in vivo cross-linking with formaldehyde, preparations of chromatin and immunoprecipitations were performed as described previously (4) using 7 μl of α-Rpb3 antibodies (Neoclonal Biotechnology) preincubated with protein A/G PLUS-agarose beads (sc-2003; Santa Cruz Biotechnology).

NMR Spectroscopy and Structure Determination—NMR studies of CstF-64 used 1 mM protein in 25 mM potassium phosphate (pH 6.0), unless otherwise stated. Much lower concentrations (0.3 mM) were used for Rna15 because of poor protein expression. Spectra were recorded at 600 or 500 MHz on Bruker DMX or Avance spectrometers. 13N-Labeled samples of CstF-64-(504–574) and CstF-64-(531–574) were used to record homonuclear (NOESY and TOCSY spectra) and 15N-edited heteronuclear data (15N-HSQC and three-dimensional 15N-NOESY-HSQC). Mixing times for the NOESY experiments used for extraction of structural constraints were 100 ms. A doubly labeled sample (15N, 13C) of the shorter construct (531–574) was used to record triple (HNCA, HN(CO)CA, and CBCA(CO)NH) and double resonance experiments (HCCH-COSY and 13C-NOESY-HSQC). NMR data were processed with nmrPipe (48) and analyzed with ANSIG (49) or Sparky (50). NOE-derived distance constraints were extracted from the two- and three-dimensional NOESY (15N and 13C edited) spectra and automatically calibrated with the program CYANA (51, 52). Torsional angles constraints for the backbone angles \( \psi \) and \( \varphi \) were obtained from the statistical analysis of 13C chemical shifts as implemented in TALOS (53). Structure calculations were executed by torsion dynamics in CYANA (version 1.06). All structures having no distance restraint violation larger than 0.2 Å or no violation of dihedral angle constraints larger than 5° were accepted (90% of all calculated structures). Rna15 and Pti1 C-terminal domains were modeled from the atomic structure of the homologous CstF64 domain using the in silico mutagenesis tools of the program MOLMOL (54).

Interaction between PC4 and CstF-64—NMR was also used to study the interaction of the C-terminal domain of human CstF-64 with various forms of unphosphorylated and phosphorylated human PC4. HSQC spectra were typically recorded at 300 K, but in several cases data were duplicated at 290 K as well. Titration experiments were conducted using the sample combinations shown in supplemental Table II at about 0.3 mM 15N-labeled protein by recording an HSQC of the labeled free protein. A second HSQC was then recorded under identical conditions after adding the unlabeled binding partner in slight excess. Buffers were 50 mM sodium acetate, 300 mM NaCl, 4 mM DTT (pH 5), except in one case when the following buffer was used to exactly duplicate the conditions reported by Calvo and Manley (36): 20 mM HEPES-Na (pH 7.9), 100 mM NaCl, 10% glycrol, 0.5 mM DTT, 0.2 mM EDTA, 0.2 mM PMSF.

RESULTS

The C-terminal Region of Rna15/CstF-64 Is Highly Conserved—The goal of this study was to determine the function of the domains of Rna15/CstF-64 that lie beyond the N-terminal RM responsible for recognition of the polyadenylation signal. The N- and C-terminal sequences of Rna15/CstF-64 orthologues (corresponding to residues 1–110 and 529–577 of hCstF-64) constitute the most highly conserved regions of the protein (Fig. 1B). Only a few variations in the C-terminal region are observed in evolutionarily distant vertebrates (fish, birds, or amphibian), and homology remains remarkably high even for more distant metazoans like insects (Drosophila, 63% identity), worms, and other invertebrates. Indeed seven positions appear totally conserved (Fig. 1B) as follows: Gln543, Leu545, Leu547, Gln551, Ile552, Leu555, and Pro556 (numbering of hCstF-64). This domain is also highly conserved in S. cerevisiae (Rna15, 31% identity and Ptl1, 23% identity) as well as in S. pombe (Ctf1, 31% identity) and in the green plants, although their sequences seem to be more evolutionarily distant from the mammalian proteins. A tissue-specific variant of CstF-64 found in mammals (55) contains variable length insertions just before, but not within, the C-terminal domain (Fig. 1B). Altogether, these observations point to the existence of an independent domain within the last 50 residues of CstF-64/Rna15.

The C-terminal Domain of Rna15/CstF-64 Forms a Novel All-helical Structure—In pursuing structural investigations of the Rna15/CstF-64 C-terminal domain, we focused on the human protein because expression levels were considerably higher compared with Rna15. Poor expression of this region of Rna15 has made it very difficult to obtain sufficient amounts to determine its structure by NMR. Although the sequence analysis shows that only the last 40 or 50 residues of CstF-64/Rna15 are conserved (Fig. 1B), we started our structural investigation with a slightly larger region (504–577 of the human sequence) based on biochemical studies of the interaction with PC4.

FIGURE 1. Sequence conservation of the C terminus of CstF-64/Rna15. A, domain organization of human CstF-64 and Rna15. The domain structure of hCstF-64 is shown, with the RM, the hinge domain responsible for interaction with CstF-77 and Symplekin (32, 34), and two progly-rich domains flanking 12 repeats of the MEAR(A/G) sequence specific to most metazoan CstF-64 proteins (8). For Rna15, the amino acids demarking the RM, the opa-like sequence containing glutamine-rich repeats (43), and the C-terminal domain are indicated. The end points of deletions in Rna15 created in this study are specified by the bars, and the locations of point mutations in the C-terminal domain (Q261A/Q269A) are denoted by asterisks. Growth of strains containing these mutations after forcing loss of an RNA15 covering plasmid on 5’-fluoroorotic acid-containing media is shown on the right. B, sequence alignment of the C-terminal regions of CstF-64/Rna15 homologues from the following: H.s. Homo sapiens, B.t. Bos taurus, M.m. Mus musculus, X.l. Xenopus laevis, B.r. Brachydanio rerio, A.g. Anopheles gambiense, D.m. Drosophila melanogaster, S.c. S. cerevisiae, and S.p. S. pombe. Sequence numbering corresponds to the human protein. For insertions found in the H. sapiens, M. musculus, and B. taurus brain- and testis-specific variant of CstF-64, the sequences are not shown in the alignment, and the number of residues in the insertion is given instead. Residues conserved in 12 of the 13 homologues are shaded in red; residues conserved in at least 7 sequences are in blue; and the less well conserved positions are in yellow. Asterisks indicate the seven invariant residues. Secondary structure prediction of hCstF-64 amino acids 531–577 provided by the algorithms JUFO, PSIPRED, and Jpred are shown above the sequence (in black), together with the experimentally determined secondary structure. Conformational Cα chemical shifts (\( \Delta \delta^C = \alpha + \beta \alpha + \gamma \)) observed for the construct CstF-64-(531–577) are plotted below the sequence; high positive values indicate helical conformation.
When a similar fragment was removed from CstF-64, the interaction with PC4 was abolished (36). The $^{15}$N-HSQC spectrum of this construct shows that approximately half of the signals have sharp line widths, negative $^1$H-$^{15}$N heteronuclear NOE (data not shown), and very poor dispersion (Fig. 2A). The remaining signals show instead broader line widths, positive NOEs, and much better dispersion. These observations suggest the presence of a structured region next to an unstructured N-terminal fragment. Because the characteristic Gly signals largely belong to the unstructured part of the domain, it is clear that the 504–529 fragment of the construct, containing nearly all of the Gly residues, is the unfolded region. These domain boundaries agree well with the pattern of conservation, which decays sharply before position 530, and with the secondary structure prediction (Fig. 1B). Signals in the $^{15}$N-HSQC spectrum of a shorter construct (531–577) overlap with the well dispersed signals in the longer construct (Fig. 2A), confirming that the last 46 residues form a structurally independent domain and that truncation of the segment 504–529 does not affect the structure of the C-terminal domain. These results

**FIGURE 2. Structure of the C-terminal domain of CstF-64/Rna15.** A, superposition of the $^1$H-$^{15}$N HSQC NMR spectra of the CstF64-(531–577) and -(504–577) constructs (red and black, respectively) indicating that the smaller construct retains the same structure in the context of the larger fragment. B, superimposition of the backbone atoms of the 20 best conformers representing the structure of CstF-64-(530–577). C, schematic ribbon representation of the structure of hCstF-64 C-terminal domain: helix 1 is in red, helix 2 in blue, and helix 3 is in green. D, a close up view of the residues involved in the contacts between helix 1 and 3. The helix backbones are indicated in the same colors as used in B and C.
prompted us to proceed with the structural studies on the construct, including only residues 531–577. The lack of aromatic residues (Tyr, Phe, or Trp) compromises the dispersion of the NMR spectra even for this shorter construct. Thus, $^{13}$C- and $^{15}$N-edited NOESY spectra had to be used in order to assign NOE data unambiguously and generate a substantial number of medium and long range NOE constraints (Table 1).

The C-terminal domain of hCstF-64 forms a 3$\alpha$-helix orthogonal bundle (Fig. 2, B–D) with two longer helices (almost four turns in length for helices 1 and 3) connected by a shorter helix 2 (two turns). The fold of the domain is maintained by a well defined network of interactions. Helices 1 and 3 contact each other through a hydrophobic interaction between residues Leu$^{564}$ (in helix 3) and Lys$^{537}$ (in helix 1), reinforced by contacts between His$^{535}$ and Gln$^{561}$ on the opposite side of helix 1 (Fig. 2D). Although residues involved in hydrophobic interaction are totally conserved throughout all CstF-64 homologues, the hydrophilic contacts are more variable. Together, these interactions maintain the angle between the axes of the two helices at the unusual value of 50° ± 5° (Fig. 2C). Helix 2 lies on top of the forked structure defined by helices 1 and 3, and it forms a small hydrophobic cluster with them that further stabilizes the structure of the domain. Because of the small size of the protein, part of the hydrophobic core is solvent-exposed, and only a handful of residues display solvent accessibilities below 10% (mean values over the 20 structure ensemble): Ala$^{538}$ (1.6%), Ile$^{541}$ (10.8%), Met$^{542}$ (7.1%), Ile$^{552}$ (2.9%), Ser$^{562}$ (7.9%), and Ile$^{563}$ (5.1%). Ile$^{552}$ is one of the least accessible and also one of the seven residues absolutely conserved among different homologues. Positions 541 and 563 conserve their hydrophobic character, whereas the long yet linear side chain of Met$^{542}$ is replaced by Lys in Rna15 and Pti1. The introduction of a positive charge might be thought of as a destabilizing mutation, but a more detailed examination of the environment of Met$^{542}$ reveals that its methyl group is solvent-exposed and therefore can be substituted with Lys without destabilizing the protein fold.

Analysis of the sequence homology data in light of this structure strongly suggests that all CstF-64 C-terminal domain homologues will have the same structure. For example, the key contacts between helices 1 and 3 (Lys$^{537}$–Leu$^{564}$ in CstF-64) are maintained in Rna15 (Lys–Trp) and Ctf1 (Arg–Leu), but they are changed to Gln–Glu in Pti1. Although these two side chains retain the ability to interact across the helix 1-helix 3 interface (i.e. by hydrogen bond), this and other differences with CstF64, Rna15, and Ctf1 suggest that Pti1 is a more distant homologue of these proteins.

The structure of the CstF-64 C-terminal domain allowed us to precisely define domain boundaries that agree well with sequence and prediction data (Fig. 1B). By using this information to design a new expression construct, we obtained sufficient amounts of the unlabeled yeast protein to determine a NOESY spectrum of the C-terminal domain of Rna15. This spectrum contains the characteristic signature expected for a helical protein (supplemental Fig. 1) as follows: amide to aliphatic NOEs consistent with a folded domain and the characteristic N-N NOEs typical of a helical protein. The low concentration of the sample because of poor protein expression resulted in a suboptimal spectrum. Nevertheless, these results, coupled with the sequence homology, unambiguously support the structural homology between the C-terminal domains of CstF-64 and Rna15.

Remarkably, out of seven absolutely conserved residues within the C-terminal domain of CstF-64, only one has an indisputable structural role (Ile$^{552}$), implying that the remaining six will have important functional roles. A revealing picture emerges when these conserved residues are mapped onto the three-dimensional structure of the domain (Fig. 3A). Five conserved hydrophobic residues (Val$^{544}$, Leu$^{545}$, Leu$^{555}$, and Pro$^{566}$) form a continuous solvent-exposed patch flanked by two Gln residues (543 and 551), which are also conserved. A comparison of solvent accessibilities across conserved hydrophobic residues provides some insight into the possible role of these residues. As mentioned before, Ile$^{552}$ is deeply buried in the structure (solvent accessibility ~2.9%), suggesting an architectural role for this residue. The other conserved hydrophobic residues are much more exposed; for example, Leu$^{547}$ has on average 62% of its side chain exposed. The conservation of this hydrophobic patch strongly suggests that it might form a protein interaction interface. The conservation of the two Gln residues is also unrelated to any structural stabilization role, suggesting that they function in providing interaction specificity as well. Finally, it is also interesting that the only absolutely conserved positive residue, Lys$^{567}$ (see above), is placed near this putative interaction surface.

**The C Terminus of Rna15 Is Critical for Cell Growth**—The evolutionary conservation of the key residues discussed above maintains an architecture that must be important for the functional role of this domain. To probe the function of this conserved C-terminal domain, Rna15 was tagged with His$_{6}$ and sequential deletions of 17–19 amino acids were made from the C terminus (Fig. 1A). These truncations were expressed in a strain carrying a wild-type RNA15 gene on a URA3-containing plasmid and a deletion of the chromosomal copy of RNA15. These strains all grew as well as wild-type cells, indicating that none of the shorter forms of Rna15 exerted a dominant negative effect on cell fitness. All of the mutant forms were stably expressed at levels comparable with the wild-type protein (Fig. 4A).

The truncated Rna15 proteins were then tested for their ability to provide Rna15 function in vivo, using plasmid shuffling to remove the covering plasmid containing the RNA15 gene. The
first two deletions (rna15-t10 and rna15-t9), which remove helix 3 or helices 2 and 3 and destroy the C-terminal domain structure, could not support cell growth. Surprisingly, removal of the entire C-terminal domain (rna15-t8) resulted in a viable rna15 mutant, albeit one that grew slowly at 24 °C and exhibited thermosensitive lethality at 37 °C (Fig. 4B). The larger deletions found in rna15-t7 and rna15-t6 resulted in mutants that could not grow at either temperature.

The C-terminal Domain of Rna15 Is Required for mRNA 3'-End Processing but Not for Transcription Termination—To decipher how the rna15-t8 mutant exerts its phenotype, we prepared extracts from wild-type and rna15-t8 cells grown at the permissive temperature and tested them for the ability to process RNA substrates in vitro. When these extracts are incubated with ATP and a precursor containing the GAL7 poly(A) site and flanking sequences, the wild-type extract efficiently cleaves and polyadenylates the RNA, but processing is greatly reduced with rna15-t8 extract (Fig. 4C, lanes 1 and 2). There is also no accumulation of cleaved, unadenylated RNA, indicating that the cleavage step is impaired. GAL7 RNA ending at the poly(A) site was used to examine the poly(A) addition step in the absence of cleavage. Only a small fraction of this substrate receives a poly(A) tail in the mutant extract in comparison to wild type (Fig. 4C, lanes 5 and 6).

Defects in the mRNA 3′-end processing machinery can cause poor transcription termination downstream of a poly(A) site (1, 2). For example, the rna15-1 mutant that is defective for in vitro cleavage and polyadenylation (7) also shows defective termination in transcription read-through assays (17). Chromatin immunoprecipitation (ChIP) has been used recently as an assay for termination defects, using as a read-out the persistence of RNAP II in regions of a gene located beyond the poly(A) site (4). We used ChIP to determine whether the rna15-t8 allele was impaired for transcription termination. In this experiment, the
A functional consequence of C-terminal truncations of Rna15. A, immunodetection of wild-type or mutated Rna15-His6 proteins in extract by SDS-PAGE, transferred to a solid support, and probed with antibodies against Rna15 as described under "Experimental Procedures." The lane marked Input contains unreacted RNA substrate. The positions of unreacted substrate and input control. The right panel shows the input control.

B, extracts were examined for coupled cleavage/polyadenylation activity using reactions containing radioactive poly(U) (Fig. 5A) and were cross-linked to RNA by ultraviolet light and were cross-linked to RNA by ultraviolet light. The positions of unreacted substrate and input control. The right panel shows the input control.

C, the same Rna15 proteins were examined by pulldown assays for interaction with RNA and with Sub1 and Rna14, two proteins known to make direct physical contact with Rna15 (36, 39). The set of truncated Rna15 constructs was expressed as His6-tagged proteins in E. coli, and the purified proteins (Fig. 5A) were examined for RNA interaction by a poly(U)-Sepharose binding assay, which was used previously to analyze the Rna15 mutant with a defective RRM (10). As expected, all truncated constructs were wild-type for binding to poly(U) (Fig. 5B) and were cross-linked to RNA by ultraviolet light with the same efficiency (data not shown). These results are consistent with the RRM being both necessary and sufficient for RNA recognition.

The same Rna15 proteins were then used in pulldown assays using nickel-affinity beads and in vitro translated Sub1 and Rna14 (Fig. 5, C and D). In agreement with previous studies (36, 39), Sub1 and Rna14 interact with wild-type Rna15. The shortest form of Rna15 that binds to Sub1 and Rna14 is rna15-t7; bind-

but drops markedly at a position 300 bases further downstream (PCR product 3). A decline similar to what is seen with wild-type cells is observed for the rna15-t8 mutant, indicating that termination in this mutant at the nonpermissive temperature is as efficient as wild-type cells. In contrast, the termination-defective pcf11-9 mutant (44) shows an increase in signal from the PCR product 3, in agreement with other ChIP analyses using this mutant (5). From these experiments we conclude that rna15-t8 is defective for mRNA 3′-end processing but normal for poly(A)-site dependent termination.

The Interaction of Rna15 with Sub1 and Rna14 Does Not Require the Highly Conserved C-terminal Domain—The results of the previous sections indicate that the C-terminal part of Rna15 provides functions important for cell growth through an effect on mRNA 3′-end processing. To gain insight into this putative function, the mutant Rna15 proteins were examined by pulldown assays for interaction with RNA and with Sub1 and Rna14, two proteins known to make direct physical contact with Rna15 (36, 39). The set of truncated Rna15 constructs was expressed as His6-tagged proteins in E. coli, and the purified proteins (Fig. 5A) were examined for RNA interaction by a poly(U)-Sepharose binding assay, which was used previously to analyze the Rna15 mutant with a defective RRM (10). As expected, all truncations were wild-type for binding to poly(U) (Fig. 5B) and were cross-linked to RNA by ultraviolet light with the same efficiency (data not shown). These results are consistent with the RRM being both necessary and sufficient for RNA recognition.

The same Rna15 proteins were then used in pulldown assays using nickel-affinity beads and in vitro translated Sub1 and Rna14 (Fig. 5, C and D). In agreement with previous studies (36, 39), Sub1 and Rna14 interact with wild-type Rna15. The shortest form of Rna15 that binds to Sub1 and Rna14 is rna15-t7; bind-

M. Kim, L. Vasiljeva, O. J. Rando, A. Zhelkovsky, C. Moore, and S. Buratowski, submitted for publication.
Function of the Rna15 C Terminus in mRNA 3'-End Processing

A. Rna15-His$_6$ input

B. Poly(U) interaction

C. Sub1 interaction

D. Rna14 interaction

E. Rna15, Rna15/Rna14

FIGURE 5. Interaction of the C-terminal third of Rna15 with Rna14 and Sub1. A, recombinant wild-type (WT) and truncated Rna15 proteins used for pulldowns. B, interaction of Rna15 with RNA. The proteins shown in A were assayed for the ability to bind to poly(U)-Sepharose beads in the presence of 100 mM KCl and then detected by silver staining. Bovine serum albumin (BSA) was used as a control and did not bind the resin (2nd lane from left). C, pulldowns with Sub1. Assays were performed by incubating $^{35}$S-labeled in vitro translated Sub1 with indicated forms of Rna15 bound to Talon beads, and interacting proteins detected using a PhosphorImager. Five percent of the input is shown in the leftmost lane, followed by a control in which the His$_6$ peptide was bound to the beads. D, pulldowns with Rna14 using the same method described in C, E, interaction of Sub1 with the Rna15-Rna14 complex. Co-immunoprecipitation was performed by incubating Rna15-His$_6$ (lanes 2 and 3) or Rna15/His$_6$-Rna14 (lanes 3 and 6) with antibody against the V5 epitope bound to protein A-agarose in the presence (lanes 2 and 3) or absence (lanes 2 and 3) of Sub1-V5/His$_6$. The immunoblot was probed with antibody raised against Rna15-His$_6$, which also detected Sub1 and Rna14 because of anti-His$_6$ antibodies in this polyclonal serum. The input (5%) is shown in lanes 1 and 4, C, control; P, pulldown; I, induced.

shift changes indicates very strongly that these two protein domains do not physically interact and that the CstF-64/PC4 interaction previously observed in vivo or in cell extracts (36) is mediated by a different domain of CstF-64 or indirectly by other proteins.

Pcf11 Interacts with the Conserved C Terminus of Rna15—In addition to Rna14 and Sub1, the Pcf11 subunit of CF IA also makes direct contact with Rna15 (39). To map the region of Rna15 required for the interaction with Pcf11, we used recombinant MBP-tagged Pcf11 bound to amylase beads as bait for the Rna15 variants (Fig. 6A). The interaction with Pcf11 was lost with the minimal truncation, rna15-t10. This truncation removes the last 16 amino acids spanning helix 3 and destroys the C-terminal domain structure. Therefore, to directly investigate the putative interaction surface identified in the structural analysis (Fig. 3A), we changed the two strictly conserved glutamines in Rna15 at positions 261 and 269 to alanine to give rna15-CTD1. During the construction of this mutant, we also isolated a second mutant (rna15-CTD2) in which three amino acids between these glutamines (in the loop region between helix 1 and helix 2) had been replaced with a single histidine (Fig. 6B). Because these residues are surface-exposed, neither of these mutations is expected to affect the global fold of the C-terminal domain. Both mutants were defective for binding to in the millimolar range) that could become significant when proteins are assembled into physiologically relevant complexes. We probed PC4 with two different C-terminal fragments of human CstF-64, one containing the last 63 amino acids and a second containing the most highly conserved 40 amino acids at the very C terminus (the fragment used for structure determination).

The analysis of the $^{15}$N-HSQC NMR spectrum of full-length PC4 was consistent with a structured domain at the C terminus (57, 58), whereas the resonances corresponding to the N-terminal fragment indicated an unfolded polypeptide. We then mixed $^{15}$N-labeled PC4 with unlabeled CstF-64, and vice versa, in several possible combinations (supplemental Table II), but the data revealed no significant spectral changes (data not shown). Reasoning that the DNA-binding domain may perhaps inhibit CstF-64 binding, we mixed the N-terminal fragment of PC4 with human CstF-64, without observing any changes at all. Finally, we used casein kinase II to phosphorylate PC4 in vitro, but we again failed to detect any interaction with CstF-64. The absence of chemical
Function of the Rna15 C Terminus in mRNA 3′-End Processing

MBP-Pcf11 (Fig. 6B). The more severe mutation, rna15-CTD2, caused slow growth at 25 and 37 °C compared with wild-type cells, and thermosensitive growth at 38 °C (Fig. 6C). The mutant with only the E261A/E269A replacement showed wild-type growth at 25 °C but was temperature-sensitive when grown in the presence of 1.5% formamide, a reagent known to enhance defects in protein/protein interactions in yeast mutants (59). These results confirm the importance of the conserved residues in this interface for Pcf11 interaction and Rna15 function.

DISCUSSION

Rna15 and its homologues, such as the metazoan CstF-64 proteins, are essential for the synthesis of the polyadenylated 3′-ends of eukaryotic mRNA. This family of proteins uses a conserved N-terminal RRM to help direct the 3′-end processing complex to the correct site on the mRNA precursor, and the strength of this interaction affects the efficiency of poly(A) site usage (10, 11, 60 and references therein). Rna15 and CstF-64 perform this function in the context of larger protein complexes, designated CF Iα in yeast and CstF in metazoan cells (61, 62). In this study, we investigate the structure and function of the highly conserved C-terminal half of CstF-64 and Rna15. By NMR, we find that this region folds into a novel structure that interacts with the Pcf11 subunit of CF Iα but not with the Rna14 subunit or the transcription modulator PC4/Sub1. We show that the C-terminal domain is critical for the function of CF Iα in 3′-end processing but dispensable for its role in the termination of RNAP II transcription downstream of the poly(A) site. We further identify two functionally important regions that lie just outside of the highly conserved C-terminal helical structure, one of which is necessary for the interaction of Rna15 with Rna14 and Sub1.

Structure and Function of the C-terminal Domain of Rna15/CstF-64—As shown by the sequence alignment in Fig. 1A, the last 50 or so amino acids represent the most highly conserved stretch in this family of proteins (36). This C-terminal domain in CstF-64 (Fig. 2) forms a three-helix bundle, with an uncommon arrangement of the helices. The conservation of key hydrophobic residues indicates that the same structure will almost certainly form in all its homologues, and this prediction is confirmed by our structural analysis of Rna15 (supplemental Fig. 1) as well as our functional studies of Rna15 truncations. Removal of helix 3 (rna15-t10) or helices 2 and 3 (rna15-t9) destroys the domain structure and causes lethality by inactivating the protein, although complete removal of the domain (rna15-t8) gives a thermosensitive growth phenotype and defective RNA processing in vitro. The inviability of the partial truncations might be due to the remnants of the C-terminal domain interfering with other protein/protein interactions.

Destruction of the C-terminal domain of Rna15 correlates with loss of Pcf11 interaction. Pcf11 is important in the structural organization of the mRNA 3′-end processing complex because of its contacts with RNAP II (63), with the Rna14, Rna15, and Clp1 subunits of CF Iα (39), and with several com-
Function of the Rna15 C Terminus in mRNA 3′-End Processing

ponents of CPF (64–68), a multisubunit factor that works in conjunction with CF IA in the maturation of 3′-ends (62). Mutations in Pcf11 are known to inhibit 3′-end processing (44, 68). Our structural analysis of the CstF-64 C-terminal domain revealed a cluster of surface-exposed residues (Fig. 3A) that we propose to be an evolutionarily conserved protein/protein interaction surface. The defects of the rna15-CTD1 and rna15-CTD2 mutants, which specifically target this hypothetical interface, confirm the importance of this region of the protein for Pcf11 recognition and Rna15 function.

Pcf11 interacts directly with the RNAP II CTD through a conserved region of Pcf11 that maps to amino acids 1–266 (69, 70). A mutation in Pcf11 that disrupts this interaction prevents transcription termination but not 3′-end processing (44), and a fragment of Pcf11 containing its CTD-binding domain has been shown recently to dismantle the RNAP II elongating complex (41, 42). A different region of Pcf11 (residues 288–400) is required for its interaction with Rna15 (44, 68). The studies cited above have led to the conclusion that the function of Pcf11 in transcription termination is independent of its function in cleavage and polyadenylation. Our finding that loss of an Rna15/Pcf11 interaction severely cripples 3′-end processing but has no effect on termination adds additional support to this model. Cleavage-defective, termination-normal phenotypes have been found with defects in other mRNA 3′-end processing factors, such as when Ssu72 is depleted (66) or for the brr5-1, pta1-2, pcf11-2, or ssa72-2 alleles (44, 71, 72). Thus, although cleavage at the poly(A) site by the 3′-end processing machinery is thought to facilitate RNAP II termination by providing an entry site for the Rat1 5′-3′-exonuclease (73), this may be one of several convergent events leading to the release of RNAP II from the DNA template (5, 74). Our study suggests that a 3′-end complex lacking the Rna15 C-terminal domain makes contacts with the elongating RNAP II that are sufficient to signal transcription termination, even if the nascent transcript is not cleaved.

The CstF-64 C-terminal Domain Folds in Other Proteins—The unusual 50° angle between helices 1 and 3 of the CstF-64 C-terminal domain observed here has not been reported previously. However, application of the DALI algorithm (75) shows that this domain has some structural similarity with parts of two other proteins as follows: the mammalian Dia1 Formin protein involved in cytoskeletal remodeling (Protein Data Bank code 1Z2C-B; Z score = 3.1) (76) and the RNA silencing suppressor p21 protein from Beet Yellows virus (Protein Data Bank code 2CWO-A; Z score = 2.5) (77). These structurally related regions form well defined domains within these two proteins (supplemental Fig. 2) with the same helical topology and very similar interhelical angles to those observed in the C-terminal domain of CstF-64. However, the similarity is only at the structural level; only one residue (Leu947) is conserved in the sequence of these three domains. To our knowledge, there is no connection of these proteins to mRNA 3′-end processing that would suggest direct structure-function parallelism.

Interestingly, the structurally related domain of Dia1 mediates interaction with the RhoC GTPase (supplemental Fig. 2 and see Ref. 76). In p21, the homologous domain participates more marginally in a homodimer interface. Rather than being isolated, this subdomain makes contacts with other regions in the p21 monomer (77). The question of whether providing a protein interaction interface is a general characteristic of this type of structure deserves further study, but there is currently not enough structural data to make such a conclusion.

The Function of Regions of Rna15 Adjacent to the C-terminal Domain—We have shown that yeast lacking the C-terminal domain of Rna15 grow more slowly than wild-type cells and die at elevated temperatures, in agreement with the severe processing defect and corresponding to a loss of Pcf11 interaction. Our analysis has also identified two functionally important regions that are immediately adjacent to the C-terminal domain. Removal of 18 amino acids in addition to the C-terminal domain cannot be tolerated, but we do not know the function of this region. The mutant protein with this larger deletion is stable, and the interaction with a second CF IA subunit, Rna14, is not perturbed. The only components of the mRNA 3′-end processing complex known to contact Rna15 are Pcf11 and Rna14, but all possible pairings with Rna15 have not been tested, and a critical interaction with this region may yet be found. Alternatively, it may simply provide a linker between the C terminus and the region responsible for interactions with Sub1 and Rna14.

A further truncation of 17 amino acids reveals a region of Rna15 important for the interaction with Rna14 (Fig. 6D). This interaction is evolutionarily conserved, because CstF-64 of humans, fruit flies, and plants all have been shown to interact with CstF-77, the Rna14 homologue (32, 34, 78, 79). Interestingly, a 44-amino acid stretch of Rna15 (residues 165–208) can be aligned with part of the CstF-64 segment that binds to CstF-77 and known as the hinge domain (24, 34). These amino acids are just upstream of the region defined by our deletion analysis as required for recognition of Rna14, suggesting that the N-terminal boundary of the Rna14 interaction domain extends beyond amino acid 208 of Rna15.

The incorporation of Rna15 and CstF-64 into their respective complexes, CF IA and CstF, positively affects the way these proteins interact with the polyadenylation signal. For example, even though the RRM of CstF-64 on its own is somewhat specific for G/U-rich sequences, the CstF complex has greater affinity and specificity for these regulatory elements (9, 29, 80). A possible explanation for these altered properties of CstF-64 has come from the structure of its RRM. The RNA interface is occluded by a C-terminal helix in the free protein that is removed when RNA binds to the protein (28). Destabilization of this helix to uncover the RRM surface of CstF-64 may be facilitated by the interaction of CstF-77 with the nearby hinge domain. We have found that all of the C-terminal truncations of Rna15 bind RNA with similar efficiency, suggesting that Rna14 or Pcf11 binding is not needed to prevent masking of the RRM by other parts of Rna15. However, we have shown previously that Rna14 modulates Rna15 activity in a different way by forming a bridge between the Rna15 and another RNA-binding protein Hrp1, thus stabilizing the contact of Rna15 with the A-rich yeast signal (10). Formation of a complex containing only Rna14 and Rna15 increases the affinity of Rna15 for RNA (45), indicating that the interaction of just these two proteins can also have a positive effect.
Insight into the function of the Rna14/Sub1 interaction region on Rna15 is provided by the analysis of the previously characterized rna15-2 and rna15-1 alleles. The importance of Rna15 for mRNA 3′-end processing was first established with these mutants (7). The rna15-1 mutant also exhibits read-through transcription beyond the poly(A) site (17). By sequencing, we have found that both mutants contain a single amino acid replacement of L214P (Fig. 6D), which lies in the middle of a region that we have found to be critical for interactions with Sub1 and Rna14 (residues 208–225). Thus, the processing defect caused by the L214P mutation may result from a destabilization of the Rna14 interaction. Interestingly, with the program BEHAIRPRED (81), this segment is predicted to form a 2:2 β-hairpin with a type I′ β-turn, in which Leu214 would form part of the turn sequence that is crucial for the formation and stability of β-hairpin structures (82).

In an earlier study (36), an Rna15 fragment containing only amino acids 191–290 was shown to be sufficient for Sub1 binding, and our studies refine the boundaries of this interaction (Fig. 6D). Our NMR analysis indicates that PCA/Sub1 does not interact directly with the C-terminal domain of CstF-64 even at millimolar concentrations, and our biochemical studies position the Sub1-interacting region of Rna15 N-terminal to the conserved domain. Sub1 can affect transcription initiation through its interaction with TFIIB (83–85) and 3′-end processing through its interaction with the Pta1 subunit of CPF (66).

Calvo and Manley (36) proposed that the termination defect of the rna15-1 mutant is a consequence of an enhanced affinity of rna15-1 for Sub1, and that a third function of Sub1 is to prevent premature termination. We have found that Sub1 interacts equally well with Rna15 in complex with Rna14 as it does with Rna15 alone, suggesting that Sub1 does not exert its anti-termination activity by interfering with the interaction of Rna15 with Rna14. We could not detect an endogenous Rna15/Sub1 interaction in pulldown assays using wild-type yeast extract (data not shown), and such an interaction has not been reported in large scale studies using mass spectrometry to analyze yeast protein complexes (86, 87), consistent with a Sub1/Rna15 interaction occurring transiently during transcription elongation. Taking all of these findings into consideration, we propose that the function of Rna15 in processing is mediated by its interactions with both Pcf11 and Rna14, but its activity in transcription termination requires only the interaction with Rna14 and can be modulated by Sub1 binding in close proximity.

Insights into the Two CstF-64 Homologues in S. cerevisiae—Pti1 is an S. cerevisiae protein that is related to Rna15 and co-purifies with CPF (26, 66, 71, 86, 88). Defects in Pti1 affect the 3′-cleavage site choice and abundance of some mRNAs (24), but the primary function of Pti1 may be to adapt the complex to work on other RNAP II transcripts (25). It has been suggested that this protein acts to uncouple cleavage and polyadenylation (25) in coordination with Nrd1-dependent 3′-end formation of nonpolyadenylated snoRNA and small nuclear RNA transcripts (56). Whereas Rna15 has also been implicated in snoRNA/small nuclear RNA termination (56), our analysis of Rna15 supports the idea that Pti1 and Rna15 use different molecular contacts to execute their respective functions in 3′-end formation. A fragment containing the hinge region of Pti1 interacts directly with Rna14 and Pta1 (24). The interaction of Pcf11 with Pti1 is also mediated by this middle region of the protein but not by its C-terminal domain as in Rna15. Moreover, Rna15 does not bind to Pta1 (72).

The difference in Pcf11 interaction might be explained by structural differences between the C-terminal domains of Rna15 and Pti1. The C-terminal domains of all metazoan CstF-64 proteins as well as that of Rna15 contain four conserved positively charged side chains (Lys537, Lys567, Lys572, and Arg560), of which only the position equivalent to Lys567 is conserved in Pti1p (Fig. 1B). Lys572 and Arg560 are placed on the opposite side to the putative Pcf11-binding interface (Fig. 3C) and hence should not have a leading role in recognition. A more likely explanation for the functional differences relies on the topological arrangement of the helices. In Cst-F64, the angle between helix 1 and 3 is determined by the contact between the hydrophobic side chains of Lys537 and Leu566. Equivalent residues can be found in other CstF-64 homologues, such as Rna15 and Cft1, but not in Pti1 (Fig. 1B). Changing the nature of this interhelical contact could have significant structural consequences for a small protein domain like this and potentially re-shape the Pcf11 binding interface, making it unable to recognize the Pcf11 protein. It should be also noted that the crucial Lys537 is near the conserved hydrophobic patch and to Lys567; thus it may also have a role in Pcf11 binding through charge interactions mediated by its side chain amine. This comparison indicates that different protein/protein interactions formed by closely related but slightly divergent factors may provide the selectivity needed to direct mRNA and snoRNA precursors to similar but distinct 3′-end formation pathways.

In summary, we have characterized the structure and function of regions of Rna15 that lie outside of the N-terminal RRM responsible for recognition of the polyadenylation signal. We have shown that these regions mediate interactions with the Rna14 and Pcf11 subunits of CF IA that are critical for the function of CF IA in mRNA 3′-end processing. Of particular importance, our analysis of the C-terminal domain of Rna15 and CstF-64 reveals a new all-helical structure that is highly conserved across eukaryotic species. We also find that this domain mediates the interaction of Rna15 with Pcf11, probably through a highly conserved, surface-exposed protein/protein interaction patch. Structural changes in the corresponding domain from Pti1 may explain why this very similar protein does not interact with Pcf11 and appears to have a function distinct from that of Rna15.

Curiously, the metazoan 3′-end processing machinery has only one Rna15 homologue, CstF-64, whose C-terminal domain sequence is closer to that of Rna15 than Pti1. The biochemical and structural work presented here on these two proteins indicates that the C-terminal domain of CstF64 will also provide a docking platform for Pcf11. Further research will be needed to determine how Pcf11 interacts with Rna15/CstF-64 and with other CF IA subunits and to dissect the implications of these interactions in the communication between transcription termination and 3′-end processing.
Function of the Rna15 C Terminus in mRNA 3’-End Processing

Acknowledgments—We are grateful to Andrew Bohm and Xiaoyuan He for reading the manuscript, to Stefan Gross for constructing the Rna15 deletion series in pET21b, and to Dana Harmon for constructing the rna15-CTD1 and rna15-CTD2 mutants. We also thank Priti Deka for help with the heteronuclear NOE experiments, Jean Bahn and Dr. Neil Dobson for help in protein expression and purification, and Alexander Murzin for analysis of the protein structure.

REFERENCES

1. Bentley, D. L. (2005) Curr. Opin. Cell Biol. 17, 251–256
2. Buratowski, S. (2005) Curr. Opin. Cell Biol. 17, 257–261
3. Saguez, C., Olesen, J. R., and Jensen, T. H. (2005) Curr. Opin. Cell Biol. 17, 287–292
4. Kim, M., Ahn, S. H., Krogan, N. J., Greenblatt, J. F., and Buratowski, S. (2004) EMBO J. 23, 354–364
5. Luo, W., Johnson, A. W., and Bentley, D. L. (2006) Genes Dev. 20, 954–965
6. West, S., Gromak, N., and Proudfoot, N. J. (2004) Nature 432, 522–525
7. Minvielle-Sebastia, L., Preker, P. J., and Keller, W. (1994) Science 266, 1702–1705
8. Takagaki, Y., MacDonald, C. C., Shenk, T., and Manley, J. L. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 1403–1407
9. MacDonald, C. C., Wilusz, J., and Shenk, T. (1994) Mol. Cell. Biol. 14, 6647–6654
10. Gross, S., and Moore, C. L. (2001) Mol. Cell. Biol. 21, 8045–8055
11. Chennathukuzhi, V. M., Lefrancois, S., Morales, C. R., Syed, V., and Hecht, N. B. (2001) Mol. Reprod. Dev. 58, 460–469
12. Duvel, K., and Braus, G. H. (1999) Nucleic Acids Res. 27, 4751–4758
13. Mandart, E., and Parker, R. (1995) Mol. Cell. Biol. 15, 6697–6696
14. Shell, S. A., Hesse, C., Morris, S. M., Jr., and Milcarek, C. (2005) J. Biol. Chem. 280, 39950–39961
15. Takagaki, Y., and Manley, J. L. (2000) EMBO J. 19, 3470–3481
16. Zhao, J., Kessler, M. M., and Moore, C. L. (1997) J. Biol. Chem. 272, 10831–10838
17. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
18. Ezaki, T., Sadowski, M., Dichtl, B., Hubner, W., and Keller, W. (2003) EMBO J. 22, 2167–2177
19. Noble, C. G., Walker, P. A., Calder, L. J., and Taylor, I. A. (2004) Nucleic Acids Res. 32, 3364–3375
20. Chen, J., and Moore, C. (1992) Mol. Cell. Biol. 12, 3470–3481
21. Koradi, R., Biller, M., and Wuthrich, K. (1996) J. Mol. Graphics 14, 51–55
22. Huber, Z., Monarez, R. R., Dass, B., and MacDonald, C. C. (2005) Ann. N. Y. Acad. Sci. 1061, 163–172
23. Steinmetz, E. J., Conrad, N. K., Brow, D. A., and Corden, J. L. (2001) Nature 413, 327–331
24. Brandsen, J., Werten, S., van der Vliet, P. C., Meisterernst, M., Krogan, N. J., and Gross, P. (1997) Nat. Struct. Biol. 4, 900–903
25. Werten, S., Wechselberger, R., Boelens, R., van der Vliet, P. C., and Kaptein, R. (1999) J. Biol. Chem. 274, 3693–3699
26. Hampsey, M. (1997) Yeast 13, 1099–1133
27. Colgan, D. F., and Manley, J. L. (1997) Genes Dev. 11, 2755–2766
28. Edmonds, M. (2002) Proc. Natl. Acad. Sci. Mol. Biol. 71, 285–289
29. Zhao, J., Hyman, L., and Moore, C. (1999) Microbiol. Mol. Biol. Rev. 63, 405–445
30. Meinhardt, A., Kamenski, T., Hoepfner, S., Baumli, S., and Cramer, P. (2005) Genes Dev. 19, 1401–1415
31. Cheng, H., He, X., and Moore, C. (2004) Mol. Cell. Biol. 24, 2932–2943
32. Dichtl, B., Asliland, R., and Keller, W. (2004) RNA 10, 965–977
33. He, X., Khan, A. U., Cheng, H., Pappas, D. L., Jr., Hampsey, M., and Moore, C. L. (2003) Genes Dev. 17, 1030–1042
34. Kuburz, A., Sadowski, M., Dichtl, B., and Keller, W. (2003) Nucleic Acids Res. 31, 3936–3945
35. Amrani, N., Minet, M., Wyers, F., Dufour, M. E., Aggerbeck, L. P., and Lacroute, F. (1997) Mol. Cell. Biol. 17, 1102–1109
36. Barilla, D., Lee, B. A., and Proudfoot, N. J. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 445–450
37. Licitasolci, D. D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J. B., and Bentley, D. L. (2002) Mol. Cell 9, 1101–1111
38. Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H., and Keller, W. (2002) Mol. Cell 10, 1139–1150
39. Zelhkovsky, A., Takanashi, Y., Nasser, T., He, X., Sterzer, U., Jensen, T. H., Domdey, H., and Moore, C. (2006) RNA (N. Y.) 12, 435–445
40. Kim, M., Krogan, N. J., Vasiljeva, L., Rando, O. J., Nedd, E., Greenblatt, J. F., and Buratowski, S. (2004) Nature 432, 517–522
41. Luo, W., and Bentley, D. (2004) Cell 119, 911–914
42. Holm, L., and Sander, C. (1998) Science 273, 595–603
43. Rose, R., Weyand, M., Lammers, M., Ishizaki, T., Ahmadian, M. R., and Wittinghofer, A. (2005) Nature 435, 513–518
44. Ye, K., and Patel, D. J. (2005) Structure (Lond.) 13, 1375–1384
87. Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrin-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Laev, A., Mena, F., Wong, P., Starostine, A., Canete, M. M., Vlashlim, J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rilstone, J. J., Gandi, K., Thompson, N. I., Musso, G., St Onge, P., Ghanny, S., Lam, M. H., Butland, G., Altal-El, A. M., Kanaya, S., Shlafitard, A., O'Shea, E., Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2006) *Nature* **440**, 637–643

88. Walsh, E. P., Lamont, D. J., Beattie, K. A., and Stark, M. J. (2002) *Biochemistry* **41**, 2409–2420