Regulation of Polyadenylation in Eukaryotic pre-mRNA

A thesis

submitted by

Chukwudi Ezeokonkwo

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in

Biochemistry

Sackler School of Graduate Biomedical Sciences

May, 2013

ADVISER: Claire Moore

ABSTRACT

The processing of eukaryotic pre-mRNA is an essential step in the synthesis of mRNA in living cells. It involves a coupled cleavage and polyadenylation reaction, in which the pre-mRNA is cleaved at the cleavage site, followed by the processive addition of a poly(A) tail to the upstream pre-mRNA transcript. The poly(A) tail is essential for the stability and export of mature mRNAs from the nucleus, and contributes to translation initiation in the cytoplasm. The Cleavage/Polyadenylation Factor (CPF) is one of the factors responsible for 3' end processing in yeast cells, and is made up of proteins that facilitate both cleavage and polyadenylation. My dissertation focuses on the mechanisms that regulate 3' end processing in yeast cells, and how interactions between the proteins that make up CPF play a role in this process. My thesis shows that interactions at the N-terminus of Pap1, a subunit of CPF and the enzyme that catalyzes poly(A) addition, inhibit and stimulate polyadenylation, and I propose a model in which these interactions are controlled by phosphorylation. I also show that Fip1, another subunit of CPF that recruits Pap1 to CPF, has a flexible linker in the middle of the protein, and that removal or replacement of the linker affects the efficiency of polyadenylation. The flexible linker might also facilitate the degradation of Fip1 and suggests another mechanism for the regulation of the polyadenylation step in yeast cells.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to all those who gave me the opportunity to complete this thesis. First, I would like to express my sincere gratitude to my advisor, Claire Moore. Claire is the best mentor I could have wished for, when I was trying to decide which labs to join. She was very patient with me, when things were not going well with my research, and she always pushed me to do my best work, and be the best that I could be. She always expected a lot from me, and I am eternally grateful to her for teaching me how to think deeply and thoroughly about science. Without her input and guidance throughout my graduate career, it would have been impossible for me to finish this work.

I also want to thank Dr. Andrew Bohm. Andrew was like my second advisor at Tufts. When I was having trouble with my original project, he came up with the idea to study the possibility that Fip1 might be a flexible linker as a side project. I jumped at this opportunity, and it ended up getting me a research paper, and being a major part of my thesis. I am profoundly grateful for his advice and all the help analyzing data and purifying proteins. I can honestly say that I would not be here today without his help. I really appreciate all his help.

I will like to further extend my gratitude to members of my thesis committee, Andrew Bohm, Gavin Schnitzler, and carol Kumamoto. My committee was very important in my research progress through graduate school, and they helped with critical comments, scientific discussions, and helpful suggestions throughout my thesis work and writing this dissertation. I also want to thank my outside examiner, Dr. Arthur Hunt for taking the time to read my thesis and being critical and helpful during my thesis defense. I really appreciate all of them.

I also want to express my sincere thanks to the Tufts MD/PhD program for accepting me into the program and helping me throughout my stay here at Tufts. I especially want to thank Naomi Rosenberg, the Dean of the Sackler School, and the advisor for the MD/PhD program. I have been meeting with her twice every year since I have been at Tufts, and she has been instrumental in getting me to this point. I want to thank Richard Wong, who is a good friend and sounding board in the program. I also want to thank Colleen Romain. She is my mum away from home. She is always there when I had issues I needed to resolve, or when I just needed to vent. She is always a very good listener, and always has a lot of advise. I thank her immensely for being there for me.

I wish to thank all the members of the Moore and Bohm labs for all their help throughout my research work. I especially want to thank Alexander Zhelkovsky for being very helpful. I also want to Mohamed Ghazy, Xiaoyuan He, Jason Keuhner, Derick Hoskinson and other members of the lab for the advice and help. The lab was a very fun place to be, and I looked forward to coming to work everyday, and discussing our favorite TV shows like Survivor and Big Bang Theory while working. They helped me cope when experiments didn't work, and were very happy when it worked. I am very proud to have worked with every one of them.

Finally, I want to express my deepest thanks and gratitude to my family. I want to thank my parents, Israel Ezeokonkwo, and my mum, Eunice Ezeokonkwo in Onitsha, Nigeria. They let me come to the United States by myself when I was 21. They trusted that I was mature enough to take care of myself and fulfill my goals. They were always at the back of my mind, and were always supporting me with prayers and advice. I thank them for all their help and for being there. I want to thank my brothers and sisters for all their help, especially my sister Uchenna Manafa and her husband, Joseph Manafa in Los Angeles. My brother, Ubaka also deserves a lot of thanks. He paid for my education

at UCSD as a foreign student. I am eternally grateful to him. I especially want to express my warmest gratitude to Susana Santos. She has been a big help through all the hardships of grad school, and I really appreciate her.

TABLE OF CONTENTS

Table of Contents……..iv

List of Tables……viii

List of Figures……...……...viii

Chapter 1

Introduction

A. Transcription and the maturation of pre-mRNAs……2

i. Transcription of Eukaryotic pre-mRNA…….…2

a. Capping…….…….....2

b. Splicing……..……..2

c. 3' end processing……..……3

ii. Termination…….…...5

B. 3' end processing of eukaryotic pre-mRNAs……..……..6

i. Polyadenylation signals in mammals……6

ii. Cleavage and polyadenylation factors in mammals…….……..9

a. Cleavage and polyadenylation specificity factor (CPSF)…….....9

b. Cleavage stimulation factor (CstF)……..……11

c. Cleavage factors I_m (CF Im)…….……...12

d. Cleavage Factor II_m (CF II_m)……..…..13

e. Poly(A) polymerase (PAP)……...……13

f. Nuclear poly(A) binding protein (PABPN1)……..……14

g. Symplekin……..……15

C. 3' end processing of pre-mRNAs in yeast……..……...16

i. Polyadenylation signals in yeast……16

ii. Yeast Cleavage and Polyadenylation factors……..18

a. Cleavage Factor IA (CF IA)……19

b. Cleavage Factor IB (CF IB)……..21

c. Cleavage/Polyadenylation Factor (CPF)……..21

d. Poly(A) binding protein (Pab1)……...26

e. Nab2……27

iiii. Plant cleavage and polyadenylation factors……28

D. Functions of the poly(A) tail in yeast……29

i. Poly(A) tails are involved in mRNA export……29

ii. Poly(A) tails are involved in quality control……...30

iii. Poly(A) tails stimulate translation initiation……...31

E. Regulation of polyadenylation……32

i. Regulation through direct interaction with PAP……32

a. U1A……...33

b. U170K……..33

c. PABPN1……..34

d. HnRNP H……34

ii. Regulation through post-translational modification of PAP……..34

a. Phosphorylation……35

b. Ubiquitination……35

c. Sumoylation……35

d. Acetylation……..36

iii. Regulation by competition with exonucleases……36

iv. Regulation due to defective mRNP assembly……..38

F. Specific aims of this thesis..……..39

Chapter 2

A flexible linker region in Fip1 is needed for efficient mRNA polyadenylation

A. Introduction……..……..42

i. Function of Fip1……42

ii. N-terminus of Fip1 interacts with mammalian PAP……..…43

iii. Goal of the study……...44

B. The flexibility of Fip1 facilitates Pap1 recruitment to the pre-mRNA during 3‘ end processing in *Saccharomyces cerevisiae*…..……...…….……41

i. Tethering of RNA to Pap1 increases efficiency of polyadenylation……..46

ii. Pap1∆90Fip1 fusion protein restores wild type activity in yeast lacking the *PAP1* and *FIP1* genes……..…49

iii. An intrinsically unstructured linker domain is important for function in polyadenylation……..……55

Chapter 3

The Role of the N-terminus of Pap1 in polyadenylation

A. Introduction……64

i. Function of Pap1……..…..64

ii. Deletion of the first 18 amino acids of Pap1 is lethal……..…64

iii. Goal of the study……..65

B. Results……65

i. Comparison of the activity of Pap1 and ∆N-Pap1……...65

ii. General interactions at the N-terminus of Pap1……..67

iii. An N-terminal fragment of Pta1 interacts with the N-terminus of Pap1……69

iv. Cft1 interacts with the N-terminus of Pap1……...73

v. Sub1 interacts with the N-terminus of Pap1……76

iv. C-terminus of Nab6 interacts with the N-terminus of Pap1……...79

v. The role of Nab6 and Sub1 in Pap1 function…….……80

Chapter 4

General Discussion

A. A flexible linker region in Fip1 is needed for efficient mRNA polyadenylation…….……....85

B. The Role of the N-terminus of Pap1 in polyadenylation……..91

C. Model for eukaryotic regulation of polyadenylation……..…….....98

D. Parallels in the regulation of mammalian PAP activity……...101

Chapter 5

Materials and Methods

Materials and Methods…….....104

References

References…….....121

List of Tables

Table 3.1: Mass Spectroscopy Results from the Peptide Column……..78

List of Figures

Figure 1.1: Eukaryotic 3' end processing……...4

Figure 1.2: Diagram of yeast and mammalian 3' end processing machineries……...5

Figure 1.2b: Diagram of yeast 3' end processing machinery..……..……16

Figure 1.3: Diagram of plant 3' end processing machinery……..29

Figure 2.1: Structure of yeast and mammalian Fip1……..…43

Figure 2.2: Pap1 interacts with amino acids residues 81-97 of Fip1……..…..45

Figure 2.3: Model to study Pap1 processivity..……..47

Figure 2.4: Activity of Pap_Np fusion protein…..……..48

Figure 2.5: Molecular depiction of the Pap_Np-RNA interaction……49

Figure 2.6: The Pap1Δ90Fip1 fusion protein……51

Figure 2.7: 5FOA results of the Pap1Δ90Fip1 fusion protein……..52

Figure 2.8: Western Analysis of Pap1Δ90Fip1 fusion protein……53

Figure 2.9: Further analysis of Pap1Δ90Fip1 fusion protein……..……54

Figure 2.10: Schematic representation of the ΔNFip1-RPA constructs……56

Figure 2.11**:** Thermosensitivity tests on the ΔNFip1-RPA constructs……57

Figure 2.12: *In vitro* analysis of modified ΔNFip1-RPA extracts……58

Figure 2.13: Poly(A) tail length at 30°C……...59

Figure 2.14: Western Analysis of ΔNFip1 constructs……61

Figure 2.15: Pta1 and CF IA interacts with the flexible linker of Fip1……62

Figure 3.1: The N-terminus of Pap1 is required for specific polyadenylation……..…67

Figure 3.2: Interactions at the N-terminus of Pap1…….....69

Figure 3.3: Pta1 interacts with the N-terminus of Pap1……71

Figure 3.4: Pta1 interacts with the N-terminus of Pap1……73

Figure 3.5: Cft1 interacts with the N-terminus of Pap1 *in vitro*……74

Figure 3.6: Cft1 does not have an inhibitory effect on Pap1 activity……..…76

Figure 3.7: Sub1 interacts with the N-terminus of Pap1……77

Figure 3.8: The Nab6 interaction requires the N-terminus of Pap1……80

Figure 3.9: Thermosensitivity of *nab6/sub1* double deletion……81

Figure 3.10: *In vitro* poly(A) addition assay at 30˚C……..…83

Figure 4.1: Prediction of the domains compositions of Fip1……88

Figure 4.2: Schematic diagram of Fip1 showing linker interactions……..……..……90

Figure 4.3: Schematic diagram of Fip1/PAP/CPSF160……..94

Figure 4.4: Sub1 rescues tail length in Glc7-depleted extracts……95

Figure 4.5: Homology between Sub1 and the C-terminus of Nab6……..……..……98

Figure 4.6: Model of Pap1 regulation……...99

Chapter 1

Introduction

A. Introduction to Maturation of Eukaryotic pre-mRNAs

In eukaryotic organisms, protein-encoding genes are exclusively transcribed by RNA polymerase II (Pol II) into messenger RNA precursors (pre-mRNAs) (Svejstrup 2004, Kornberg, 2007). The pre-mRNAs are synthesized in the nucleus and have to undergo three modification steps to become mature and functional messenger RNAs (mRNA). These steps are capping, splicing and 3' end processing. After these modifications, the mature mRNA is exported to the cytoplasm where it is translated into proteins by the ribosomes. These processing reactions often occur co-transcriptionally in the nucleus at the same time as Pol II transcription, and are facilitated by the recruitment of mRNA processing factors to the phosphorylated C-terminal domain (CTD) of the largest subunit of Pol II (Svejstrup 2004, Maniatis and Reed, 2002, Buratowski, 2009). The CTD is made up of Y-S-P-T-S-P-S repeats, and the phosphorylation states of Serines 2 and 5 of the CTD repeats are important for the progress of the transcription cycle. Serine 5 phosphorylation is required for promoter clearance, and defines the transition from initiation to elongation, while serine 2 phosphorylation is needed for elongation (Svejstrup 2004, Maniatis and Reed, 2002, Buratowski, 2009).

Capping is the first processing event during transcription elongation, and the presence of the essential cap structure on the 5' end is important for mRNA export and stability, translation initiation, and efficient splicing (Furuichi and Shatkin, 2000; Shibagaki et al, 1992). The next processing step involves splicing, which removes non-coding introns from the newly synthesized pre-mRNAs and joins the coding exon segments together in a two-step trans-esterification reaction (Kramer, 1996). The spliceosome is the enzyme responsible for catalyzing the splicing reaction, and it is made up of a large and dynamic multi-component complex consisting of the U1, U2, U4/U6, and U5 small nuclear ribonucleoproteins (snRNPs) and a multitude of non-

snRNP proteins (Sharp, 1998; Will and Luhrmann, 2001, Ritchie et al, 2009).

The last modification step in the maturation of mRNA is the 3' end processing of the pre-mRNA transcripts. This consists of a tightly coupled two-step reaction at the 3' end of pre-mRNAs (Fig 1.1) (Mandel et al, 2008; Millevoi and Vagner, 2009). The first step is a site-specific endonucleolytic cleavage of an internal phosphodiester bond at the cleavage site on the pre-mRNA transcript (Zhao et al, 1999a; Mandel et al, 2008), which is followed by the processive synthesis of a poly(A) tail onto the 3' end of the upstream cleavage product, while the downstream product is degraded. In eukaryotic cells, the three known exceptions to this pathway of processing RNAP II transcripts are: 1) the small nucleolar and nuclear RNAs (snoRNAs and snRNAs), 2) the major histone mRNAs, which are endonucleolytically cleaved downstream of a stem-loop structure by CPSF-73, but are not polyadenylated (Dominski and Marzluff, 1999; Perumal and Reddy, 2002, Dominski et al, 2005, Wilusz and Spector, 2010) and 3) micro-RNAs (miRNA) which are derived from the processing of a larger precursor primary miRNA (pri-miRNA) by a nuclear multiprotein complex, the "microprocessor", which include Drosha, RNase III and Pasha (partner of Drosha) (Denli et al 2004, Gregory et al, 2004). The length of poly(A) tails depends on the species, averaging 250 nucleotides in mammals, while in *Saccharomyces cerevisiae* this tail is 50-90 adenosines in length (Patel and Butler, 1992). Poly(A) tails have important roles in mRNA stability and cytoplasmic export, and are also required for efficient translation initiation (Tarun and Sachs, 1996; Paillard and Osborne, 2003).

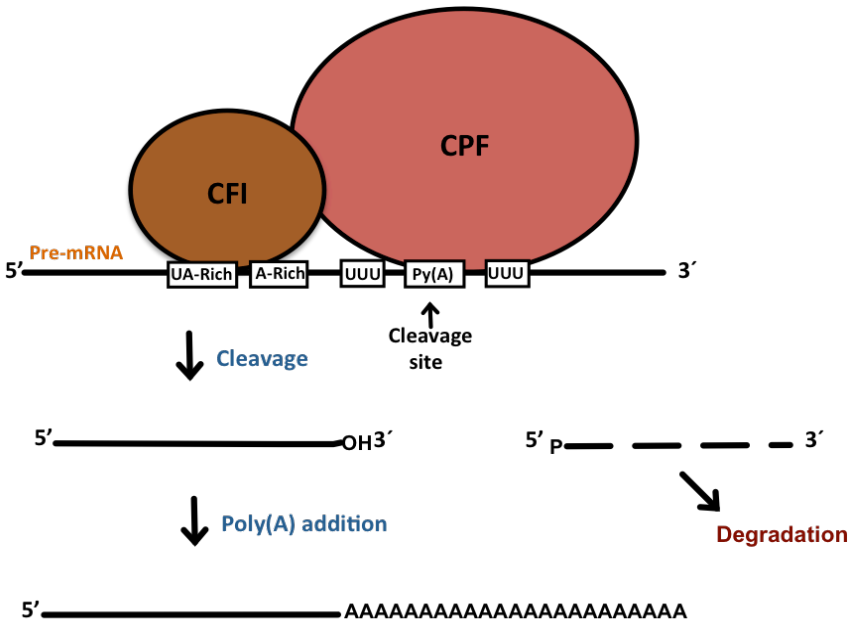


Figure 1.1: Eukaryotic 3' end processing, using the *S. cerevisiae* system for illustration. The first step is a site-specific endonucleolytic cleavage followed by the processive synthesis of a poly(A) tail onto the 3' end of the upstream cleavage product, while the downstream product is degraded.

The final stage of mRNA transcription is termination, which involves the release of the mRNA transcripts and RNA Pol II from the site of transcription on the DNA template. Termination is important in the cell because defects in termination can disrupt proper gene expression. It also affects the recycling of Pol II for subsequent rounds of transcription, and may result in read-through by the RNA polymerase II into the promoters of downstream genes (Eggermont and Proudfoot, 1993, Snyder et al, 1998; Greger and Proudfoot, 1998). The processing of mRNA 3' ends and transcription termination are generally tightly linked in eukaryotic organisms (Proudfoot et al, 2002; Proudfoot, 2004; Buratowski, 2005). The efficiency of termination is dependent on the presence of a functional poly(A) site on the pre-mRNA during transcription, and it has been shown that mutations in the poly(A) signal result in transcription beyond the

termination site in yeast cells (Connelly and Manley, 1998).

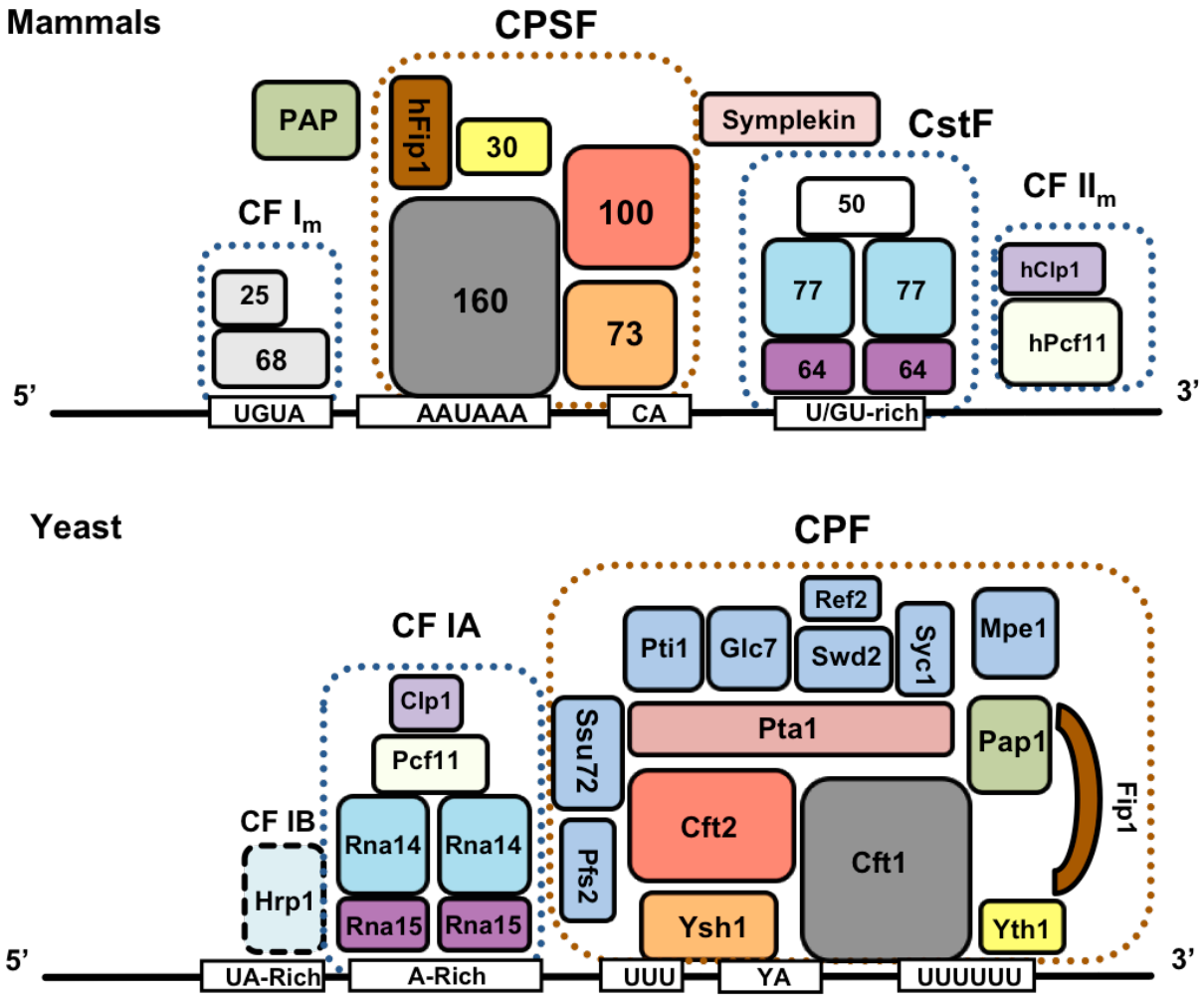


Figure 1.2: Diagram of the mammalian and yeast 3' end processing machineries. Functionally homologous factors are shown in the same colors. The known interactions between the different proteins are also indicated where possible by contact between the boxes. PAP/Pap1, depicted in green, is the polyadenylation enzyme, and Ysh1/CPSF 73, depicted in orange, is the endonuclease. The polyadenylation signals are also depicted on the pre-mRNA. CA is the mammalian cleavage site, while YA is the yeast cleavage site. Rna14/Rna15 and the mammalian homologues CstF 77/64 are known to exist as dimers in the complex.

B. 3’ end processing of eukaryotic pre-mRNAs

3’ end processing of pre-mRNAs is a tightly coupled two-step reaction that involves a site-specific endonucleolytic cleavage of the mRNA precursor followed by the addition of a poly(A) tail to the upstream cleavage product (Mandel et al, 2008; Millevoi and Vagner, 2009). The coupled cleavage and polyadenylation reaction requires the presence of polyadenylation signal sequences located in the 3’ untranslated region (3’ UTR) of the primary transcript to recruit the protein factors and guide them to the site of cleavage and polyadenylation (Zhao et al., 1999a; Mandel et al, 2008; Millevoi and Vagner, 2009).

Extensive biochemical study of this process has revealed that a complex set of proteins is responsible for performing 3’ end processing in both yeast and mammals (reviewed in Keller and Minvielle-Sebastia, 1997; Wahle and Ruegsegger, 1999; Zhao et al., 1999a; Mandel et al, 2008; Millevoi and Vagner, 2009). Both organisms require multi-protein complexes to perform 3’ end processing of pre-mRNAs and most of the mammalian polypeptides have yeast counterparts (Fig 1.2), indicating conservation during evolution (Keller and Minvielle-Sebastia, 1997). The identification of these protein complexes was made possible by the development of *in vitro* processing assays and fractionation of cell extracts from mammalian cells (Moore and Sharp, 1985) and yeast cells (Butler and Platt, 1988) to purify the active factors. The polyadenylation reaction can be reconstituted *in vitro* from the purified protein complexes, or by whole cell extracts, with suitable RNA substrates (Butler and Platt, 1988; Butler et al., 1990; Chen and Moore, 1992; Minvielle-Sebastia et al, 1994).

i. Polyadenylation signals in mammals

In mammalian cells, the polyadenylation signal is made up of the three sequence elements that indicate the polyadenylation site and two other sequence elements that enhance and regulate cleavage and polyadenylation of the pre-mRNA (Fig 1.2) (Zhao et al, 1999a; Mandel et al, 2008; Millevoi and Vagner, 2009). The first element identified was the hexanucleotide AAUAAA, which is located 10-30 nucleotides upstream of the cleavage site (Proudfoot and Brownlee, 1976). This sequence is one of the most highly conserved signal sequences, and is found in almost all mammalian pre-mRNAs (Proudfoot, 1991; Wahle and Kuhn, 1997). Mutagenesis studies have established the essential requirement of this sequence in both cleavage and polyadenylation (Manley et al., 1985; Wahle and Keller, 1992). The AUUAAA signal is the most common variant of this signal found in mammalian pre-mRNAs (Graber et al, 1999a). The alternative A-rich polyadenylation signal, which is reminiscent of 3′ end processing signals in plants and yeast, makes up a new class of non-canonical mammalian polyadenylation signals. It is comprised of a critical U/GU-rich downstream core element and an upstream A-rich tract (Nunes et al, 2010; Dickson and Wilusz, 2010). It lacks an AAUAAA hexamer, and bioinformatic analysis indicate that these A-rich signals may account for up to a third of non-canonical 3′ end processing events in mammalian cells (Nunes et al, 2010; Dickson and Wilusz, 2010).

The downstream element (DSE) is the second element of the core polyadenylation signal, and it was identified because a deletion more than 10 nucleotides downstream of the cleavage site reduced polyadenylation by three fold (Simonsen and Levinson, 1983). It is located about 30 nucleotides downstream of the poly(A) site (Zhao et al., 1999a), but has been observed further downstream in a few instances (Takagaki and Manley, 1997). The DSE is less conserved than the AAUAAA

sequence, and has been observed in two different forms, a GU-rich element with the consensus sequence YGUGUUYY (Y = pyrimidine) and a U-rich sequence (UUUUU) (Gil and Proudfoot, 1987; Chou et al., 1994). Point mutations or small deletions in the DSE have only weak effects, and larger deletions are required to abolish function. Nevertheless, the proximity of the DSE to the poly(A) site can affect the cleavage site position (MacDonald et al., 1994) and the efficiency of cleavage (Gil and Proudfoot, 1987). The DSE element is removed during cleavage, and this ensures that the pre-mRNA is not cleaved more than once.

The third mammalian polyadenylation signal is the UGUA sequence, which is found upstream of the cleavage site, and contributes to binding specificity with the CF I_m through an intramolecular G:A Watson–Crick/sugar-edge base interaction (Yang et al, 2010). Analysis of CF I_m-selected RNAs exhibited a highly related set of sequences of the form: 5′-UGUAA-UGUA(C/G/U)-GC-rich-3′, which make up the UGUA element. Three UGUAA elements are required for CF I_m-mediated regulation of poly(A) site cleavage *in vitro*, and these elements are conserved within the primary human and mouse CF I_m 68 kDa subunit poly(A) sites (Brown and Gilmartin, 2003).

The fourth mammalian polyadenylation signal is the cleavage site that is situated between the AAUAAA and the DSE (Zhao et al., 1999a; Mandel et al, 2008; Millevoi and Vagner, 2009). The distance between the hexanucleotide AAUAAA sequence and the DSE usually determines the selection of the cleavage site (Chen et al., 1995). In 70% of vertebrate pre-mRNAs, endonucleolytic cleavage occurs downstream of an adenosine residue (Sheets et al, 1990). The preceding nucleotide was a cytosine in 59% of 269 pre-mRNA sequences examined, which defines CA as the consensus cleavage site, with cleavage occurring immediately 3’ of the A residue (Sheets et al, 1990).

In addition to these signals, other non-essential sequences called auxiliary sequences can modulate the efficiency of processing (Gilmartin et al, 1995). The

auxiliary upstream element does not have a consensus sequence, but usually consists of a U-rich element (UUUU), which is located upstream of the AAUAAA signal. It enhances the efficiency of polyadenylation by facilitating binding of other polyadenylation factors to the cleavage site (Brackenridge and Proudfoot, 2000). Downstream auxiliary elements have also been identified in mammals, but are less common. They are generally G-rich elements, and lack both a conserved sequence and distance from the cleavage site. (Arhin et al, 2002).

ii. The Mammalian Cleavage and Polyadenylation Factors

The 20 proteins that are responsible for 3' end processing in mammals are organized into four sub-complexes and associated proteins, which were identified through the fractionation of cell extracts (Moore and Sharp, 1984). The mammalian protein factors are depicted in Figure 1.2. The cleavage and polyadenylation specificity factor (CPSF), the cleavage stimulation factor (CstF), the Cleavage factor I (CF I_m), and the cleavage factor II (CF II_m). Other associated proteins which are part of this machinery include poly(A) polymerase (PAP), poly(A) binding protein (PABP) and symplekin (Zhao et al., 1999a; Mandel et al, 2008; Millevoi and Vagner, 2009). Cleavage of the mammalian pre-mRNA substrates requires a combination of CPSF, CstF, CF I_m, CF II_m, while the CPSF, PAP and PABP are required for polyadenylation (Christofori and Keller, 1988; Takagaki et al, 1989). These factors are described in more detail below.

a) Cleavage and Polyadenylation Specificity Factor (CPSF)

CPSF is made up of five subunits: CPSF 160, CPSF 100, CPSF 73, CPSF 30 and hFip1. All the CPSF subunits also have yeast homologues, and are required for

cleavage and polyadenylation (Barabino et al, 1997; Bienroth et al, 1991; Jenny et al, 1994; Murthy and Manley, 1995; Kaufmann et al, 2004). CPSF specifically binds to the highly conserved hexanucleotide AAUAAA signal located upstream of the cleavage site. The largest subunit, CPSF-160, binds to the AAUAAA sequence on the pre-mRNA to direct cleavage (Jenny et al., 1994), and also plays a role in polyadenylation though its interaction with PAP (Murthy and Manley, 1995). CPSF 160 also interacts with the 77-kDa subunit of CstF, and plays a key role in coordinating cleavage and polyadenylation (Murthy and Manley, 1995; Colgan and Manley, 1997; Zhao et al, 1999). CPSF-73 is the endonuclease responsible for cleavage of pre-mRNAs (Ryan et al., 2004; Mandel et al, 2006). It contains a metallo-β-lactamase domain at its N-terminus, and is a member of the β-CASP superfamily of proteins (Mandel et al, 2008; Millevoi and Vagner, 2009). It has significant homology to the yeast protein Ysh1/Brr5, which also serves as the pre-mRNA cleaving nuclease. Another protein subunit in the CPSF is CPSF-100, which is also essential for viability. It has significant sequence homology to CPSF-73, and both proteins contain metallo-β-lactamase domains (Mandel et al, 2008; Millevoi and Vagner, 2009). However, CPSF- 100 has no endonucleolytic activity, since a part of the conserved motif necessary for the activity is lacking. The function of CPSF-100 in 3' end processing is still unknown.

CPSF-30 contains five CCCH zinc-finger domains and a zinc knuckle domain (DuBois et al, 1990), which are important for RNA-binding (Barabino et al, 1997). CFSF-30 has also been shown to bind preferentially to poly(U) sequences of RNA by *in vitro* RNA binding assays (Barabino et al, 1997), interacts with hFip1, and has significant homology to Yth1, which is part of the yeast 3' end-processing complex (Kaufmann et al., 2004; Mandel et al, 2008; Millevoi and Vagner, 2009). The last CPSF component is human Factor interacting with PAP (hFip1) (Kaufmann et al., 2004). It has some homology to the yeast Fip1 protein, and directly interacts with CPSF-30, CPSF 160,

CstF-77 and PAP. It binds preferentially to U-rich sequences on the pre-mRNA through an arginine-rich domain at the C-terminus of the protein, and stimulates PAP activity by recruiting PAP to the pre-mRNA during 3' end processing in mammalian cells (Kaufmann et al., 2004). It also forms a ternary complex with CPSF 160 and PAP in vitro, which suggests that CPSF 160 and hFip1 act together in mammalian poly(A) site recognition, and facilitate recruitment of PAP to the pre-mRNA during 3' end processing (Kaufmann et al., 2004).

b) Cleavage stimulation Factor (CstF)

The CstF is composed of three proteins, CstF-77, CstF-64 and CstF-50 (Mandel et al, 2008; Millevoi and Vagner, 2009), and binds to the U-rich or GU-rich sequence elements downstream of the cleavage site of the pre-mRNA (MacDonald et al., 1994; Takagaki et al., 1990). CstF-77 has been shown to bridge the interaction with the other CstF subunits, and to directly interact with CPSF-160, possibly stabilizing the CPSF-CstF-RNA complex (Murthy and Manley, 1995). It is homologous to the yeast Rna14 (Mandel et al, 2008; Millevoi and Vagner, 2009). The next subunit that makes up CstF is CstF-64, which was one of the first components of the 3' end-processing complex to be identified (Wilusz and Shenk, 1998). It has significant homology with Rna15, and binds the G/U-rich DSE of pre-mRNAs (MacDonald et al, 1994). It also interacts with symplekin through a highly conserved 100-residue hinge in the protein (Takagaki and Manley, 2000). The final component of the CstF complex is CstF-50, which is required for cleavage *in vitro* (Takagaki and Manley, 1994). It is composed of seven WD-40 repeats, which are required for binding with CstF-77 (Takagaki and Manley, 2000). Both CstF-77 and Cst-50 have also been shown to bind the phosphorylated CTD (Ser 2) of RNA Pol II, but CstF-50 binds with a higher efficiency (McCracken et al., 1997b).

c) Mammalian Cleavage Factor I_m (CF I_m)

The mammalian cleavage factor (CF I_m) is composed of the 25 kDa, 59 kDa and 68 kDa subunits as well as a less abundant fourth subunit of 72 kDa, which co-purifies with CF I_m activity from HeLa extracts (Ruegsegger et al, 1996). CF I_m activity can be reconstituted with only the 25 and 68 kDa subunits. CF I_m is required for cleavage *in vitro*, but has no clear sequence homologue in yeast, though Hrp1 may be a functional homolog (Takagaki et al, 1989). The CF I_m 59 kDa and CF I_m 68 kDa subunits contain three distinct domains: a RNP-type RNA-binding domain at the N-terminus, a central proline-rich domain and a C-terminal RS-like domain similar to that of the SR proteins involved in splicing. The 25-kDa subunit binds to RNA and was shown to associate with the largest subunit of CF I_m, with PAP, and with PABP (Dettwiler et al, 2004; Kim and Lee, 2001). The N-terminal RNA binding domain of the 68-kDa subunit forms a heterodimer with the 25-kDa subunit. This heterodimer also binds the UGUAA sequence on the pre-mRNA, which facilitates the ability of CF I_m to regulate cleavage during the processing of pre-mRNA (Brown and Gilmartin, 2003). The nature of the 72-kDa subunit of CF I_m is still unknown. The CF I_m has been shown to increase stability of the CPSF-CstF-RNA complex (Ruegsegger et al., 1996), and it may contribute to recognition of the pre-mRNA and identification of the poly(A) site (Mandel et al, 2008). CF I_m is also the first component of the mammalian 3′ end processing complex found to be capable of regulating the processing of the pre-mRNA of one of its subunits, both positively and negatively, at the level of poly(A) site cleavage (Brown and Gilmartin, 2003; Yang et al, 2010).

d) Mammalian Cleavage Factor II_m (CF II_m)

The mammalian cleavage factor II_m (CF II_m) is made up of two subunits, which are hPcf11 and hClp1. Both proteins were initially identified through their yeast homologues (Pcf11p and Clp1p) (de Vries et al, 2000). The yeast homologue of hPcf11 is equivalent to the N-terminal half of hPcf11. It also contains a CTD interacting domain (CID), which binds to the CTD in yeast (Sadowski et al, 2003). hPcf11 also contains two zinc finger motifs and 30 repeats of the consensus sequence LRFDG. In addition, hPcf11 was shown to play a role in cleavage, but is dispensable for polyadenylation. The function of the C-terminal domain of hPcf11 is unknown. hClp1 is evolutionarily conserved across all eukaryotes, and has 23% identity with the yeast Clp1 protein (de Vries et al, 2000). The structure of hClp1 shows similarity to ATPases, and it contains Walker A and B motifs, suggesting that hClp1 is an ATP/GTP binding protein (de Vries et al, 2000). hClp1 is involved in the cleavage reaction, and might play a role in tRNA splicing and as an RNA-kinase that activates synthetic short interfering RNAs (Leuschner et al, 2006).

e) Poly(A) polymerase (PAP)

PAP is the enzyme that catalyzes the polyadenylation reaction on cleaved pre-mRNAs, and plays a key role during 3' end processing of eukaryotic pre-mRNAs. It is required for cleavage and polyadenylation in mammals (Mandel et al, 2008; Millevoi and Vagner, 2009). The N-terminal part of the protein is highly conserved and contains the catalytic domain that is characteristic for the superfamily of nucleotidyltransferases (Holm and Sander, 1995; Martin and Keller, 1996). The structures of the bovine PAP and the yeast homologue have been solved, and both structures indicate three

functional domains of the protein, the N-terminal, the middle and the C-terminal domain (Bard et al, 2000; Martin et al., 2000). The N-terminal domain coordinates the two metal ions (Mg^{2+} and Mn^{2+}) required for catalysis, while the catalytic site contains three conserved aspartate residues that are essential for activity (Bard et al, 2000; Martin et al., 2000). The C-terminal domain is required for hFip1 and CPSF 160 binding (Martin et al., 2000; Kaufmann et al., 2004), and contains an RNA binding domain (Thuresson et al, 1994). It also has a bipartite nuclear localization signal (NLS), and a region rich in serine and threonine residues that is important for the regulation of PAP activity by phosphorylation (Abuodeh et al, 1998; Colgan et al, 1996). PAP alone has a low level of activity, which is enhanced when manganese is substituted for magnesium (Wahle, 1991). It can add poly(A) tails to RNAs containing the AAUAAA sequence in the presence of CPSF, (Murthy and Manley, 1995), or in an AAUAAA-independent manner in the presence of CF I_m and the hFip1 subunit of CPSF (Gilmartin et al, 1995, Brown and Gilmartin, 2003).

f) Poly(A) binding protein (PABPN1)

The nuclear poly(A) binding protein PABPN1 is required for efficient polyadenylation in the cell. It has also been shown that the efficiency of polyadenylation is increased 80-fold by the binding of the PABPN1 to the pre-mRNA during polyadenylation (Kerwit et al, 2003). CPSF and PAP are sufficient to allow poly(A) addition onto the pre-cleaved RNA substrate, but fast and processive activity requires PABPN1, which also plays a key role in controlling poly(A) tail length (Bienroth et al, 1993). PABPN1 interacts directly with PAP in the presence of CPSF and RNA, and strongly increases the RNA affinity of the polymerase, thereby stimulating PAP activity, without increasing the catalytic activity of PAP (Kerwit et al, 2003). Processivity of PAP is

high in the presence of both factors, which facilitate poly(A) tail synthesis without dissociation of the polyadenylation complex in *in vitro* studies (Bienroth et al, 1993). PABPN1 has been proposed to coat the growing poly(A) tail, with only one PABPN1 directly interacting with PAP. This would result in a contiguous coverage of the growing poly(A) tail by PABPN1, which is needed for measurement of the tail length (Kerwit et al, 2003). Once the poly(A) tail has reached full length (250 residues in mammalian cells), the polyadenylation reaction becomes slow and distributive (Wahle, 1995). PABPN1 is therefore proposed to control poly(A) tail length by regulating the interaction between poly(A) polymerase and the CPSF (Kühn et al, 2009). PABPN1 is a 33-kDa protein, and contains an acidic N-terminal domain, a basic C-terminal domain and a single RNP domain in the center, and interacts with CPSF-30 *in vitro* (Chen et al, 1999).

g) Symplekin

Symplekin was originally identified as part of the tight junctions in mammalian cells (Keon et al, 1996), but was later shown to play a role in mammalian 3' end processing (Takagaki and Manley, 2000). It shares homology with Pta1, which is a yeast protein that plays a major role in cleavage and polyadenylation as a scaffold protein that organizes the 3' end-processing complex (Zhao et al, 1999b, Ghazy et al, 2009). Symplekin interacts with the CTD Serine 5 phosphatase, Ssu72, the hinge region of CstF-64, and has also been shown to form a stable complex with CPSF and CstF (Takagaki and Manley, 2000; Xiang et al, 2010). It might also play a role as a scaffold protein in the 3' end processing of the non-polyadenylated histone pre-mRNAs, and in other polyadenylation reactions that occur in the cytoplasm (Barnard et al, 2004; Kolev and Steitz, 2005). A recent crystal structure shows the N-terminal domain of Symplekin in a ternary complex with Ssu72 and a CTD Ser5 phosphopeptide (Xiang et al, 2010).

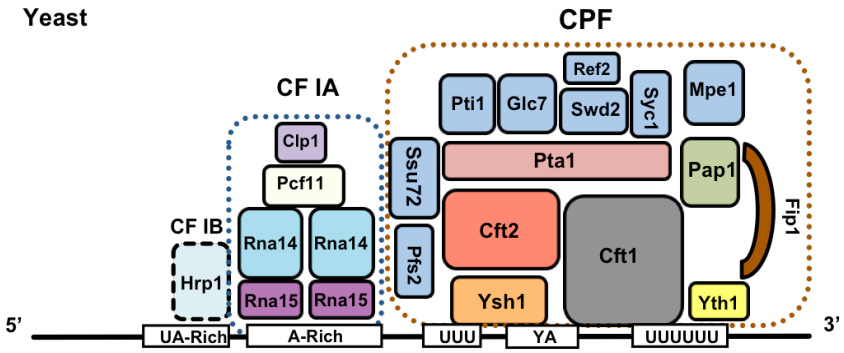


Figure 1.2b: Diagram of the yeast 3' end processing machinery

C. 3' end processing of pre-mRNAs in yeast

i. Polyadenylation signals in yeast

In contrast to the mammalian polyadenylation signals that are highly conserved, the polyadenylation signals in yeast *S. cerevisiae* are more degenerate and redundant (Zhao et al., 1999a). The yeast signals are defined by five sequence elements: the efficiency element (EE), the positioning element (PE), the upstream U-rich element (UUE), the downstream U-rich element (DUE), and the poly(A) site/cleavage site (Figure 1.2). These sequence elements were identified through mutagenesis studies and computational analyses of a large set of known and predicted 3' end processing sites (Graber et al, 1999b). Deletion or mutation of some individual yeast signal sequences has only a slight effect or no effect on yeast 3' end processing, which suggests a high degree of redundancy (Beyer et al, 1997; Zhao et al, 1999a). Typically, the presence of one weak signal in the 3' UTR of a transcript can be compensated for by stronger signals in other sequence elements.

The EE is located furthest upstream from the cleavage site, between 35 and 70 nucleotides from the cleavage site, and is an AU-rich element (Zhao et al, 1999a). It is the most efficient sequence to direct 3' end processing (Guo et al, 1995), with the U residues at the first and fifth positions being the most critical nucleotides in the sequence. The EE is not always required for cleavage, but it enhances the efficiency of cleavage (Guo and Sherman, 1996). The EE is specifically bound by CF IB, which acts in cleavage site selection (Kessler et al, 1997; Minvielle-Sebastia et al, 1998). The PE is an A-rich element that is located between 10 and 30 nucleotides upstream of the cleavage site. The most effective motifs for the PE are AAUAAA and AAAAAA, but just like the EE, it is highly variable (Zhao et al, 1999a; Mandel et al, 2008). The PE is present in all organisms and directs the position of the cleavage site. It also contributes to the efficiency of the processing, as single point mutations or deletion of the motif result in the reduction in the use of the cleavage site (Abe et al, 1990; Hyman et al, 1991). It resembles the mammalian AAUAAA signal, but does not play the same critical role in 3' end processing.

The upstream and downstream elements occur within 10 nucleotides upstream and downstream of the cleavage site, and are best characterized as U-rich elements (Graber et al, 1999a). Mutations in the DUE and UUE reduce cleavage activity, but do not stop cleavage. They were found to contribute to cleavage site selection and to enhance cleavage *in vitro* (Dichtl and Keller, 2001), and are not found in mammals (Graber et al, 1999a). Three components of the CPF - Cft1, Cft2 and Yth1 have been shown to bind to these U-rich elements (Barabino et al, 2000; Dichtl et al, 2002a; Dichtl and Keller, 2001; Keller et al, 1991).

The consensus poly(A) site was identified from mapping of the poly(A) sites of numerous genes and consists of a pyrimidine followed by one or more A residues: $Y(A)_n$ (Heidmann et al, 1994; Russo et al, 1993). Cleavage takes place at the 3' side of an

adenosine residue, which is similar to the mammalian reaction (Mandel et al, 2008). In contrast to animal genes that use a single poly(A) site downstream of the hexanucleotide AAUAAA, yeast genes often use a cluster of poly(A) sites or adjacent sub-optimal poly(A) sites, also called cryptic cleavage sites (Zhao et al, 1999a). U-rich motifs are frequently found immediately before and after the poly(A) site (Graber et al, 1999b), and in some transcripts, like *ADH2*, a downstream sequence is required for efficient 3' end processing (Zhao et al, 1999a). The yeast signal sequences are believed to act in cooperation, as shown for some genes (like *CYC1*) that use multiple weak signals to act additively to provide a strong signal (Guo et al, 1995). The presence of secondary structures in the primary transcripts may also be important in some genes (Sadhale and Platt, 1992).

ii. The yeast cleavage and polyadenylation factors

In the yeast *S. cerevisiae*, 3' end processing requires the presence of CPF, CF IA and CF IB, and Pab1. Although the polyadenylation signals used by mammals and yeast are different in their sequence and organization, protein factors of the 3' end processing apparatus of both organisms exhibit high conservation (Figure 1.2). The yeast cleavage and polyadenylation factors were identified from the fractionation of yeast cell extracts (Chen and Moore, 1992; Dichtl et al, 2002b; Kessler et al, 1996; Nedea et al, 2003), and the reconstitution of 3' end processing assays with partially purified chromatographic fractions initially defined five protein factors, the Cleavage Factors I and II (CF I and CF II), Polyadenylation Factor PF I, poly(A) polymerase (Pap1) and poly(A) binding protein (Pab1). The CF II and PF I had several polypeptides in common and were later shown to be part of the same functional unit renamed the CPF (Ohnacker et al, 2000). Further purification of the CF I also identified the different

components that make up the complex (CF IA and CF IB) (Kessler et al, 1996). The large majority of genes that encode for the yeast 3' end processing polypeptides are essential for cell viability, confirming the essential function of their products in pre-mRNA 3' end formation.

a) The Cleavage Factor IA (CF IA)

The Cleavage Factor I (CF I) was originally identified as a factor required for both the cleavage and poly(A) addition reactions (Chen and Moore, 1992), and further purification was used to separate it into two factors, CF IA and CF IB (Kessler et al, 1996). CF IA is made up of four essential protein subunits - Rna14 (76 kDa), Pcf11 (72 kDa), Clp1 (50 kDa) and Rna15 (38 kDa). CF IB consists of the single polypeptide Hrp1. There are many protein-protein interactions between the CF IA and CPF, which suggest a high degree of cooperation in the activity of these two complexes in 3' end processing. Rna14 and Rna15 are tightly associated, forming a heterodimeric complex required for RNA-binding (Noble et al, 2004), and are essential for cleavage and polyadenylation (Kessler et al., 1997). Rna14 has sequence homology to mammalian CstF-77 (24% identity) and interacts with Pcf11, Hrp1, Pfs2 and Cft1 (Gross and Moore, 2001; Kyburz et al, 2003; Ohnacker et al, 2000). Rna14 binds to the unphosphorylated CTD of RNA Pol II, and this binding increases in the presence of CTD phosphorylation at Ser 2 (Sadowski et al, 2003).

Rna15 is the yeast homologue of CstF-64 (Takagaki and Manley, 1994) and contains a single RNA recognition motif (RRM)-type RNA binding domain at its N-terminus and a C-terminal region that is essential for polyadenylation and transcription termination (Minvielle-Sebastia et al, 1994; Qu et al, 2007). Rna15 can be UV cross-linked to RNA substrates, but its exact contact with the mRNA precursor is not certain.

Structural analysis of the RRM of Rna15 shows that the protein has affinity for GU-rich sequences, and discriminates against C and A bases on the RNA (Kessler et al, 1996; Minvielle-Sebastia et al, 1991; Takagaki and Manley, 1994; Pancevac et al, 2010). However, the UV cross-linking experiments suggest that Rna15 binds to the A-rich PE in the presence of Rna14 and Hrp1, and that this requires the Rna15 RRM (Gross and Moore, 2001). Structural studies of the RRMs of Rna15 and Hrp1 also show interaction between the proteins when both are bound to RNA, and confirmed the preference of Rna15 for the A-rich positioning element (Leeper et al, 2010). Rna15 interacts with Pcf11 and Rna14 in CF IA, and it, along with Rna14, is present as a dimer in CF IA (Dr. Bohm, unpublished results). It also interacts with Sub1, which might function as an anti-terminator during transcription (Calvo and Manley, 2001).

Pcf11 is involved in cleavage and polyadenylation, and was initially identified in two-hybrid screens as a polypeptide interacting with Rna14 and Rna15 (Amrani et al, 1997). A CTD-interaction domain (CID) located within about 130 residues of the N-terminus of Pcf11 is responsible for the direct interaction of Pcf11 with the CTD phosphorylated on serine 2, and plays a role in the coupling of 3' end processing and transcription termination (Barilla et al, 2001; Licatalosi et al, 2002; Sadowski et al, 2003). Pcf11 is also involved in the dismantling of the RNAP II elongation complex (Zhang et al, 2005). Mutations in the CID reduce or eliminate CTD interaction, which results in cell death, but surprisingly does not affect 3' end processing. This suggests that the pre-mRNA 3' end processing activities and CTD binding of Pcf11 can be functionally uncoupled from each other (Sadowski et al, 2003). Pcf11 has putative homologues in many species that all contain a N-terminal CID (Barilla et al, 2001; Steinmetz and Brow, 1996; Yuryev et al, 1996). Clp1 is the yeast homologue of the mammalian hClp1, and is conserved throughout eukaryotes (de Vries et al, 2000). Clp1 shares 23% identity with hClp1, and binds to Pcf11 but not to the other CF IA subunits (Gross and Moore, 2001).

It also binds strongly to Ysh1/Brr5, and weakly to Cft2 (Kyburz et al, 2003) and Pta1 (Ghazy et al, 2009).

b) The Cleavage Factor IB (CF IB)

CF IB is made up of the single 73 kDa Nab4/Hrp1 polypeptide (Kessler et al, 1997; Kessler et al, 1996). It regulates cleavage site selection by suppressing the use of cryptic poly(A) sites (Minvielle-Sebastia et al, 1998) and specifically binds the pre-mRNA at the AU-rich EE (Chen and Hyman, 1998; Valentini et al, 1999). The central region of Hrp1p contains two consecutive RRM-type RNA-binding domains (RBDs) arranged in tandem (Henry et al, 1996). The C-terminal part of Hrp1 is rich in arginine and glycine, with methylation sites. Similar domains usually mediate protein-protein interactions (Zhao et al, 1999a), and the methylation of Hrp1 controls shuttling of the protein between the nucleus and the cytoplasm (Henry et al, 1996; Kessler et al, 1997; Shen et al, 1998).

c) The Cleavage/Polyadenylation Factor (CPF)

The CPF is made up of 15 proteins, and includes Ysh1/Brr5, the endonuclease, and Pap1, which is the enzyme that catalyzes poly(A) addition (Mandel et al, 2006; Linger et al, 1991, Bard et al, 2000). The CPF proteins were identified through affinity purification and mass spectrometry of the purified proteins (Dichtl et al, 2002b; Gavin et al, 2002; Nedea et al., 2003), and contain the following proteins, from largest to smallest: Yhh1/Cft1 (150 kDa), Ydh1/Cft2 (105 kDa), Ysh1/Brr5 (100 kDa), Pta1 (85 kDa), Pap1 (64 kDa), Ref2 (60 kDa), Mpe1 (58 kDa), Fip1 (55 kDa), Pfs2 (53 kDa), Pti1 (47 kDa), Swd2 (37 kDa), Glc7 (36 kDa), Yth1 (26 kDa), Ssu72 (23 kDa), Syc1 (20 kDa). Swd2, Glc7, Ref2, Pti1, Ssu72, and Syc1 are part of a sub-complex called the APT complex

(Associated with Pta1), and play a role in the processing of snoRNAs, which are also transcribed by RNA Pol II (Nedea et al, 2003). Cft1, Cft2, Ysh1/Brr5 and Pta1 also form a sub-complex called CF II, which along with the CF I is sufficient for cleavage, but not polyadenylation (Zhao et al, 1997).

The largest subunit of CPF is Cft1, which is the homologue of the mammalian CPSF-160. It is essential for cleavage and polyadenylation (Stumpf and Domdey, 1996). Cft1 is an RNA-binding protein that binds the pre-mRNA at U-rich sequences surrounding the poly(A) site and is involved in poly(A) site recognition (Dichtl et al, 2002a). The RNA-binding domain in the middle of the protein is composed of β-propeller repeats (Dichtl et al, 2002a). Cft1 has been shown to interact with the phosphorylated CTD during transcription, and probably has a role in the coupling of 3' end formation and transcription (Dichtl et al, 2002a). It also interacts directly with Fip1, Cft2, and Swd2, and has a putative interaction with Pap1, which might play a role in the regulation of yeast poly(A) addition (Mandel et al, 2008; Millevoi and Vagner, 2009).

Cft2 shares significant homology to CPSF-100, and is also related to Brr5 and CPSF-73 (Jenny et al, 1996). Cft2 does not contain any known RNA-binding domains, but can be UV cross-linked to wild type pre-mRNAs in the presence of ATP, but not to pre-cleaved RNAs or substrates lacking the EE (Zhao et al, 1997). Cft2 has also been shown to bind with high affinity to the U-rich sequences surrounding the cleavage site and to the EE to a lesser extent (Dichtl and Keller, 2001). Cft2 also interacts with several 3' end-processing polypeptides, which include Ysh1/Brr5, Cft1, Pta1, Pfs2, Ssu72, Pcf11, and also to the CTD of RNA Pol II (Kyburz et al, 2003, Mandel et al, 2008).

Ysh1/Brr5 is the yeast homologue of mammalian CPSF-73 (Chanfreau et al, 1996; Jenny et al, 1996). It is essential for both cleavage and polyadenylation activities *in vivo* and *in vitro*. The mammalian homologue of Ysh1, CPSF-73 has been reported to be the mammalian 3' end processing endonuclease (Mandel et al, 2006; Ryan et al,

2004), and Ysh1 is presumed to be the endonuclease that acts on yeast pre-mRNAs. The C-terminus of Ysh1 has sequence homology to Syc1, which is proposed to act as a negative regulator of Ysh1 activity (Zhelkovsky et al, 2006).

Pta1 was originally identified as a protein that played a role in processing of pre-tRNAs (O'Connor and Peebles, 1992), but was later shown to be essential for eukaryotic mRNA 3' end processing. It is involved in both cleavage and poly(A) addition reactions (Zhao et al, 1999b). The activity of Pta1 during 3' end processing is regulated by the phosphorylation state of the protein, which is controlled by the phosphatase, Glc7. The depletion of Glc7 from yeast extract inhibited the polyadenylation of a *GAL7* RNA substrate, which suggests that phosphorylated Pta1 inhibits poly(A) addition (He and Moore, 2005). Furthermore, addition of non-phosphorylated Pta1 can restore polyadenylation activity in defective Glc7-depleted extracts (He and Moore, 2005). Pta1 also functions as a scaffold through its interaction with many of the proteins that make up the complex. It interacts with CPF subunits Ysh1, Pti1, Ssu72, Glc7, Syc1, Cft2, Fip1, Pfs2, and Pap1 and the CF IA subunits Pcf11 and Clp1 (Ghazy et al, 2009).

Pap1 catalyzes the polyadenylation of cleaved pre-mRNAs and was the first 3' end processing factor to be purified (Lingner et al, 1991a; Lingner et al, 1991b). The mammalian and yeast proteins share 47% identity in the first 400 amino acids, a region that comprises the catalytic domain containing three aspartate residues highly conserved across species (Martin et al, 2000). The yeast Pap1 lacks two C-terminal domains present in mammalian PAP: a domain that carries a nuclear localization signal and a serine/threonine rich domain. An RNA-binding site (C-RBS) at the C-terminus is essential for activity (Zhelkovsky et al, 1998). In contrast to PAP, yeast Pap1 is responsible for polyadenylation, but is not involved in cleavage (Mandart and Parker, 1995). Pap1 has a Fip1 binding site at the C-terminus of the protein, which overlaps with the C-RBS domain and allows Fip1 to regulate the processivity of the polymerase in vitro

(Zhelkovsky et al, 1998; Meinke et al, 2008). The structures of Pap1 alone and in complex with the Pap1-interacting domain of Fip1 have been reported (Bard et al, 2000, Meinke et al, 2008), and the mechanism of action of the enzyme during poly(A) addition has also been described (Balbo et al, 2005; 2007; Balbo and Bohm, 2007; 2009). Pap1 undergoes phosphorylation and ubiquitination during the S-G2 transition of the cell cycle, which inactivates the enzyme (Mizrahi and Moore, 2000).

Fip1 (Factor interacting with Pap1) was originally identified in yeast as a protein interacting with Pap1, and is exclusively involved in polyadenylation (Preker et al, 1995). Fip1 recruits Pap1 into the CPF complex through its interaction with Yth1 (Barabino et al, 2000). Fip1 inhibits the non-specific polyadenylation activity of Pap1 through multiple interactions, and interacts with Pap1 through amino acids residues 80-105 of the protein, which are essential for cell viability (Helmling et al, 2001). Furthermore, amino acid residues 206-220 are responsible for interaction with Yth1, and a deletion of Fip1 from the C-terminus up to amino acid 220 (but no further) is functional in polyadenylation (Helmling et al, 2001). Fip1 has an extended conformation in solution and when bound to Pap1, suggesting that very little of Fip1 folds into an ordered structure, and it is highly susceptible to proteasomal degradation in the cell (Meinke et al, 2008; Saguez et al, 2008). It interacts with Ysh1/Brr5, Pta1, Pfs2 and Rna14 (Barabino et al, 2000; Ohnacker et al, 2000; Preker et al, 1995, He and Moore, 2005; Ghazy et al, 2009).

Yth1 is a CPF protein that binds to U-rich sequences around the cleavage site on the pre-mRNA (Barabino et al., 2000). It is homologous to the mammalian CPSF-30 (Barabino et al, 1997). Yth1 binds directly to Fip1 and Brr5 and is essential for cleavage and polyadenylation (Barabino et al, 1997; Barabino et al, 2000). Deletion of the C-terminus of Yth1 results in loss of both Fip1 and Pap1 from the CPF complex, suggesting that both proteins are tethered to the complex through Yth1, and binding of Yth1 to Fip1 inhibits Yth1-RNA binding (Barabino et al, 2000; Tacahashi et al, 2003).

Glc7 is an essential CPF protein, which is required for cell viability (Gavin et al, 2002). It is part of the APT complex in yeast, and is the only isoform of type 1 serine/threonine protein phosphatase in yeast (Feng et al, 1991; Nedea et al, 2003). During pre-mRNA 3' end processing, Glc7 is required for polyadenylation, but is dispensable for cleavage. Depletion or inhibition of Glc7 activity in yeast extracts resulted in defective polyadenylation in *in vitro* assays, while cleavage was normal (He and Moore, 2005). Furthermore, phosphorylated Pta1 was found to be a substrate of Glc7 within the CPF complex, which suggests that phosphorylation plays a major role in the regulation of CPF activity (He and Moore, 2005).

Mpe1 is an evolutionary conserved protein that is exclusively involved in polyadenylation of pre-mRNAs (Vo et al, 2001). Mpe1 was identified through its suppression of a temperature-sensitive Pcf11 mutation (Vo et al, 2001). Pfs2 is an essential 53-kDa CPF protein that contains seven WD-40 repeats between amino acids 90 and 380 (Ohnacker et al, 2000, Wang et al, 2005). Pfs2 interacts with the CPF proteins Fip1, Brr5, and Swd2, and the CF IA protein, Rna14, which suggests that it might play a role in the assembly and stabilization of both factors during 3' end processing (Ohnacker et al, 2000; Dichtl et al, 2004). It is essential for both cleavage and polyadenylation.

The remaining five components of the CPF along with Glc7 make up the yeast APT (Associated with Pta1) complex, which is important for 3' end processing of small nucleolar RNAs (snoRNA). They include Ssu72, Pti1, Swd2, Syc1 and Ref2 (Nedea et al, 2003). Ssu72 is involved in transcription termination of both pre-mRNAs and snoRNAs (Ganem et al, 2003; Nedea et al, 2003; Steinmetz and Brow, 2003) and affects start site selection during transcription initiation. Ssu72 interacts with the N-terminus of Pta1, and is directly involved in pre-mRNA cleavage but not in polyadenylation (He et al, 2003; Ghazy et al, 2009). It also interacts with Cft2, and the

Rpb2 subunit of RNA Pol II (Pappas and Hampsey, 2000). Pti1 shows significant sequence homology to Rna15 and CstF-64, and is not essential for mRNA 3' end processing. Defects or mutations in the protein affect cleavage site choice, and the polyadenylation length of only some pre-mRNAs (Dheur et al, 2003; Skaar and Greenleaf, 2002). Swd2 was identified through its interaction with Pta1, and has no direct effect on mRNA 3' end cleavage *in vitro* (Cheng et al, 2004). Mutations or a depletion of the protein however disrupt transcription termination (Cheng et al, 2004; Dichtl et al, 2004). Swd2 interacts with many components of the CPF, as well as Pcf11 in CF IA, and might play a role in tethering the complex to other proteins.

Ref2 is one of the two non-essential proteins in the CPF. It is an RNA-binding protein that is required for the efficient 3' end processing of weak poly(A) sites (Russnak et al, 1995). Syc1 is the second non-essential protein in the CPF (Gavin et al., 2002; Nedea et al., 2003). Its amino acid sequence is highly similar to the C-terminal region of Ysh1, which is essential for cell viability. The presence of Syc1 was shown to exacerbate the processing and growth defects of many processing mutants (Zhelkovsky et al, 2006), and it is required for efficient 3' end processing of poly(A) sites with poor consensus sequences (Russnak et al, 1995).

d) The poly(A) binding protein (Pab1)

Pab1 is a highly conserved 70-kDa RNA binding protein that is associated with polyadenylated mRNAs in both the nucleus and cytoplasm (Setyono and Greenburg, 1981; Adam et al, 1986). Pab1 contains four N-terminal RNA recognition motifs (RRMs), which are connected to a C-terminal helical domain via a segment rich in proline and methionine (Mangus et al, 2004; Sachs et al, 1986). The polyadenylation reaction can occur without Pab1, but Pab1 is needed for poly(A) tail length control (Kessler et al,

1997; Minvielle-Sebastia et al, 1997). Pab1 interacts directly with Rna15, and regulates poly(A) tail length by limiting the access of Pap1 to the pre-mRNA, which inhibits Pap1 activity (Amrani et al, 1997). It also functions in recruiting a poly(A)-specific nuclease, PAN, that is composed of two subunits, Pan2 (127 kDa) and Pan3 (76 kDa) (Deardorff and Sachs, 1997; Lowell et al, 1992). A minimum of 12 adenosine nucleotides is required for Pab1 interaction during polyadenylation, and Pab1 has also been proposed to coat the poly(A) to facilitate tail length control (Adam et al, 1986; Setyono and Greenberg, 1981; Kerwitz et al, 2003). The association of Pab1 with the poly(A) tail also leads to stimulation of mRNA translation *in vivo* and *in vitro* by interacting with eIF4G, a component of the yeast translation initiation factor (Tarun and Sachs, 1996; Tarun et al, 1997). Pab1 also has inhibitory functions in mRNA decapping and deadenylation (Caponigro and Parker, 1995).

e) Nab2

Nab2 is a nuclear polyadenylated RNA-binding protein required for nuclear mRNA export and poly(A) tail length control (Hector et al, 2002, Dheur et al, 2005). Nab2 is required for mRNA poly(A) tail length control *in vivo* and *in vitro,* and it plays an essential and direct role in limiting tail length during polyadenylation . Depletion of Nab2 *in vivo* leads to the appearance of hyperadenylated mRNA in the cell which cannot be rescued by the over-expression of *PAB1* or expression of nuclear-targeted Pab1. It also leads to nuclear accumulation of mRNA. These defects however can be uncoupled in a *nab2* mutant, in contrast to other nuclear export mutants (Hector et al, 2002). Nab2 also stimulates polyadenylation while inhibiting hyperadenylation *in vitro* in the absence of Pab1; and does not show poly(A) tail trimming activity *in vitro* (Hector et al, 2002). Nab2 binds to the Nuclear Transport Factor Kap104 *in vitro* assays, and this binding

specifically inhibits poly(A) tail length control by Nab2 (Dheur et al, 2005). The regulation of *NAB2* mRNA 3′-end formation requires the exosome and Trf4 of the TRAMP complex, and the level of this mRNA is controlled by competition between polyadenylation and Rrp6p-dependent degradation (Roth et al, 2009). Nab2 contains a zinc finger domain that interacts with poly(A) RNA, and is critical for the proper control of poly(A) tail length (Kelly et al, 2010), and it interacts with the nuclear pore associated protein, Mlp1 to facilitate mRNA export from the nucleus (Fasken et al, 2008).

iii. A brief overview of plant cleavage and polyadenylation factors

The components that make up the plant cleavage and polyadenylation system are not as well studied as the mammalian and yeast systems. Recent work on the plant 3' end processing machinery (reviewed in Hunt, 2008), however, shows that it contains homologs of the yeast and mammalian cleavage and polyadenylation factors except for CF I_m, which is also absent in yeast (Hunt, 2008) and Hrp1, which is part of the yeast machinery. The plant poly(A) signal consists of three poorly conserved sequence elements, which include one or more A-rich near-upstream elements (NUE), each directing processing at a poly(A) site a short distance downstream of it, and an extensive far-upstream U-rich element (FUE) that enhances processing efficiency at all sites, and a U-rich element (CE) that contains a pyrimidine-adenosine dinucleotide that functions as the cleavage site (Hunt, 2008; Millevoi and Vagner, 2009). Just like the mammalian system, *Arabidopsis* contains a CPSF complex (AtCPSF), which includes AtCPSF160, AtCPSF100, AtCPSF73-I, AtCPSF30, AtFIPS5, AtCPSF73-II and AtFY (Hunt, 2008; Millevoi and Vagner, 2009). It also contains AtCPSF73-II and AtFY, which are plant specific proteins involved in plant female development and flowering (Millevoi and

Vagner, 2009). The *Arabidopsis* CstF (AtCstF) contains homologs of the human CstF64 and CstF77 subunits, but lacks a functional CstF50 homologue (Yao et al, 2002). AtCstF is linked to the AtCPSF through its interaction with AtFip and AtPAP, which is the enzyme responsible for polyadenylation. The rest of the factors include AtCLPS3, AtPCFS4, and AtSYM5, which is a homolog of symplekin and Pta1 (Zhao et al, 2009) (Fig 1.3).

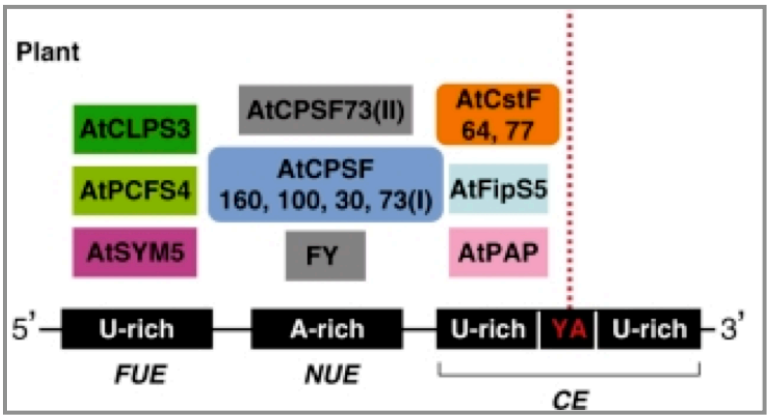


Figure 1.3: The Plant cleavage and polyadenylation machinery (Adapted from Millevoi and Vagner, 2009).

D. The functions of the poly(A) tail in eukaryotes

Poly(A) tails are synthesized in the nucleus of all eukaryotic cells, and range in length from 50-90 nucleotides in yeast to 250 nucleotides in mammals (Patel and Butler, 1992). The length of the poly(A) tract added to the messengers is tightly regulated by the poly(A) binding proteins Pab1 and Nab2 (Kessler et al, 1997; Minvielle-Sebastia et al, 1997). The roles that poly(A) tails play in eukaryotic cells are described below.

i. Poly(A) tails are involved in mRNA export

The poly(A) tails are also involved in the export of mRNA transcripts from the nucleus to the cytoplasm (Huang and Carmichael, 1996; Long et al, 1995). After complete maturation, mRNA molecules are translocated into the cytoplasm through the nuclear pore complex to serve as templates for translation into proteins. Pab1, which coats the nascent poly(A) tails, has been shown to interact with specific nucleoporins and the nuclear export signal export receptor Xpo1 (Allen et al, 2001; Hammell et al, 2002). Nab2 also plays a major role in the process through the interaction with Mlp1 (Fasken et al, 2008). Immuno-fluorescence studies have shown that Pab1 is found in both the nucleus and the cytoplasm, which supports the observation that the protein shuttles between both compartments of the cell (Anderson et al, 1993). The nuclear accumulation of abnormally long mRNA observed in *nab2* and *pab1* mutants constitute a form of quality control in the cell, which retains incorrectly expressed mRNAs in the nucleus, and helps the mature mRNAs to reach the cytoplasm (Hammell et al, 2002; Hilleren et al, 2001, Hector et al, 2002).

ii. Poly(A) tails are involved in RNA quality control

The involvement of the poly(A) tail in mRNA quality control is also due to the actions of Trf4, which is another type of nuclear poly(A) polymerase in yeast belonging to the polymerase β-type nucleotidyltransferase superfamily (Aravind and Koonin, 1999). Trf4 exhibits a slow distributive mode of poly(A) tail addition, which is quite different from the fast processive activity of Pap1, and is conserved throughout eukaryotes (LaCava et al, 2005; Vanacova et al, 2005). Genetic and biochemical studies of Trf4 demonstrated a new function for poly(A) tails in the degradation of aberrant RNAs (tRNAs, rRNAs, and

snRNAs). Trf4 stably associates with an RNA helicase, Mtr4, and with one of two RNA-binding proteins, Air1 or Air2, to form the Trf4/Air1/Mtr4 polyadenylation complex (TRAMP). The TRAMP complex has polymerase activity, and can polyadenylate small RNAs *in vitro*, such as tRNAs. It can also recognize structural features induced by misfolded RNAs, which are recruited to the complex by interacting with Air2p. The slow addition of the poly(A) tails to these misfolded RNAs by Trf4 has been shown to stimulate degradation by the exosome, a 3'-5' exonuclease complex that degrades numerous RNAs (Mitchell et al, 1997), revealing a novel involvement of the poly(A) tails in RNA quality control (LaCava et al, 2005; Vanacova et al, 2005). This is different from Pap1 activity, where access of the exosome to the RNA is prevented by the processivity of the enzyme, and the addition of poly(A) binding proteins. Poly(A) tails are also important for mRNA stability by preventing degradation of the mRNA by exonucleases (Beelman and Parker, 1995; Caponigro and Parker, 1996; Sachs and Wahle,1993). The process of mRNA decay can be initiated by three distinct events - the removal of the 5' cap, endonucleolytic cleavage, and poly(A) shortening (Jacobson and Peltz, 1996). Poly(A) tail shortening in yeast is followed by decapping of the transcript by the Dcp1-Dcp2 complex, and subsequent 5' to 3' exonucleolytic degradation by Xrn1 or the 3'-5' action of the exosome (Decker and Parker, 2002).

iii. Poly(A) tails stimulate translation initiation

One of the major roles of the poly(A) tail in eukaryotic cells is to stimulate the initiation of translation of pre-mRNA transcripts (Iizuka et al, 1994; Jacobson and Favreau, 1983; Sachs and Davis, 1989). This is mediated by Pab1, which acts synergistically with the 5' end cap structure to enhance the binding of the 40S small ribosomal subunit to the mRNA (Tarun and Sachs, 1995; Tarun and Sachs, 1996). The

recruitment of the 40S small ribosomal subunit to the mRNA is mediated by a physical interaction between Pab1 and the translation initiation factor eIF4F. This interaction with Pab1 occurs through a contact to the eIF4G component of the eIF4F complex and results in the circularization of the mRNA molecule (Tarun and Sachs, 1996; Tarun et al, 1997). These combined cooperative interactions enhance the affinity of eIF4F for the 5' cap of the mRNA (Borman et al., 2000; Haghighat and Sonenberg, 1997; Luo and Goss, 2001) and the RNA-binding capacity of Pab1 (Le et al, 1997; Munroe and Jacobson, 1990a; Munroe and Jacobson, 1990b). Translation initiation enhancement by poly(A) tails is also found in higher eukaryotes (Munroe and Jacobson, 1990a). Poly(A) tails also have a role in development by regulating the expression of certain gene products, and poly(A) tail shortening or extension was shown to regulate the translation level of some mRNAs (Curtis et al, 1995). In summary, the roles of poly(A) tails illustrate the interdependence of the different steps in gene expression.

E. Regulation of Polyadenylation

i. Regulation through direct interaction with PAP

There are different regulatory mechanisms that control polyadenylation in the eukaryotic cell. In many genes, regulation is due to alternative polyadenylation that might result from a tandem array of poly(A) signals within a single 3'-untranslated region (UTR) or alternative 3'-terminal exons, the first of which can be skipped in some cells (Edwalds-Gilbert et al, 1995, Zhao et al, 1999). For other genes however, regulation involves post-translational modification of components of the polyadenylation machinery or interactions between the polyadenylation factors and other proteins that can inhibit polyadenylation activity. A good example of such an outside factor is U1A, a small

nuclear RNP (snRNP) A protein that is part of the cellular splicing machinery (Gunderson et al, 1994; 1997).

Expression of the U1A mRNA is negatively auto-regulated when two molecules of U1A protein bind to a polyadenylation inhibitory element RNA (PIE RNA) on the U1A pre-mRNA, resulting in an inhibition of polyadenylation (Boelens et al. 1993). U1A is thus able to limit its own production when there is excess U1A in the cell. Studies on U1A activity have shown that cleavage was not affected by bound U1A, but poly(A) addition was strongly inhibited, even in non-specific polyadenylation assays (Gunderson et al, 1994; 1997; Colgan and Manley, 1997). U1A interacts directly with PAP through a 14 amino acid residue domain, the PAP regulatory domain (PRD) in U1A that is necessary and sufficient for PAP interaction and inhibition (Gunderson et al, 1997; Klein Gunnewiek et al, 2000)**.** The U1A interaction site on PAP is a 20 amino acid region at the C-terminal domain of the protein. *In vitro* studies have shown that this region is sufficient for inhibition of PAP activity, as it can also confer U1A-mediated inhibition onto the yeast enzyme (Gunderson et al, 1997). U1A also uses a similar mechanism to selectively control post-cleavage expression of μ secretory mRNA in B-cells (Philips et al, 2001).

U1-70K is very similar to U1A, and is another protein that can regulate polyadenylation by interacting directly with PAP (Gunderson et al, 1998). It is a component of the U1 snRNP that has been implicated in the control of late gene expression of bovine papillomavirus (BPV) in cattle (Ko and Gunderson, 2002). *In vitro* studies have also shown that just like U1A, U1-70K contains a regulatory domain within the protein that can directly bind to PAP, and inhibit polyadenylation activity (Gunderson et al, 1998). Analysis of the PAP inhibitory domains of both proteins have also revealed a consensus PRD sequence, which has been found in a number of proteins, and the PRD has been shown in *in vitro* studies to be sufficient for PAP inhibition (Gunderson et al, 1998; Ko and Gunderson, 2002). Human splicing proteins like U2AF65 and U2B, a

U2 snRNP-associated protein, contain these PRD sequences. They are also found in a small subset of SR proteins, which are splicing regulatory proteins, containing stretches of alternating serine and arginine residues. They function both in spliceosome assembly (Zahler et al, 1992) and in controlling alternative splice site selection (Caceres et al, 1994; Manley and Tacke, 1996; Graveley, 2000). The U2AF65 PRD physically and specifically interacts with PAP, thereby coupling 3′ end processing and splicing in HeLa nuclear extract, while SRP20 modulates the activity of an upstream polyadenylation site in the calcitonin gene (Lou et al, 1998; Vagner et al, 2000).

Another protein that plays a role in the regulation of polyadenylation activity is PABPN1, which greatly stimulates the activity of PAP by binding PAP and increasing PAP affinity for the RNA as described above (Kerwitz et al, 2003). It can also terminate poly(A) tail elongation through a PABPN1-dependent remodeling of the polyadenylation complex that blocks PAP stimulation, and depends on the complete coating of the poly(A) tail by PABPN1 (Kerwitz et al, 2003).

HnRNP H also plays a role in the regulation of PAP activity by stimulating PAP activity through direct interaction with PAP (Millevoi et al, 2009). This is through the action of the polypyrimidine tract-binding protein (PTB), which is a splicing regulator that plays a positive role in pre-mRNA 3′ end processing when bound upstream of the polyadenylation signal. Binding of PTB upstream of poly(A) sites directly recruits hnRNP H to G-rich sequences associated with several poly(A) signals, with hnRNP H subsequently interacting directly with PAP to stimulate polyadenylation (Millevoi et al, 2009).

ii. Regulation through post-translation modification of PAP

Post-translation modification by phosphorylation, acetylation, ubiquitination and sumoylation also play a major role in the regulation of PAP activity (Shimazu, et al, 2007). Mammalian poly(A) polymerase (PAP) is inactivated during the cell cycle by hyper-phosphorylation, which occurs specifically during mitosis (M) at the cyclin-dependent kinase consensus sites located in the Ser/Thr-rich region of PAP (Colgan et al, 1996). The down-regulation of PAP might be used in the cell to prevent polyadenylation of defective pre-mRNA transcripts, and might affect all eukaryotic polyadenylation. It also contributes to the silencing of gene expression during mitosis in the cell (Colgan et al, 1996; Colgan and Manley, 1997). The phosphorylation of PAP also facilitates interaction with 14-3-3ε, which is a protein factor that affects both PAP localization and activity (Kim et al, 2003). In yeast, Pap1 phosphorylation begins at S phase and persists through G_2, but is not observed during nuclear division (Mizrahi and Moore, 2000). Phosphorylation of yeast Pap1 does not require the activity of the Cdc28 kinase, unlike mammalian PAP, which is phosphorylated by the Cdc28 homologue (Colgan et al, 1996). Even though Cdc28 activity is not directly required for Pap1 phosphorylation, it may be involved in the correct timing of the modification (Mizrahi and Moore, 2000). The difference in the timing of the modification notwithstanding, phosphorylation plays a major role in the cell cycle regulation of mRNA polyadenylation in the yeast and mammalian systems.

Ubiquitination may also play a role in the regulation of Pap1 action during the cell cycle. Two-hybrid analysis shows that Pap1 interacts with Ufd1 (ubiquitin fusion degradation), a protein that functions at a post-ubiquitination step in the yeast UFD pathway (del Ohmo et al, 1997; Johnson et al, 1995). Just like with phosphorylation, the ubiquitination of yeast Pap1 also occurs during the late S and G_2 phases of the cell

cycle, suggesting tight regulation of a subset of yeast mRNAs during that phase of the cell cycle (Mizrahi and Moore, 2000). Only a small amount of Pap1 is ubiquitinated in the cell, and this modification does not alter the stability of the bulk of cellular Pap1 (Mizrahi and Moore, 2000).

Another post-translational modification that regulates PAP activity is sumoylation (SUMO). Two of the six sumoylation sites on mammalian PAP overlap one of two nuclear localization signals (NLS) on the protein. Mutation of this NLS or the expression of a SUMO protease with wild-type PAP, caused localization of PAP to the cytoplasm, which suggests that sumoylation facilitates PAP nuclear localization (Vethantham et al, 2007). Furthermore, down-regulation of sumoylation led to decreases in PAP levels, suggesting that sumoylation helps to stabilize PAP. Conversely, the activity of purified PAP was shown to be inhibited by *in vitro* sumoylation, which suggests that sumoylation regulates multiple functions of mammalian PAP (Vethantham et al, 2007). Two-hybrid analysis shows that Pap1 interacts with Uba2, a homolog of ubiquitin-activating (E1) enzyme involved in the addition of SUMO, but a regulatory function for SUMO in yeast Pap1 has not been defined.

Finally, PAP is also acetylated at its C-terminal region, which is a modification that does not affect its polyadenylation activity. Acetylation of PAP inhibits its binding with CF I_m25, which is also acetylated via an interaction between the 68-kDa subunit of CF I_m and the CREB-binding protein (CBP) acetytransferase (Vethantham et al, 2007). Acetylation of PAP also inhibits its nuclear localization, by blocking its binding to the importin α/β complex.

iii. Regulation by competition with exonucleases

Cytoplasmic polyadenylation is used to activate the translation of dormant

maternal RNA transcripts in oocytes, and requires the presence of two *cis*-acting elements in the 3' untranslated region (UTR) of mRNAs, which include a nuclear polyadenylation signal and a UA-rich cytoplasmic polyadenylation element (CPE), which interacts with the CPE binding protein (CPEB) (Fox et al, 1989; McGrew and Richter, 1990; Hake and Richter, 1994). After transport into the cytoplasm, some mRNAs synthesized in oocytes are deadenylated and stored as dormant messages without translation (Huarte et al, 1992). These dormant mRNA contain short poly(A) tails, which are translationally repressed, and are elongated to 150 nucleotides or greater following induction of oocyte maturation (Simon and Richter, 1994). Repression of polyadenylation is due a large CPE masking element, which spans more than 400 nucleotides and contains dodecauridine as an integral part of the element. Activation of cytoplasmic polyadenylation is however dependent on dodecauridine and the hexanucleotide AAUAAA, indicating that CPE plays a dual role in repression and activation of poly(A) elongation (Simon et al, 1992). The poly(A) tails are shortened in the cytoplasm through a mechanism controlled by CPEB and PARN, a poly(A)-specific ribonuclease (Hake and Richter, 1994; Kim and Richter, 2006). Symplekin mediates the interaction between the CPEB, CPSF and Gld2, which is a poly(A) polymerase that is part of the CPEB-containing complex. The CPEB-containing complex also contains PARN (Kim and Richter, 2006). PARN is more active than Gld2, which results in short poly(A) tails in the cytoplasm, but when oocytes mature, CPEB phosphorylation causes PARN to be expelled from the ribonucleoprotein complex, which allows Gld2 to elongate poly(A) by default (Kim and Richter, 2006).

During DNA damage in the cell, mRNA 3' end cleavage is strongly but transiently inhibited through a mechanism mediated by an interaction between CstF-50 and BRCA1-associated protein BARD1 (Kleiman and Manley, 2001; Mirkin et al, 2008). CstF-50 provides a link between transcription-coupled RNA processing and DNA repair

in the cell, and also interacts with PARN in the extracts of UV-exposed cells. The formation of the CstF-50/PARN complex is important for the inhibition of 3′ cleavage and activation of deadenylation upon DNA damage (Mirkin et al, 2008), and requires the tumor suppressor BARD1. PARN activity in the cell is inhibited by cap-binding protein-80 (CBP80), and this inhibition is reversed by the CstF-50/BARD complex (Cevher et al, 2010). Regulation of polyadenylation/deadenylation in the nucleus by the CstF-50/PARN complex plays a role in the regulation of gene expression during DNA damage, and helps prevent the expression of defective mRNAs in the cell (Cevher et al, 2010).

iv. Regulation due to defective mRNP assembly

Throughout their lifetime, mRNAs are associated with protein factors, which can either be stably bound or subject to dynamic exchange. These proteins factors along with their associated mRNAs are called mRNA ribonucleoprotein particles or mRNPs (Moore, 2005: Shyu et al, 2008). They may serve as adaptors that allow mRNAs to interface with the machinery mediating their subcellular localization, translation, and decay. They also play important roles in the stability of the mRNA and possibly in the nucleocytoplasmic transport of mRNA. They are made up of a wide range of protein factors that bind RNA and include the nuclear poly(A) binding protein in mammalian cells, and Nab2 in yeast (Moore, 2005; Shyu et al, 2008). Studies using the mutants of yeast THO complex and its associated RNA helicase Sub2 show premature transcription termination and poorly polyadenylated and unstable pre-mRNA due to defective assembly of the mRNPs (Saguez et al, 2008). The regulation of polyadenylation observed in these mutants is due to the specific down-regulation of Fip1, a co-factor of Pap1, by the nuclear exosome and its activator, Trf4 (Saguez et al, 2008). The down-

regulation of polyadenylation in these mutants is used to slow the production of aberrant mRNA due to defective mRNP assembly.

In conclusion, there are many diverse mechanisms used by the cell to regulate cleavage and polyadenylation in the eukaryotic cell. The different cellular mechanisms that regulate polyadenylation include direct interaction with PAP by proteins like U1A and U1-70K. It also includes post-translational modification of PAP by phosphorylation, ubiquitination, sumoylation or acetylation or by the down-regulation of CPF factors like Fip1 due defective mRNP formation. Regulation can also be due to competition with exonucleases like the CstF-50/PARN complex, which inhibits cleavage during periods of DNA damage in the cell or masking elements like the UA-rich CPE, which repress dormant mRNA in oocytes. The use of alternative poly(A) sites by the cell also plays a critical role in the regulation of polyadenylation in the cell. The regulation of polyadenylation is therefore of critical importance to cell, and enables the cell to stop mRNA production due to defects in the transcription machinery or during stress situations like cellular DNA damage.

F. Specific Aims of this thesis

In recent years, great progress has been made in the study of the factors involved in 3' end processing, their biochemistry and their connection with other aspects of mRNA synthesis. Despite all of these efforts, however, some of the molecular mechanisms that regulate 3' end processing in eukaryotic cells remain unclear. The goal of my thesis is study the function and regulation of yeast poly(A) polymerase (Pap1) as part of the CPF complex during 3' end processing. For polyadenylation to take place in the yeast or mammalian cell, Pap1 must be incorporated into a competent 3' end-

processing complex so that it can act on the 3' ends of cleaved pre-mRNAs before they are degraded by exonucleases. Pap1 also needs to access the new 3' end of the pre-mRNA without being impeded by cleavage factors, and must initiate and processively elongate the poly(A) tail, and terminate its activity once the poly(A) tail reaches the correct size. Finally, under special situations in the cell, PAP activity needs to be inhibited even though appropriate cleavage has occurred. To understand the control of the poly(A) addition step, I have used yeast as a model eukaryote because of the ease of biochemical and genetic studies in this organism.

The first aim of my thesis focuses on the recruitment of Pap1 into the CPF complex through its interaction with Fip1, and on the properties of Fip1 that facilitate this process. The goal of this aim is to better understand how Pap1 gains access to the 3' end of the pre-mRNA through its interaction with Fip1. The second aim focuses on the mechanisms that regulate the actions of Pap1 during 3' end processing. Studies on Pap1 activity have shown that it changes from a distributive to processive enzyme when incorporated into the CPF complex. The goal of this study is to identify the factors within the CPF complex that might interact with Pap1, and to characterize how these interactions regulate Pap1 activity during 3' end processing.

Chapter 2

A flexible linker region in Fip1 is needed for efficient mRNA polyadenylation

INTRODUCTION

Factor interacting with poly(A) polymerase (Fip1) is a 36-kDa protein, which was first identified in yeast through its interaction with poly(A) polymerase (Pap1) (Preker et al, 1995). It is part of the Cleavage/Polyadenylation Factor (CPF), which along with the Cleavage Factor I (CFI) make up the 3' end-processing complex in yeast (Mandel et al, 2008; Millevoi and Vagner, 2009) (Fig 1.2). Fip1 interacts with CPF subunits Pap1 and Yth1 (Fig 2.1), and tandem affinity tags placed on Fip1 have been shown to pull down the CPF complex (Preker, 1995; Zhelkovsky et al, 1998; Helmling et al, 2001; Dheur et al, 2005). Deletion of the *FIP1* gene is lethal in yeast, but yeast viability is not affected by C-terminal deletions up to amino acid 220 (Preker et al, 1995, Helmling et al, 2001). Deletions beyond this point result in temperature sensitivity and defects in polyadenylation (Helmling et al, 2001). Yeast Fip1 contains a serine-rich acidic domain at the N-terminus, and a C-terminal stretch of 37 amino acids containing 16 proline residues (Fig 2.1) (Preker et al, 1995; Zhao et al, 1999). This domain organization is shared in Fip1 proteins across species from yeast to mammals (48% identity, 72% similarity over 60 amino acids) (Kaufmann et al, 2004; Mandel et al, 2008). The human homolog of Fip1 (hFip1) is a 66-kDa protein, which contains a charged RNA-binding domain at the C-terminus. This domain, which is absent in yeast, is made up of a mixed arginine-aspartate (RD) domain and an arginine-rich region (Fig 2.1) (Kaufmann et al, 2004).

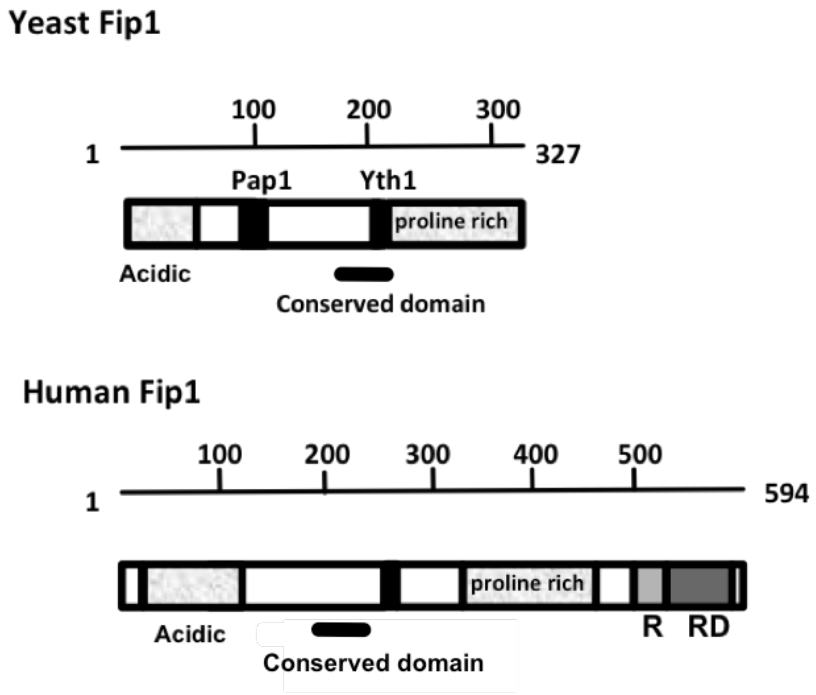


Figure 2.1: Schematic diagram of yeast and human Fip1 showing the acidic, conserved and proline rich domains of both proteins. The Pap1 and Yth1 interaction domains are shown on yeast Fip1. Also shown are the charged arginine rich (R) and the mixed charged domain (RD) at the C-terminus of human Fip1.

The acidic N-terminus of hFip1 (residues 1-111) mediates the interaction with human poly(A) polymerase (PAP), and the conserved segment (residues 137-243) interacts with CPSF 30, which is a homolog of yeast Yth1 (Kaufman et al, 2008). Purification of HeLa cell extracts also show that hFip1 is found in a stable ternary that includes CPSF 160 and PAP (Kaufmann et al, 2004, Mandel et al, 2008). In yeast, Pap1 is linked to the CPF through a direct interaction with Fip1 (Preker et al, 1995). Fip1 in turn is held in the 3' end complex through its interaction with Yth1 (Barabino et al, 2000). Analysis of the region of yeast Fip1 that interacts with Yth1 shows that the interaction domain of Fip1 is localized to residues 206-220, which is part of the conserved domain of Fip1 (Helmling et al, 2001; Mandel et al, 2008).

Isolated Pap1 is a distributive enzyme that releases its RNA substrate after each round of catalysis, but after recruitment by Fip1 into CPF, Pap1 behaves processively and selectively elongates pre-mRNAs that contain mRNA processing signal sequences (Preker et al, 1997). Little is known about how Pap1 gains access to the 3' end of the pre-mRNA immediately after cleavage, and the role that Fip1 plays in this process. While the structure of Pap1 and the Pap1-interaction domain of Fip1 have been determined (Meinke et al, 2008), the conformation of the rest of Fip1 when it is bound to Pap1 as part of CPF is unknown.

The goal of the current study is to better understand how Pap1 gains access to the 3' end of pre-mRNA through its interaction with Fip1 and why it changes from a distributive to processive enzyme when incorporated into the CPF. We also studied the role of Fip1 in 3' end processing through its interaction with Pap1 and how the unstructured nature of Fip1 is important for Fip1 function during 3' end processing. We report here that polyadenylation is much more efficient when Pap1 is artificially tethered to the RNA substrate. We also show that a fusion protein of Pap1 and Fip1 can rescue loss of both proteins in yeast, and finally, we show that Fip1 contains a flexible linker that bridges the Pap1 and the Yth1 interaction sites of the protein.

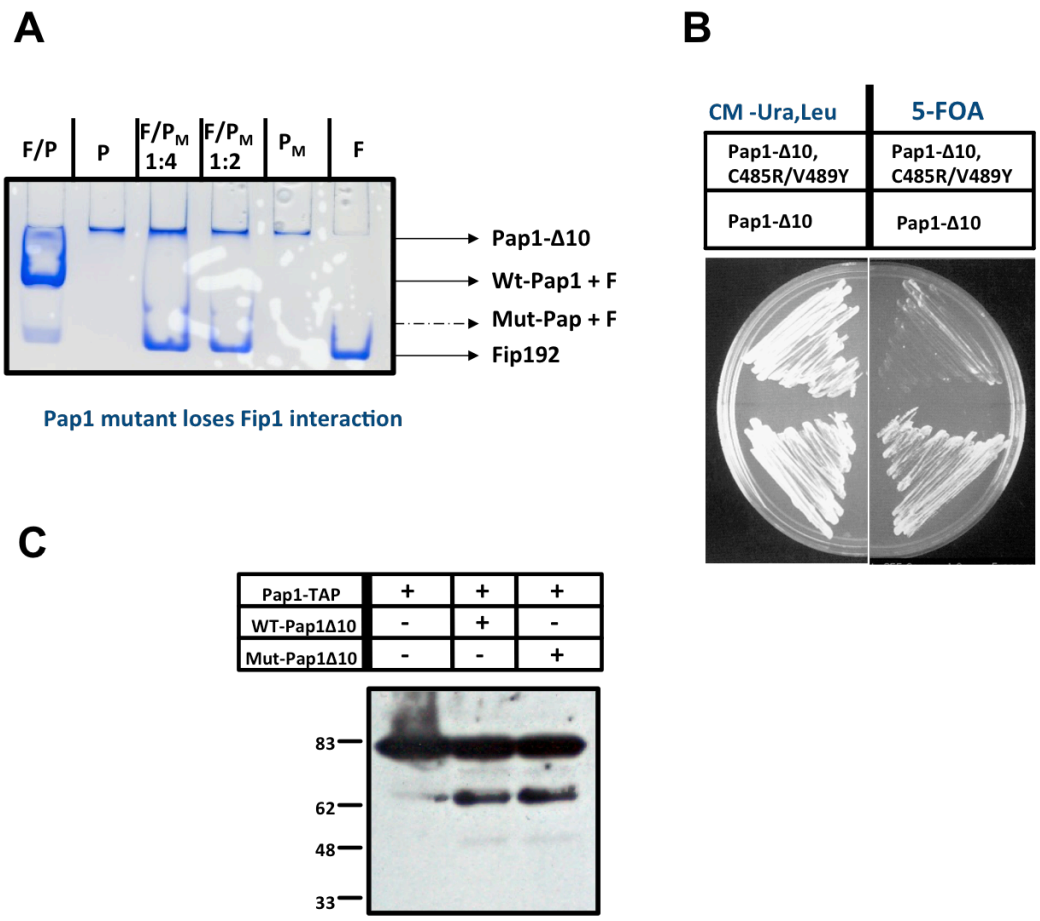


Figure 2.2: Pap1 interacts with amino acids residues 81-97 of Fip1 (A) C485R/V489Y Pap1 loses Fip1 interaction (courtesy of Gretchen Meinke). Wild type Pap1 is labeled P, Fip1 is labeled F, and C485R/V489Y Pap1 is labeled P_M. To form the complexes, 2.5 µg of bacterially expressed, purified Fip192 was combined with wild-type and mutant Pap1 as indicated. Samples were run on a 10% polyacrylamide gel. Pap1Δ10 is yeast Pap1 with a 32 amino acid truncation at the C-terminal end and contains a C-terminal His_6 tag for purification purposes. Deletion of the C-terminus and addition of the His_6 tag does not compromise enzymatic activity of Pap1. (B) Mutations in the Fip1 interaction domain of Pap1 are lethal in yeast. Yeast containing a *PAP1* chromosomal deletion and the *PAP1* gene on a *URA3* plasmid were transformed with a *LEU2*+ plasmid containing the *PAP1-Δ10* or *pap1-Δ10, C485R/V489Y* allele. Transformants were grown on complete media minus leucine and uracil (CM − Leu, Ura) or on media containing 5-fluoroorotic acid (5FOA), which forces loss of the *URA3* plasmid. Growth on 5FOA is observed only with cells containing *PAP1-Δ10*. (c) The Pap1-Δ10, C485R/V489Y mutant protein is expressed in yeast at the same level as wild-type Pap1. Western blot was performed on extracts from cells expressing only a TAP-tagged Pap1 or the TAP-tagged Pap1 and either pap1-Δ10, C485R/V489Y or Pap1-Δ10 using a Pap1-specific antibody. The upper bands are the TAP-tagged Pap1. The lower ones are the native and mutant proteins without the TAP tag.

RESULTS

Tethering of Pap1 to Fip1 is essential for viability.

Based on the structure of Pap1 and the Pap1-interaction domain of Fip1, we showed that residues 81-97 of Fip1 are needed for interaction with Pap1 in yeast (Fig 2.2A) (Meinke et al, 2008). As part of this study, I showed that mutations in the Fip1 interaction domain of Pap1 are also lethal in yeast (Fig 2.2B), even though the mutant Pap1 had normal activity on oligo(A) substrate, and are expressed normally in the protein (Fig 2.2C) (Meinke et al, 2008). We also found that Fip1 adopts an extended conformation in solution and maintains this conformation when bound to Pap1, indicating that very little of Fip1 folds into an ordered structure (Meinke et al, 2008). These results suggested that a primary function of Fip1 is to link Pap1 to CPF and thus to the substrate RNA so that Pap1 can act processively.

Tethering of RNA to Pap1 increases the efficiency of polyadenylation.

To further study the role of the CPF complex in holding Pap1 on the RNA during polyadenylation, we generated a simplified model system where the RNA substrate is tethered directly to Pap1. The model is comprised of a bacteriophage λ N protein fused to the C terminus of Pap1 (Fig 2.3A). Tethering models of this sort have been successfully used in the study of other mRNA-protein interactions (Coller and Wickens, 2007).

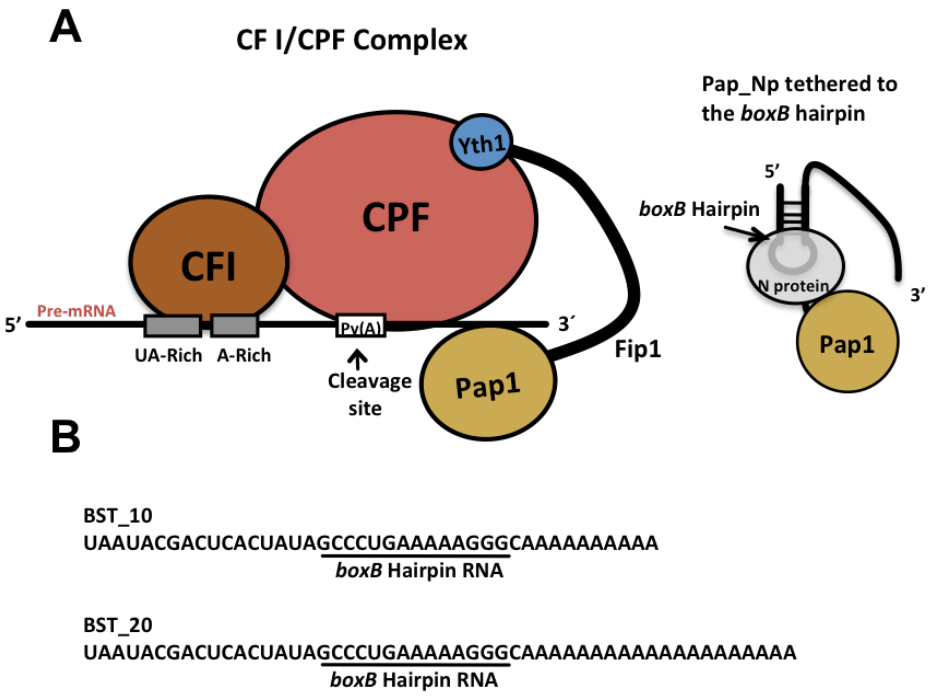


Figure 2.3: A simplified model system to study Pap1 processivity. (A) The yeast 3' end-processing complex showing the Fip1 as a flexible protein and the tethered Pap-N model. (B) RNA sequences of the substrates used in poly(A) addition reactions. The *box B* hairpin sequence is underlined.

The bacteriophage λ N protein has an arginine-rich motif that binds to the *box B* hairpin RNA sequence (Scharpf et al., 2000). The Pap1-bacteriophage λ N fusion protein (Pap_Np) exhibited polyadenylation activity comparable to the untagged enzyme on an oligo A_{18} substrate (data from Rosanna Lee and Paul Balbo), indicating that the N-protein tag did not compromise enzyme function. To determine the effect of tethering on Pap1 activity, the rate of polyadenylation was measured by the incorporation of radioactive ATP onto an RNA substrate, BST_20, which contains a *box B* hairpin sequence inserted directly upstream of a 20 adenosine tract (Fig 2.3B). Pap_Np added adenosines to this substrate at a rate that was 20-fold greater than that of Pap1 (Fig

2.4A). This is consistent with Pap_Np having a higher affinity for the *box B* hairpin loop on the BST_20 RNA substrate.

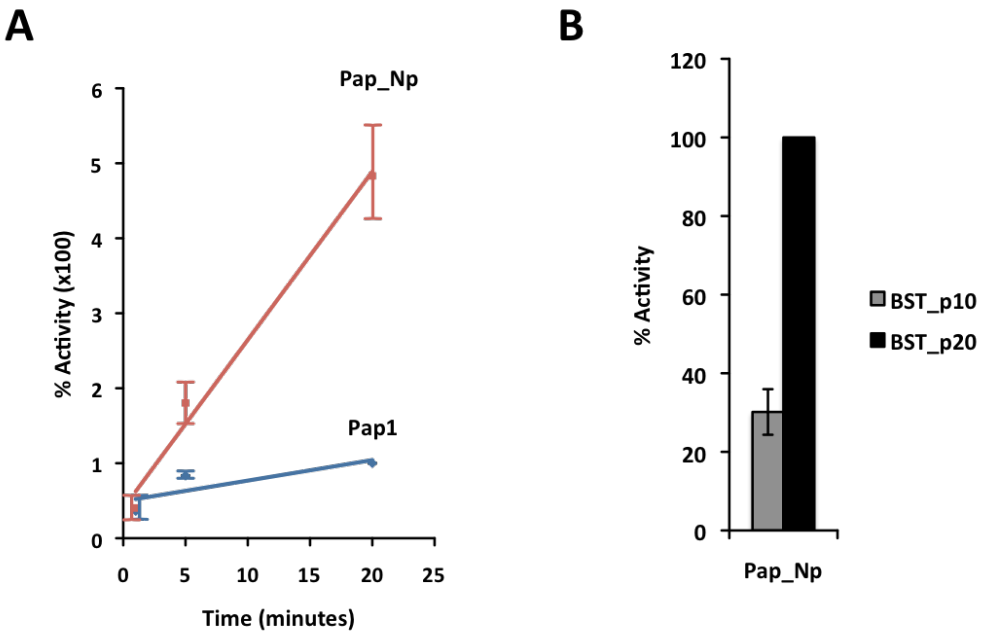


Figure 2.4: Activity of the Pap_Np fusion protein (A) The rate of polyadenylation using BST_p20 RNA as substrate. Polyadenylation activity of Pap1 and Pap_Np was determined by incorporation of [$\propto$-^{32}p] ATP into the growing poly(A) at 30°C, with time points taken at 1, 5 and 20 minutes, followed by measurement of TCA-precipitable counts. The experiments were done in triplicate, and standard error is indicated. 100% Activity is the activity of Pap1 after 20 minutes of incubation at 30°C. (B) Polyadenylation activity of the Pap_Np on *BST_p10* and *BST_p20* RNA substrates performed as described in Fig 1C. 100% Activity is the activity of Pap_Np on the BST_20 substrate after 20 minutes of incubation at 30°C.

When an RNA substrate containing only 10 adenosines beyond the *box B* hairpin (BST_10, Fig. 2.4B), is used, the efficiency of polyadenylation by Pap_Np is greatly reduced. As both substrates contain the *box B* hairpin sequence, they should bind with the same affinity to the enzyme. Thus, the difference observed in the polyadenylation

activity is most likely due to the difference in length between the hairpin sequence and the 3' end of the RNA. To determine if this difference in activity makes physical sense, we built a molecular model of the Pap_Np fusion protein with a bound *box B* hairpin and a 3' poly(A) sequence (Fig 2.5). We placed the existing structure of the RNA-bound N-protein (Scharpf et al., 2000) near the C-terminus of the structure of Pap1 in complex with its RNA substrate (Balbo and Bohm, 2007). In this model, fifteen nucleotides are needed to span the distance between the Pap1 active site and the N-protein domain to which the RNA is bound. Thus, the model helps explain the biochemical observation that BST_20 is a much better substrate for Pap_Np than BST_10, which has only 10 adenosines beyond the binding site.

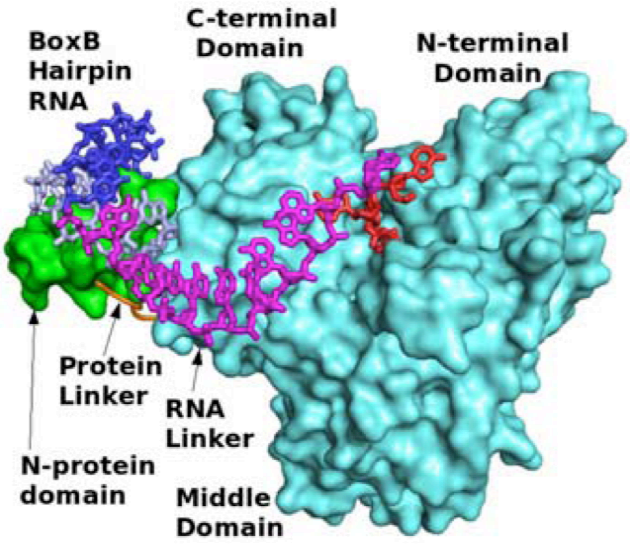


Figure 2.5: Molecular depiction of the Pap_Np-RNA interaction showing the crystal structure of the *boxB RNA*/N-protein complex (blue/green) (Scharpf et al, 2000). This complex was modeled at the C-terminal end of the Pap1 structure (cyan). Five nucleotides were present in the crystal structure of Pap1 with RNA (red) but not all are visible in this orientation (Balbo and Bohm, 2007) The nucleotides colored purple were modeled to span the space between those observed bound to the N-protein and Pap1 (Courtesy of Gretchen Meinke).

The Pap1Δ90Fip1 fusion protein restores wild type activity in yeast strains lacking the PAP1 and FIP1 genes.

To investigate the role of tethering in Pap1 activity in the context of the CPF complex, we created a yeast strain in which Pap1 is fused covalently to a truncated version of Fip1 and expressed as one protein in a yeast strain background where the individual *PAP1* and *FIP1* genes have been disrupted. Fip1 binds to the outer surface of the C-terminal domain of Pap1 through an interaction interface located within amino acids 80 to 105 of Fip1, with amino acid 105 of Fip1 positioned near the top of the Pap1 C-terminal domain (Fig 2.6B) (Meinke et al, 2008). We designed a plasmid encoding a truncated Fip1 protein lacking the N-terminal 90 amino acids (Δ90Fip1), which includes the non-essential first 80 amino acids and part of the essential Pap1 interaction domain (Figure 2.6A) (Helmling et al, 2001). Like the Fip1 truncation lacking the first 105 amino acids, this construct could not rescue a *FIP1* gene deletion (unpublished result from Dr. Zhelkovsky), and confirmed the prediction from the structure that amino acids between positions 80-90 would be crucial for Pap1 interaction. The Δ90Fip1 was fused to the C-terminus of Pap1, which is located near the bottom of the C-terminal domain (Fig 2.6B), and tested for viability in yeast strains, which contained either a disrupted *PAP1* or *FIP1* gene. Both knockouts were rescued by this construct.

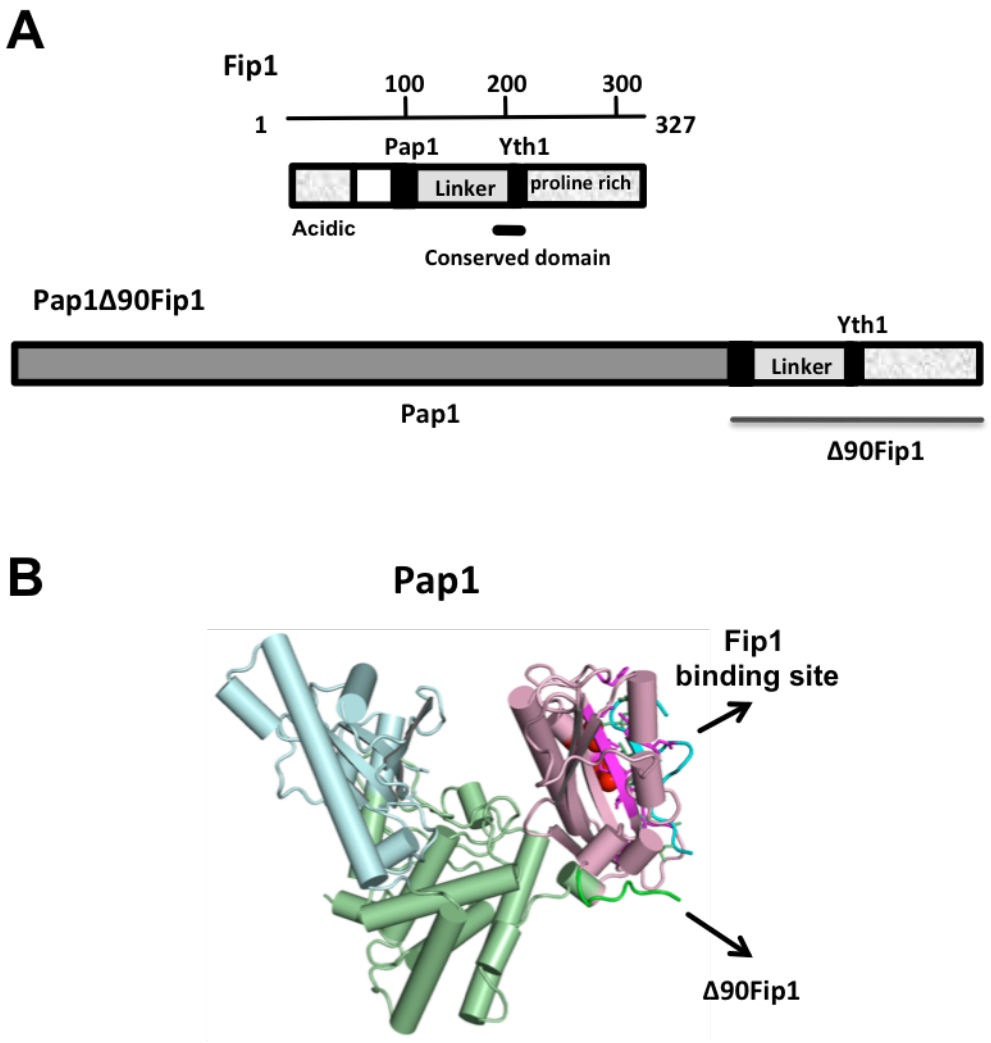


Figure 2.6: (A) Schematic diagram of Fip1 showing the Pap1 and Yth1 interaction domains, the conserved region of the protein and the flexible linker. Also shown is the Pap1Δ90Fip1 fusion protein, with Δ90Fip1 attached covalently at the C terminus of Pap1. (B) Structure of Pap1 adapted from Meinke et al, 2008 showing the Fip1 binding site, and the Δ90Fip1 fusion site at the C-terminus.

To test whether this plasmid can cover the loss of the chromosomal copies of the essential *PAP1* and *FIP1* genes together, yeast strains containing the double knockout of both genes were made by mating individual haploid strains with each single gene

disrupted by a different selectable marker. The resulting diploid cells were transformed with an *URA3*-marked plasmid carrying the *PAP1Δ90FIP1* construct, and then sporulated and tetrads dissected. After sporulation, the viable haploid cells with both the *FIP1* and *PAP1* gene disruptions all contained the *PAP1Δ90FIP1* plasmid. No cells survived the double gene disruption without a plasmid, indicating as expected that a double disruption of *PAP1* and *FIP1* is lethal to yeast. To confirm that the *PAP1Δ90FIP1* construct rescues the double deletion of *PAP1* and *FIP1*, the double knockout cells were tested for growth on medium containing 5-FOA, a reagent that is toxic to yeast cells expressing *URA3* (Fig 2.7). The results show that the yeast strains carrying the *PAP1Δ90FIP1* construct cannot grow on 5-FOA, when loss of the covering plasmid is forced.

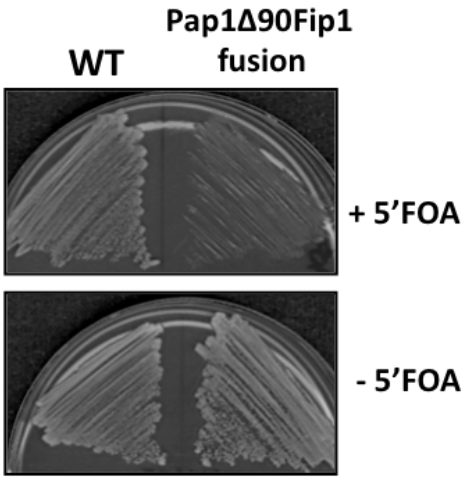


Figure 2.7: Pap1Δ90Fip1 rescues a double deletion of *FIP1* and *PAP1* genes. The yeast strain containing Pap1Δ90Fip1 on a *URA3* plasmid and disruption of the chromosomal *FIP1* and *PAP1* genes was grown on medium in the presence (+) or absence (-) of 5-FOA, which forces the loss of *URA3* plasmid. No growth is observed on 5-FOA in the absence of Pap1Δ90Fip1. As would be expected, a *ura3* strain with intact *FIP1* and *PAP1* genes (WT) grows in both conditions.

To test the phenotype of this yeast strain, a wild type control strain was made by replacing the *PAP1Δ90FIP1* construct with two yeast plasmids carrying copies of *PAP1* and *FIP1*. Western blot analysis with a monoclonal Pap1 antibody (Fig 2.8A) and a polyclonal Fip1 antibody (Fig 2.8B) showed that the individual Pap1 and Fip1 proteins found in wild-type cells had been replaced by the fusion protein in the Pap1Δ90Fip1 strain.

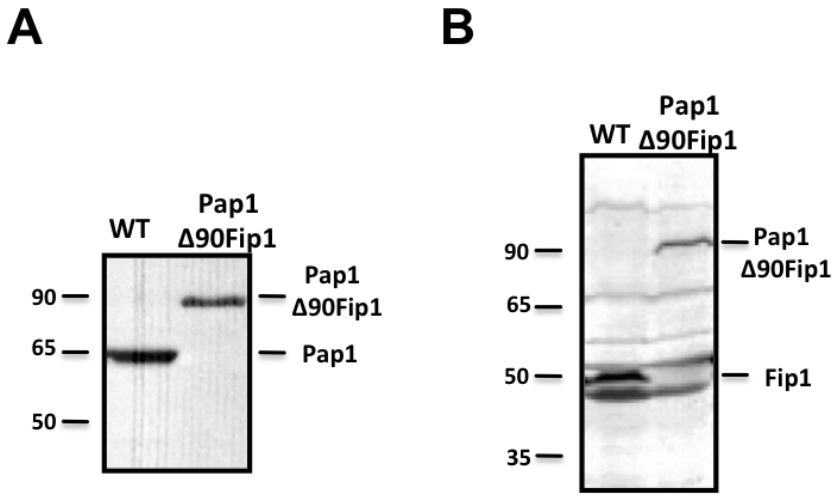


Figure 2.8: Western analysis of Pap1Δ90Fip1 fusion protein (A) The Pap1Δ90Fip1 fusion protein is expressed in the yeast strain. Western blot using Pap1 monoclonal antibody performed on yeast extracts shows the loss of individual Pap1 protein, and the presence of the fusion protein. (B) Western blot using Fip1 polyclonal antibody on yeast extracts shows the loss of individual Fip1 protein and the presence of the fusion protein in Pap1Δ90Fip1 extract.

Thermosensitivity testing showed no growth defects at different temperatures compared to the wild type control (Fig 2.9A), confirming that having Pap1 and Fip1 combined into one polypeptide is functionally equivalent to the individual Pap1 and Fip1 proteins found in wild-type cells. In support of this conclusion, analysis of the *in vitro* processing activity of extract made from cells expressing Pap1Δ90Fip1 showed slightly

reduced polyadenylation efficiency, but normal tail length, when compared to the wild-type (Fig 2.9B). In these assays, radiolabeled *GAL7* RNA substrate ending at the *GAL7* poly(A) site was incubated with extract and ATP. Analysis of total RNA from these yeast strains using a previously described method to specifically visualize poly(A) tails (Preker et al, 1995, Kessler et al, 1997) also showed that poly(A) tails formed *in vivo* were normal in length (Fig 2.9C, courtesy of Dr. Zhelkovsky). From these results, we conclude that the Fip1/Pap1 interaction does not need to be disrupted at any time during 3' end processing, and that the precise point of contact on the Pap1 C-terminal domain is not important.

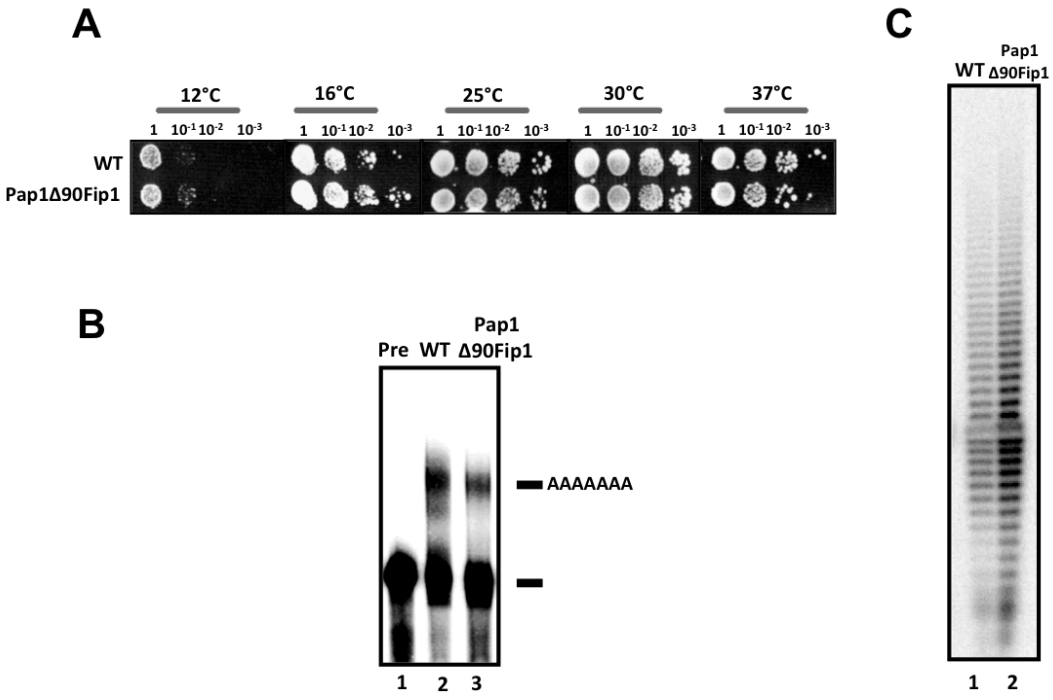


Figure 2.9: (A) Thermosensitivity tests on the wild type and yeast strains containing the Pap1∆90Fip1 fusion protein. Serial dilutions of liquid cultures were spotted on YPD medium and incubated at the indicated temperatures for up to 120 hrs. (B) *In vitro* poly(A) addition assay at 30˚C using pre-cleaved *GAL7* RNA substrate ending at the poly(A) site. Unreacted precursors are shown in the lane marked "Pre", and the position of polyadenylated RNA indicated on the right. (C) Poly(A) tail length in Pap1∆90Fip1 yeast strains were normal when grown at 30°C (Courtesy of Dr. Zhelkovsky).

An intrinsically unstructured linker domain is important for Fip1 function in polyadenylation.

Our previous biochemical data have shown that Fip1 adopts an extended conformation in solution and also maintains this conformation when bound to Pap1 in solution (Meinke et al, 2008). These observations suggest that very little of Fip1, except the small parts directly contacting Pap1 and Yth1, folds into an ordered structure, which is characteristic of intrinsically unstructured proteins (IUPs) (Fink, 2005). IUPs are proteins or regions of proteins that have no compact globular domains. In one class of these proteins, an intrinsically unstructured linker domain (IULD) tethers flanking folded domains and allows the mobility needed for these domains to interact with other proteins and nucleic acids (Jacobs et al, 1999). These proteins show no sequence conservation in the linker, but mobility is conserved across different species (Daughdrill et al., 2007). The resemblance of Fip1 to these proteins can help explain how Pap1 gains access to the 3' end of the pre-mRNA after cleavage.

To test if Fip1 acts as a flexible tether when bound on one end to Pap1 and on the other to Yth1 in the intact CPF complex (Fig. 2.6A), amino acid residues 106 – 190 of Fip1, which is the region between the Pap1 and Yth1 interaction sites, was replaced with a similar sized segment containing the IULD of the human replication protein A (RPA70) (Fig 2.10). Previous studies have used the 72-amino acid IULD of RPA70 as a model for studying intrinsic unstructured proteins (Daughdrill et al., 2007). For this study, we used the ΔNFip1 construct, which lacks the non-essential 80 amino acids at the N terminus, and behaves like full-length Fip1 (Helmling et al., 2001). Extra copies of the RPA70 IULD linker were also introduced to study the effect of increasing the length of this putative flexible linker, and a Fip1 construct lacking the linker region and juxtaposing

the Pap1 and Yth1 interaction sites was made to test the consequence of losing the linker all together (Fig 2.10).

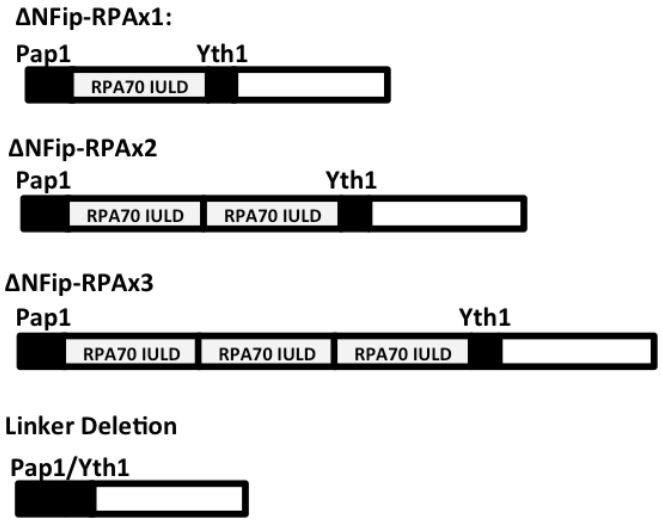


Figure 2.10: Schematic diagram of modified ΔNFip1p showing the substitution of Fip1 sequence between the Pap1 and Yth1 interaction domains with copies of the RPA70 IULD.

All of the ΔNFip1 constructs were viable in the absence of the wild-type Fip1 protein (Fig 2.11). Thermosensitivity tests on the constructs showed normal growth at all temperatures, except 40°C where slower growth is observed in all constructs except wild type Fip1 and ΔNFip1 (Fig 2.11). In the presence of 1.5% formamide, which is thought to further destabilize protein/protein interactions within cells (Hampsey, 1997), thermosensitivity tests showed a growth defect at higher temperatures (37°C and 40°C) in the absence of the Fip1 flexible linker, which persisted if the linker was replaced with one or two copies of the RPA70 IULD (Fig 2.11). The growth defect was however rescued in the presence of three copies of the RPA70 IULD. These results suggest that

the linker region of Fip1 is important for cell growth, but that it cannot simply be replaced by an IULD of similar length.

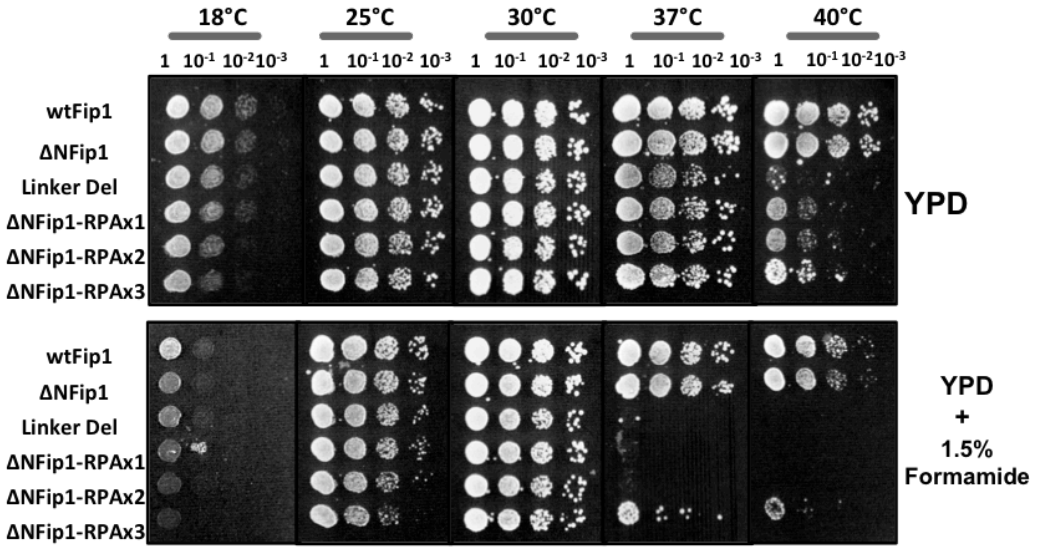


Figure 2.11: Thermosensitivity tests on the wild type Fip1 (wtFip1), ΔNFip1 and the modified ΔNFip1-RPA70 yeast strains on YPD media, and YPD media supplemented with 1.5% formamide, performed as described for Fig 2G. Plasmids carrying the indicated constructs of *FIP1* were introduced into the PJP22 yeast strain by plasmid shuffling (Helmling et al, 2001).

To study the role of the Fip1 linker in polyadenylation, *in vitro* assays were performed using radiolabeled *GAL7* RNA substrates ending at the *GAL7* poly(A) site. Extract from the yeast strain expressing a deletion of the Fip1 linker showed very low polyadenylation activity at 30°C and 37°C, and few of the tails reached the normal length seen in the ΔNFip1 extract (Fig 2.12A, lanes 3 and 8). Low polyadenylation activity is also observed in extract from yeast strains in which the Fip1 linker was replaced with one or two copies of RPA70 (Fig 2.12A, lanes 4-5 and 9-10), while full polyadenylation activity is restored when the linker is substituted with three copies of RPA70 (Fig 2.12A,

lanes 6 and 11). In this case, the size of the poly(A) tail is normal, indicating that the tether is probably not functioning as a ruler for tail length.

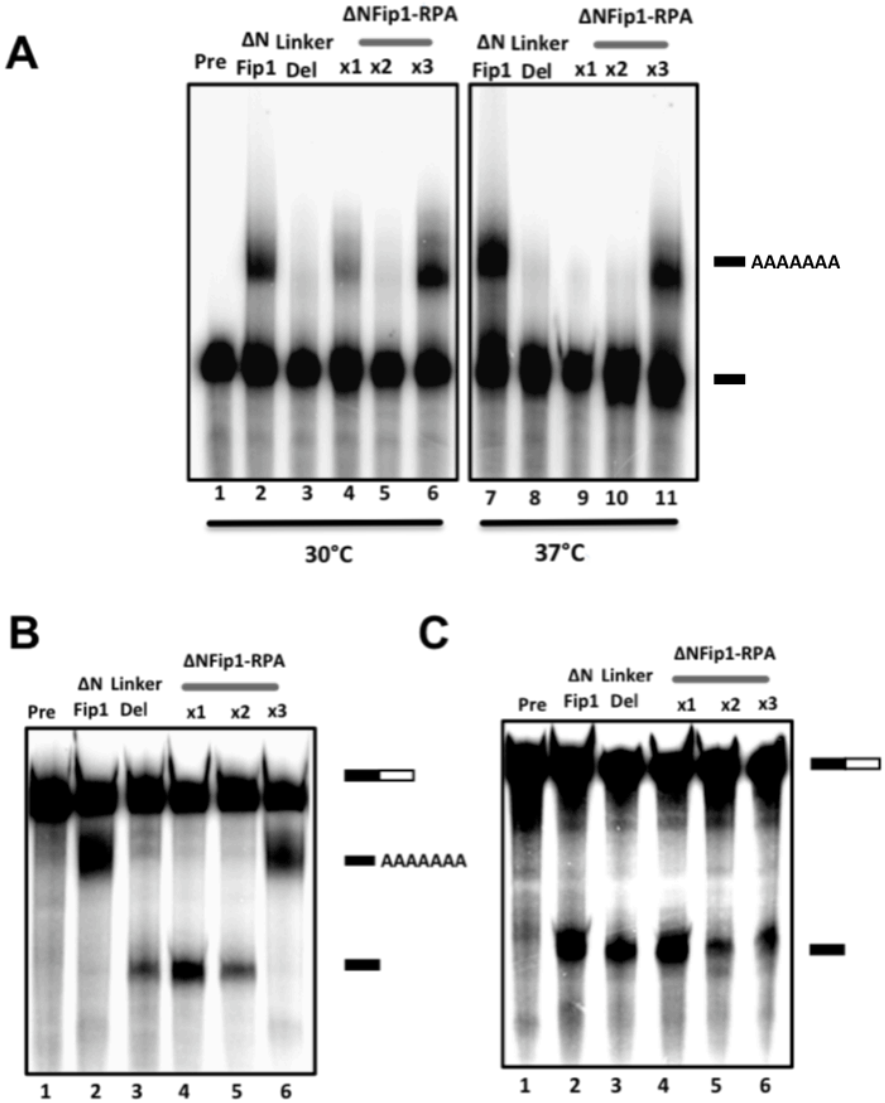


Figure 2.12: *In vitro* analysis of ΔNFip1 and modified ΔNFip1-RPA extracts (A) *In vitro* poly(A) addition assays at 30˚C and 37˚C using pre-cleaved ^{32}P-labeled *GAL7* RNA ending at the poly(A) site and extracts from the indicated strains were performed as described (Helmling et al, 2001). (B) *In vitro* coupled cleavage and poly(A) addition assays at 30˚C using uncleaved ^{32}P-labeled *GAL7* RNA substrate which contains a poly(A) site and flanking sequence and the same extracts used in Panel A. (C) *In vitro* cleavage assays at 30˚C using uncleaved ^{32}P-labeled *GAL7* RNA substrate which contains a poly(A) site and flanking sequence and the same extracts used in Panel C.

An *in vitro* cleavage and polyadenylation assay using uncleaved radiolabeled *GAL7* RNA was used to test the effect of the Fip1 modifications on the cleavage step, and when cleavage and polyadenylation are coupled. A cleaved, unadenylated product accumulates in extracts containing constructs with the linker deletion, or with the one and two RPA70 replacements, while polyadenylation is restored with three RPA70 replacements (Fig 2.12B). Cleavage was also normal in all constructs (Fig 2.12C). This result confirms that the CF I and CPF polyadenylation factors in the extracts are competent for cleavage, that the Fip1 linker is not involved in cleavage, and that modification of the Fip1 linker has inhibited polyadenylation. Analysis of the *in vivo* poly(A) tails made from yeast strains containing the ΔNFip1, ΔNFip1-RPAx1, ΔNFip1-RPAx2, ΔNFip1-RPAx3 and the Linker Deletion forms of Fip1 were analyzed and showed normal poly(A) tail length at 30°C in all the samples (Fig 2.13). This is consistent with the lack of any growth defect at this temperature (Fig 2.11), even though extracts made from cells expressing ΔNFip1-RPAx1, ΔNFip1-RPAx2, and the Linker Deletion forms of Fip1 are defective in *in vitro* assays and at high temperatures in the presence of formamide *in vivo*.

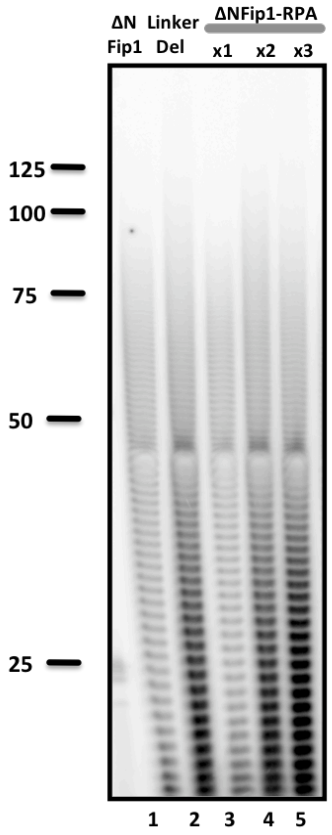


Figure 2.13: Poly(A) tail length in ΔNFip1p and the modified ΔNFip1p-RPA70 yeast strains were normal when grown at 30°C. ΔNFip1 yeast strains grown overnight in YPD were diluted to an OD of 0.1, and grown at 30°C for approximately 6 hrs. Cells were harvested by centrifugation at OD 0.8 and total yeast RNA was isolated from cells using a modification of the hot phenol method (Köhrer and Domdey, 1991). Labeling, RNAse digestion, and visualization of poly(A) tails were performed as described previously (Preker et al, 1995, Kessler et al, 1997).

Western analysis was used to detect the modified forms of Fip1 in the extract, the absence of the normal Fip1 protein, and the presence of Pap1, showing that the low polyadenylation efficiency observed in some of the extracts in the *in vitro* assays was not due to a lack of Pap1 or poor expression of the mutant Fip1 proteins (Fig 2.14A). To confirm that the Fip1 modifications did not affect recruitment of Pap1 into the CPF complex in these samples, immunoprecipitation using anti-Fip1 antibodies showed that the association of Pap1 and other selected subunits of CPF with Fip1 did not change as

a consequence of alterations which caused poor poly(A) addition (Fig 2.14B). This set of experiments shows that the linker region of Fip1 between the Pap1 and Yth1 interaction sites is important for polyadenylation *in vitro*, and that three copies of a flexible RPA 70 IULD can reconstitute the polyadenylation function of this linker.

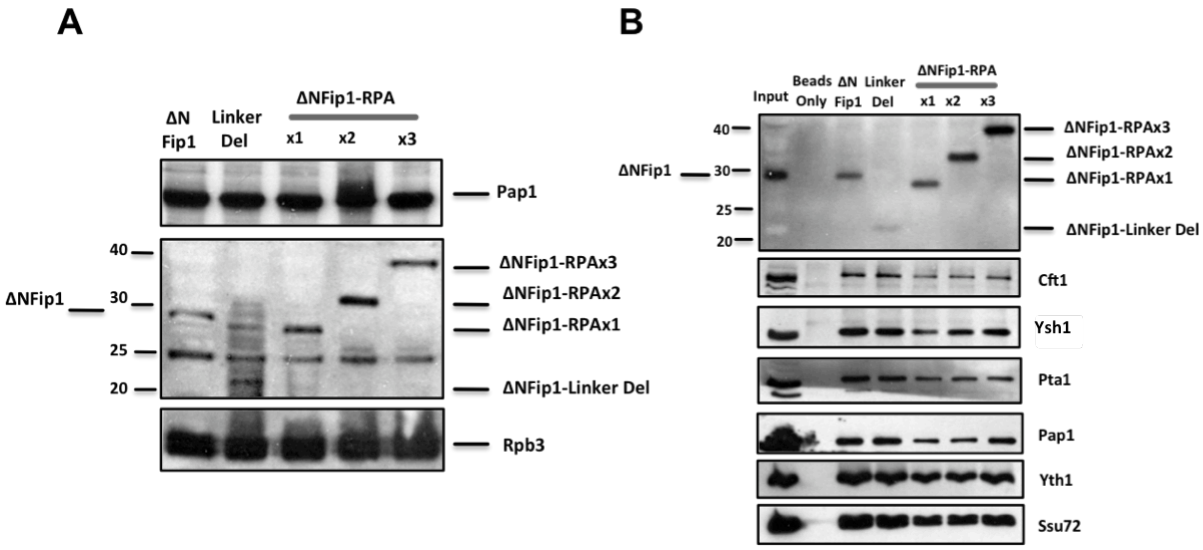


Figure 2.14: (A) Western blot analysis showing the presence of Pap1 and the mutant Fip1 proteins in yeast extract. Rpb3 was used as a loading control. The band at 25kDa is a non-specific protein, which is detected with the polyclonal Fip1 antibody. (B) Immunoprecipitation of Fip1 and associated proteins from extract with anti-Fip1 rabbit polyclonal antibody followed by Western blot detection of Fip1 and the indicated CPF proteins. For immunoprecipitation reactions, 100 µl of protein A-agarose was incubated with extracts (200 µg total protein) from ΔNFip1 yeast strains and incubated overnight at 4°C as described (Helmling et al, 2001). Input contains 20 µg total protein of ΔNFip1 yeast extract.

The function of the Fip1 linker region could not be restored by one or two copies of the RPA IULD, suggesting that this region may have additional roles, such as interaction with other subunits of the processing complex. Since Fip1 has been reported to interact with Rna14 of CF IA and Pta1 of CPF *in vitro* (Preker et al, 1995; Ghazy et al, 2009), pull-down experiments using recombinant proteins were used to test whether

interactions between Pta1 and CF IA were dependent on the flexible linker of Fip1. For these assays, we used two C-terminal truncations of Fip1 – one which included the linker region (Fip1, 1-206) and one which completely lacked it (Fip1, 1-105). The interaction between Pta1 and between CF IA was detected with the longer Fip1 construct, but lost with the 1-105 form (Fig 2.15). This suggests that the linker interacts with other subunits of mRNA 3' end processing complex to appropriately position Pap1 for efficient poly(A) addition. These results support a model in which the Pta1 and CF IA interactions are lost when the linker is substituted with the RPA70 IULD, but that a long enough replacement linker, as with 3 copies of RPA70, allows enough flexibility for Pap1 to find the 3' end during processing.

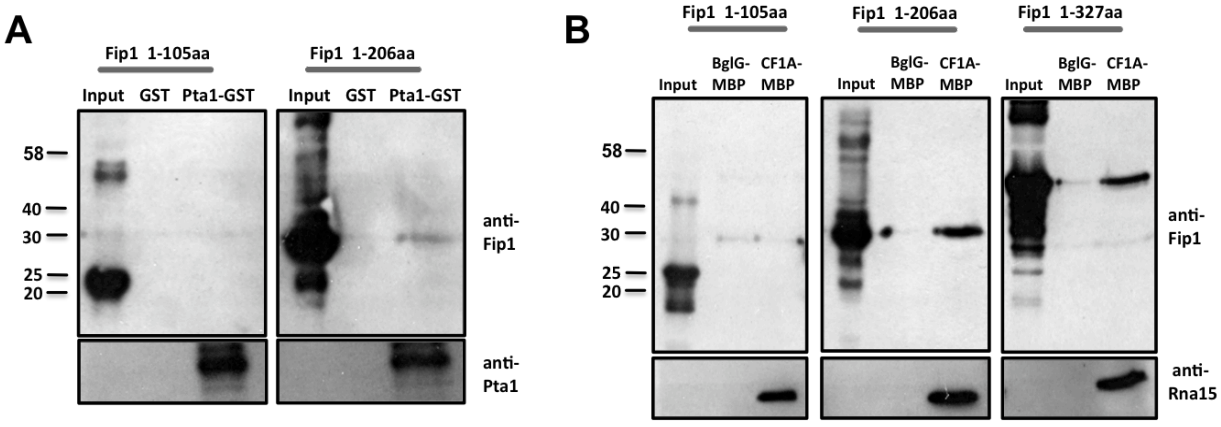


Figure 2.16: Pta1 and CF IA interact with the flexible linker of Fip1. (A) The flexible linker of Fip1 interacts with Pta1. Western blot analysis of pull-downs using purified recombinant GST-Pta1 and purified recombinant Fip1 truncations containing 1-105 amino acids (without Fip1 linker) and 1-206 amino acids (with Fip1 linker). Control pull-downs used GST only. (B) The flexible linker of Fip1 interacts with CF 1A. Western blot analysis of pull-downs using CF IA made from recombinant Rna14, Rna15, Pcf11, and MBP-tagged Clp1. CF IA was mixed with the Fip1 truncations or full-length Fip1 (1-327), followed by purification of complexes on amylose resin. MBP-tagged BglG, an *E. coli* protein, was used in control pull-downs. Western blot of Rna15 is used to detect the amount of CF IA in each pull-down.

Chapter 3

The role of the N-terminus of Pap1 in polyadenylation

INTRODUCTION

Poly(A) polymerase (Pap1) retains catalytic activity and nucleotide specificity when isolated away from the CPF complex (Edmonds, 1990). The structure of Pap1 is composed of three globular domains, which surround a central RNA binding cleft (Bard et al, 2000; Balbo and Bohm, 2007). As discussed in the previous chapter, the C-terminal domain is required for interaction with Fip1, which tethers Pap1 to CPF during 3' end processing (Preker et al, 1995, Helmling et al, 2001; Meinke et al., 2008). Fip1 is also involved in the regulation of Pap1 activity during polyadenylation (Helmling et al, 2001). However, other factors that directly regulate the activity of the enzyme during polyadenylation have not been described.

Previous studies have shown that a deletion of the first 18 amino acids at the N-terminus of Pap1 is lethal in yeast cells, and causes a defect in the specific polyadenylation of RNA precursors when assayed *in vitro* using purified complexes (Zhelkovsky et al, 1995). However, a recombinant Pap1 lacking the 18 amino acids at the N-terminus is not defective for the non-specific addition of poly(A) to oligo(A) primers (Zhelkovsky et al, 1995). Since, the interaction of Fip1 with Pap1 does not require the first 18 amino acids of Pap1 (Preker et al, 1995, Helmling et al, 2001), these findings suggest a novel mechanism for the regulation of polyadenylation, which involves the interaction of the N-terminus of Pap1 with other unidentified proteins in the 3' end-processing complex.

One of these proteins might be Pta1, which functions in pre-mRNA cleavage, poly(A) addition, and transcription termination (Zhao et al, 1999b; Millevoi and Vagner, 2009). Recombinant Pta1 interacts strongly with the *in vitro* translated products of Ssu72, Pti1, and Ysh1, which are CPF proteins and weakly with Fip1 (Ghazy et al, 2009). It also interacts weakly, but specifically with Pap1 (and thus might have a role in

the regulation of Pap1 activity, since phosphorylation of Pta1 is known to prevent polyadenylation but not cleavage (He and Moore, 2005). Cft1 of CPF might also play a role in the regulation of Pap1 activity. CPSF 160, which is the mammalian homolog of Cft1, interacts selectively with the AAUAAA region of the pre-mRNA transcript and with mammalian PAP. Recombinant CPSF 160 was shown to inhibit non-specific polyadenylation, and increasing its concentration in an AAUAAA-dependent polyadenylation assay using CPSF and PAP inhibits specific polyadenylation (Murthy and Manley, 1995). This interaction and effect has not been described in yeast. Yeast Cft1 interacts with the RNA transcript and the RNA polymerase II (Pol II) CTD, and may be one of the subunits communicating poly(A) site recognition to Pol II to initiate termination (Dichtl et al, 2002).

The goal of this study is to identify the yeast proteins that interact at the N-terminus of Pap1 and characterize how these interactions play a role in the regulation of Pap1 activity during 3' end processing.

RESULTS

Comparison of the activity of Pap1 and ΔN-Pap1

Deletion of the first 18 amino acids of Pap1 (ΔN Pap1) causes a defect in the specific polyadenylation of radiolabeled *GAL7* RNA precursors when assayed *in vitro* in the presence of Mg^{2+} and other purified factors (Fig 3.1A and Zhelkovsky et al, 1995). However, non-specific polyadenylation assays in the presence of Mn^{2+} show that recombinant ΔN Pap1 is not defective for the addition of poly(A) to oligo(A) primers. To confirm that the difference in polyadenylation activity observed in both assays is not due

to the metal cation used in the reaction, the activity of ΔN-Pap1 was tested in the presence of Mg^{2+} and Mn^{2+}. We show that the ΔN-Pap1 enzyme has activity with Mg^{2+} and Mn^{2+} (Fig 3.1B) comparable to the wild type enzyme (personal communication from Dr. Balbo). This confirms that the difference observed in the previous experiment cannot be attributed to the metal, but is due to external factors such as interaction with other proteins that affect the activity of the enzyme.

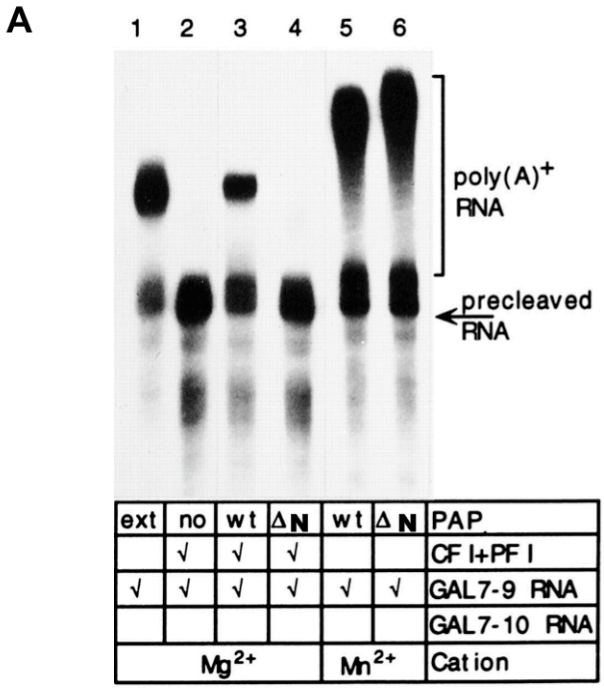
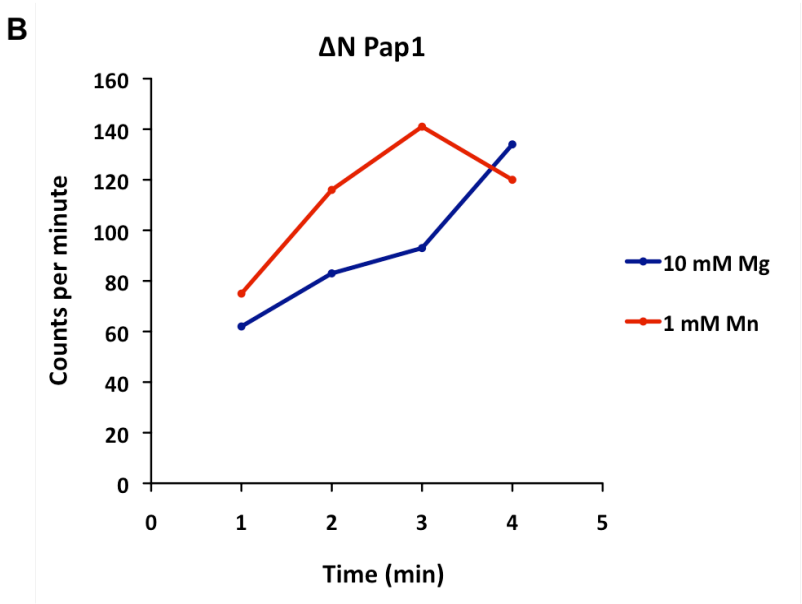


Figure 3.1: The N-terminus of Pap1 is required for specific polyadenylation. (A) *In vitro* poly(A) addition assay at 30˚C using pre-cleaved *GAL7* RNA substrate ending at the poly(A) site. CF I and PF I were partially purified factors that were needed in addition to Pap1 to reconstitute poly(A) addition. GAL7-9 RNA is wild type in sequence and GAL7-10 is deleted for the UA-rich polyadenylation signal (Adapted from Zhelkovsky et al, 2005).



(B) Activity of ΔN Pap1 in the presence of Mg^{2+} and Mn^{2+} at 30°C determined by measuring the incorporation of [^{3}H]AMP into oligo A_{18} RNAs by ΔN Pap1. 1 µl of ΔN Pap1 was used in both experiments. Counts per minute indicate the amount of tritium label collected at the indicated times.

General interactions at the N-terminus of Pap1

Fip1 tethers Pap1 to the CPF complex during 3' end processing, and the interaction domain for Fip1 has been mapped to the C-terminus of Pap1 (Zhelkovsky et al, 1998; Meinke et al., 2008). To study general interactions at the amino terminus of Pap1, pull downs were performed using full length Pap1 and a Pap1 truncation lacking the first 18 amino acids (ΔNPap1) from yeast nuclear extracts. A glutathione-S-transferase (GST) tag was added at the C-terminal domain of both Pap1 constructs to allow for interactions at the N-terminus of Pap1, and the pull down results were visualized by silver stain (Fig 3.2A, lane 2 and 4). Analysis by western blot using

antibodies against Pta1 and Yth1 showed that some of the proteins observed were components of the CPF complex (Fig 3.2B). This result suggests that the inactivity of ΔN-Pap1 in specific poly(A) addition may not be due to a failure to be incorporated into CPF. However, to be certain of such a conclusion, it would be necessary to determine if the full spectrum of subunits is present in similar stoichiometry, or to show that the complex assembled with ΔN-Pap1 was active in the cleavage step.

To eliminate interactions due to Fip1, truncations of Pap1 lacking 206 amino acids at the C-terminus of Pap1 (Pap1ΔC), which include the Fip1 binding site, were made to give Pap1ΔC-GST and ΔNPap1ΔC-GST. These were used in pull-downs from yeast nuclear extracts (Fig 3.2A). Pull-downs with Pap1 lacking the C-terminus (Pap1ΔC-GST) showed a profile of protein interactions (Fig 2A, lane 6), which looks very different from that of the CPF proteins observed using full length Pap1 (Fig 3.2A, lane 2). This suggests the presence of Fip1-independent interactions at the N-terminus of Pap1, which may play a role in Pap1 function. Consistent with this idea, some of the bands are no longer present when truncated Pap1 lacking 18 amino acids at the N-terminus and the C-terminal Fip1 interacting domain (ΔNPap1ΔC-GST) is used for the pull-downs (Fig 3.2A, lane 8). Since the pull downs were performed only once, additional pull downs using Pap1ΔC-GST and ΔNPap1ΔC-GST are needed verify the results obtained and to better study proteins that interact at the N-terminus of Pap1.

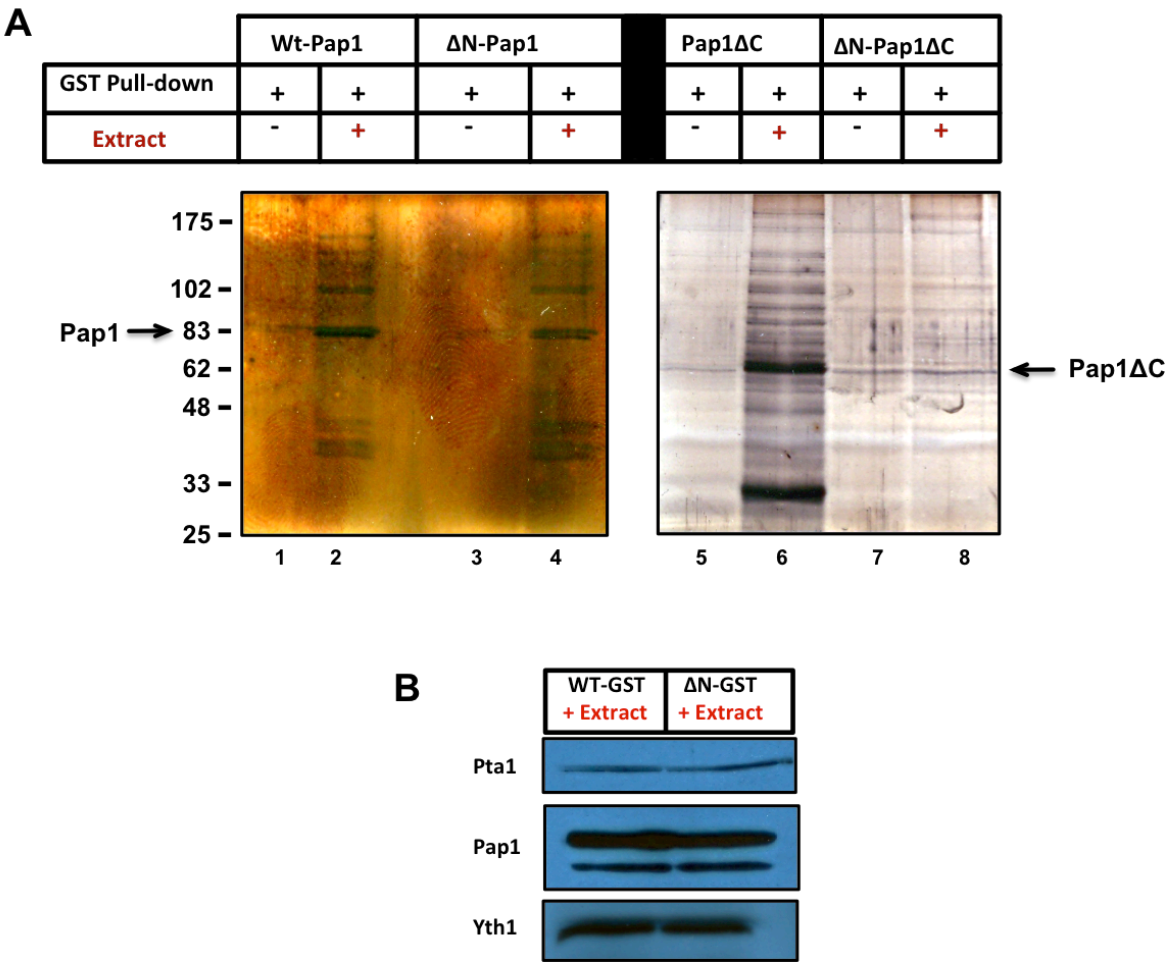


Figure 3.2: Interactions at the N-terminus of Pap1. (A) SDS PAGE and silver stain analysis of proteins that interacted with Pap1 after pull downs from nuclear extract. Lane 6 shows non-specific binding of proteins to Pap1 lacking the C-terminus. Pap1ΔC-GST and ΔNPap1ΔC-GST pull downs need to be repeated to confirm the results. (B) Western blot analysis of CPF proteins from Fig 3.2A using antibodies against Pta1, Yth1 and Pap1.

An N-terminal fragment of Pta1 interacts with the N-terminus of Pap1

A previous report had shown some interaction between Pta1 and Pap1 by using pull down of *in vitro* translated proteins (Ghazy et al, 2009). We used a *GAL4*-based two-hybrid system to identify interaction domains on Pta1 and Pap1. Full length and truncations of Pta1 were tested for interaction with full length Pap1 and ΔNPap1. The Pta1 constructs used for this analysis are described in Figure 3.3A and have been previously shown to express protein and in most cases interact with other proteins (Ghazy et al, 2009).

Consistent with the presence of the Fip1 interaction domain, both Pap1 and ΔNPap1 interact with Fip1 (Fig 3.3B). However, the two-hybrid results showed no interaction between full length Pta1 and Pap1. Analysis using truncations of Pta1 showed interaction between the Δ300-785 amino avid truncation of Pta1 and full length Pap1. No interaction was observed between this truncation and ΔNPap1. This suggests that the 300 amino acids at the N-terminus of Pta1 interact with the N-terminus of Pap1.

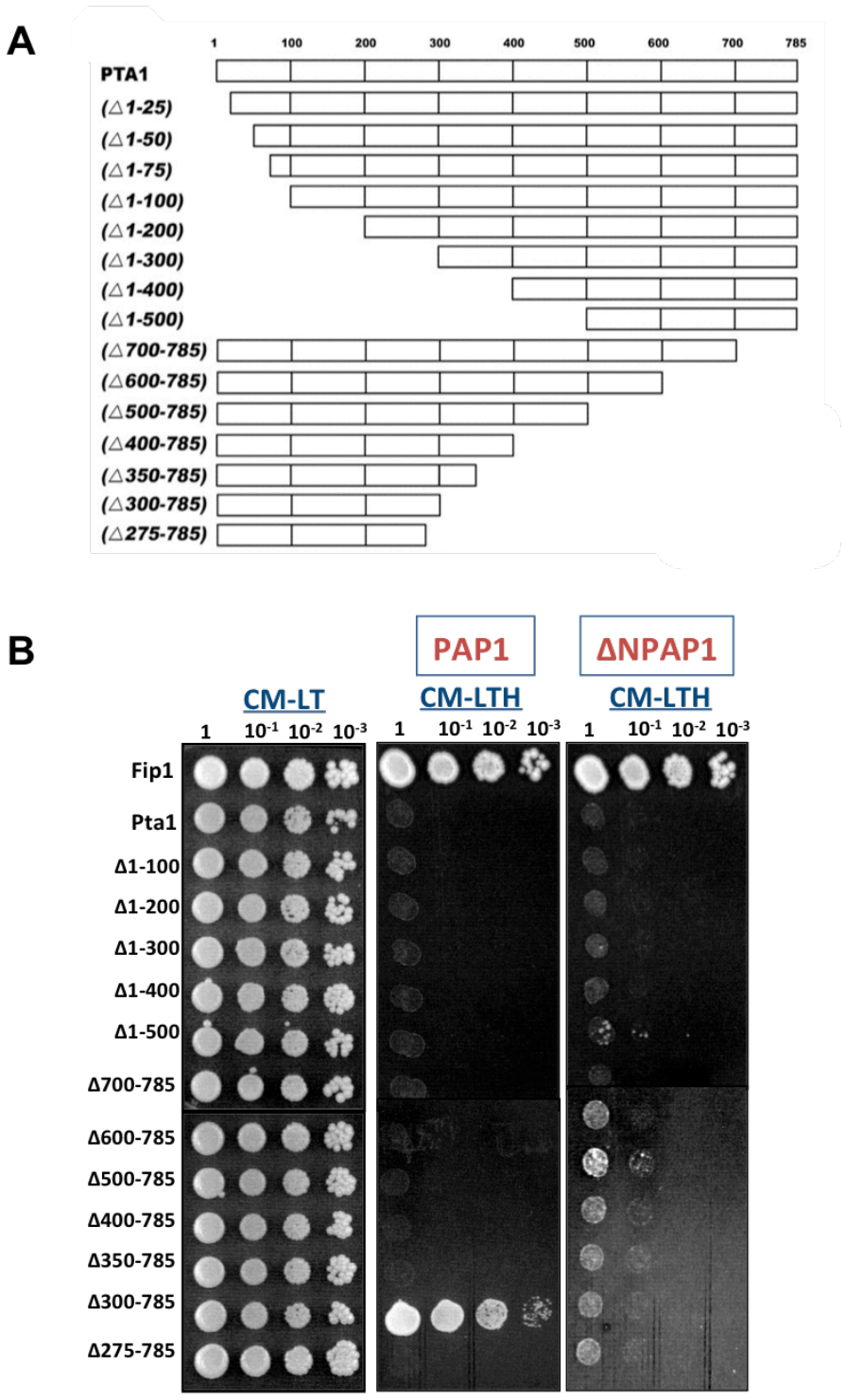


Figure 3.3: Pta1 interacts with the N-terminus of Pap1 (A) Full length Pta1 and Pta1 truncations, used for two-hybrid analysis. (B) 1-300 amino acid truncation of Pta1 interacts with the N-terminus of Pap1. Fip1 was used as a control. CM = Complete Media, L = Leu, T = Trp, H = His.

To further analyze this interaction, pull down experiments using recombinant proteins fused to GST were used to test for direct interaction between Pta1 and Pap1. For this experiment, constructs of Pta1 containing the full-length protein or fragments containing the first 1-300 and 1-275 amino acids of the protein fused to a GST tag were used to pull down His_6-tagged Pap1 or ΔN-Pap1. The purified proteins are shown in Figure 3.4A. For pull downs, equal amounts of Pta1 were bound to GST beads, and incubated with Pap1- His_6 and ΔN-Pap1- His_6. Bound Pap1 was visualized by a western blot using monoclonal Pap1 antibody (Fig 3.4B).

The GST pull-downs showed interaction between Pap1 and full-length Pta1. The 1-300 amino acid fragment of Pta1 also interacts with Pap1, while there is no interaction between full length Pap1 and the 1-275 fragment of Pta1. This confirms the 2-hybrid results, which showed interactions between amino acids 1-300 of Pta1 and Pap1. The interaction between full length Pta1 and Pap1 evident from the pull down assays was not detected by two-hybrid analysis. This may be due to steric constraints that prevent a two-hybrid signal when both proteins are part of CPF, or due to interaction between Pta1 and another protein *in vivo* in the two-hybrid, which blocks its interaction with Pap1. A small interaction is also observed between GST and full length Pap1. Since the GST concentration used for the pull downs were determined visually, a higher concentration of GST might have used for the GST control. This however this does not affect the results, since a higher concentration of Pap1 was pulled down by full length Pta1 and the 1-300 amino acid truncation of Pta1.

Interestingly, ΔN-Pap1 does not interact with Pta1 or the 1-300 Pta1 fragment *in vitro* (Fig 3.4B). The ΔN-Pap1 protein is active in the nonspecific poly(A) addition assay (Fig 3.1), indicating that unfolding of the protein is not an explanation for lack of interaction. Moreover, it can pull down Pta1 from yeast whole cell extract (Fig 3.2B). In

conclusion, these findings suggest a direct interaction between Pap1 and Pta1 through the Pap1 N-terminus, and indirect interactions through other CPF subunits, probably through the Fip1 interaction at the Pap1 C-terminus.

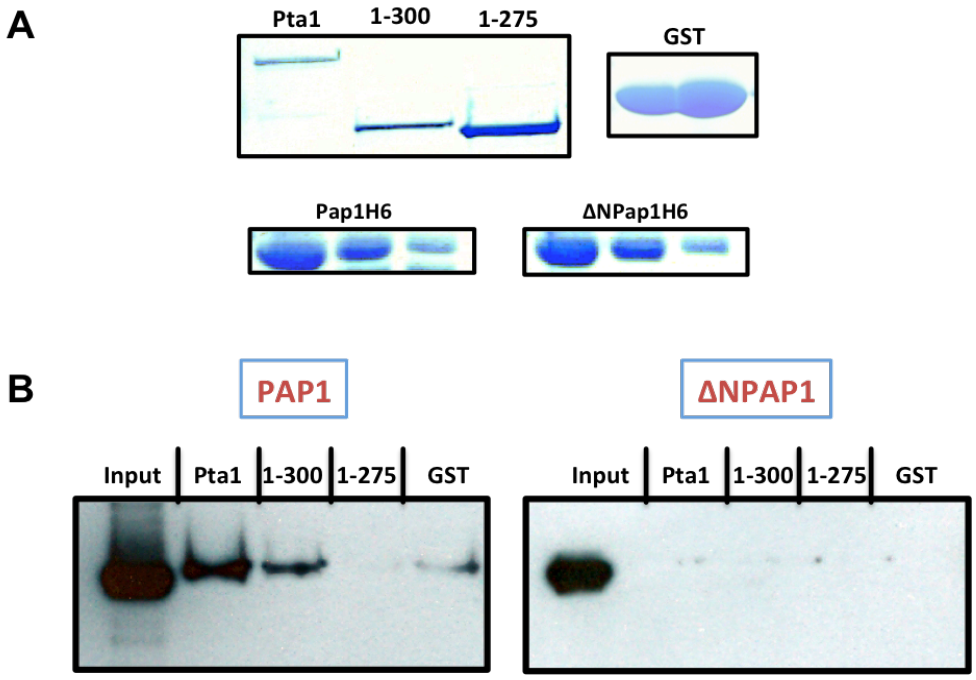


Figure 3.4: Pta1 interacts with the N-terminus of Pap1 (A) Purified recombinant proteins used for GST pull down. The three bands of purified Pap1 protein represent different elutions of the purified Pap1 protein. (B) Western blot analysis showing interaction between full length Pta1 and the 1-300 amino acid fragment of Pta1 with the N-terminus of Pap1.

Cft1 interacts with the N-terminus of Pap1

CPSF160, which is the mammalian homologue of Cft1, has been shown to interact with mammalian poly(A) polymerase (Murthy and Manley, 1995). To ask if this interaction is conserved in yeast, two-hybrid analysis was performed between full length Cft1, full length Pap1 and a truncation of Pap1 missing the N-terminus (ΔNPap1). Two-

hybrid analysis using *GAL4* activation domain fused to Cft1 (GAD-Cft1) and a *GAL4* binding domain fused to Pap1 (GBD-Pap1) showed no interaction between the proteins in the two-hybrid system (Fig 3.5A).

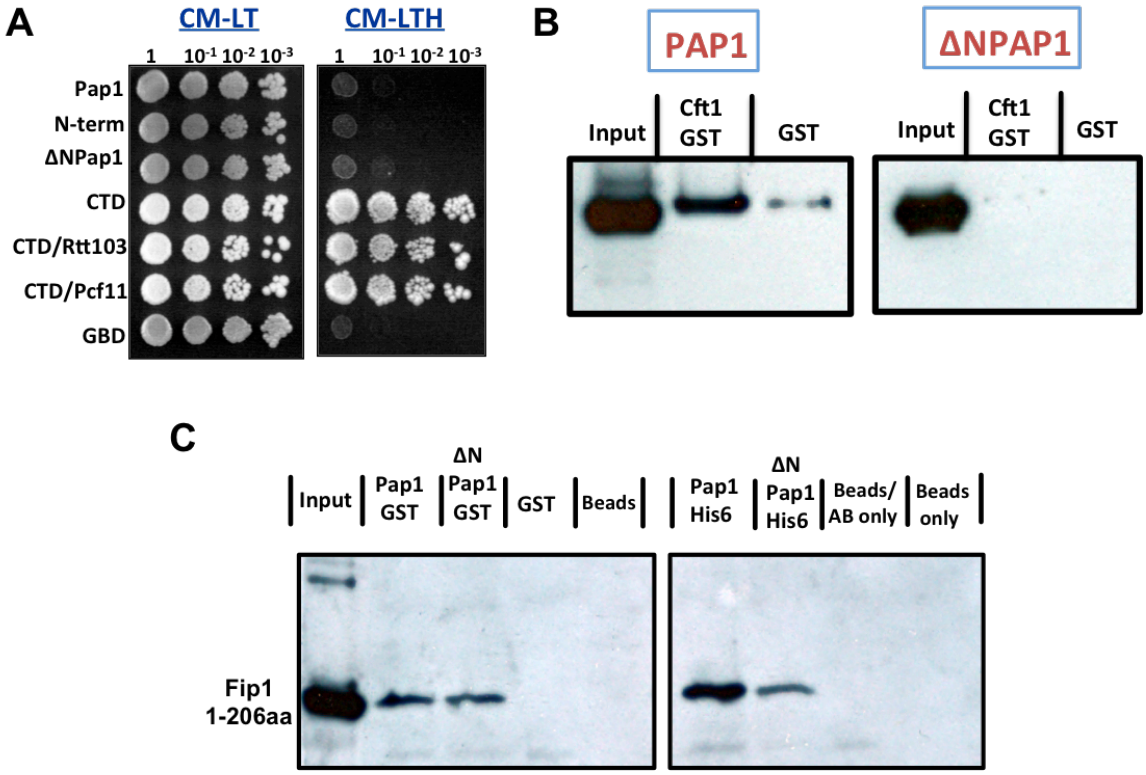


Figure 3.5: Cft1 interacts with the N-terminus of Pap1 *in vitro*. (A) Two-hybrid analysis between the indicated proteins and full length Cft1. A CTD construct, which previously shown to interact with Cft1, Rtt103, and Pcf11 (Kim et al, 2004; Luo and Bentley, 2004; Noble et al, 2005) was used as a control. (CTD, C-terminal domain of Rpb1, a subunit of RNA pol II; N-term – the 18 amino acids at the N-terminus of Pap1) (B) GST pull-downs showing interaction between full length Cft1 and Pap1. Pap1 was detected by western blot using monoclonal Pap1 antibody. (C) Western Blot analysis using polyclonal Fip1 antibody showing that full length Pap1 and ΔNPap1 interacts with a truncation of a 1-206 amino acid truncation of Fip1, which contains the Pap1 binding domain. The left panel shows Pap1-GST and ΔNPap1-GST pulldowns of 1-206aa Fip1. The right panel shows immunoprecipitation of 1-206aa Fip1 using Pap1-His6, ΔNPap1-His6 immobilized on protein A beads using monoclonal Pap1 antibody. Beads = Protein A Agarose beads, A = Monoclonal Pap1 antibody, B = Protein A Agarose beads.

To rule out the possibility of a false negative, which was also observed with the Pta1 two-hybrid, pull downs using recombinant GST-Cft1 and Pap1-His6 was used to test a direct interaction between both proteins. The GST pull down showed interaction between Cft1 and full length Pap1, but no interaction with the Pap1 truncation lacking the N-terminus (Fig 3.5B). This shows that like its mammalian homologue, Cft1 interacts with Pap1, and that this interaction is localized to the N-terminus of Pap1.

To test the stability of the recombinant ΔNPap1 constructs used in our experiments, full length Pap1 and ΔNPap1 constructs used in the pull down experiments were used to pull down a 1-206 amino acid truncation of Fip1, which contains the Pap1 interaction domain of the protein. Pap1-GST constructs were immobilized on GST beads, while Pap1-His6 constructs were immunoprecipitated using protein A Agarose beads. The pull downs showed interaction between the Pap1 constructs and the 1-206 amino acid truncation of Fip1 (Fig 3.5C). This shows that the ΔNPap1 constructs are stable, and can interact with proteins in the pull down experiments. It also confirms the observation that the N-terminus of Pap1 interacts with Pta1 and Cft1.

Recombinant CPSF 160 was shown to inhibit both specific and non-specific polyadenylation, and increasing its concentration in an AAUAAA-dependent polyadenylation assay using calf thymus CPSF and HeLa PAP inhibits polyadenylation (Murthy and Manley, 1995). To study the role of Cft1 in yeast polyadenylation, polyadenylation assays in the presence of Mg^{2+} or Mn^{2+} were performed using ATP and radiolabeled *GAL7* RNA substrates ending at the *GAL7* poly(A) site. Purified recombinant Cft1 was mixed with Pap1 using a molar ratio of 1:1 in the presence of the *GAL7* RNA substrate to study the effect of Cft1 on polyadenylation. A recombinant Fip1 truncation, containing amino acids 1-192, which encompasses the Pap1 interaction domain and the flexible linker region of Fip1, was added to some of the reactions to

study the effect of Cft1 in the presence of Fip1 and Pap1. Fip1 has been shown to inhibit Pap1 activity in nonspecific polyadenylation assays (Zhelkovsky et al, 1998).

As expected, a three-fold higher concentration of Fip1 strongly inhibited Pap1 in the presence of either Mg^{2+} or Mn^{2+} (Fig 3.6, lanes 3 and 7). However, Cft1 does not inhibit the activity of Pap1 with either cation at the tested molar ratio of 1:1 (Fig 3.6, lanes 4 and 8). Murthy and Manley showed that CPSF 160 inhibited mammalian PAP at a molar ratio of 3:1, and this inhibition increased as the amount of CPSF 160 is increased in the reaction (Murthy and Manley, 1995). A shorter poly(A) tail is observed when Pap1 and Cft1 are mixed in the presence of Mg^{2+} (Fig 3.6, lane 8), which suggests that Cft1 is affecting the activity of Pap1 in an undefined manner. We also show that Cft1 does not alleviate the inhibition of Pap1 activity in the presence of a three-fold higher molar concentration of Fip1. This suggests that Cft1, through direct interaction with Pap1, might play a role in yeast polyadenylation that is different from the role of CPSF 160 in mammalian polyadenylation. Further analysis of the effect of Cft1 on Pap1 alone and in the presence of Fip1, using higher molar concentrations of Cft1 and a GST control is needed, to better understand the role of Cft1 on Pap1 action.

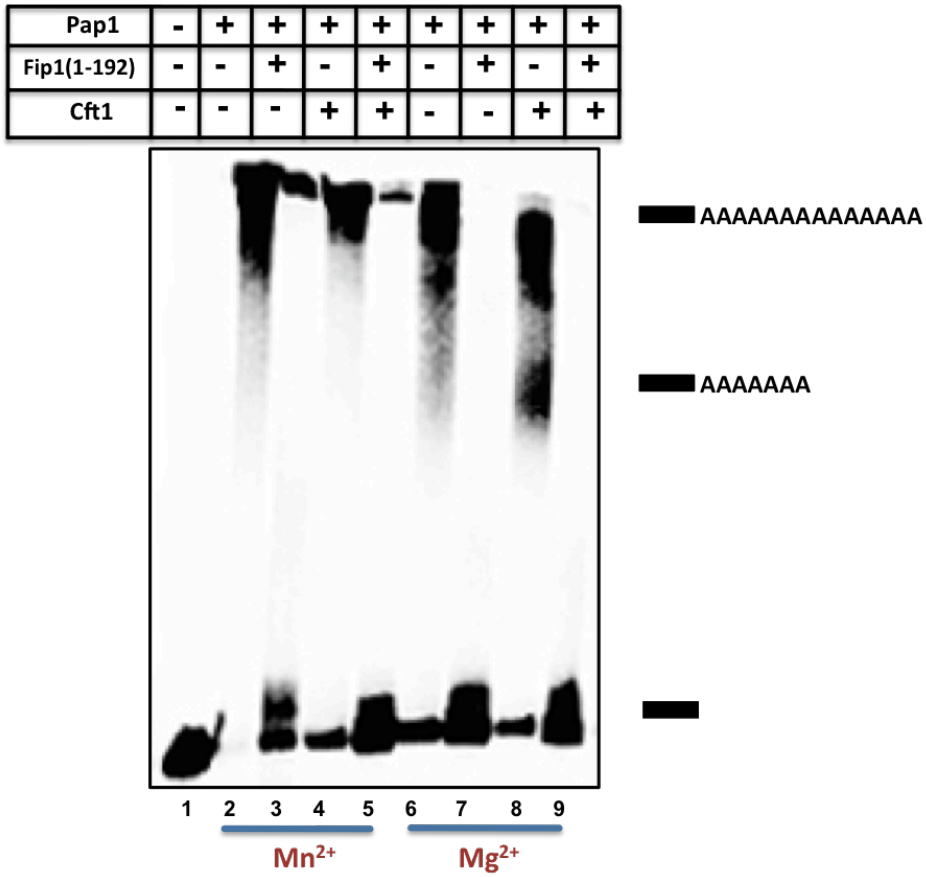


Figure 3.6: Cft1 might not have an inhibitory effect on Pap1 activity. Cft1 was mixed with Pap1 and Fip1 in the presence of Mg^{2+} and Mn^{2+} at a 1:1 molar ratio to determine the effect of the protein on Pap1 activity.

The non-essential yeast protein, Sub1, interacts with the N-terminus of Pap1

To find other proteins that interact with the N-terminus of Pap1, a peptide comprising the 18 amino acids at the N-terminus of Pap1 was coupled to a peptide column, and used to pull down interacting proteins from yeast nuclear extract. An eleven amino acid peptide from the C-terminus of Npl3 was used as control peptide (Fig 3.7A) (Dermody et al, 2008). The interacting proteins were run on a gel, cut into gel slices (Fig

3.7B) and analyzed by mass spectroscopy. Table 3.1 shows the list of proteins obtained from this analysis.

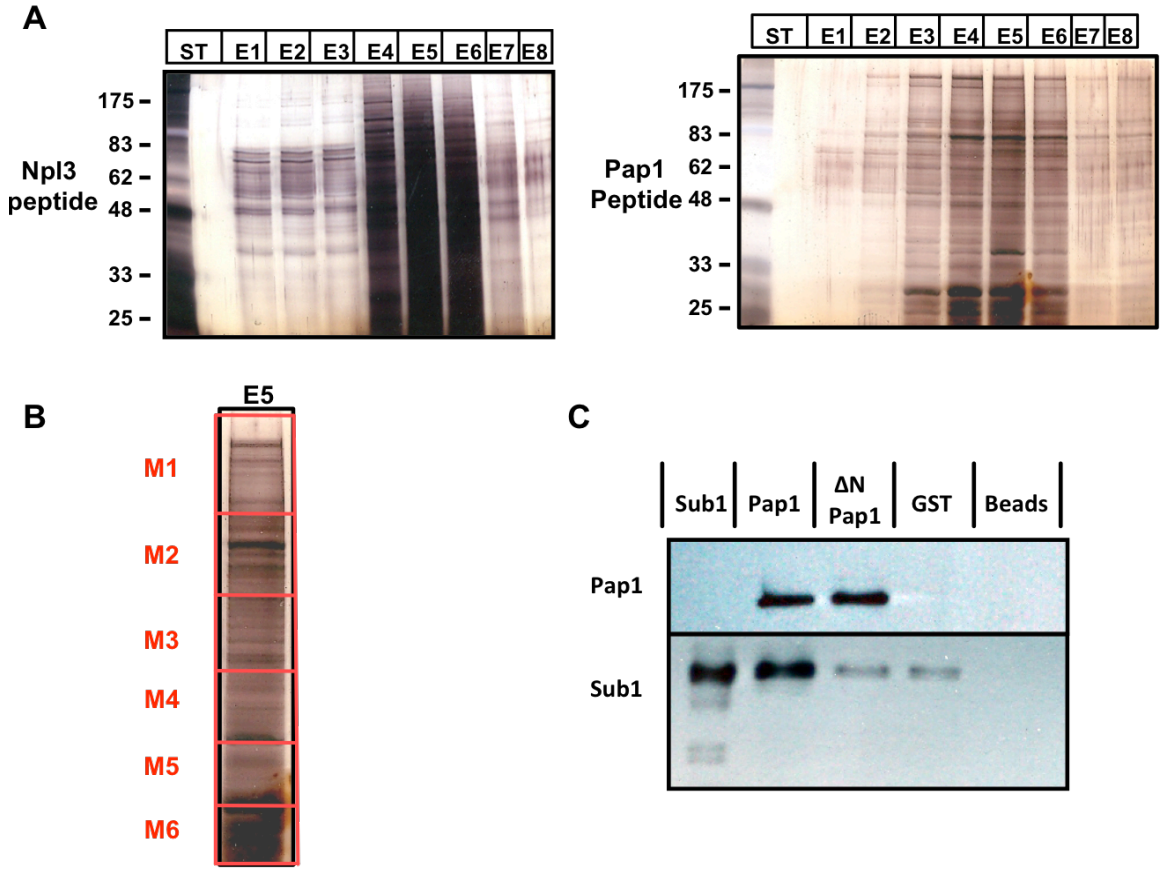


Figure 3.7: Sub1 interacts with the N-terminus of Pap1. (A) Silver stain results of peptide pull downs. Npl3 peptide was used as a control. (B) Lane 5 of the Pap1 peptide pull down was cut as indicated, and analyzed by mass spectroscopy (C) Western blot analysis using monoclonal V5 antibody for Sub1 and monoclonal Pap1 antibody for Pap1 shows interaction between full length Pap1 and Sub1.

Table 3.1: Mass Spectroscopy Results from the Peptide Column.

Protein	Peptides matched	Protein Description
Adh1p	4	Alcohol dehydrogenase
Anc1p	1	Transcription initiation factor TFIIF small subunit
Cdc33p	3	mRNA cap binding protein eIF-4E
Chc1p	2	Presumed vesicle coat protein
Cue5	3	Coupling of Ubiquitin conjugation to ER degradation
Efb1p	1	Translation elongation factor EF-1beta
Fas1p	3	Acetyl transferase, enoyl reductase and malonyl/palmityl transferase
Fas2p	4	Alpha subunit of fatty acid synthase
Hsc82p	11	Constitutively expressed heat shock protein
Hsp104p	11	104 kDa heat shock protein
Hsp60p	4	Mitochondrial chaperonin, homolog of E. coli groEL protein
Hsp82p	6	Homolog of mammalian Hsp90
Iki1p	1	Sensitivity to pGKL killer toxin
Loc1p	1	Double-stranded RNA-binding protein
Nop1	2	Nucleolar protein
Nop6	3	Ribosome biogenesis
Pdr13p	2	Hsp70 Protein
Pfk2p	6	Phosphofructokinase beta subunit
Pil1	1	Endocytosis
Rpg1p	3	Translation initiation factor eIF3
Rpl11bp	3	60S ribosomal protein L11B
Rpl13ap	1	Ribosomal protein L13A
Rpl19bp	2	Ribosomal protein L19B
Rpl20ap	2	Ribosomal protein L20A
Rpn1p	2	Subunit of 26S Proteasome
Rpp0p	2	60S ribosomal protein P0
Rps6bp	2	40S ribosomal gene product S6B
Rps9bp	3	Ribosomal protein S9B
Rps13p	1	Ribosomal protein S13
Rps14ap	3	Ribosomal protein S14A
Rps15p	2	40S ribosomal protein S15
Rps17bp	3	Ribosomal protein S17B
Rps17bp	2	Ribosomal protein S17B (rp51B)
Ser3p	2	3-phosphoglycerate dehydrogenase
Ssa1p	6	Heat shock protein of HSP70 family
Sub1p	**6**	**Transcriptional co-activator, 3' end processing**
Tfp1p	2	Vacuolar ATPase V1 domain subunit A
Ygl242cp	1	

Sub1, which is a non-essential yeast protein that interacts with Rna15 and Pta1 (Calvo and Manley et al, 2001; He et al, 2003), was one of the proteins identified from the mass spectroscopy analysis of the bound proteins from the column. The analysis of the yeast nuclear extract proteins bound to the Npl3 control column did not show the presence of Sub1 (Dermody et al, 2008). Since Sub1 has been implicated in 3' end processing (He et al, 2003), it was chosen for further analysis. Pull downs were performed using recombinant Sub1-V5-$His_{6,}$ and Pap1-GST or ΔN-Pap1-GST. Sub1 interacts with full length Pap1, and this interaction is lost in the absence of the N-terminus of Pap1 (Fig 3.7C). This confirms the interaction observed with the peptide column, and shows that Sub1 has a direct interaction with the N-terminus of Pap1.

A C-terminal truncation of Nab6 interacts with the N-terminus of Pap1

Nab6 is a non-essential-RNA binding protein of undefined function in yeast. Nab6 was identified as a protein that interacted with Pap1 in a two-hybrid screen performed in the Moore lab by Dr. Zhelkovsky. The interaction was localized to last 210 amino acids of Nab6. To further explore this interaction, another two-hybrid analysis was performed using the C-terminal truncation of Nab6 and either full length Pap1 or ΔN-Pap1. The interaction of Nab6 with full length Pap1 is lost in the absence of the N-terminus of Pap1 (Fig 3.8A). A weaker interaction was also observed with a GBD fusion containing only the 18 amino acids at the N-terminus of Pap1. Pull downs using Pap1-GST and Nab6-V5 confirmed that the Nab6 interaction requires the N-terminus of Pap1 (Fig 3.8B).

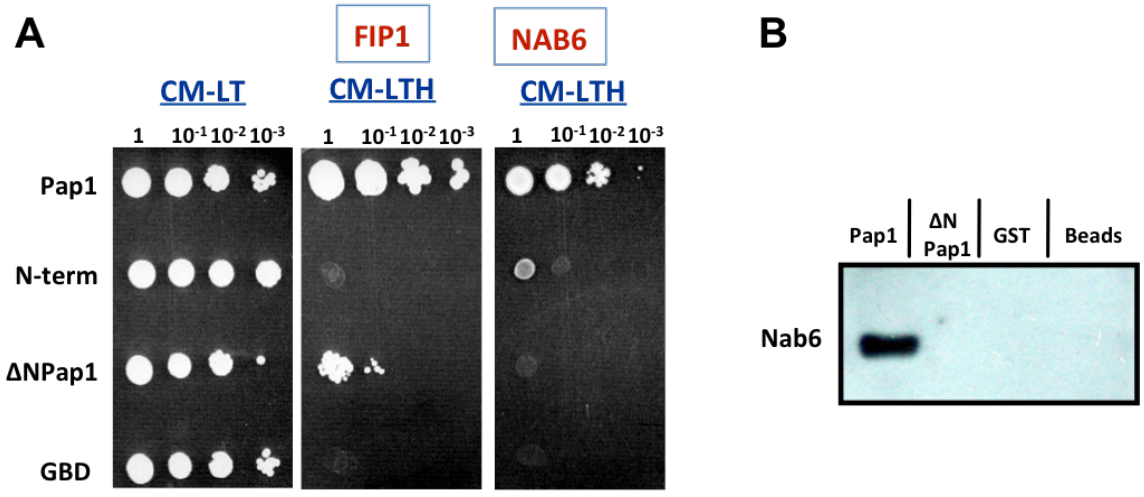


Figure 3.8: The Nab6 interaction requires the N-terminus of Pap1. (A) Two-hybrid analysis between the indicated proteins and a C-terminal truncation of Nab6. (B) Western blot analysis of GST pull downs using monoclonal V5 antibody shows an interaction between full length Pap1 and the last 210 amino acids of Nab6 but not with ΔN-Pap1.

The Role of Nab6 and Sub1 in Pap1 function

Nab6 and Sub1 are both non-essential yeast proteins. To study the role that these proteins play in Pap1 activity, genes encoding both proteins were deleted in W303 and BY4742 yeast strains, and the single and double deletions studied for their effects on growth and poly(A) addition. Both yeast strains were viable at 30°C with a double deletion of the *SUB1* and *NAB6* genes (Fig 3.9A and 3.9B). Analysis of the thermosensitivity of the double deletion in the W303 strain showed a mild cold sensitive phenotype at 18°C, and *Δsub1* alone is temperature sensitive at 40°C (Fig 3.9A). This shows that Nab6 and Sub1 might be playing a redundant role in yeast survival at 18°C. No thermosensitive phenotypes were observed in the BY4742 yeast strain (Fig 3.9B). The reason for the different response observed in W303 and BY4742 is not known.

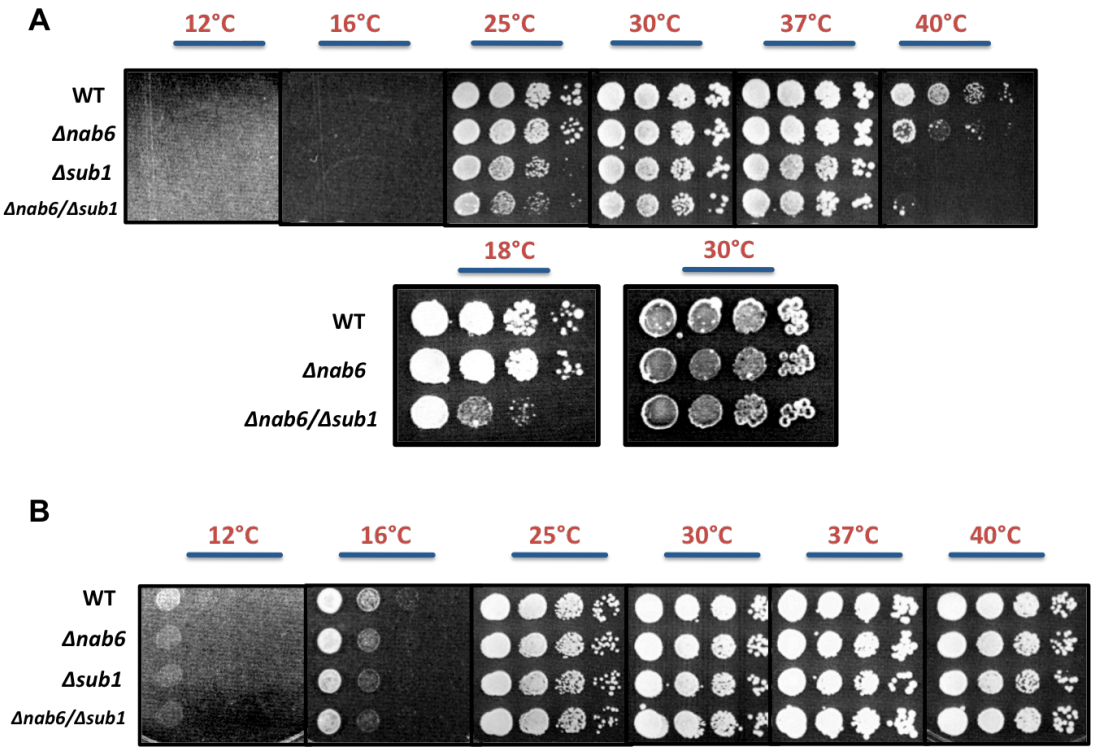


Figure 3.9: Thermosensitivity of *nab6/sub1* double deletion.
(A) Thermosensitivity tests on the W303 wild type and deletion yeast strains lacking the individual and the combined genes. Serial dilutions of liquid cultures were spotted on YPD medium and incubated at the indicated temperatures for up to 120 hrs. (B) Thermosensitivity tests on the BY4742 wild type and deletion yeast strains lacking the individual and the combined genes.

To study the role of Nab6 and Sub1 in Pap1 activity, *in vitro* assays were performed using radiolabeled *GAL7* RNA substrates ending at the *GAL7* poly(A) site and extracts from the W303 yeast strain expressing a deletion of the individual *NAB6* and *SUB1* genes, and a double deletion of both genes (Fig 3.10A). The data shows no change in the poly(A) tail length of the *Δnab6* W303 yeast strain, and a slight decrease in the poly(A) tail length observed in the *Δsub1* W303 yeast strains. *In vitro* assays using the double deletion of both the *SUB1* and *NAB6* genes in W303 yeast strains shows a

slight increase in poly(A) tail length suggesting a role for Nab6 and Sub1 in modulating the poly(A) tail length. The deletion strains are also less efficient for processing, even though this has to be repeated to confirm this behavior. Further *in vitro* assays using the deletion yeast strains are needed to confirm the effects of Nab6 and Sub1 on poly(A) tail length. It will be also interesting to see if reactions performed at different temperatures or with extracts made from yeast grown at different temperatures will reveal differences from wild type processing.

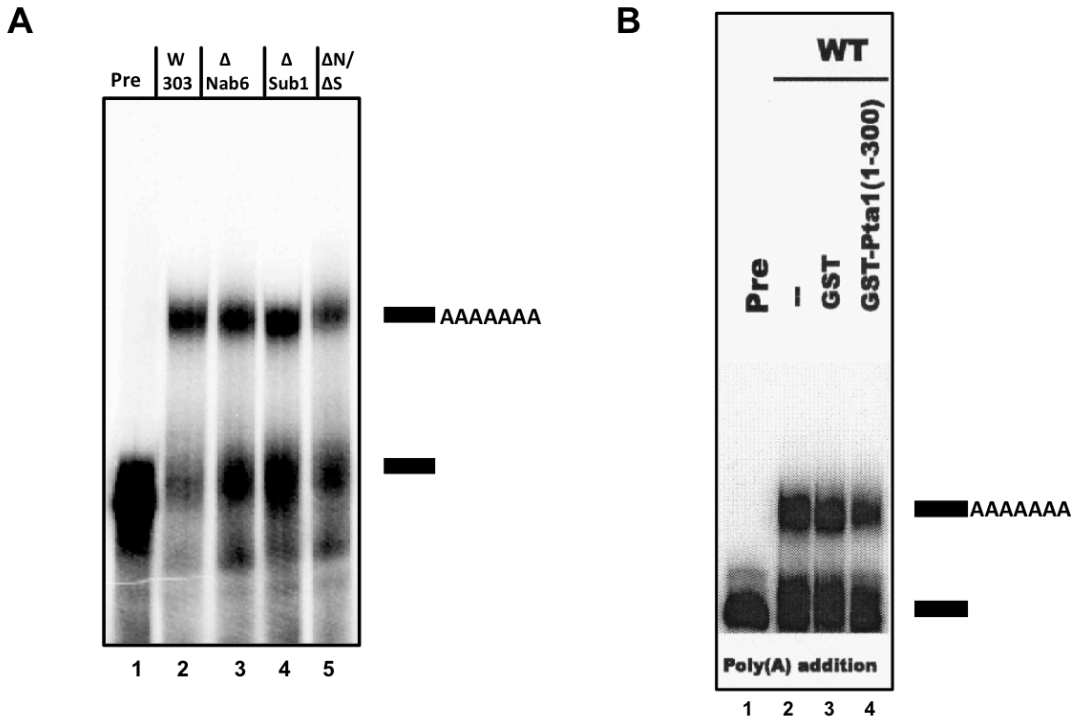


Figure 3.10: *In vitro* poly(A) addition assay at 30˚C using pre-cleaved *GAL7* RNA substrate ending at the poly(A) site. Unreacted precursors are shown in the lane marked "Pre", and the position of polyadenylated RNA indicated on the right. (A) Polyadenylation activity in W303 yeast strains with containing a deletion of the individual *NAB6* or *SUB1* genes, and a double deletion of *NAB6* and *SUB1* genes (Courtesy of Pei-Chun Yeh) (B) 1-300 truncation of Pta1 inhibits poly(A) addition *in vitro* (Courtesy of Mohamed Ghazy).

Chapter 4

General Discussion

A flexible linker region in Fip1 is needed for efficient mRNA polyadenylation

In this study, we investigated how Pap1 gains access to the 3' end of pre-mRNA through its interaction with Fip1 and how it changes from a distributive to processive enzyme when part of the CPF complex. We created a simplified model of CPF by fusing Pap1 to the λ phage N-protein domain, which then recognized the *boxB* hairpin on the RNA substrates used in our experiments. By increasing the affinity of Pap1 for the RNA, this artificial tethering increased the efficiency of polyadenylation (Figure 2.4A). This is consistent with the processive activity observed when human PAP is in a complex with hFip1 and RNA substrate (Kaufman et al, 2004). Like its yeast homolog, hFip1 binds to PAP, but unlike the yeast Fip1, hFip1 has an arginine-rich RNA binding domain at the C-terminus of the protein that is sufficient to tether PAP to an RNA precursor and initiate polyadenylation of the precursor (Kaufman et al, 2004). Together, these observations suggest that in yeast, the CPF complex, through a Fip1 tether, facilitates processive polyadenylation by locking Pap1 onto the RNA substrate. In the case of CPF, the connection to RNA is most likely through interactions with CF I bound to polyadenylation signals upstream of the poly(A) site. CPF is thought to contact U-rich stretches flanking the poly(A) site, but as discussed below, contact with the upstream U-rich element after cleavage may not allow sufficient distance for the new 3' end to engage with Pap1.

We also found that the length of the RNA substrate affects the efficiency of polyadenylation when Pap1 is tethered to the RNA. Our results show a three-fold decrease in polyadenylation activity with a substrate that has only 10 nucleotides between the tethering site and the 3' end, compared to one in which this distance is extended to 20 nucleotides. The 3' end of an RNA with only 10 nucleotides from the binding site probably cannot easily reach the Pap1 active site, thereby significantly reducing the initial extension of the poly(A) tail, when compared to a longer substrate.

This distance is consistent with our molecular model of Pap_Np-RNA interaction (Figure 2.4B), which suggests that 15 nucleotides are needed to span this distance. Our findings also predict that CPF will not contact the RNA during the initiation of poly(A) addition, as the U-rich stretch that it is thought to recognize lies within 10 nucleotides of the cleavage site (Graber et al, 1999; Preker et al, 1997).

We also examined parameters that might affect Pap1 activity at the 3' end of the pre-mRNA when it is complexed with CPF. We showed that Pap1 covalently linked to a shortened, unviable form of Fip1 was able to rescue a combined deletion of the individual *PAP1* and *FIP1* genes. This yeast strain was also active for polyadenylation *in vitro* (Figure 2.9B). Previous studies have shown that Pap1 interaction with the 3' end of the cleaved pre-mRNA is mediated by Fip1, which holds Pap1 in CPF (Barabino et al, 2000; Helmling et al, 2001). Our data suggests that the attachment point of Pap1 to Fip1 is not as important in the cell as the recruitment of Pap1 into the CPF complex, which allows initiation of polyadenylation by tethering Pap1 to the RNA substrate. This is consistent with our simplified CPF model and with the mammalian and plant systems, where Fip1 contains RNA-binding domains (Kaufman et al, 2004; Forbes et al, 2005.)

Fip1 has been shown to inhibit the activity of Pap1 on oligo(A) (Zhelkovsky et al, 1998), and this inhibition, which depends on the region between the Pap1 and Yth1 interaction sites, was proposed to have a regulatory function. However, in the context of CPF, the activity of Pap1 is enhanced rather than inhibited (Preker et al, 1997), and the inhibition seen with purified proteins is relieved by RNA, which is expected to be at a high local concentration during mRNA polyadenylation (Meinke et al, 2008). Our data suggests that Fip1 might have a different type of regulatory role on the action of Pap1 through its flexible linker. The ability of Pap1 to gain access to the 3' end of the freshly cleaved RNA substrate in the CPF complex has always been a puzzle. The complex of proteins needed for pre-mRNA cleavage is quite large, and one might expect that this

bulky complex would occlude Pap1 from the 3' end. One solution to this problem might lie in the flexibility conferred by the extended linker between the Pap1 and Yth1 interaction sites on Fip1, which behaves like an Intrinsically Unstructured Linker Domain (IULD). An analogous protein is RPA70, whose IULD tethers two high-affinity single-stranded DNA (ssDNA) binding domains to an amino terminal protein interaction domain in RPA70. This tethering provides mobility, which enables the protein interaction domain to interact with ssDNA and other proteins, and facilitates binding of the amino terminus of RPA70 with ssDNA in competition with p53 (Miller et al., 1997; Jacobs et al., 1999; Daughdrill et al., 2001). Studies on the behavior of IULDs in proteins have shown that even though their sequence is not conserved, their dynamic behavior as tethers, which offer backbone flexibility to proteins, is evolutionarily conserved across most species (Daughdrill et al., 2007).

Close examination of human Fip1 reveals properties that might also be consistent with an IULD-containing protein. Like Fip1, hFip1 displays very low sequence conservation across species, has an apparent size by gel mobility that is higher than that predicted from amino acid sequence, and parts of the protein are susceptible to proteolysis (Kaufmann et al, 2004). Furthermore, prediction of the domain structure of yeast Fip1, hFip1 and *Arabidopsis* Fip1 using IUPred (Dosztányi et al, 2005), indicate a high degree of unstructured regions in the proteins (Fig 4.1). The predicted globular domains in the Fip1 proteins correspond to the conserved regions, which interact with CPSF 30 in hFip1 and its homolog, Yth1, in Fip1.

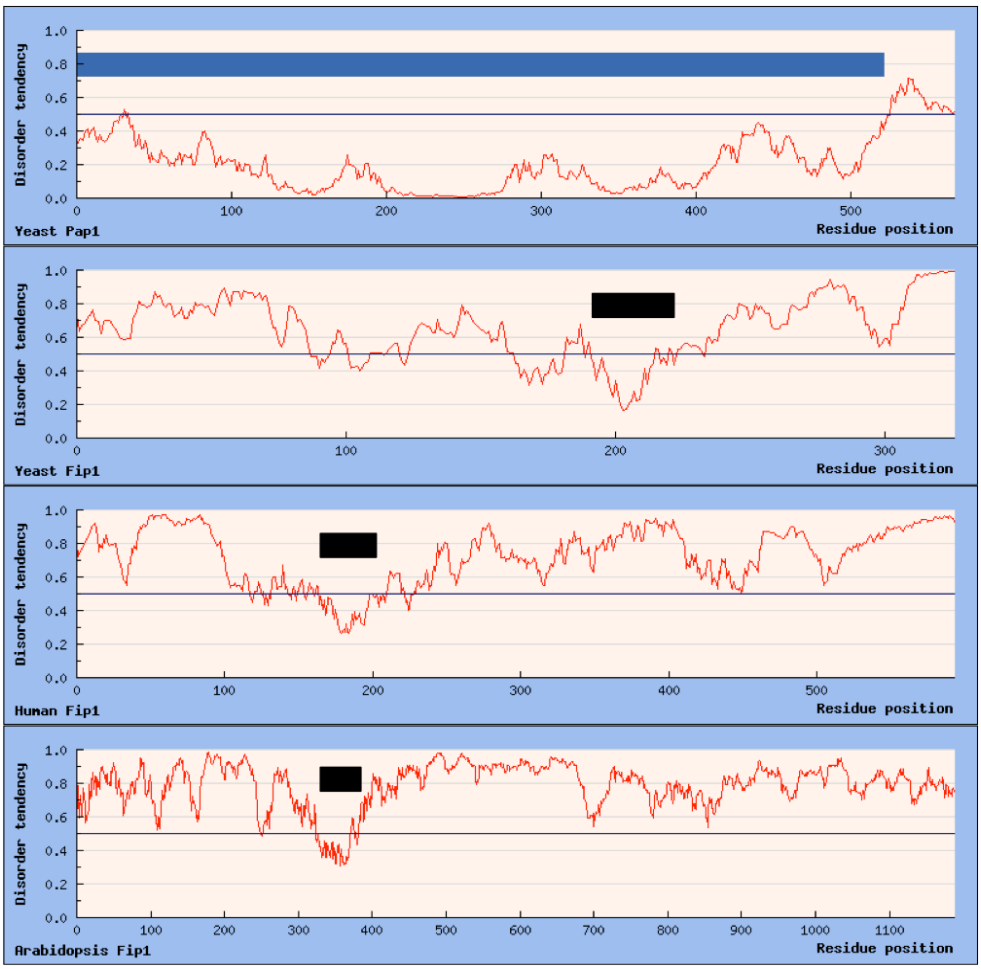


Figure 4.1: Prediction of the domain compositions of yeast Pap1, yeast Fip1, human Fip1 and *Arabidopsis* Fip1 using IUPred (Dosztányi et al, 2005). Protein sequences were entered into the IUPred web server, which generated the disorder tendency of different regions of the proteins. Yeast Pap1 was used to test the accuracy of the predicted domains. The blue box represents the predicted globular domain of yeast Pap1, which corresponds to the published Pap1 structure (Bard et al, 2000), and the short region at the C-terminus known to be unstructured. The black boxes represent the predicted globular domains of yeast, human Fip1 and *Arabidopsis* Fip1.

Surprisingly, a yeast strain expressing a mutant Fip1 completely missing the linker region exhibited wild-type growth except at high temperature (Figure 2.11). This is

consistent with our results showing that initiation of polyadenylation *in vivo* depends mostly on adequate tethering of Pap1 to the RNA, and that once primed, mRNAs can become sufficiently polyadenylated for cell survival. This observation might be due to some residual Fip1 linker, which is available in the Fip1 linker deletion construct, since the reported flexible linker region of Fip1 (Helmling et al, 2001) used for removing the linker region of Fip1 was determined by deletion of Fip1 domains, and not very precise. This minimal Fip1 however, functioned poorly in polyadenylation *in vitro*, with little precursor receiving a tail, and with few of those tails reaching a normal length, suggesting that the linker is necessary for efficient initiation of poly(A) addition in our processing extract. This corresponds with a severe temperature-dependent growth defect *in vivo* in the presence of 1.5% formamide (Figure 2.11). We propose that the linker in Fip1 not only tethers Pap1 to the CPF complex, but also provides the flexibility needed to most effectively access the 3' end of the cleaved pre-mRNA. However, the observation that two copies of the RPA70 IULD (144 amino acids) cannot restore the function of the 84 amino acid Fip1 linker and causes the same growth defect in the presence of 1.5% formamide suggests that this region has additional roles. We have shown that the Fip1 linker interacts with other subunits of mRNA 3' end processing complex including the CF IA complex and Pta1 of the CPF complex *in vitro* (Figure 4.2), and it is likely that these interactions facilitate Fip1 activity in the cell. This interaction between Fip1 and components of the CF IA complex is conserved across species. CstF 77, which is the homolog of Rna14, has been shown to interact with Fip1 in humans (Kaufmann et al, 2004), and plants (Forbes et al, 2005; Addepalli and Hunt, 2008).

When the normal nuclear environment is disturbed by the presence of formamide and high temperature or the preparation of processing extracts, such interactions may become more critical, leading to the thermosensitive growth and the poor polyadenylation that we have observed *in vitro*. However, three copies of the RPA70

IULD were sufficient to restore thermosensitive growth, and gave wild-type activity in processing extract, indicating that the need for these additional interactions can be partially overcome if the linker is sufficiently long.

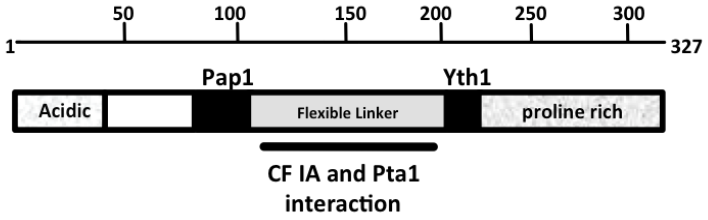


Figure 4.2: Schematic diagram of Fip1 showing the Pap1 and Yth1 interaction domains, and the flexible linker of Fip1, which interacts with CF IA and Pta1.

The Fip1 linker might also be an important target for the nuclear mRNA surveillance machinery. Saguez *et al.* 2008 have shown that the accumulation of defective mRNPs in THO/Sub2 mutant strains leads to the degradation of Fip1 in the cell through a mechanism dependent on the proteasome and the Rrp6 component of the nuclear exosome (Saguez et al, 2008). Another study found that prolonged depletion of the Glc7 phosphatase caused loss of poly(A) addition as well as Fip1 depletion (He and Moore, 2005). In both cases, Pap1 levels were normal. These findings suggest that Fip1 is particularly susceptible to degradation. One of the features of intrinsically disordered proteins is their potential for rapid degradation (Fink, 2005), and indeed, natively unfolded proteins like tau and unmodified alpha-synuclein can be directly processed by the 20S proteasome without a requirement for ubiquitination (Tofaris et al., 2001;David et al., 2002). Other evidence indicates that increased phosphorylation of CPF causes a weakened association of Fip1, which might expose the linker and facilitate its degradation (He and Moore, 2005). Pap1 is a target for posttranslational modifications

such as phosphorylation and sumoylation, which either inhibit its activity or localize it to the nucleus (Colgan et al, 1996, 1998; Mizrahi and Moore, 2000; Vethantham et al, 2008). Together, the research on Fip1 suggests that controlling Fip1 levels may be another way for the cell to regulate its polyadenylation capacity.

Our results extend our understanding of the functions of Pap1 and Fip1 as part of the 3' end-processing complex. We also show that Fip1 is a flexible linker which facilitates the access of Pap1 to the 3' end of substrate mRNA. Further work on this aspect of Fip1 function is needed to define the role that its unstructured domain plays in the regulation of 3' end processing and mRNA surveillance.

The Role of the N-terminus of Pap1 in polyadenylation

In this study, we investigated how Pap1 is regulated through interactions at the N-terminus of the protein and discovered novel interactions at the N-terminus of Pap1. We first showed that ΔN-Pap1, in spite of being inactive for specific poly(A) addition, could still interact with CPF subunits in extract, suggesting that it can be incorporated into a CPF complex (Fig. 3.2). We also found that Pta1 and Cft1, which are essential components of the CPF complex, interact with the N-terminus of Pap1. As part of the CPF complex, Pta1 functions in pre-mRNA cleavage, poly(A) addition, and transcription termination (Mandel et al, 2008; Millevoi and Vagner, 2009). Pta1 also acts as a scaffold in CPF, and different regions of the protein interact with Ssu72, Pti1, and Ysh1, which are CPF subunits (Nedea et al, 2003, Ghazy et al, 2009). We show here that the Pap1 interaction domain on Pta1 is the essential 1-300 amino acids at the N-terminus of the protein. This region of Pta1 is dispensable for polyadenylation in *in vitro* cleavage and polyadenylation assays, and interacts with Ssu72, which might block an inhibitory activity of this domain on 3' end processing (Ghazy et al, 2009). A fragment of Pta1 containing

the first 300 amino acids suppresses polyadenylation in wild type extract, suggesting that this region of Pta1 inhibits Pap1 activity (Figure 3.10B, unpublished data courtesy of Dr. Ghazy) and supporting a direct interaction between this region of Pta1 and Pap1.

Previous data have implicated a regulatory role for Pta1 on Pap1 activity during 3' end processing. Extract depleted of Glc7, a phosphatase that dephosphorylates Pta1, accumulates cleaved, unadenylated products due to defective poly(A) addition (He and Moore, 2005). Addition of Pta1, but not phosphorylated Pta1, to this extract rescued poly(A) addition, which suggests that phosphorylated Pta1 inhibits Pap1 activity (Fig 4.3). As this inhibition must be relieved after cleavage, we propose that there is a transient interaction between Pta1 and Pap1 mediated by phosphorylation.

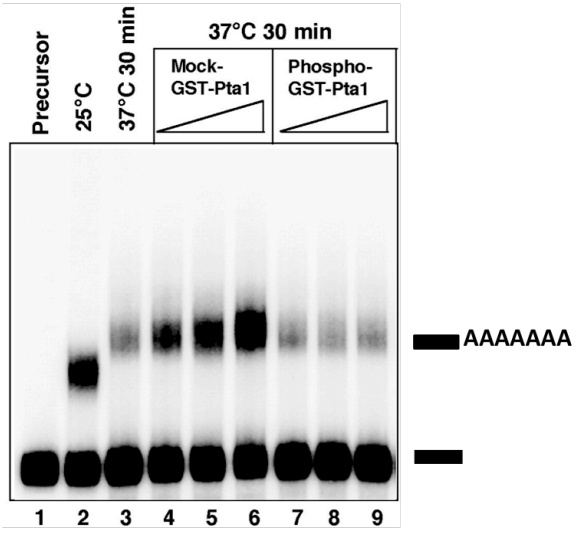


Figure 4.3: Phosphorylated Pta1 inhibits the activity of Pap1 (Courtesy of Xiaoyuan He, He and Moore, 2005).

In this scenario, during the cleavage step of 3' end processing, phosphorylated Pta1 binds to Pap1 and inhibits Pap1 activity. This inhibition is subsequently released by the action of the phosphatase Glc7, which dephosphorylates Pta1. This would explain

the weak *in vitro* interaction observed between Pta1 and Pap1 (Ghazy et al, 2009), and the difficulty in identifying Pap1 interaction by two-hybrid analysis using full length Pta1 (Fig 3.3). This transient interaction alone is apparently not sufficient to hold Pap1 in the CPF complex, as disruption of the Fip1/Yth1 interaction results in loss of Pap1 from the CPF complex (Preker et al, 1995; Barabino et al 2000). However, when the CPF is subjected to prolonged phosphorylation in Glc7-depleted extracts, the interaction of Fip1, but not Pap1, with the CPF complex is weakened (He and Moore, 2005). We propose that the continued phosphorylation of Pta1 seen in these extracts leads to an interaction between Pap1 and Pta1, which is sufficient to maintain Pap1 in the CPF complex in the absence of Fip1.

We also show that the N-terminus of Pap1 interacts with Cft1, which is an essential RNA binding protein in CPF required for cleavage and polyadenylation (Mandel et al, 2008; Millevoi and Vagner, 2009). The mammalian homologue of Cft1, CPSF 160, also interacts with mammalian poly(A) polymerase (PAP), and inhibits both specific and non-specific polyadenylation (Murthy and Manley, 1995). We found that Cft1 caused a reduction in tail length when mixed with Pap1, suggesting that Cft1 also directly restricts the activity of Pap1 in an undefined manner. However, other data suggests that the natural role of Cft1/CPSF160, especially in the context of the larger CPF and CPSF complexes, is likely to enhance poly(A) polymerase activity. In the human CPSF, both CPSF160 and hFip1 interact with PAP, and both proteins bind RNA (Fig. 4.4). CPSF160 contacts the AAUAAA signal sequence, and hFip1 preferentially binds to U-rich sequence elements on the pre-mRNA. Kaufmann et al, 2004 showed that CPSF160, mammalian Fip1 and PAP form a ternary complex *in vitro*, and suggested that CPSF 160 and mammalian Fip1 act together in poly(A) site recognition and recruitment of PAP to the RNA. However, they also found that hFip1 and CPSF160 together were not sufficient for reconstitution of AAUAAA-dependent polyadenylation with PAP *in vitro*.

Therefore, additional proteins must be needed to complement the activities of these CPSF subunits.

Like its human homologue, yeast Fip1 binds to Pap1 in the CPF complex, but unlike its human counterpart, it does not interact directly with the pre-mRNA. In the yeast CPF, another protein with RNA binding properties must hold Pap1 onto the RNA substrate. Cft1, which binds RNA through β-propeller repeats in the middle of the protein, has been proposed to play a role in poly(A) site recognition during yeast 3' end processing (Dichtl et al, 2002), in conjunction with other RNA binding proteins in CPF and in CF I. We propose that the network of RNA interactions, combined with Cft1 interaction at the N-terminus of Pap1, provides specificity and stimulates Pap1 activity, in the absence of a Fip1/RNA interaction in yeast. We believe the interaction of Cft1 with the N-terminus of Pap1 is also transient, since it is not sufficient to tether Pap1 to the CPF in the absence of a Fip1 interaction. A stimulatory role for the interaction of the N-terminus of Pap1 with Cft1 is supported by the observation that deletion of this region prevents specific poly(A) addition in *in vitro* assays, and is lethal in yeast cells (Zhelkovsky et al, 1995). Ghazy et al, 2009 also showed that the N-terminal 300 amino acids of Pta1, which is the Pap1 interaction site, is dispensable for 3' end processing, ruling out a stimulatory function for Pta1 during poly(A) addition. They also showed that the N-terminus of Pta1 interacts with Ssu72, which is a CPF phosphatase essential for cleavage, but not polyadenylation during 3' end processing. This suggests a mechanism for communicating the status of 3' end processing to Pta1, whereby loss of interaction with Ssu72 might be a signal that cleavage is complete, which leads to the dephosphorylation of Pta1, and subsequent polyadenylation. Further work is needed to better understand the mechanism of action of the N-terminus of Pta1. The interactions between Pap1 and Cft1 in yeast cells, and hFip1, CPSF160 and RNA in mammalian cells are summarized in Figure 4.4.

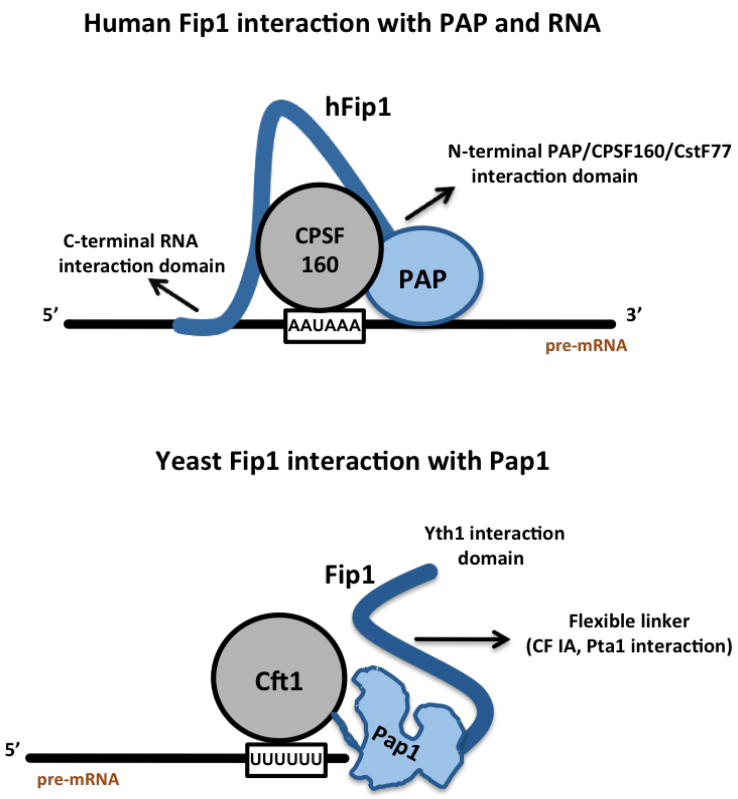


Figure 4.4: Schematic diagram of interactions between Fip1 and CPSF160 and PAP. Mammalian Fip1 contains a charged RNA interaction domain at the C-terminus of the protein. It interacts with CPSF160, CstF77 and PAP at the N-terminus, and also shows some interaction with CPSF160 at the C-terminus (Kaufmann et al, 2004). Yeast Fip1 shows no RNA interaction, and interacts with Pap1 and Yth1 through the previously described interaction domains (Fig 4.2). The flexible linker of yeast Fip1 interacts with CF IA and Pta1. Post-cleavage interaction of Pap1 and Cft1 occurs through the 18 amino acids at the N-terminus of Pap1.

We also report an interaction between the N-terminus of Pap1 and the non-essential yeast proteins, Sub1 and Nab6. Sub1 is a transcriptional co-activator in yeast, and is a homolog of the mammalian PC4 protein (Ge et al. 1994; Henry et al. 1996). It plays a role in yeast 3' end processing and termination through its interaction with Rna15, which is a part of the CF IA factor of the yeast polyadenylation machinery. It has also been reported to interact with Pta1, in competition with Ssu72, and its over-

expression suppresses the processing defects of a *pta1* mutant (He et al, 2003). Furthermore, the Moore lab has shown that while addition of Pta1 can rescue the defect in poly(A) addition caused by Glc7 depletion, it cannot restore physiological poly(A) length in Glc7-depleted extracts (He and Moore, 2005). However, in support of a role for Sub1 in regulating poly(A) addition, our unpublished work shows that addition of Sub1 does lead to normal-length tails on the fraction of RNA that does get polyadenylated in Glc7-depleted extract (Fig 4.5, lanes 7-9).

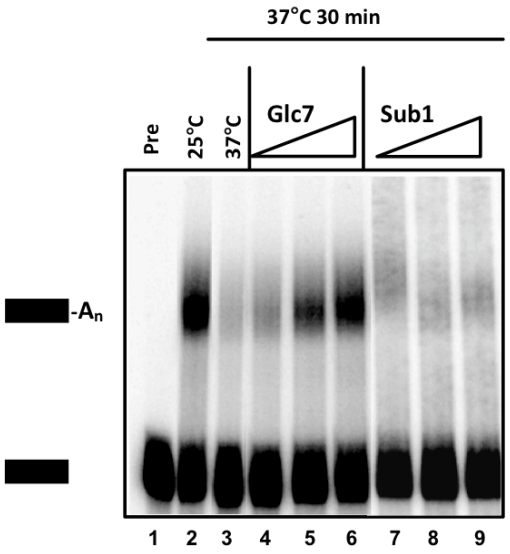


Figure 4.5: Addition of Sub1 to Glc7-depleted extract restores normal tail length (Courtesy of Xiaoyuan He).

Nab6 is an RNA-binding protein that associates with mRNAs encoding cell wall proteins in high-throughput studies (Samantha and Liang, 2003). It also binds to polyadenylated RNA and its over-expression suppresses growth defects caused by the deletion of *RRP6* in yeast cells *(*Abruzzi et al, 2007). Rrp6 is part of the nuclear exosome complex, and is involved in the degradation of improperly processed RNAs, non-coding

RNAs, and byproducts of RNA processing in the nucleus (Allmang et al, 1999; van Hoof et al, 2000; Wyers et al, 2005; Abruzzi et al, 2007). We show that a deletion of both *NAB6* and *SUB1* causes growth defects at 18°C in yeast, while a severe growth defect is observed in yeast lacking the *SUB1* gene at 40°C (Fig 3.9), suggesting a role for both proteins in Pap1 activity during high stress situations in the cell. A role for components of the cleavage and polyadenylation system in stress response in the cell has also been described in *Arabidopsis*, where an orthologue of CPSF30 plays a role in plant response to oxidative stress (Zhang et al, 2008).

The double deletion of *NAB6* and *SUB1* genes show a slight increase in poly(A) tail length in *in vitro* polyadenylation assays (Fig 3.10A), which is indicative of an inhibitory role for both proteins on poly(A) addition in the cell. This is consistent with the observation that over-expression of Nab6 suppresses the growth defect observed in the *Δrrp6* yeast strain, which exhibit longer poly(A) tails *in vivo* and *in vitro* (Abruzzi et al, 2007, and unpublished data from Mathias Poulsen, Jensen lab). Over-expression of Nab6 in this strain might suppress the growth defect by inhibiting Pap1 activity through Nab6 interaction at the N-terminus of Pap1, thereby restoring normal poly(A) tail length. A role for Nab6 in down-regulating polyadenylation is further supported by genetic experiments, which show that a *NAB6* deletion partially rescues growth defects due to the *rna14-3* and *pcf11-2* mutations, which are components of the yeast 3' end processing complex needed for both cleavage and poly(A) addition (unpublished data from Abruzzi et al, 2007). The PAN nuclease, which specifically degrades poly(A), trims tails by about 10-20 n (Brown and Sachs, 1998), and is important for efficient release of mRNA from the site of transcription (Dunn et al, 2005). Perhaps a physiological role for inhibition of Pap1 by Nab6 and Sub1 is to prevent Pap1 from competing with PAN in trimming tails to the mature length.

A redundant role for Nab6 and Sub1 in poly(A) activity is consistent with the fact that some homology (about 12% identity) exists between Sub1 and the Pap1 interacting C-terminal domain of Nab6 (Fig 4.6), and that both proteins interact genetically with components of the High Osmolarity Glycerol (HOG) osmoresponse pathway (Constanzo et al, 2010; Mager and Siderius, 2002; Rosonina et al, 2009). Sub1 facilitates recruitment of RNA polymerase II to osmo-response genes, thereby enabling osmo-response gene transcription during osmotic shock (Rosonina et al, 2009), while *NAB6* interacts genetically with *NBP2*, which is a protein that negatively regulates Hog1 by recruitment of phosphatase Ptc1 to the Pbs2-Hog1 complex (Constanzo et al, 2010).

Our results show that the N-terminus of Pap1 plays a key role in the regulation of Pap1 activity during yeast 3' end processing. We show that this region of Pap1 interacts with Pta1 and Cft1, which are essential components of the CPF complex. In the model discussed below, we propose that Pap1 activity is inhibited by the interaction with Pta1, and stimulated by the subsequent interaction with Cft1. We also show that this region of Pap1 interacts with Nab6 and Sub1, and that a deletion of both proteins yields temperature sensitive phenotypes in yeast. Further work on the N-terminus of Pap1 is needed to understand the role that it plays in Pap1 activity, especially its interaction with key subunits of CPF such as Pta1 and Cft1.

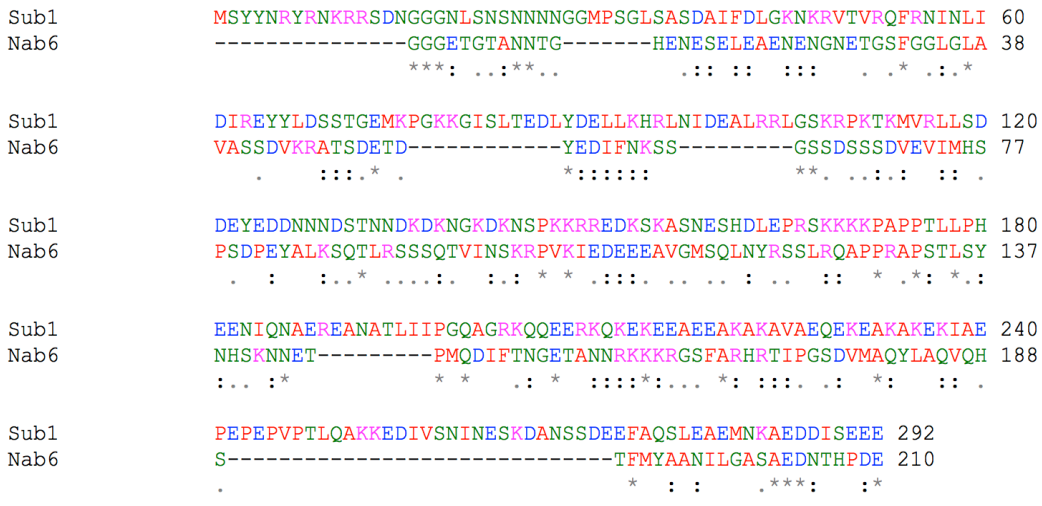


Figure 4.6: Homology between Sub1 and the C-terminal Pap1 interacting domain of Nab6. Clustal W was used for sequence alignment of both full length Sub1, and the 210 amino acid C-terminus of Nab6.

A model for eukaryotic regulation of polyadenylation

The processing of pre-mRNA is clearly of crucial importance to the cell, and disruption of this process can have adverse effects on cell growth and viability. Regulation of polyadenylation is therefore not only important in maintaining efficient cellular RNA synthesis, but it also plays a major role in the ability of the cell to control gene expression. Our studies on the regulation of eukaryotic polyadenylation have identified two new factors that could have a crucial role in the ability of the cell to control polyadenylation. The first major regulatory factor is the essential N-terminus of Pap1 which might sequentially interact with CPF subunits Pta1 and Cft1 during 3' end processing, and the second factor is the flexible linker of Fip1 which is important for Pap1 activity in the cell, and facilitates degradation of Fip1 and the subsequent inhibition of poly(A) addition.

Based on the work described in Chapters 2 and 3, I propose a model for the regulation of Pap1 activity in yeast (Figure 4.7). In our model, Pta1 phosphorylation controlled by the Glc7 phosphatase regulates poly(A) addition during 3' end processing by blocking the critical interaction between Pap1 and Cft1. Pta1 is present in the CPF complex in its phosphorylated form during complex assembly and cleavage, and inhibits Pap1 activity by interacting with the N-terminus of Pap1. Following cleavage, the activity of the phosphatase, Glc7 is stimulated in CPF by an unknown mechanism, and dephosphorylates Pta1, thereby releasing its interaction with the N-terminus of Pap1. The N-terminus of Pap1 is now available to interact with the pre-mRNA bound Cft1, which facilitates Pap1 recognition of the poly(A) site, and initiation of poly(A) addition. In this model, subsequent rephosphorylation of Pta1 restores inhibition of Pap1 activity and helps restrict the length of the poly(A) tails.

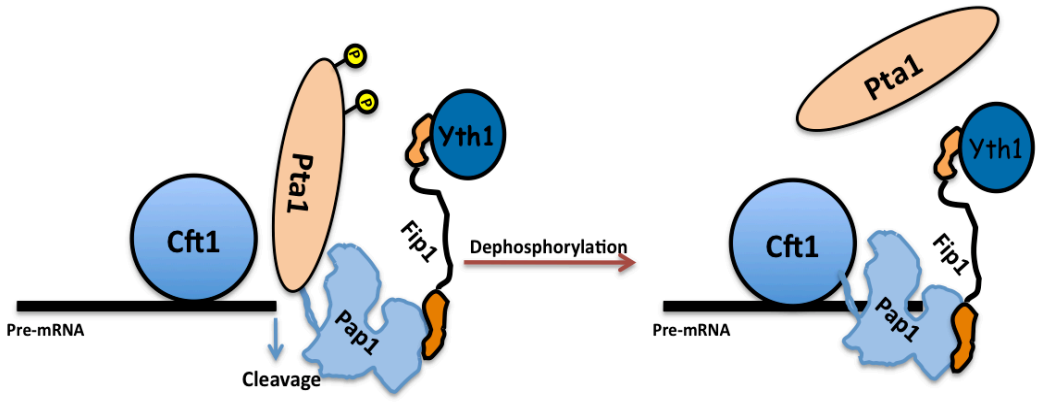


Figure 4.7: Model of Pap1 regulation through interactions at the N-terminus of Pap1. Phosphorylated Pta1 inhibits binding of Pap1 to Cft1. Dephosphorylation following pre-mRNA cleavage relieves the inhibition of Pap1, and facilitates binding of Pap1 and Cft1, and subsequent poly(A) addition. Rephosphorylation of Pta1 may also help terminate poly(A) addition.

The second crucial aspect of our model is the ability of Fip1 and its flexible linker to facilitate the interaction of Pap1 with Pta1 and Cft1. Pap1 is recruited and tethered to the CPF complex through its interaction with Fip1 (Preker et al, 1997). Our studies show that Fip1 has an intrinsic unstructured linker domain (IULD) in the middle of the protein between its Pap1 interaction site and its Yth1 interaction site. We show that the flexibility of this IULD domain of Fip1 allows Pap1 to access the 3' end of the cleaved pre-mRNA during 3' end processing. The flexible linker of Fip1 may allow Pap1 to 'swing' from Pta1, which when phosphorylated would inhibit Pap1 activity, to Cft1, which would stimulate Pap1 activity, perhaps by positioning it to act at the new 3' end formed by cleavage at the poly(A) site (Fig 4.7). The interaction that we have discovered between Pta1 and the Fip1 linker may further restrict the movement of the linker.

Intrinsically disordered proteins like Fip1 are prone to degradation by the proteasome in the cell without a requirement for ubiquitination (Fink, 2005; Tofaris et al, 2001; David et al, 2002), and the degradation of Fip1 is a key step in a form of nuclear surveillance, which inhibits 3' end processing by facilitating Fip1 degradation in the presence of defective mRNP assembly (Saguez et al, 2008). He and Moore, 2005 showed that prolonged depletion of Glc7 in yeast whole cell extracts also results in the reduction of cellular levels of Fip1. They also showed that the phosphorylation of the CPF complex causes a weakened association of Fip1 with the complex, which might expose the Fip1 linker and facilitate its degradation (He and Moore, 2005). Based on our model of Pap1 regulation in Figure 4.7, the depletion of Glc7 or other defects in cellular mRNP assembly will prevent dephosphorylation of Pta1 and result in persistent inhibition of poly(A) tail addition to cleaved pre-mRNA transcripts due to continued binding of Pta1 at the N-terminus of Pap1. This leads to Fip1 degradation, which helps remove aberrant mRNA by shutting down poly(A) addition.

Parallels in the regulation of plant and mammalian PAP activity

Our model is consistent with the regulation of polyadenylation in mammalian cells (Gunderson et al, 1997; 1998; Klein Gunnewiek et al, 2000; Ko and Gunderson, 2002**)**. Proteins like U1A and U1-70K, which are components of the splicing machinery, have been shown to down-regulate production of their individual pre-mRNAs by inhibiting polyadenylation. They directly interact with mammalian PAP, using a PAP regulatory domain (PRD) to inhibit poly(A) addition, but not cleavage (Gunderson et al, 1998; Ko and Gunderson, 2002). The interaction site on PAP of the PRD is a 20 amino acid regulatory domain at the C-terminus of PAP, which is absent in yeast Pap1, but like the 18 amino acids at the N-terminus of Pap1, is sufficient for inhibition of PAP activity (Gunderson et al, 1997). This suggests that the N-terminus of Pap1 is a regulatory domain that plays a similar role in yeast Pap1 as the C-terminal PAP regulatory domain in mammalian PAP, and that the 300 amino acids at the N-terminus of Pta1 functions like a mammalian PAP regulatory domain (PRD). In both systems, inhibition of poly(A) polymerase activity through interaction with these domains affects poly(A) addition, but not cleavage. Our yeast model for the regulation of polyadenylation is different from the reported regulation of PAP activity in plant 3' end processing (Forbes et al, 2005). Human Fip1, unlike yeast Fip1, has an arginine-rich RNA binding domain at its C terminus, which stimulates PAP activity by recruiting it to the pre-mRNA. This domain is sufficient to overcome inhibition of PAP activity by the mammalian CPSF 160 (Kaufmann et al, 2004). In *Arabidopsis*, the N-terminus of AtFip1(V) is sufficient to stimulate the activity of *Arabidopsis* PAP, without a need for the RNA binding domain at the C-terminus of the protein (Forbes et al, 2005). However, analysis of *Arabidopsis* Fip1(V) and mammalian Fip1 shows the presence of highly unstructured regions in both proteins, which suggests that just like yeast Fip1, both *Arabidopsis* Fip1(V) and

mammalian Fip1 might be susceptible to degradation by the proteasome as a means of regulating poly(A) addition (Figure 4.1). Further studies on *Arabidopsis* Fip1(V) and mammalian Fip1 are needed to understand the roles IULDs in these proteins might play in the regulation of poly(A) addition.

Our studies on the regulation of polyadenylation in yeast have shed light on two novel mechanisms that could facilitate the regulation of poly(A) addition in yeast cells. The first mechanism involves the direct inhibition and stimulation of Pap1 activity through interactions at the N-terminus of the protein with other components of the CPF complex that play essential roles in the regulation of Pap1 activity in the cell. In the second mechanism, the flexibility of the Fip1 IULD plays a key role in facilitating the down-regulation of polyadenylation due to an increased susceptibility to destruction by the proteasome. This study enhances our understanding of the functions of two of the key protein factors that make up the yeast 3' end processing complex, but also highlights our lack of complete information on the roles of the individual protein factors involved in eukaryotic 3' end processing. This is due to the difficulty of studying large proteins complexes like the CF I and CPF complexes in yeast, and the CPSF complex in mammals. Obtaining structural information on the components of the mammalian and yeast complexes, and the complexes as a whole will facilitate our understanding of the role of individual proteins and protein/protein interactions in the mechanism and regulation of polyadenylation. Further studies on the function of the N-terminus of Pap1 are needed to better understand the hypothesis that Cft1 interaction stimulates polyadenylation during 3' end processing, while phosphorylated Pta1 inhibits polyadenylation. The role of the interactions of the Pap1 N-terminus with the non-essential yeast proteins Nab6 and Sub1 also needs further investigation. Finally, further work is needed to better understand the regulation of 3' end processing through nuclear

mRNA surveillance, and the role that the unstructured linker domain of Fip1 and other CPF proteins might play in this process.

Chapter 5

Materials and Methods

Chapter 2

Yeast strains and plasmid preparation

All the yeast strains used in this study were derived from strain PJP22 (*MAT∝ leu2-3, 112, ura3-53 trp1 his4 fip1::LEU2* pIA34 [*CEN4 URA3 FIP1*]), which was described previously (Helmling et al, 2001). The Pap1Δ90Fip1::PAP1/FIP1 strain was made by mating a SH1A yeast strain (PJP22 fip1::LEU2, pRS314/FIP1(TRP)) with ATCC 200868 (*MATa ade2Δ::hisG his3Δ200 leu2Δ0 lys2Δ0 met15Δ0 trp1Δ63*) to create an appropriately marked strain. The cells were then sporulated, dissected, and selected for a Leu+, Trp+, Ura-, His- phenotype, and then mated with W303 (*MAT*α {*leu2-3, 112 trp1-1 can1-100 ura3-1 ade2-1 his3-11, 15*} [*phi*$^{+}$]). The resulting diploid cells were transformed with a SacI/PvuII Pap1 fragment in which the portions from SnaBI to StuI were replaced with a *HIS3* gene to disrupt the *PAP1* gene. Disruption of the *PAP1* gene was confirmed by Southern blot analysis. These diploid cells, which contained a disruption of one copy of *FIP1* and one copy of *PAP1,* were transformed with the pHCP50/PAP1Δ90FIP1 (*URA3*) plasmid, sporulated and dissected to identify haploid cells with the knockouts of *FIP1* and *PAP1* covered by the pHCP50/PAP1Δ90FIP1 (*URA3*) plasmid. An isogenic wild type strain was made by transforming the SH1A strain with two plasmids containing wild type *FIP1* and *PAP1* (YCp50/PAP1 and RS314/FIP1), and plating on selective medium containing 5-FOA. The yeast colonies that grew on this medium were then tested on medium lacking tryptophan and leucine, which confirmed the presence of the two plasmids. ΔNFip1-RPA constructs were made by introducing a 72-amino acid section of RPA 70 into p314ΔNFip1 (Helmling et al, 2001). The RPA sequence was obtained by PCR from the p11dtRPA plasmid, which was a kind gift from

P. Bullock at Tufts University. The three RPA strains were made sequentially from 1-3 using 3 sets of primers for PCR on the p11dtRPA plasmid. They are as follows: For ΔNFip1-RPAx1, RPA1-5': CGGCGGGTAGCTAGC GGAGTGAAGATTGGCAATCCAGTG, and RPA1-3': ATAATCACTTAAGTTGGCCC CCGGTTGCCTCCACATATGCCATGGCTGTGTTCCCC CAGAAGTGTG. For ΔNFip1-RPAx2, RPA2-5': CGGCGGGTACCATGGGGAGTGAAGA TTGGCAATCCAGTG, RPA2-3':ATAATCACTTAAGTTGGCCCCCGGTTGCCTCCACTG TGTTCCCCCAGAAGTGTG. For ΔNFip1-RPAx3, RPA3-5':CGGCGGGTACATATGGGA GTGAAGATTGGCAATCCAGTG and RPA3-3': ATAATCACTTAAGTTGGCCCCCGGTT GCCTCCACTGTGTTCCCCCAG AAGTGTG. ΔNFip1-RPAx1 was made by ligation, after cutting p314ΔNFip1 and the PCR #1 with NheI and NdeI. ΔNFip1-RPAx3 was made by ligation after cutting ΔNFip1-RPAx1 and the PCR #3 with NdeI and AflII. ΔNFip1-RPAx2 was made by ligation after cutting ΔNFip1-RPAx3 and PCR #2 with NcoI and NdeI. The ΔNFip1 linker-deletion construct was made by introducing an NheI site at position 340 of the ΔNFip1-RPAx1 sequence by site-directed mutagenesis using the Quick-change protocol (Stratagene, La Jolla, CA). The plasmid was cut with NheI, separated from the released linker region fragment by gel electrophoresis and ligated to itself.

The ΔNFip1-RPA yeast strains were obtained by plasmid shuffling. Plasmids containing variations of ΔNFip1-RPA were introduced into PJP22 by transformation using the lithium acetate method (Becker and Guarente, 1991), and plated on complete medium lacking uracil and tryptophan. Transformants were isolated, grown in liquid media, and plated on synthetic complete medium containing 1 mg/ml of 5-fluoroorotic acid (5-FOA). Colonies growing on 5-FOA were re-examined to confirm loss of the *URA3* plasmid, and the presence of the correct construct was verified by Western

blotting using polyclonal Fip1 antibody as described (Helmling et al, 2001). Monoclonal Pap1 antibody at a 1:100 dilution was used for detecting Pap1 (Kessler et al, 1995). To compare growth rates, cells were grown overnight at 30°C in liquid media. Cell densities were normalized by dilution with fresh media, and equal volumes of 10-fold serial dilutions were spotted on YPD plates or on YPD plates supplemented with 1.5% formamide, and incubated as indicated in the figure legends.

Yeast poly(A) polymerase tethered to bacteriophage λ N protein (Pap_Np) was constructed using the PAP1Δ10-His_6 plasmid (Balbo et al., 2005). The PAP1Δ10-His_6 plasmid encodes yeast Pap1 with a 32 amino acid truncation at the C-terminal end and a C-terminal His_6 tag for purification purposes. Deletion of the C-terminus and addition of the His_6 tag does not compromise enzymatic activity and results in a more soluble, better behaved protein than the wild-type enzyme. PCR was used to generate a bacteriophage λ N protein DNA with a C-terminal His_6 tag, which contained the BlpI/BpuI restriction site at the 5' end, and the XhoI restriction site at the 3' end. The primers used for the PCR were 5' TATTATCTCGAGGCCGCCATG GATGCACAAACACGCCGCCGCG- 3' (NPFXX1) and, 5' AATATTGCTCAGCGGAGGTGGT GTGGTGGTGGTGGGGATTTGCTGCT TTCCAT TGAGCCTG-3' (NPRXX1). The PCR product and the PAP1Δ10- His_6 plasmid were cleaved with Blp1 and XhoI, and ligated to create the PAP1_N-His_6 plasmid. Sequencing confirmed the presence of the *PAP1* sequence followed by a short double alanine linker, the N protein and a His_6 tag (Pap_Np).

In vitro mRNA 3' end processing assays

Preparation of yeast whole cell extract was done as described (Zhao et al.,

1999). After ammonium sulfate precipitation, the protein was resuspended in 200 μl of Buffer D (20 mM Tris-HCl, ph 7.9, 50 mM KCl, 0.2 mm EDTA, 0.5 mM DTT, 20% glycerol, 2 μM pepstatin A, 0.6 μM leupeptin), and dialyzed twice against 2 liters of the same buffer for 2 hours, and then overnight. Processing assays using full length and pre-cleaved ^{32}P-labeled RNA were done as described (Helmling et al., 2001). Immunoprecipitation using polyclonal Fip1 antibody was performed as described (He et al., 2005), and the results were analyzed by Western blot using monoclonal Pap1 and Pta1 antibody at a 1:10,000 dilution.

In vivo RNA analysis

For the analysis of poly(A) tails, ΔNFip1 yeast strains were grown in YPD overnight to an approximate OD_{600} of 26.0. The yeast strains were diluted using pre-warmed YPD to an OD of 0.1, and grown at 30°C for approximately 6 hrs. Cells were harvested by centrifugation at OD 0.8 and stored at -20°C. For analysis at 39°C, the diluted cells were grown at 30°C until an O.D of 5, and then at 39°C for approximately 3 hrs, and then harvested. Total yeast RNA was isolated from cells using a modification of the hot phenol method (Köhrer and Domdey, 1991). Labeling, RNAse digestion, and visualization of poly(A) tails were performed as described previously (Preker et al, 1995, Kessler et al, 1997).

Protein expression and purification

The Pap_Np- His_6 plasmid was transformed into *E. coli* BL-21 (DE3) cells, over-expressed and purified as described for PAP1Δ10 (Balbo et al., 2005). The purified protein was concentrated using a Millipore 30kDa centrifugation filter. A buffer (40 ml)

consisting of 15% glycerol, 30 mM potassium phosphate of pH 6.8, 20 mM NaCl, and 10 mM 2-mercaptoethanol was used to wash and replace the previous buffer. Pap_Np was further concentrated using a 10-kDa Millipore centrifugation filter at 12,000 rpm until the volume was 300 µl. Small aliquots of 50 µl were frozen in liquid nitrogen and stored at -80°C.

Protein expression and in vitro protein-protein interaction assays

The Fip1 1-105 amino acid and the Fip1 1-206 amino constructs have been described previously (Helmling et al, 2001). All protein expression was performed in *Escherichia coli* Rosetta (DE3), and expression and purification was performed as described (Zhelkovsky et al, 1995; Chung et al, 2003). The BglG-MBP protein was a kind gift from Dr. Andrew Wright in the Microbiology Department at Tufts University. The purified CF IA complex was also a kind gift from Dr. James Gordon in the Biochemistry Department at Tufts University. GST pull downs were performed using a modification of the protocol described by Ghazy et al, 2009. GST tagged proteins were incubated with previously washed glutathione-Sepharose beads in 200 µl of buffer IP-150 (10 mM Tris [pH 7.9], 150 mM NaCl, 0.25M EDTA, 10% Glycerol, 1M DTT, 1% NP-40 and protease inhibitors) for 2 hours at 4°C, with gentle shaking. After incubation, the beads were pelleted at slow speed and washed four times with buffer IP-150. For pull downs from whole cell extract, protein from 50 µl of beads eluted with buffer IP-150 containing 50 mM glutathione at 4°C for 1 hour. The remaining samples were incubated with either whole cell extract or the recombinant protein of interest in buffer IP-150 for 2 hours at 4°C. After incubation, the beads were pelleted at slow speed and washed four times with buffer IP-150. To elute the proteins, beads were incubated with buffer IP-150 containing 50 mm glutathione at 4°C for 1 hour with gentle shaking. The eluted samples were either

frozen in liquid nitrogen and stored at -80°C or run on SDS PAGE gels, and analyzed by western blot. The CF IA complex was reconstituted after incubation in buffer IP-300 on ice for 30 minutes, and incubated with MBP beads at 4°C for 2 hours. The incubation was performed in buffer IP-300, which contains all the components of buffer IP 150 except for 300 mM NaCl. After incubation, the beads were pelleted at slow speed and washed four times with buffer IP-300. Pull downs with the wild type Fip1 and Fip1 truncations were performed for 2 hrs at 4°C with gentle shaking. After incubation, the beads were pelleted at slow speed and washed four times with buffer IP-150. To elute the proteins, the sample containing the beads were resuspended in 1X SDS buffer, boiled for 5 minutes and pelleted. The supernatant which contained the eluted proteins were beads were either frozen in liquid nitrogen and stored at -80°C or run on SDS PAGE gels, and analyzed by western blot.

Western blot analyses

Proteins expression and protein-protein interactions were analyzed using Western blotting. Extracts used for western blots were made by growing the yeast in 10 ml of YPD overnight. The sample was then pelleted, and resuspended in 300 µl of lysis buffer (50 mm HEPES at pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 10% glycerol and protease inhibitors) in a 1.5 ml eppendorf tube. 300 µl of glass beads were added sample, and vortexed for 5 minutes at maximum speed 4°C. After vortexing, the samples were pelleted at maximum speed, and the supernatant containing the protein was removed, and used for western analysis. Extracts were resolved on a SDS-10% polyacrylamide gel, and proteins were transferred onto Immobilon-P polyvinylidene difluoride (Millipore) membranes. A 1:100 dilution of monoclonal mouse Pap1 antibody was used for analyzing Pap1 samples. For Fip1 analysis, a 1:100,000 dilution of

monoclonal mouse Fip1 antibody was used. The rest of the antibody dilutions are as follows: Cft1 – 1:6000 dilution of polyclonal rabbit antibody; Brr5 – 1:6000 dilution of polyclonal rabbit antibody; Pta1 – 1:16000 dilution of monoclonal mouse antibody; Cft1 – 1:6000 dilution of polyclonal rabbit antibody; Yth1 – 1:6000 dilution of polyclonal rabbit antibody; Ssu72 – 1:6000 dilution of polyclonal rabbit antibody. Glutathione *S*-transferase (GST) antibody was obtained from BD Biosciences-Pharmingen, and GAL4-DBD (RK5C1) antibody was obtained from Santa Cruz Biotechnology.

Enzyme activity measurement

The specific activity of the Pap_Np was determined as described by Balbo et al, 2005. Briefly, a polyadenylation assay that incorporated tritium AMP for adenylation was used to monitor the activity of the Pap_Np at 30°C. Reaction reagents with the exception of Pap_Np protein were mixed to a total volume of 18 μL and equilibrated to reaction temperature for 20 minutes. 150 ml of chilled 15% trichloroacetic acid was applied to halt aliquots of the reaction and precipitate the RNA products at intervals of 1 minute for the first 4 minutes. The reactions were incubated on ice for 30 minutes, and centrifuged at maximum speed for 20 minutes at 4°C. The pellets were washed with cold 15% TCA twice and re-dissolved in 150 ml 1% SDS and 10 mM NaOH. Liquid scintillation counting was used to determine the amount of tritium used in each reaction.

Polyadenylation kinetics

Box B hairpin loop RNA (Fig 1A) were obtained from Dharmacon, Inc (Thermo Scientific) and de-protected according to directions. All reactions were conducted at

30°C using an assay that measured incorporation of [$\propto$-^{32}p] ATP (PerkinElmer) into the growing poly(A) tail. The components of the reaction were equilibrated at room temperature for 15 minutes, and each sample contained 2 μl of 5x T3/T7 buffer solution (Promega), 2 μl of BSA (10 mg/ml), 0.2 μl MgCl2 (1 M), 0.2 μl spermidine (0.1 M), 1 mM ATP, 20 μCi of [$\propto$-^{32}P] ATP and autoclaved distilled water for a total of 10 μl. The reaction was initiated by adding a 10 μl mix containing Pap1 or Pap_Np (50 ng) and the RNA substrates (0.015 μM). After incubation for 20 minutes at 30°C, the reaction was stopped with 150 μl of ice cold 15% TCA. To determine the rate of reaction, the reactions were stopped after 1 minute, 5 minutes and 20 minutes. A control reaction that contained the same components without poly(A) polymerase was performed to determine the background signal. Samples were then incubated on ice for 30 minutes and centrifuged at 16,000 rpm for 30 minutes at 4°C. The precipitated pellet was washed with ice cold 15% TCA and dissolved in 10 mM NaOH and 0.1% SDS. Using liquid scintillation counting, [$\propto$-^{32}P] levels were measured in counts per minute. The experiments were replicated twice, and the final results were obtained by subtracting the background signal.

Chapter 3

Yeast strains and cultures

The following yeast strains were used for this study. BY 4742 (*MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*) was obtained from the American Type Culture Collection (ATCC), W303 (*MATa leu2-3 112 trp1-1 can1-100 ura3-1 ade2-1 his3-11 15 [phi+]*), and PJ69-4A (*MAT***a** leu2-3,112 ura3-52 trp1-901 his3-200 gal4Δgal80Δ GAL-ADE2

lys2::GAL1-HIS3 met2::GAL7-LacZ) used for two-hybrid analysis was obtained from Dr. Craig (James et al, 1996). For the two-hybrid constructs, the *PAP1* wild-type gene and the *ΔN PAP1* truncation were described previously (Zhelkovsky et al, 1998). To generate the 18-amino-acid N-terminal peptide of Pap1 for two-hybrid, a PCR product containing the peptide was generated from the pJPAP1 plasmid (Zhelkovsky et al, 1995), and ligated in frame into the *GAL4* binding domain vector using the Sma I restriction site at the N-terminus and a Cla I restriction site at the C-terminus. To make the *GAL4* binding domain-CTD construct, a PCR product containing the *CTD* gene was ligated in frame into the *GAL4* binding domain vector using the Sma I restriction site at the N-terminus and a Cla I restriction site at the C-terminus. The Pta1-*GAL4* binding domain plasmid was described previously (Ghazy et al, 2009), and the Cft1- *GAL4* activation domain plasmid was a kind gift from Dr. Keller, and was described previously (Dichtl et al, 2002). *SUB1* knockout was performed in the W303 yeast background, with a *NAB6* gene deletion performed by Dr. Zhelkovsky (*Δnab6* W303). The *SUB1* gene was replaced using a PCR product obtained from amplifying the region between 150 nucleotides upstream and downstream of the *SUB1* ORF in the *SUB1* ORF deletion yeast strain obtained from the American Type Culture Collection (ATCC). The deletion strain was in the W303 background, and contained a *KANMX* replacement of the *SUB1* gene. The *NAB6* knockout was performed in the *SUB1* ORF deletion yeast strain on a BY 4742 background (*Δsub1* BY4742). The *NAB6* gene was replaced using a PCR product obtained from amplifying the region between 300 nucleotides upstream and downstream of the *NAB6* ORF in the *NAB6* ORF deletion yeast strain obtained from Dr. Zhelkovsky, which contained a *NAB6* gene replaced with *HIS3*. Confirmation of the gene deletion was done by observing a shift in size of the PCR product from genomic DNA using primers around either the *NAB6* or *SUB1* genes in yeast (Figure 12A, lanes 3-5 for *NAB6* and Figure 12B, lane 8 for *SUB1*). Growth properties of yeast strains were

analyzed by growing the strains in liquid yeast extract-peptone-dextrose (YPD) medium at room temperature to an optical density at 600 nm of 1.0, spotting 5 µl of 10-fold serial dilutions onto YPD plates, and incubating the plates for 3 days at 13°C 16°C 18°C, 24°C, 30°C, 37°C, or 40°C.

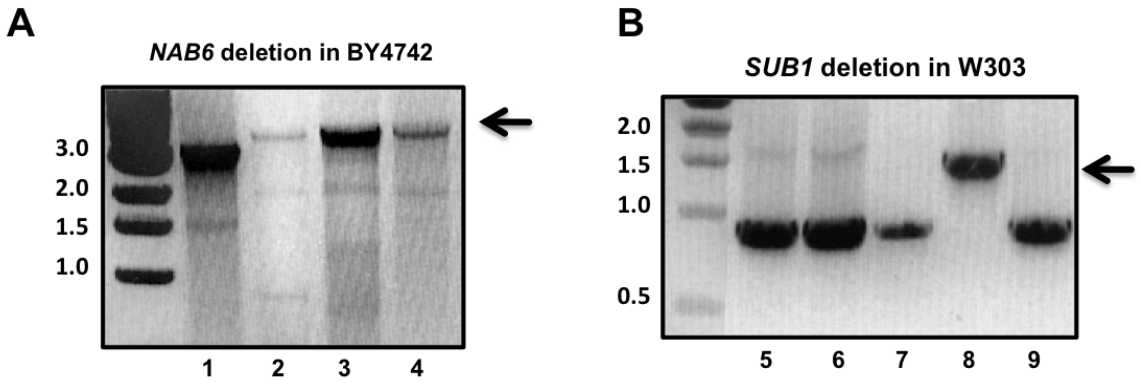


Figure 3.12: Confirmation of gene deletion. (A) PCR product obtained from PCR performed on genomic DNA isolated from BY4742 transformants. The arrow indicates the shift in size caused by insertion of the *HIS3* gene at the *NAB6* locus (B). PCR product obtained from PCR performed on genomic DNA isolated from W303 transformants. The arrow indicates the shift in size caused by insertion of the *KANMX* gene at the *SUB1* locus. Lanes 1 and 5 indicates strains without the deletion.

Two-hybrid analysis

Two-hybrid analysis was performed as described (James et al, 1996). Briefly, two hybrid constructs were transformed into PJ69-4A yeast strain. Transformants were selected on complete medium lacking leucine and tryptophan (CM−LT) to ensure that both the *GAL4* activation domain and *GAL4* binding domain plasmids were present. Protein-protein interactions were scored by the ability of cells to grow on complete medium lacking leucine, tryptophan, and histidine (CM−LTH). Tenfold serial dilution of

the cultures was made, and 5 µl of each dilution was plated onto CM−LT as a growth control or CM−LTH for scoring two-hybrid interactions. Plates were incubated at 30°C or 37°C for 2 to 3 days.

Peptide Column

The UltraLink EDC/DADPA Immobilization Kit (Pierce) was used for peptide column analysis. The Pap1 Peptide contained 18 amino acids, MSSQKVFGITGPVSTVGA (synthesized by the Tufts Core Facility), while the Npl3 control peptide had 11 amino acids: RGGYDSPRGGY (BCMP/HMS Biopolymers Lab). Peptides were normalized by OD at 280 nm. The peptides were coupled to the columns according to directions at room temperature. For pull downs, 100 ng of nuclear extract was incubated with each of the conjugated peptides for 2 hours at room temperature in IPP-150 buffer (10 mM Tris [pH 7.9], 150 mM NaCl, 1 mM MgOAc, 2 mM $CaCl_2$, 0.1% NP-40, 1 mM DTT and the protease inhibitors PMSF, leupeptin, antipain and pepstatin-A), and washed using 12 column volumes of buffer IPP-150. The column was drained by gravity, and eluted using glycine buffer (100 mM glycine, pH 3). Eluted samples were either frozen at -20°C, or run on SDS PAGE gel, and analyzed by silver stain and mass spectroscopy.

Plasmid preparation

All yeast expression constructs were made using the YCp50 *URA* vector. The Pap1 yeast expression vector has been described (Zhelkovsky et al, 1995; 1998). To make the expression vector, a PCR product containing *SUB1*, generated from yeast genomic DNA, was ligated into YCp50 *URA* vector using the Sac I restriction site at

sequence encoding the N-terminus of Sub1, and the Pvu II restriction site at the C-terminus of Sub1. For the *NAB6* expression vector, a PCR product containing *NAB6*, generated from yeast genomic DNA, was ligated into YCp50 *URA3* vector using the Sac I restriction site at the N-terminus of Nab6, and the Pvu II restriction site at the C-terminus of Nab6. For protein expression in *E. coli*, the pJPAP1 expression vector was used to make the Pap1-GST, ΔN-Pap1-GST, Pap1-His6, ΔN-Pap1-His6 and Nab6-V5-His6 expression plasmids used in this study (Zhelkovsky et al, 1995). The Pap1-His6, ΔN-Pap1-His 6 expression plasmids are described previously (Zhelkovsky et al, 1995). To make the Pap1-GST plasmid, a GST tag at the C-terminus was introduced into the pJPAP1 expression plasmid. For the ΔN-Pap1-GST, a PCR product containing a Pap1 truncation without the 16 amino acid N-terminus of Pap1 was generated, with a Sac I restriction site at the N-terminus, and a Pvu II restriction site at the C-terminus, and introduced into pJPAP1. To add the GST tag at the C-terminus of both constructs, a PCR product containing GST with a Pvu II restriction site at both ends was generated from pGEX-6P-2 GST vector, and ligated in frame into the Pvu II site at the C-terminus of both constructs. Sequencing was used to verify both constructs. The Sub1 plasmid was described previously (He et al, 2003). For the Nab6 plasmid, a PCR product containing the Pap1-interacting C-terminal domain of Nab6 and a V5 tag at the N-terminus of the protein, generated from yeast genomic DNA was ligated into a pJPAP1 vector, from which the *PAP1* sequence has been removed using a Sac I site introduced at the N-terminus and a Pvu II site introduced at the C-terminus. The Pta1 and Pta1 truncation expression plasmids were described previously (Ghazy et al, 2009), and the Cft1-GST construct was a kind gift from Dr. Keller, and was described previously (Dichtl et al, 2002).

Protein expression and in vitro protein-protein interaction assays

All protein expression was performed in *Escherichia coli* Rosetta (DE3), and expression and purification was performed as described (Zhelkovsky et al, 1995; Chung et al, 2003). GST pull downs were performed using a modification of the protocol described by Ghazy et al, 2009. GST tagged proteins were incubated with previously washed glutathione-Sepharose beads in 200 µl of buffer IP-150 for 2 hours at 4°C, with gentle shaking. After incubation, the beads were pelleted at slow speed and washed four times with buffer IP-150. For pull downs, protein bound to 50 µl of was incubated with either whole cell extract or the recombinant protein of interest in buffer IP 150 for 2 hours at 4°C. After incubation, the beads were pelleted at slow speed and washed four times with buffer IP-150. To elute the proteins, beads were incubated with buffer IP-150 containing 50 mM glutathione at 4°C for 1 hour with gentle shaking. The eluted samples were either frozen in liquid nitrogen and stored at -80°C or run on SDS PAGE gels, and analyzed by either silver stain or western blot.

Western blot analyses

Proteins expression and protein-protein interactions were analyzed using Western blotting. Extracts used for western blots were made by growing the yeast in 10 ml of YPD overnight. The sample was then pelleted, and resuspended in 300 µl of lysis buffer (50 mm HEPES at pH 7.4, 100 mM potassium acetate, 2 mM magnesium acetate, 10% glycerol and protease inhibitors) in a 1.5 ml eppendorf tube. 300 µl of glass beads were added sample, and vortexed for 5 minutes at maximum speed 4°C. After vortexing, the samples were pelleted at maximum speed, and the supernatant containing the protein was removed, and used for western analysis. Extracts were resolved on a SDS-10% polyacrylamide gel, and proteins were transferred onto Immobilon-P polyvinylidene

difluoride (Millipore) membranes. A 1:100 dilution of monoclonal Pap1 antibody was used for analyzing Pap1 samples. For Sub1 and Nab6 detections, V5 antibody (Invitrogen) at a 1:5000 dilution was used. Glutathione *S*-transferase (GST) antibody was obtained from BD Biosciences-Pharmingen, and GAL4-DBD (RK5C1) antibody was obtained from Santa Cruz Biotechnology.

In vitro mRNA 3' end processing assays

Preparation of yeast whole cell extract was done as described (Zhao et al., 1999). After ammonium sulfate precipitation, the protein was resuspended in 200 μl of Buffer D (20 mM Tris-HCl, ph 7.9, 50 mM KCl, 0.2 mm EDTA, 0.5 mM DTT, 20% glycerol, 2 μM pepstatin A, 0.6 μM leupeptin), and dialyzed twice against 2 liters of the same buffer for 2 hours, and then overnight. Processing assays using full length and pre-cleaved ^{32}P-labeled RNA were done as described (Helmling et al., 2001).

Enzyme activity measurement

The specific activity of the ΔN-Pap1 was determined as described by Balbo et al, 2005. Briefly, a polyadenylation assay that incorporated tritium AMP for adenylation was used to monitor the activity of the ΔN-Pap1 at 30°C. Reaction reagents with the exception of ΔN-Pap1 were mixed to a total volume of 18 μL and equilibrated to reaction temperature for 20 minutes. 150 μl of chilled 15% trichloroacetic acid was applied to halt aliquots of the reaction and precipitate the RNA products at intervals of 1 minute for the first 4 minutes. The reactions were incubated on ice for 30 minutes, and centrifuged at maximum speed for 20 minutes at 4°C. The pellets were washed with cold 15% TCA

twice and re-dissolved in 150 μl 1% SDS and 10 mM NaOH. Liquid scintillation counting was used to determine the amount of tritium used in each reaction.

Primers used in this study

For Pap-GST, 5' primer:

CGACGACGAATGAGCTCTCAAAAGGTTTTTGGTATTACTGGACCTGTT.

For the ΔPap1-GST,

CGACGACGAATGAGCTCAGCAGCAGAAAATAAATTAAATGATAGTTTA

3' primer:

CAGCAGCAGCAGCTGTGAGTGGTGGTGGTGGTGGTGCTCGAGATTCTTCCTTT.

For the C-terminal truncation, 5' primer:

CGACGACGAATGAGCTCTCAAAAGGTTTTTGGTATTACTGGACCTGTT.

3' primer:

CAGCAGCAGCAGCTGTTTTTCGAATAAATTGGCCCAGGACTTCTTATT.

For Sub1 integration into the YCp50 *URA3* plasmid,

5' primer:

CGACGACGACGACGAATGAGCTCTTCATATTACAACAGGTATAGGAACAAAA

3' primer:

CGACGACGACGAAACAGCTGCTTATTCTTCTTCACTTATGTCGTCTTCAGC

For the Nab6-V5 construct

5' primer:

CGGCGGCGGGAATTCGGTAAACCAATTCCAAATCCATTGTTGGGTTTGGATTCTAC

TGGCGGTGGTGAAACAGGAACT

3' primer:

CGGCGGCGGAAGCTTCTCGTCAGGATGCGTGTTGTCTTCC

For the CTD two-hybrid,

5’ primer:

GCGGCGGCTCCCGGGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT

3’ primer:

CGGCGGCGGATCGATTCATCTGGAATTTTCATTTTCATTATGCTTTT

For the Sub1 knockout,

5’ primer: (-150—120 bp upstream of ORF)

ATAGGAAGACAAGAGAAGAGGTTGTAAACG

3’ primer: (+120-150 bp downstream of ORF)

TATAAATCACATATTTGATTATATGAATTG

For testing the knockout of Sub1,

5’ primer:

TACACATCAATTTTTCGACATATATACAAACACAAGCGCT

3’ primer:

TGGAAGACGTTGACATAAGCAAGCTCAACTTCCAGGACTA

For making Nab6 Knockout,

5’ primer: (300 bp upstream of ORF)

TCACTTTCACCTTATTCACAGCTTTTACGT

3’ primer: (300 bp downstream of ORF)

GAAACGAAGTAGAAAAAAATTAACAGGTTT

For testing the knockout of Nab6

5’ primer:

TTCACGTCGTTCGTTTATATTATCAAGGAA

3’ primer:

AAAAGGGAGAAACACCTTGGCAGGCCGCCT

References

Abe, A., Hiraoka, Y. and Fukasawa, T. 1990. Signal sequence for generation of mRNA 3' end in the *Saccharomyces cerevisiae GAL7* gene. *EMBO J*, **9**, 3691-3697.

Abruzzi K, Denome S, Olsen JR, Assenholt J, Haaning LL, Jensen TH, Rosbash M. 2007. A novel plasmid-based microarray screen identifies suppressors of rrp6Delta in Saccharomyces cerevisiae. *Mol Cell Biol.* **27**: 1044-55.

Abuodeh, R., Wei, H. and Yuan, D. 1998. Effect of upstream RNA processing on selection of mu S versus mu M poly(A) sites. *Nucleic Acids Res*, **26**, 5417-5424.

Adam, S.A., Nakagawa, T., Swanson, M.S., Woodruff, T.K. and Dreyfuss, G. 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol Cell Biol*, **6**, 2932-2943.

Addepalli B, Hunt AG. 2008. The interaction between two Arabidopsis polyadenylation factor subunits involves an evolutionarily-conserved motif and has implications for the assembly and function of the polyadenylation complex. *Protein Pept Lett.* **15**: 76-88.

Allen, N.P., Huang, L., Burlingame, A. and Rexach, M. 2001. Proteomic analysis of nucleoporin interacting proteins. *J Biol Chem*, **276**, 29268-29274.

Allmang, C., J. Kufel, G. Chanfreau, P. Mitchell, E. Petfalski, and D. Tollervey. 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J*. **18**:5399-5410.

Amrani N, M Minet, M Le Gouar, F Lacroute, F Wyers. 1997. Yeast Pab1 interacts with Rna15 and participates in the control of the poly(A) tail length in vitro. *Mol. Cell. Biol.* 17. 3694-3701.

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.

Anderson, J.T., Paddy, M.R. and Swanson, M.S. 1993. PUB1 is a major nuclear and cytoplasmic polyadenylated RNA-binding protein in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **13**, 6102-6113.

Ansari, A. and Hampsey, M. 2005. A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. *Genes Dev*, **19**, 2969-2978.

Aravind, L. and Koonin, E.V. 1999. DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res*, **27**, 1609-1618.

Arhin G, Boots M, Bagga P. Milcarek C. and Wilusz,J. 2002. Downstream sequence elements with different affinities for the hnRNP H/H_ protein influence the processing efficiency of mammalian polyadenylation signals. *Nucleic Acid Res.* **30**: 1842–1850.

Balbo PB, Bohm A. 2007. Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure.* **15**: 1117-31.

Balbo PB, Bohm A. 2009. Proton transfer in the mechanism of polyadenylate polymerase. *Biochem J.* **420**: 229-38.

Balbo PB, Meinke G, and Bohm A. 2005. Kinetic studies of yeast poly(A) polymerase indicate an induced fit mechanism for nucleotide specificity. *Biochemistry.* **44**: 7777-86.

Balbo PB, Toth J, Bohm A. 2007. X-ray crystallographic and steady state fluorescence characterization of the protein dynamics of yeast polyadenylate polymerase. *J Mol Biol.* **366**: 1401-15.

Barabino, S.M.L., Hubner, W., Jenny, A., Minvielle-Sebastia, L. and Keller, W. 1997. The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev*, **11**, 1703-1716.

Barabino, S.M., Ohnacker, M. and Keller, W. 2000. Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs. *EMBO J*, **19**, 3778-3787.

Bard J, Zhelkovsky AM, Helmling S, Earnest TN, Moore CL, and Bohm A. 2000. Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. *Science.* **289**: 1346-9.

Barilla, 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*. *PNAS*, **98**, 445-450.

Barnard, D. C., Ryan, K., Manley, J. L. and Richter, J. D. 2004. Symplekin and xGLD-2 are required for CPEB mediated cytoplasmic polyadenylation. Cell 119, 641–651.

Becker D, Guarente L. 1991. High efficiency transformation of yeast by electroporation. *Methods Enzymol.* **194**:183-187

Beelman, C.A. and Parker, R. 1995. Degradation of mRNA in eukaryotes. *Cell*, **81**, 179-183.

Beyer, K., Dandekar, T. and Keller, W. 1997. RNA ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3'-end processing of pre-mRNA. *J Biol Chem*, **272**, 26769-26779.

Bienroth S, Keller W and Wahle E. 1993. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J.* **12**: 2800-2808.

Bienroth, S., Wahle, E., Suter-Crazzolara, C. and Keller, W. 1991. Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. *J. Biol. Chem.*, **266**, 19768-19776.

Borman, A.M., Michel, Y.M. and Kean, K.M. 2000. Biochemical characterisation of cap-poly(A) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5'- end. *Nucleic Acids Res*, **28**, 4068-4075.

Brackenridge S, Proudfoot N. J. 2000. Recruitment of a basal polyadenylation factor by the upstream sequence element of the human lamin B2 polyadenylation signal. *Mol. Cell. Biol*. **20:** 2660–2669.

Brown CE, Sachs AB. 1998. Poly(A) Tail Length Control in *Saccharomyces cerevisiae* Occurs by Message-Specific Deadenylation. *Mol Cell Biol.* **18.** 6548-6559.

Brown CE, Tarun, S.Z., Jr., Boeck, R. and Sachs AB. 1996. PAN3 encodes a subunit of the Pab1p-dependent poly(A) nuclease in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **16**, 5744-5753.

Brown, K.M. and Gilmartin, G.M. 2003. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. *Mol Cell*, **12**, 1467-1476.

Buratowski S. 2005. Connections between mRNA 3′ end processing and transcription termination, *Curr. Opin. Cell Biol.* **17:** 257–261.

Buratowski S. 2009. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* 36: 541–46

Butler, J.S. and Platt, T. 1988. RNA processing generates the mature 3' end of yeast *CYC1* messenger RNA *in vitro*. *Science*, **242**, 1270-1274.

Butler, J.S., Sadhale, P.P. and Platt, T. 1990. RNA processing *in vitro* produces mature 3' ends of a variety of *Saccharomyces cerevisiae* mRNAs. *Mol Cell Biol*, **10**, 2599-2605.

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

Calvo, O. and Manley, J.L. 2001. Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. *Mol Cell*, **7**, 1013-1023.

Calvo, O. and Manley, J.L. 2005. The transcriptional co-activator PC4/Sub1 has multiple functions in RNA polymerase II transcription. *EMBO J*, **24**, 1009-1020.

Caponigro, G. and Parker, R. 1995. Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes Dev*, **9**, 2421-2432.

Caponigro, G. and Parker, R. 1996. Mechanisms and control of mRNA turnover in *Saccharomyces cerevisiae*. *Microbiol Rev*, **60**, 233-249.

Cevher MA, Zhang X, Fernandez S, Kim S, Baquero J, Nilsson P, Lee S, Virtanen A,

Kleiman FE. 2010. Nuclear deadenylation/polyadenylation factors regulate 3' processing in response to DNA damage. *EMBO J.* **29:** 1674-87.

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

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

Chen, F., MacDonald, C.C. and Wilusz, J. 1995. Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res*, **23**, 2614-2620.

Chen, S. and Hyman, L.E. 1998. A specific RNA-protein interaction at yeast polyadenylation efficiency elements. *Nucleic Acids Res*, **26**, 4965-4974.

Chen, Z., Li, Y. and Krug, R.M. 1999. Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *EMBO J*, **18**, 2273-2283.

Cheng, H., He, X. and Moore, C. 2004. The essential WD repeat protein Swd2 has dual functions in RNA polymerase II transcription termination and lysine 4 methylation of histone H3. *Mol Cell Biol*, **24**, 2932-2943.

Chou, Z.F., Chen, F. and Wilusz, J. 1994. Sequence and position requirements for uridylate-rich downstream elements of polyadenylation signals. *Nucleic Acids Res*, **22**, 2525-2531.

Christofori, G. and Keller, W. 1988. 3' cleavage and polyadenylation of mRNA precursors in vitro requires a poly(A) polymerase, a cleavage factor and a snRNP. *Cell*, **54**, 875-889.

Colgan DF, Murthy KG, Prives C, Manley JL. 1996. Cell-cycle related regulation of poly(A) polymerase by phosphorylation. *Nature.* **384**: 282-5.

Colgan DF, Murthy KG, Zhao W, Prives C, Manley JL. 1998. Inhibition of poly(A) polymerase requires p43cdc2/cyclin B phosphorylation of multiple consensus and non-consensus sites. *EMBO J.* **17**:1053-62.

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

Coller, J. and M. Wickens. 2007. Tethered function assays: an adaptable approach to study RNA regulatory proteins. *Methods Enzymol*. **429**: 299-321.

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, Prinz J, St Onge RP, VanderSluis B, Makhnevych T, Vizeacoumar FJ, Alizadeh S, Bahr S, Brost RL, Chen Y, Cokol M, Deshpande R, Li Z, Lin ZY, Liang W, Marback M, Paw J, San Luis BJ, Shuteriqi E, Tong AH, van Dyk N, Wallace IM, Whitney JA, Weirauch MT, Zhong G, Zhu H, Houry WA, Brudno M, Ragibizadeh S, Papp B, Pál C, Roth FP, Giaever G, Nislow C,

Troyanskaya OG, Bussey H, Bader GD, Gingras AC, Morris QD, Kim PM, Kaiser CA, Myers CL, Andrews BJ, Boone C. 2010. The genetic landscape of a cell. *Science.* **327**: 425-31.

Curtis, D., Lehmann, R. and Zamore, P.D. 1995. Translational regulation in development. *Cell*, **81**, 171-178.

Connelly S, Manley JL. 1998. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* **2**: 440-52.

Dantonel JC, Murthy KG, Manley JL and Tora L. 1997. Transcription factor TFIID recruits factor CPSF for formation of 3_ end of mRNA. *Nature* **389:** 399–402.

Daughdrill GW, Ackerman J, Isern NG, Botuyan MV, Arrowsmith C, Wold MS, Lowry DF. 2001. The weak interdomain coupling observed in the 70 kDa subunit of human replication protein A is unaffected by ssDNA binding. *Nucleic Acids Res.* **29**: 3270-6.

Daughdrill GW, Narayanaswami P, Gilmore SH, Belczyk A, and Brown CJ. 2007. Dynamic behavior of an intrinsically unstructured linker domain is conserved in the face of negligible amino acid sequence conservation. *J Mol Evol.* **65**: 277-88.

David DC, Layfield R, Serpell L, Narain Y, Goedert M, Spillantini MG. 2002. Proteasomal degradation of tau protein. *J Neurochem.* **83**: 176-85.

Deardorff, J.A. and Sachs, A.B. 1997. Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast poly(A)-binding protein. *J Mol Biol*, **269**, 67-81.

Decker, C.J. and Parker, R. 2002. mRNA decay enzymes: decappers conserved between yeast and mammals. *Proc Natl Acad Sci U S A*, **99**, 12512-12514.

Dehe, P.M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodriguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A., Shevchenko, A., Rosaleny, L.E., Tordera, V., Chavez, S., Stewart, A.F. and Geli, V. 2006. Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *J Biol Chem*, **281**, 35404-35412.

Deka, P., Rajan, P. K., Canadillas, J. M. P. and Varani, G. 2005. Protein and RNA dynamics play key roles in determining the specific recognition of GU-rich polyadenylation regulatory elements by human Cstf-64 protein. *J. Mol. Biol.* **347**: 719–733.

del Olmo M, Mizrahi N, Gross S, Moore C. 1997. The Uba2 and Ufd1 proteins of *S. cerevisiae* interact with poly(A) polymerase and affect the polyadenylation activity of extracts. *Mol Gen Genet.* **255**: 209–218.

Denli AM, B.B. Tops, R.H. Plasterk, R.F. Ketting, G.J. Hannon. 2004. Processing of primary microRNAs by the microprocessor complex, *Nature* **432:** 231–235.

Dermody JL, Dreyfuss JM, Villén J, Ogundipe B, Gygi SP, Park PJ, Ponticelli AS, Moore CL, Buratowski S, Bucheli ME. 2008. Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. *PLoS One*. **3**: e3273.

Dettwiler, S., Aringhieri, C., Cardinale, S., Keller, W. and Barabino, S.M. 2004. Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J Biol Chem*, **279**, 35788-35797.

de Vries H, Ruegsegger U, Hubner W, Friedlein A, Langen H and Keller W. 2000. Human pre-mRNA cleavage factor II (m) contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J*, **19**, 5895-5904.

Dheur S, Nykamp KR, Viphakone N, Swanson MS, 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.* 280:24532-8.

Dheur, S., Vo le, T.A., Voisinet-Hakil, F., Minet, M., Schmitter, J.M., Lacroute, F., Wyers, F. and Minvielle-Sebastia, L. 2003. Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. *EMBO J*, **22**, 2831-2840.

Dichtl, B. and Keller, W. 2001. Recognition of polyadenylation sites in yeast premRNAs by cleavage and polyadenylation factor. *EMBO J*, **20**, 3197-3209.

Dichtl, B., Blank, D., Sadowski, M., Hubner, W., Weiser, S. and Keller, W. 2002a. Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. *EMBO J*, **21**, 4125-4135.

Dichtl, B., Blank, D., Ohnacker, M., Friedlein, A., Roeder, D., Langen, H. and Keller, W. 2002b. A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. *Mol Cell*, **10**, 1139-1150.

Dichtl, B., Aasland, R. and Keller, W. 2004. Functions for *S. cerevisiae* Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. *RNA*, **10**, 965-977.

Dickson AM, Wilusz J. 2010. Polyadenylation: alternative lifestyles of the A-rich (and famous?). *EMBO J.* **29**: 1473

Dominski, Z. and Marzluff, W.F. 1999. Formation of the 3' end of histone mRNA. *Gene*, **239**, 1-14.

Dominski, Z., Yang, X.C., Purdy, M., Wagner, E.J. and Marzluff, W.F. 2005. A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100. *Mol Cell Biol*, **25**, 1489-1500.

Dominski Z, Yang XC, Marzluff WF. 2005. The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing. *Cell.* **123**: 37-48.

Dosztányi Z, Csizmok V, Tompa P, Simon I. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on energy content. *Bioinformatics.* **21**: 3433-4.

DuBois, R.N., McLane, M.W., Ryder, K., Lau, L.F. and Nathans, D. 1990. A growth factor-inducible nuclear protein with a novel cysteine/histidine repetitive sequence. *J Biol Chem*, **265**, 19185-19191.

Dunn EF, Hammell CM, Hodge CA, Cole CN. 2005. Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev.* **19**: 90-103.

Edmonds, M. 1990. Polyadenylate polymerase. *Methods Enzymol.* **181**: 161-170.

Eggermont J, Proudfoot NJ. 1993. Poly(A) signals and transcriptional pause sites combine to prevent interference between RNA polymerase II promoters. *EMBO J.* **12**: 2539-48.

Fasken MB, Stewart M, Corbett AH. 2008. Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear pore-associated protein, Mlp1, in mRNA export.*J Biol Chem.* **283**: 27130-43

Feng, Z.H., Wilson, S.E., Peng, Z.Y., Schlender, K.K., Reimann, E.M. and Trumbly, R.J. 1991. The yeast GLC7 gene required for glycogen accumulation encodes a type 1 protein phosphatase. *J Biol Chem*, **266**, 23796-23801.

Fink AL. 2005. Natively unfolded proteins. *Curr Opin Struct Biol.* **15**: 35-41.

Fox, C. A., M. D. Sheets, and M. Wickens. 1989. Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. *Genes Dev.* **3**: 2151-2162.

Forbes KP, Addepalli B, Hunt AG. 2006. An *Arabidopsis* Fip1 homolog interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. *J Biol Chem.* **281**: 176-86.

Furuichi, Y. and Shatkin, A.J. 2000. Viral and cellular mRNA capping: past and prospects. *Adv Virus Res*, **55**, 135-184.

Ganem, C., Devaux, F., Torchet, C., Jacq, C., Quevillon-Cheruel, S., Labesse, G., Facca, C. and Faye, G. 2003. Ssu72 is a phosphatase essential for transcription termination of snoRNAs and specific mRNAs in yeast. *EMBO J*, **22**, 1588-1598.

Garcia-Gimeno, M. A., Munoz, I., Arino, J. and Sanz, P. 2003. Molecular characterization of Ypi1, a novel *Saccharomyces cerevisiae* type 1 protein phosphatase inhibitor. J. Biol. Chem. 278, 47744–47752.

Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dümpelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B,

Bork P, Russell RB, Superti-Furga G. 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature.* **440** : 631-6.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. and Superti-Furga, G. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, **415**, 141-147.

Ge H, Zhao Y, Chait BT, Roeder RG. 1994. Phosphorylation negatively regulates the function of coactivator PC4. Proc Natl Acad Sci. **91**:12691–12695.

Ghazy MA, He X, Singh BN, Hampsey M, Moore C. 2009. The essential N terminus of the Pta1 scaffold protein is required for snoRNA transcription termination and Ssu72 function but is dispensable for pre-mRNA 3'-end processing. *Mol Cell Biol.* **29**: 2296-307.

Gil, A. and Proudfoot, N.J. 1987. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit beta-globin mRNA 3' end formation. *Cell*, **49**, 399-406.

Gilbert, W. and Guthrie, C. 2004. The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Mol Cell*, **13**, 201- 212.

Gilmartin, G.M., Fleming, E.S., Oetjen, J. and Graveley, B.R. 1995. CPSF recognition of an HIV-1 mRNA 3'-processing enhancer: multiple sequence contacts involved in poly(A) site definition. *Genes Dev*, **9**, 72-83.

Gonzalez, C.I., Ruiz-Echevarria, M.J., Vasudevan, S., Henry, M.F. and Peltz, S.W. 2000. The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Mol Cell*, **5**, 489-499.

Graber, J.H., Cantor, C.R., Mohr, S.C. and Smith, T.F. 1999a. *In silico* detection of control signals: mRNA 3'-end-processing sequences in diverse species. *PNAS*, **96**, 14055-14060.

Graber, J.H., Cantor, C.R., Mohr, S.C. and Smith, T.F. 1999b. Genomic detection of new yeast pre-mRNA 3'-end-processing signals. *Nucleic Acids Res*, **27**, 888- 894.

Graber, J.H. 2003. Variations in yeast 3'-processing cis-elements correlate with transcript stability. *Trends Genet*, **19**, 473-476.

Greger lH, Aranda A, Proudfoot N. 2000. Balancing transcriptional interference and initiation on the GAL7 promoter of Saccharomyces cerevisiae. *PNAS.* **97:** 8415-20.

Greger IH, Proudfoot NJ. 1998. Poly(A) signals control both transcriptional termination and initiation between the tandem GAL10 and GAL7 genes of Saccharomyces cerevisiae. *EMBO J.* **17**: 4771-9.

Gregory RI, K.P. Yan, G. Amuthan, T. Chendrimada, B. Doratotaj, N. Cooch, R. Shiekhattar. 2004. The microprocessor complex mediates the genesis of microRNAs, *Nature* **432:** 235–240.

Gross S, Moore CL. 2001. Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. *Mol Cell Biol.* **21**: 8045-55.

Grunberg-Manago, M. 1999. Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annu Rev Genet*, **33**, 193-227.

Gu, M. and Lima, C.D. 2005. Processing the message: structural insights into capping and decapping mRNA. *Curr Opin Struct Biol*, **15**, 99-106.

Gunderson SI, Beyer K, Martin G, Keller W, Boelens WC, Mattaj LW. 1994. The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell.* **76**: 531-41.

Gunderson SI, Vagner S, Polycarpou-Schwarz M, Mattaj IW. 1997. Involvement of the carboxyl terminus of vertebrate poly(A) polymerase in U1A autoregulation and in the coupling of splicing and polyadenylation. *Genes Dev.* **11**: 761-73.

Gunderson SI, Polycarpou-Schwarz M, Mattaj IW. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell.* **1**: 255-64.

Guo, Z., Russo, P., Yun, D.-F., Butler, J.S. and Sherman, F. 1995. Redundant 3' end forming signals for the yeast *CYC1 mRNA*. *PNAS*, **92**, 4211-4214.

Guo, Z. and Sherman, F. 1995. 3'-end-forming signals of yeast mRNA. *Mol Cell Biol*, **15**, 5983-5990.

Guo, Z. and Sherman, F. 1996. 3'-end-forming signals of yeast mRNA. *Trends in Biochemical Sciences*, **21**, 477-481.

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.

Hampsey M. 1997. A review of phenotypes in Saccharomyces cerevisiae.*Yeast.* **13**: 1099-133.

Haghighat, A. and Sonenberg, N. 1997. eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J Biol Chem*, **272**, 21677-21680.

Hake LE, Richter JD. 1994. CPEB is a specificity factor that mediates cytoplasmic

polyadenylation during Xenopus oocyte maturation. *Cell.* **79**: 617-27.

He X, Khan AU, Cheng H, Pappas DL Jr, Hampsey M, Moore CL. 2003. Functional interactions between the transcription and mRNA 3' end processing machineries mediated by Ssu72 and Sub1. *Genes Dev.* **17**: 1030-42.

He X, and Moore C. 2005. Regulation of yeast mRNA 3' end processing by phosphorylation. *Mol Cell.* **19**: 619-29.

Hector RE, Nykamp KR, Dheur S, Anderson JT, Non PJ, Urbinati CR, Wilson SM, Minvielle-Sebastia L, Swanson MS. 2002. Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *EMBO J.* **21**: 1800-1810.

Heidmann, S., Schwindewolf, C., Stumpf, G. and Domdey, H. 1994. Flexibility and interchangeability of polyadenylation signals in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **14**, 4633-4642.

Helmling S, Zhelkovsky A, and Moore CL. 2001. Fip1 regulates the activity of Poly(A) polymerase through multiple interactions. *Mol Cell Biol.* **21**: 2026-37.

Henry, M., Borland, C.Z., Bossie, M. and Silver, P.A. 1996. Potential RNA binding proteins in *Saccharomyces cerevisiae* identified as suppressors of temperature sensitive mutations in *NPL3*. *Genetics*, **142**, 103-115.

Henry NL, Bushnell DA, Kornberg RD. 1996. A yeast transcriptional stimulatory protein similar to human PC4. J Biol Chem. **271**: 21842–21847.

Hernandez, H., Dziembowski, A., Taverner, T., Seraphin, B. and Robinson, C.V. 2006. Subunit architecture of multimeric complexes isolated directly from cells. *EMBO Rep*, **7**, 605-610.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R. and Jensen, T.H. 2001. Quality control of mRNA 3'-end processing is linked to the nuclear exosome. *Nature*, **413**, 538-542.

Hofmann, I., Schnolzer, M., Kaufmann, I. and Franke, W.W. 2002. Symplekin, a constitutive protein of karyo- and cytoplasmic particles involved in mRNA biogenesis in Xenopus laevis oocytes. *Mol Biol Cell*. **13**: 1665-1676.

Holm, L. and Sander, C. 1995. DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem Sci*, **20**, 345-347.

Huang, Y. and Carmichael, G.G. 1996. Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol Cell Biol*, **16**, 1534-1542.

Huarte J, Stutz A, O'Connell ML, Gubler P, Belin D, Darrow AL, Strickland S, Vassalli JD. 1992. Transient translational silencing by reversible mRNA deadenylation. *Cell.* **69:** 1021-30.

Hunt AG. 2008. Messenger RNA 3' end formation in plants. *Curr Top Microbiol Immunol.* **326:** 151-77

Hyman, L.E., Seiler, S.H., Whoriskey, J. and Moore, C.L. 1991. Point mutations upstream of the yeast ADH2 poly(A) site significantly reduce the efficiency of 3'-end formation. *Mol Cell Biol*, **11**, 2004-2012.

Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. 2004. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **32**: 1037-49.

Iizuka, N., Najita, L., Franzusoff, A. and Sarnow, P. 1994. Cap-dependent and cap independent translation by internal initiation of mRNAs in cell extracts prepared from *Saccharomyces cerevisiae*. *Mol Cell Biol*, **14**, 7322-7330.

Jacobs DM, Lipton AS, Isern NG, Daughdrill GW, Lowry DF, Gomes X, Wold MS. 1999. Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker. *J Biomol NMR* **14**: 321-331.

Jacobson, A. and Favreau, M. 1983. Possible involvement of poly(A) in protein synthesis. *Nucleic Acids Res*, **11**, 6353-6368.

Jacobson, A. and Peltz, S.W. 1996. Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annu Rev Biochem*, **65**, 693-739.

Jenny, A., Hauri, H.P. and Keller, W. 1994. Characterization of cleavage and polyadenylation specificity factor and cloning of its 100-kilodalton subunit. *Mol Cell Biol*. **14**: 8183-8190.

Jenny, A., Minvielle-Sebastia, L., Preker, P.J. and Keller, W. 1996. Sequence similarity between the 73-kilodalton protein of mammalian CPSF and a subunit of yeast polyadenylation factor I. *Science*. **274:** 1514-1517.

Jensen TH, Neville M, Rain JC, McCarthy T, Legrain P, Rosbash M. 2000. Identification of novel Saccharomyces cerevisiae proteins with nuclear export activity: cell cycle-regulated transcription factor ace2p shows cell cycle-independent nucleocytoplasmic shuttling. *Mol Cell Biol.* **20**: 8047-58.

Jenuwein, T. and Allis, C.D. 2001. Translating the histone code. *Science*, **293**, 1074-1080.

Johnson E, Ma P, Ota I, Varshavsky A. 1995. A proteolytic pathway that recognizes ubiquitin as a degradation signal. *J Biol Chem.* **270:**17442–17456.

Jove, R. and Manley, J.L. 1984 *In vitro* transcription from the adenovirus 2 major late promoter utilizing templates truncated at promoter-proximal sites. *J Biol Chem*, **259**, 8513-8521.

Kaufmann I, Martin G, Friedlein A, Langen H, Keller W. 2004. Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *EMBO J.* **23**: 616-26.

Keller, W., Bienroth, S., Lang, K.M. and Christofori, G. 1991. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *EMBO J*, **10**, 4241-4249.

Keller, W. and Minvielle-Sebastia, L. 1997. A comparison of mammalian and yeast pre-mRNA 3'-end processing. *Current Opinion in Cell Biology*, **9**, 329-336.

Kelly SM, Leung SW, Apponi LH, Bramley AM, Tran EJ, Chekanova JA, Wente SR, Corbett AH. 2010. Recognition of polyadenosine RNA by the zinc finger domain of nuclear poly(A) RNA-binding protein 2 (Nab2) is required for correct mRNA 3'-end formation. *J Biol Chem.* **285:** 26022-32

Keon, B. H., Schafer, S., Kuhn, C., Grund, C. and Franke, W.W. 1996. Symplekin, a novel type of tight junction plaque protein. J. Cell. Biol. 134, 1003–1018.

Kessler MM, Zhao J, Moore CL.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**: 27167-75.

Kessler MM, Henry MF, Shen E, Zhao J, Gross S, Silver PA, Moore CL. 1997. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes Dev.* **11**: 2545-56.

Kessler MM, Zhelkovsky AM, Skvorak A, and Moore CL. 1995. Monoclonal antibodies to yeast poly(A) polymerase (PAP) provide evidence for association of PAP with cleavage factor I. *Biochemistry.* **34**: 1750-9.

Kerwitz, Y., Kuhn,U., Lilie,H., Knoth, A., Scheuermann, T., Friedrich, H., Schwarz, E. and Wahle,E. 2003. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. EMBO J. 22, 3705–3714.

Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, Greenblatt JF, Buratowski S. 2004. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature.* **432**: 517-22.

Kim, H. and Lee, Y. 2001. Interaction of poly(A) polymerase with the 25-kDa subunit of cleavage factor I. *Biochem Biophys Res Commun*, **289**, 513-518.

Kim H, Lee JH, Lee Y. 2003. Regulation of poly(A) polymerase by 14-3-3epsilon. *EMBO J.* **22**:5208–5219.

Kim JH, Richter JD. 2006. Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol Cell.* **24:** 173-83.

Kleiman FE, Manley JL. 2001. The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression. *Cell.* **104**: 743-53.

Ko B, Gunderson SI. 2002. Identification of new poly(A) polymerase-inhibitory proteins

capable of regulating pre-mRNA polyadenylation. *J Mol Biol.* **318:** 1189-206.

Köhrer K, Domdey H. 1991. Preparation of high molecular weight RNA. *Methods Enzymol.* **194**: 398-405.

Kolev, N. G. and Steitz, J. A. 2005. Symplekin and multiple other polyadenylation factors participate in 3' end maturation of histone mRNAs. Genes Dev. 19, 2583–2592.

Kornberg RD. 2007. The molecular basis of eukaryotic transcription. *PNAS.* **104**:12955-61.

Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C. and Hampsey, M. 2004. Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol Cell*, **14**, 387-394.

Kühn U, Gündel M, Knoth A, Kerwitz Y, Rüdel S, Wahle E. 2009. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *J Biol Chem.* **284**: 22803-14.

Kyburz, A., Sadowski, M., Dichtl, B. and Keller, W. 2003. The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. *Nucleic Acids Res*, **31**, 3936-3945.

Kyburz, A., Friedlein, A., Langen, H. and Keller, W. 2006. Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3' end processing and splicing. *Mol Cell*, **23**, 195-205.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A. and Tollervey, D. 2005. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell*, **121**, 713-724.

Le, H., Tanguay, R.L., Balasta, M.L., Wei, C.C., Browning, K.S., Metz, A.M., Goss, D.J. and Gallie, D.R. 1997. Translation initiation factors eIF-iso4G and eIF-4B interact with the poly(A)-binding protein and increase its RNA binding activity. *J Biol Chem*, **272**, 16247-16255.

Leeper TC, Qu X, Lu C, Moore C, Varani G. 2010. Novel protein-protein contacts facilitate mRNA 3'-processing signal recognition by Rna15 and Hrp1. *J Mol Biol*. **401**: 334-49

Leuschner, P.J., Ameres, S.L., Kueng, S. and Martinez, J. 2006. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep*, **7**, 314-320.

Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL. 2002. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol Cell.* **9**: 1101-11.

Lingner, J., Kellermann, J. and Keller, W. 1991a. Cloning and expression of the essential gene for poly(A) polymerase from *S. cerevisiae*. *Nature*, **354**, 496-498.

Lingner, J., Radtke, I., Wahle, E. and Keller, W. 1991b. Purification and characterization of poly(A) polymerase from *Saccharomyces cerevisiae*. *J Biol Chem*, **266**, 8741-8746.

Long, R.M., Elliott, D.J., Stutz, F., Rosbash, M. and Singer, R.H. 1995. Spatial consequences of defective processing of specific yeast mRNAs revealed by fluorescent in situ hybridization. *RNA*, **1**, 1071-1078.

Lowell, J.E., Rudner, D.Z. and Sachs, A.B. 1992. 3'-UTR-dependent de-adenylation by the yeast poly(A) nuclease. *Genes Dev*, **6**, 2088-2099.

Luo, Y. and Goss, D.J. 2001. Homeostasis in mRNA initiation: wheat germ poly(A)-binding protein lowers the activation energy barrier to initiation complex formation. *J Biol Chem*, **276**, 43083-43086.

Luo, W. and Bentley, D. 2004. A ribonucleolytic rat torpedoes RNA polymerase II. *Cell*, **119**, 911-914.

Luo, W., Johnson, A.W. and Bentley, D.L. 2006. The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev*, **20**, 954-965.

MacDonald, C.C., Wilusz, J. and Shenk, T. 1994. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol Cell Biol*, **14**, 6647-6654.

Mager WH, Siderius M. 2002. Novel insights into the osmotic stress response of yeast. *FEMS Yeast Res.* **2**: 251-7.

Mandart, E. and Parker, R. 1995. Effects of mutations in the *Saccharomyces cerevisiae RNA14, RNA15,* and *PAP1* genes on polyadenylation *in vivo*. *Mol Cell Biol*, **15**, 6979-6986.

Mandel CR, Bai Y, Tong L. 2008. Protein factors in pre-mRNA 3'-end processing. *Cell Mol Life Sci.* **65**: 1099-122.

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'-endprocessing endonuclease. *Nature*, **444**, 953-956.

Mangus, D.A., Smith, M.M., McSweeney, J.M. and Jacobson, A. 2004. Identification of factors regulating poly(A) tail synthesis and maturation. *Mol Cell Biol*, **24**, 4196-4206.

Manley, J.L., Yu, H. and Ryner, L. 1985. RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation *in vitro*. *Mol Cell Biol*, **5**, 373-379.

Manley, J.L. and Takagaki, Y. 1996. The end of the message - another link between yeast and mammals. *Science*, **274**, 1481-1482.

Maniatis T, Reed R. 2002. An extensive network of coupling among gene expression machines. *Nature.* 416: 499-506.

Martin, G. and Keller, W. 1996. Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and catalytic domain, homologous to the family X polymerases, and to other nucleotidyltransferases. *EMBO J*, **15**, 2593-2603.

Martin, G., Keller, W. and Doublie, S. 2000. Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J*, **19**, 4193-4203.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S. and Bentley, D.L. 1997a. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev*, **11**, 3306-3318.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. 1997b. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, **385**, 357-361.

McGrew, L. L., and J. D. Richter. 1990. Translational control by cytoplasmic polyadenylation during Xenopus oocyte maturation: characterization of cis and trans elements and regulation by cyclin/MPF. *EMBO J.* **9:** 3743-3751.

Meinhart A, Cramer P. 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature.* **430:** 223-6.

Meinke G, Ezeokonkwo C, Balbo P, Stafford W, Moore C, and Bohm A. 2008. Structure of yeast poly(A) polymerase in complex with a peptide from Fip1, an intrinsically disordered protein. *Biochemistry.* **47**: 6859-69.

Millevoi S, Decorsière A, Loulergue C, Iacovoni J, Bernat S, Antoniou M, Vagner S. 2009. A physical and functional link between splicing factors promotes pre-mRNA 3' end processing. *Nucleic Acids Res.* **37**: 4672-83.

Millevoi S, Vagner S. 2009. Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res.* **38:** 2757-74.

Miller SD, Moses K, Jayaraman L, Prives C. 1997. Complex formation between p53 and replication protein A inhibits the sequence-specific DNA binding of p53 and is regulated by single-stranded DNA. Mol Cell Biol. **17**: 2194-201.

Minvielle-Sebastia, L., Winsor, B., Bonneaud, N. and Lacroute, F. 1991. Mutations in the yeast *RNA14* and *RNA15* genes result in an abnormal mRNA decay rate; Sequence analysis reveals an RNA-binding domain in the RNA15 protein. *Mol Cell Biol*, **11**, 3075-3087.

Minvielle-Sebastia, L., Preker, P.J. and Keller, W. 1994. RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. *Science*, **266**, 1702-1705.

Minvielle-Sebastia, L., Preker, P.J., Wiederkehr, T., Strahm, Y. and Keller, W. 1997. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in pre-messenger RNA 3'-end formation. *PNAS*, **94**, 7897-7902.

Minvielle-Sebastia, L., Beyer, K., Krecic, A.M., Hector, R.E., Swanson, M.S., Keller, W. (1998) Control of cleavage site selection during mRNA 3' end formation by a yeast hnRNP. *EMBO J*, **17**, 7454-7468.

Mirkin N, Fonseca D, Mohammed S, Cevher MA, Manley JL, Kleiman FE. 2008. The 3' processing factor CstF functions in the DNA repair response. *Nucleic Acids Res.* **36:** 1792-804.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. and Tollervey, D. 1997. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. *Cell*, **91**, 457-466.

Mizrahi N, Moore C. 2000. Posttranslational phosphorylation and ubiquitination of the Saccharomyces cerevisiae Poly(A) polymerase at the S/G(2) stage of the cell cycle. *Mol Cell Biol.* **20**: 2794-802.

Moore, C.L. and Sharp, P.A. 1985. Accurate cleavage and polyadenylation of exogenous RNA substrate. *Cell*, **41**, 845-855.

Moore MJ. 2005. From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309:** 1514–1518

Munroe, D. and Jacobson, A. 1990a. mRNA poly(A) tail, a 3' enhancer of translational initiation. *Mol Cell Biol*, **10**, 3441-3455.

Munroe, D. and Jacobson, A. 1990b. Tales of poly(A): a review. *Gene*, **91**, 151-158.

Murthy, K.G. and Manley, J.L. 1995. The 160-kD subunit of human cleavage polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev*. **9**: 2672-2683.

Nedea, E., He, X., Kim, M., Pootoolal, J., Zhong, G., Canadien, V., Hughes, T., Buratowski, S., Moore, C.L. and Greenblatt, J. 2003. Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *J Biol Chem*, **278**, 33000-33010.

Nemeth, A., Krause, S., Blank, D., Jenny, A., Jeno, P., Lustig, A. and Wahle, E. 1995. Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II. *Nucleic Acids Res*, **23**, 4034-4041.

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.

Noble, C.G., Hollingworth, D., Martin, S.R., Ennis-Adeniran, V., Smerdon, S.J., Kelly, G.,

Taylor, I.A. and Ramos, A. 2005. Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat Struct Mol Biol*, **12**, 144-151.

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.

Nunes NM, Li W, Tian B, Furger A. 2010. A functional human Poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J.* **29**:1523-3

O'Connor, J.P. and Peebles, C.L. 1992. *PTA1*, an essential gene of *Saccharomyces cerevisiae* affecting pre-tRNA processing. *Mol Cell Biol*, **12**, 3843-3856.

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'-endprocessing complex. *EMBO J*, **19**, 37-47.

Paillard L, Osborne HB. 2003. East of EDEN was a poly(A) tail. *Biol Cell.* 95: 211-9.

Pappas, D.L., Jr. and Hampsey, M. 2000. Functional interaction between Ssu72 and the Rpb2 subunit of RNA polymerase II in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **20**, 8343-8351.

Patel D, Butler JS. 1992. Conditional defect in mRNA 3' end processing caused by a mutation in the gene for poly(A) polymerase. *Mol Cell Biol.***12**: 3297-304.

Peggie, M.W., MacKelvie, S.H., Bloecher, A., Knatko, E.V., Tatchell, K. and Stark, M.J. 2002. Essential functions of Sds22p in chromosome stability and nuclear localization of PP1. *J Cell Sci*, **115**, 195-206.

Pancevac C, Goldstone DC, Ramos A, Taylor IA. 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**: 3119-32.

Perumal, K. and Reddy, R. 2002. The 3' end formation in small RNAs. *Gene Expr*, **10**, 59-78.

Preker PJ, Lingner J, Minvielle-Sebastia L, Keller W. 1995. The FIP1 gene encodes a component of a yeast pre-mRNA polyadenylation factor that directly interacts with poly(A) polymerase. *Cell.* **81**: 379-89.

Preker PJ, Ohnacker M, Minvielle-Sebastia L, Keller W. 1997. A multisubunit 3' end-processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor. *EMBO J.* **16**: 4727-37.

Proudfoot, N. 1991. Poly(A) signals. *Cell*, **64**, 671-674.

Proudfoot N. 2004. New perspectives on connecting messenger RNA 3′ end formation to transcription, *Curr. Opin. Cell Biol.* **16:** 272–278

Proudfoot NJ, Brownlee GG. 1976. 3' non-coding region sequences in eukaryotic messenger RNA. *Nature.* **263**: 211-4.

Proudfoot NJ, Furger A, Dye MJ. 2002. Integrating mRNA processing with transcription. *Cell.* **108**: 501-12.

Qu X, Perez-Canadillas JM, Agrawal S, De Baecke J, Cheng H, Varani G, 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**: 2101-15

Rasmussen, E.B. and Lis, J.T. 1993. *In vivo* transcriptional pausing and cap formation on three Drosophila heat shock genes. *PNAS*, **90**, 7923-7927.

Rigo F, Martinson HG. 2009. Polyadenylation releases mRNA from RNA polymerase II in a process that is licensed by splicing. *RNA.* **15:** 823-36.

Ritchie DB, Schellenberg MJ, MacMillan AM. 2009. Spliceosome structure: piece by piece. *Biochim Biophys Acta.* **1789**: 624-33.

Rosonina E, Willis IM, Manley JL. 2009. Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. *Mol Cell Biol.* **29**: 2308-21.

Roth KM, Byam J, Fang F, Butler JS. 2009. Regulation of NAB2 mRNA 3'-end formation requires the core exosome and the Trf4p component of the TRAMP complex. *RNA.* **15:** 1045-58.

Ruegsegger, U., Beyer, K. and Keller, W. 1996. Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. *J Biol Chem*, **271**, 6107-6113.

Russnak, R., Nehrke, K.W. and Platt, T. 1995. REF2 encodes an RNA-binding protein directly involved in yeast mRNA 3'-end formation. *Mol Cell Biol*, **15**, 1689-1697.

Russo, P., Li, W.Z., Guo, Z. and Sherman, F. 1993. Signals that produce 3' termini in CYC1 mRNA of the yeast Saccharomyces cerevisiae. *Mol Cell Biol*, **13**, 7836-7849.

Ryan, K., Calvo, O. and Manley, J.L. 2004. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *RNA*, **10**, 565-573.

Sachs, A.B., Bond, M.W. and Kornberg, R.D. 1986. A single gene from yeast for both nuclear and cytoplasmic polyadenylate-binding proteins: domain structure and expression. *Cell*, **45**, 827-835.

Sachs, A.B. and Davis, R.W. 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell*, **58**, 857-867.

Sachs, A. and Wahle, E. 1993. Poly(A) tail metabolism and function in eukaryotes. *J Biol Chem*, **268**, 22955-25968.

Sadhale, P.P. and Platt, T. 1992. Unusual aspects of *in vitro* RNA processing in the 3' regions of the GAL1, GAL7, and GAL10 genes in *Saccharomyces cerevisiae*. *Mol Cell Biol*, **12**, 4262-4270.

Sadowski, M., Dichtl, B., Hubner, W. and Keller, W. 2003. Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. *EMBO J*, **22**, 2167-2177.

Saguez C, Schmid M, Olesen JR, Ghazy MA, Qu X, Poulsen MB, Nasser T, Moore C, Jensen TH. 2008. Nuclear mRNA surveillance in THO/sub2 mutants is triggered by inefficient polyadenylation. *Mol Cell.* **31**: 91-103.

Samanta MP and Liang S. 2003. Predicting protein functions from redundancies in large-scale protein interaction networks. *Proc Natl Acad Sci.* **100**: 12579-83

Schärpf M, Sticht H, Schweimer K, Boehm M, Hoffmann S, and Rösch P. Antitermination in bacteriophage lambda: The structure of the N36 peptide-boxB RNA complex. *Eur J Biochem.* **267**: 2397-408.

Setyono, B. and Greenberg, J.R. 1981. Proteins associated with poly(A) and other regions of mRNA and hnRNA molecules as investigated by crosslinking. *Cell*, **24**, 775-783.

Simon R, Richter JD. 1994. Further analysis of cytoplasmic polyadenylation in Xenopus embryos and identification of embryonic cytoplasmic polyadenylation element-binding proteins. *Mol Cell Biol.* **14:** 7867-75.

Simon, R., J.-P. Tassan, and J. D. Richter. 1992. Translational control by poly(A) elongation during Xenopus development: differential repression and enhancement by a novel cytoplasmic polyadenylation element. *Genes Dev.* **6:** 2580-2591.

Simonsen, C. C. and Levinson, A. D. 1983. Analysis of processing and polyadenylation signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. *Mol. Cell. Biol.* **3:** 2250–2258.

Sharp PA. 1998. RNA splicing and genes. *JAMA.* **260**: 3035-41.

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

Shen, E.C., Henry, M.F., Weiss, V.H., Valentini, S.R., Silver, P.A. and Lee, M.S. 1998. Arginine methylation facilitates the nuclear export of hnRNP proteins. *Genes Dev*, **12**, 679-691.

Sheets, M.D., Ogg, S.C. and Wickens, M.P. 1990. Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic Acids Res*, **18**, 5799-5805.

Shibagaki, Y., Itoh, N., Yamada, H., Nagata, S. and Mizumoto, K. 1992. mRNA capping enzyme. Isolation and characterization of the gene encoding mRNA guanylyltransferase subunit from *Saccharomyces cerevisiae*. *J Biol Chem*, **267**, 9521-9528.

Shimazu T, Horinouchi S, Yoshida M. 2007. Multiple histone deacetylases and the CREB-binding protein regulate pre-mRNA 3'-end processing. *J. Biol. Chem.* **282**:4470–4478.

Shyu AB, Wilkinson MF, van Hoof A. 2008. Messenger RNA regulation: to translate or to degrade. *EMBO J.* **27:** 471-81.

Skaar, D.A. and Greenleaf, A.L. 2002. The RNA Polymerase II CTD Kinase CTDKI Affects Pre-mRNA 3' Cleavage/Polyadenylation through the Processing Component Pti1p. *Mol Cell*, **10**, 1429-1439.

Smith, T.F., Gaitatzes, C., Saxena, K. and Neer, E.J. 1999. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci*, **24**, 181-185.

Snyder M, Sapolsky RJ, Davis RW. 1988. Transcription interferes with elements important for chromosome maintenance in Saccharomyces cerevisiae. *Mol Cell Biol.* **8**: 2184-94.

Stark, M.J. 1996. Yeast protein serine/threonine phosphatases: multiple roles and diverse regulation. *Yeast*, **12**, 1647-1675.

Steinmetz, E.J. and Brow, D.A. 1996. Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Mol Cell Biol*, **16**, 6993-7003.

Steinmetz, E.J. and Brow, D.A. 2003. Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription. *Mol Cell Biol*, **23**, 6339-6349.

Stumpf, G. and Domdey, H. 1996. Dependence of yeast pre-mRNA 3'-end processing on CFT1: a sequence homolog of the mammalian AAUAAA binding factor. *Science*, **274**, 1517-1519.

Sun, Z.W. and Hampsey, M. 1996. Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection in vivo. *Mol Cell Biol*, **16**, 1557-1566.

Svejstrup JQ. 2004. The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim Biophys Acta.* 1677: 64-73.

Tacahashi, Y., Helmling, S. and Moore, C.L. 2003. Functional dissection of the zinc finger and flanking domains of the Yth1 cleavage/polyadenylation factor. *Nucleic Acids Res*, **31**, 1744-1752.

Takagaki, Y., Manley, J.L. 1994. A polyadenylation factor subunit is the human

homologue of the *Drosophila suppressor of forked* protein. *Nature*, **372**, 471-474.

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

Takagaki, Y. and Manley, J. L. 2000. Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* **20**: 1515–1525.

Takagaki, Y., Manley, J.L., MacDonald, C.C., Wilusz, J. and Shenk, T. 1990. A multi subunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. *Genes Dev*, **4**, 2112-2120.

Takagaki, Y., Ryner, L.C. and Manley, J.L. 1989. Four factors are required for 3'- end cleavage of pre-mRNAs. *Genes Dev*, **3**, 1711-1724.

Tan, Y.S., Morcos, P.A. and Cannon, J.F. 2003. Pho85 phosphorylates the Glc7 protein phosphatase regulator Glc8 in vivo. *J Biol Chem*, **278**, 147-153.

Tarun, S.Z., Jr. and Sachs, A.B. 1995. A common function for mRNA 5' and 3' ends in translation initiation in yeast. *Genes Dev*, **9**, 2997-3007.

Tarun, S.Z., Jr. and Sachs, A.B. 1996. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J*, **15**, 7168-7177.

Tarun, S.Z., Jr., Wells, S.E., Deardorff, J.A. and Sachs, A.B. 1997. Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *PNAS*, **94**, 9046-9051.

Thuresson, A.C., Astrom, J., Astrom, A., Gronvik, K.O. and Virtanen, A. 1994. Multiple forms of poly(A) polymerases in human cells. *PNAS*, **91**, 979-983.

Tofaris GK, Layfield R, Spillantini MG. 2001. Alpha-synuclein metabolism and aggregation is linked to ubiquitin-independent degradation by the proteasome. *FEBS Lett.* **509**: 22-6.

Valentini, S.R., Weiss, V.H. and Silver, P.A. 1999. Arginine methylation and binding of Hrp1p to the efficiency element for mRNA 3'-end formation. *RNA*, 5, 1-9.

van Hoof, A., P. Lennertz, and R. Parker. 2000. Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.* **2**: 441-452.

Venters BJ, Pugh BF. 2009. How eukaryotic genes are transcribed. *Crit Rev Biochem Mol Biol.* 44: 117-41.

Vethantham V, Rao N, Manley JL. 2008. Sumoylation regulates multiple aspects of mammalian poly(A) polymerase function. *Genes Dev.* **22**: 499-511.

Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G. and Keller, W. 2005. A new yeast poly(A) polymerase complex involved

in RNA quality control. *PLoS Biol*, 3, e189.

Vethantham V, Rao N, Manley JL. 2007. Sumoylation modulates the assembly and activity of the pre-mRNA 3' processing complex. *Mol Cell Biol.* 27: 8848-58.

Vethantham V, Rao N, Manley JL. 2008. Sumoylation regulates multiple aspects of mammalian poly(A) polymerase function. *Genes Dev.* 22: 499-511.

Vo, L.T., Minet, M., Schmitter, J.M., Lacroute, F. and Wyers, F. 2001. Mpe1, a zinc knuckle protein, is an essential component of yeast cleavage and polyadenylation factor required for the cleavage and polyadenylation of mRNA. *Mol Cell Biol*, 21, 8346-8356.

Wahle, E. 1991. Purification and characterization of a mammalian polyadenylate polymerase involved in the 3' end processing of messenger RNA precursors. *J Biol Chem*, **266**, 3131-3139.

Wahle, E. and Keller, W. 1992. The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. *Annu Rev Biochem*, **61**, 419-440.

Wahle, E. and Kuhn, U. 1997. The mechanism of 3' cleavage and polyadenylation of eukaryotic pre-mRNA. *Prog Nucleic Acid Res Mol Biol*, **57**, 41-71.

Wahle, E. and Ruegsegger, U. (1999) 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev*, **23**, 277-295.

Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. 1982. Distantly related sequences in the alpha- and betasubunits of ATP synthase, myosin, kinases and other ATPrequiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951.

Wang, S.W., Asakawa, K.,Win, T. Z., Toda, T. and Norbury, C. J. 2005. Inactivation of the pre-mRNA cleavage and polyadenylation factor Pfs2 in fission yeast causes lethal cell cycle defects. Mol. Cell. Biol. 25, 2288–2296.

West S, Gromak N, Proudfoot NJ. 2004. Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*. 432: 522-5.

Will, C.L. and Luhrmann, R. 2001. Spliceosomal UsnRNP biogenesis, structure and function. *Curr Opin Cell Biol*, 13, 290-301.

Williams-Hart, T., Wu, X. and Tatchell, K. 2002. Protein phosphatase type 1 regulates ion homeostasis in *Saccharomyces cerevisiae*. *Genetics*, 160, 1423-1437.

Wilusz J. and Shenk T. 1988. A 64 kDa nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. Cell 52, 221–228.

Wilusz JE, Spector DL. 2010. An unexpected ending: noncanonical 3' end processing mechanisms. *RNA.* **16**: 259-66

Wyers, F., M. Rougemaille, G. Breard, J.-C. Rousselle, M.-E. Dufour, J. Boulay, F. Devaux, A. Namane, D. Libri, B. Séraphin, and A. Jacquier. 2005. Cryptic Pol II transcript degradation by nuclear quality control pathway involving a new poly(A) polymerase. *Cell*.121: 725-737.

Xiang K, Nagaike T, Xiang S, Kilic T, Beh MM, Manley JL, Tong L. 2010. Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. *Nature.* **467**: 729-33.

Yang Q, Gilmartin GM, Doublié S. 2010. Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3' processing. *PNAS*. **107**: 10062-7.

Yao Y, Song L, Katz Y, Galili G. 2002. Cloning and characterization of Arabidopsis homologues of the animal CstF complex that regulates 3' mRNA cleavage and polyadenylation. *J Exp Bot.* **53:** 2277-8.

Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M. and Corden, J.L. 1996. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *PNAS*, 93, 6975-6980.

Zhang, J. and Corden, J.L. 1991. Phosphorylation causes a conformational change in the carboxyl-terminal domain of the mouse RNA polymerase II largest subunit. *J Biol Chem*, 266, 2297-2302.

Zhang, Z., Fu, J. and Gilmour, D.S. 2005. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. *Genes Dev*, 19, 1572-1580.

Zhang J, Addepalli B, Yun KY, Hunt AG, Xu R, Rao S, Li QQ, Falcone DL. 2008. A polyadenylation factor subunit implicated in regulating oxidative signaling in Arabidopsis thaliana. *PLoS One.* **3**: e2410.

Zhao J, Hyman L, Moore C. 1999a. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev.* 63: 405-45.

Zhao J, Kessler MM, Moore CL. 1997. Cleavage factor II of Saccharomyces cerevisiae contains homologues to subunits of the mammalian Cleavage/ polyadenylation specificity factor and exhibits sequence-specific, ATP-dependent interaction with precursor RNA. *J Biol Chem.* 272: 10831-8.

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

Zhao H, Xing D, Li QQ. 2009. Unique features of plant cleavage and polyadenylation specificity factor revealed by proteomic studies. *Plant Physiol.* **151**: 1546-56

Zhelkovsky AM, Kessler MM, Moore CL. 1995. Structure-function relationships in the *Saccharomyces cerevisiae* poly(A) polymerase. Identification of a novel RNA binding site and a domain that interacts with specificity factor(s). *J Biol Chem.* 270: 26715-20.

Zhelkovsky A, Helmling S, Moore C. 1998. Processivity of the *Saccharomyces cerevisiae* poly(A) polymerase requires interactions at the carboxyl-terminal RNA binding domain. *Mol Cell Biol.* 18: 5942-51.

Zhelkovsky, A., Tacahashi, Y., Nasser, T., He, X., Sterzer, U., Jensen, T.H., Domdey, H. and Moore, C. 2006. The role of the Brr5/Ysh1 C-terminal domain and its homolog Syc1 in mRNA 3'-end processing in *Saccharomyces cerevisiae*. *RNA*, 12, 435-445.

Zielinski, R., Hellman, U., Kubinski, K. and Szyszka, R. 2006. Fip1-an essential component of the *Saccharomyces cerevisiae* polyadenylation machinery is phosophorylated by protein kinase CK2. *Mol Cell Biochem*, 286, 191-197.