



University of Kentucky
UKnowledge

University of Kentucky Doctoral Dissertations

Graduate School

2005

CHARACTERIZATION OF PLANT POLYADENYLATION TRANSACTING FACTORS-FACTORS THAT MODIFY POLY(A) POLYMERSE ACTIVITY

Kevin Patrick Forbes
University of Kentucky

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Forbes, Kevin Patrick, "CHARACTERIZATION OF PLANT POLYADENYLATION TRANSACTING FACTORS-FACTORS THAT MODIFY POLY(A) POLYMERSE ACTIVITY" (2005). *University of Kentucky Doctoral Dissertations*. 444.

https://uknowledge.uky.edu/gradschool_diss/444

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

ABSTRACT OF DISSERTATION

Kevin Patrick Forbes

The Graduate School
University of Kentucky
2004

CHARACTERIZATION OF PLANT POLYADENYLATION TRANS-
ACTING FACTORS-FACTORS THAT MODIFY POLY(A)
POLYMERASE ACTIVITY

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the College of Agriculture
at the University of Kentucky

By

Kevin Patrick Forbes

Lexington, Kentucky

Director: Dr. Arthur G. Hunt, Professor of Agronomy

Lexington, Kentucky

2004

Copyright © Kevin Patrick Forbes 2004

ABSTRACT OF DISSERTATION

CHARACTERIZATION OF PLANT POLYADENYLATION TRANS-ACTING FACTORS-FACTORS THAT MODIFY POLY(A) POLYMERASE ACTIVITY

Plant polyadenylation factors have proven difficult to purify and characterize, owing to the presence of excessive nuclease activity in plant nuclear extracts, thereby precluding the identification of polyadenylation signal-dependent processing and polyadenylation in crude extracts. As an alternative approach to identifying such factors, a screen was conducted for activities that inhibit the non-specific activity of plant poly(A) polymerases (PAP). One such factor (termed here as Putative Polyadenylation Factor B, or PPF-B) was identified in a screen of DEAE-Sepharose column fractions using a partially purified preparation of a plant nuclear poly(A) polymerase. This factor was purified to near homogeneity. Surprisingly, in addition to being an effective inhibitor of the nuclear PAP, PPF-B inhibited the activity of a chloroplast PAP. In contrast, this factor stimulated the activity of the yeast PAP. Direct assays of ATPase, proteinase, and nuclease activities indicated that inhibition of PAP activity was not due to depletion of substrates or degradation of products of the PAP reaction. The major polypeptide component of PPF-B proved to be a novel linker histone (RSP), which copurified with inhibitory activity by affinity chromatography on DNA-cellulose. The association of inhibitory activity with a linker histone and the spectrum of inhibitory activity, raise interesting possibilities regarding the role of PPF-B in nuclear RNA metabolism. These include a link between DNA damage and polyadenylation, as well as a role for limiting the polyadenylation of stable RNAs in the nucleus and nucleolus.

The *Arabidopsis* genome possesses genes encoding probable homologs of most of the polyadenylation subunits that have been identified in mammals and yeast. Two of these reside on chromosome III and V and have the potential to encode a protein that is related to the yeast and mammalian Fip1 subunit (AtFip1-III and AtFip1-V). These genes are universally expressed in *Arabidopsis* tissues. AtFip1-V stimulates the non-specific activity of at least one *Arabidopsis* nuclear PAP, binds RNA, and interacts with other polyadenylation homologs AtCstF77 and AtCPSF30. These studies suggest that AtFip1-V is an authentic polyadenylation factor that coordinates other subunits and plays a role in regulating the activity of PAP in plants.

KEYWORDS: Fip1, Poly(A) Polymerase, mRNA 3' end formation, RNA Binding,
RNA-processing

Kevin Patrick Forbes

February 2, 2005

CHARACTERIZATION OF PLANT POLYADENYLATION TRANS-
ACTING FACTORS-FACTORS THAT MODIFY POLY(A)
POLYMERSE ACTIVITY

By

Kevin Patrick Forbes

Arthur G. Hunt

Director of Dissertation

Arthur G. Hunt

Director of Graduate Studies

February 2, 2005

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name

Date

DISSERTATION

Kevin Patrick Forbes

The Graduate School
University of Kentucky
2004

CHARACTERIZATION OF PLANT POLYADENYLATION TRANS-
ACTING FACTORS-FACTORS THAT MODIFY POLY(A)
POLYMERASE ACTIVITY

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the College of Agriculture
at the University of Kentucky

By

Kevin Patrick Forbes

Lexington, Kentucky

Director: Dr. Arthur G. Hunt, Professor of Agronomy

Lexington, Kentucky

2004

Copyright © Kevin Patrick Forbes 2004

ACKNOWLEDGEMENTS

This dissertation work was carried out under the guidance of my major advisor Dr. Arthur G. Hunt. Without his advice, patience and friendship off and on the basketball court, this dissertation would not have been accomplished. I would like to extend my sincere appreciation to my committee members, Drs. Joseph Chappell, Robert L. Houtz, and Brian C. Rymond, for their advice and guidance throughout my graduate career. I would also like to acknowledge Dr. Deane Falcone for offering counsel and encouragement. Dr. Falcone was initially on my committee and has since then left for the University of Massachusetts-Lowell.

I would like to express my gratitude to the present members of the Hunt lab, Carol von Lanken, Lisa Meeks, and Dr. Balasubrahmanyam Addepalli (Balu) and past members Barb Elliott, Drs. Quinn Li, Tomal Dattaroy, and Jaydip Das Gupta. I have truly enjoyed their technical support, stimulating discussion and friendship. My special appreciation is also given to everyone in the labs of Dr. Chappell, Dr. Glenn Collins and Dr. Randy Dinkins, for their sharing of materials and insightful discussions.

I would also like to acknowledge my close friends here in Lexington, KY and in Morgan Hill, CA. They have all contributed to my well-being and have shown me that there is more to life than science.

Last but not least, it is the support of my parents Russell and Sally Forbes and my sister Jennifer and brother-in-law Darrin Tornberg that has made it possible for me to accomplish such a feat. Without their love and understanding, I would not be where I am today.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	iii
LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
CHAPTER ONE: LITERATURE REVIEW	1
CHAPTER TWO: CHARACTERIZATION OF PUTATIVE POLYADENYLATION FACTOR-B (PPF-B), AN INHIBITOR OF NON-SPECIFIC POLY(A) POLYMERASE ACTIVITY	23
Introduction.....	24
Results	26
Discussion.....	34
Materials and Methods	39
CHAPTER THREE: CHARACTERIZATION OF <i>ARABIDOPSIS</i> GENES	
ENCODING FIP1 HOMOLOGS	63
Introduction.....	64
Results	66
Discussion.....	73
Materials and Methods	77
CHAPTER FOUR: FUTURE DIRECTIONS	106
REFERENCES	113
VITA	134

LIST OF TABLES

Table I. List of oligonucleotides used for PCR and sequencing.....	86
---	----

LIST OF FIGURES

Figure 1.1. Eukaryotic Poly(A) Signals.	20
Figure 1.2. Mammalian Cleavage and Polyadenylation Complex.....	21
Figure 1.3. Yeast Cleavage and Polyadenylation Complex.	22
Figure 2.1. Purification and Activity of PPF-B.	46
Figure 2.2. Effects of PPF-B on Different Poly(A) Polymerases.....	47
Figure 2.3. PPF-B does not Promote the Hydrolysis of ATP.....	48
Figure 2.4. PPF-B does not act Proteolytically on Plant PAPs.....	49
Figure 2.5. PPF-B does not Degrade Pre-formed Products of the PAP Reaction.	50
Figure 2.6. PPF-B RNA is Distinct from U1 snRNP.....	52
Figure 2.7. PPF-B does not appear to be Immunologically Related to Other Plant Poly(A) Factors.....	53
Figure 2.8. Association of PPF-B with a Linker Histone.....	56
Figure 2.9. Sub-Cellular Localization of RSP with an <i>Arabidopsis</i> PAP.....	57
Figure 2.10. Pure PPF-B Contains an AMP-Forming Activity.	58
Figure 2.11. AMP-Forming Activity can be Eluted From DNA-Cellulose.	61
Figure 2.12. <i>Arabidopsis</i> PAPs Contain Potential Sites for Sumolation.....	62
Figure 3.1. Amino acid sequences of <i>Arabidopsis thaliana</i> Fip1 proteins.	88
Figure 3.2. Alignment of Peptide Domains Found in Fip1 Proteins.....	89
Figure 3.3. AtFip1 Genes are Expressed in all <i>Arabidopsis</i> Tissues.	90
Figure 3.4. Polypeptides Related to the <i>Arabidopsis</i> Fip1 are Present in Plant Extracts. 91	
Figure 3.5. Qualitative Interactions Between AtFip1-V and Other <i>Arabidopsis</i> 3'end Processing Homologs in yeast.	92
Figure 3.6. AtFip1 Stimulates the Non-Specific Activity of AtPAP(IV).....	95
Figure 3.7. The C-terminal Domain of AtFip1 (FipC) Binds RNA.....	97
Figure 3.8. The C-terminal Domain of AtFip1 (FipC) Binds Polyadenylation Sites. ...	100
Figure 3.9. The C-terminal Domain of AtFip1 (FipC) has a low affinity for RNA.	102
Figure 3.10. Model for the Organization of the Plant Cleavage and Polyadenylation Apparatus Around Fip1.	105

CHAPTER ONE

Literature Review

Functions of mRNA 3'end Formation and Poly(A) Tail

In eukaryotes, mature messenger RNAs are generated from larger RNA polymerase II generated precursors, by the processes of splicing, 5' capping and 3' polyadenylation that take place in the nucleus of the cell. Virtually all eukaryotic mRNAs possess a polyadenylate tract that is added post-transcriptionally in the nucleus as a part of the 3' end (with the best known exception being cell-cycle regulated histone mRNAs). These poly(A) tails are essential structural elements of eukaryotic mRNAs and have been suggested to influence most aspects of mRNA function.

Poly(A) tails have been shown to function in the initiation and efficiency of translation. Protein-mediated interactions between the cap of the 5' end and poly(A) tail of the RNA helps to load the 40S ribosome subunit onto the message resulting in enhancement of translation (Gallie, 1991; Sachs et al., 1997). The 5' terminal cap and poly(A) tail contact one another through an elongation initiation factor (eIF4G)-mediated interaction between cap binding protein eIF4E and poly(A) binding protein. This interaction helps recruit the translation initiation complex at the cap and helps load the 40S ribosome subunit onto the message. Although the poly(A) tail is not essential for translation (Proweller and Butler, 1994; Searfoss and Wickner, 2000), polyadenylated mRNAs are translated much more efficiently than mRNAs lacking a poly(A) tail (Proweller and Butler, 1997; Preiss and Hentze, 1998).

In addition, the poly(A) tail plays a role in the regulation of mRNA stability. Messenger RNA levels are regulated by a relatively rapid turnover rate in order to adjust for the physiological needs of the cell. The general pathways of mRNA decay in yeast are highly dependent on the initial steps of poly(A) tail shortening (Caponigro and Parker, 1996; Tucker and Parker, 2000; Parker and Song, 2004). This deadenylation is often the rate limiting step of mRNA decay since the ensuing steps of turnover are dependent on the length of the shortened poly(A) tail (Beelman and Parker, 1995). The proteins and enzymes involved in mRNA decay in yeast are conserved in higher eukaryotes and the pathway is then likely to also be conserved.

3' end formation is thought to contribute in the export of mRNA from the nucleus into the cytoplasm. Several studies in both yeast and metazoans have concluded that a poly(A) tail is essential for efficient mRNA export (Eckner et al., 1991; Fortes et al.,

1994; Huang and Carmichael, 1996; Brodsky and Silver, 2000). In contrast, it has been reported that mRNAs lacking a poly(A) tail can also be exported. In yeast strains in the absence of a functional poly(A) polymerase or when mRNA 3' ends have been formed by a self-cleaving ribozyme and thus lacking a poly(A) tail, mRNAs can be exported (Proweller and Butler, 1994; Duvel et al., 2002). Mutations affecting yeast 3' end processing factors lead to export defects (Hammell et al., 2002). There are significant defects in polyadenylation *in vivo* and *in vitro* in these mutant strains, however the mRNAs that are polyadenylated are not exported. In addition, the recruitment of export factors is less efficient in yeast strains with defects in the 3' end processing machinery (Lei and Silver, 2002). These studies imply a functional interaction between 3' end processing machinery and mRNA export factors. Therefore, 3' end formation and export of mRNA are suggested to be a mechanistically coupled process.

RNA Sequences Which Specify Cleavage and Polyadenylation

Sequences on RNA precursors determine the processing efficiency of cleavage and polyadenylation. These *cis*-acting elements specify where on the RNA precursor the cleavage and polyadenylation apparatus binds, where cleavage occurs and the site of poly(A) synthesis. These polyadenylation signals [poly(A) signals] have been suggested to be required for effective splicing, transcription termination, and possibly translation termination (Yeung et al., 1998; Cooke et al., 1999; Hilleren and Parker, 1999; Lei and Silver, 2002). Although there is growing information on poly(A) signals in several eukaryotic systems, in this review we will focus on the mammalian poly(A) signals that are well defined and the studies that have led to a better understanding of poly(A) signals in yeast, and then compare them to what is known about poly(A) signals in plants.

Mammalian polyadenylation signals are defined by three highly conserved elements: the poly(A) signal AAUAAA upstream of the cleavage site, U/GU-rich elements downstream of the cleavage site, and the cleavage poly(A) site itself, referred to as the point of poly(A) addition (references mentioned below; Figure 1.1). The AAUAAA sequence motif found roughly 10-30 nucleotides upstream of the cleavage site is one of the most highly conserved sequence elements known (Proudfoot, 1991; Graber et al., 1999a). Early studies suggested that AAUAAA was found in most if not all

mammalian mRNAs and it had been established that this hexanucleotide is essential for both cleavage and polyadenylation (Manley, 1988; Wickens, 1990; Wahle and Keller, 1992). There are few variants to AAUAAA. The most frequent variant motif of AAUAAA is AUUAAA, whose activity in cleavage and polyadenylation is comparable to AAUAAA (Graber et al., 1999a; MacDonald and Redondo, 2002; Zarudnaya et al., 2003). This variant and other non-AAUAAA polyadenylation signals have an unusually high incidence in male germ cell mRNAs (Wallace et al., 1999; MacDonald and Redondo, 2002). Although AAUAAA is the predominant poly(A) signal, these variants suggest that polyadenylation signal usage might differ in different cell types.

The downstream element (DSE; Figure 1.1) is less conserved than the AAUAAA motif and two main types have been described, an element either U- or GU-rich that is located within approximately 30 nucleotides downstream of the poly(A) site (Proudfoot, 1991; Colgan and Manley, 1997). The hexanucleotides UGUUUU and UGUGUU are the predominant consensus sequence found in DSEs as determined by computational analysis of animal ESTs (Graber et al., 1999a). A polyadenylation signal may contain one or more DSEs (Legendre and Gautheret, 2003; Zarudnaya et al., 2003). DSEs are present in a large number of genes and their deletion has been shown to suppress polyadenylation (Zhao et al., 1999a). This suggests that DSEs are important for proper cleavage and polyadenylation.

The site of cleavage and polyadenylation, the poly(A) site, is determined by the AAUAAA sequence and DSE(s) (Chen et al., 1995). The sequences surrounding the cleavage site are not conserved, although the dinucleotide sequence CA or UA often defines the poly(A) site for most genes (Sheets et al., 1990; Chen et al., 1995; Graber et al., 1999a). The rate of cleavage at this site proceeds very slowly, although cleavage and polyadenylation complexes (see below) form rapidly (Wahle and Ruegsegger, 1999).

A secondary sequence element, upstream sequence element (USE), has been found to modulate the efficiency of 3' end processing in a positive fashion (Figure 1.1). USEs are located upstream of AAUAAA and are often U-rich with no consensus sequence. USEs were originally identified in animal virus poly(A) signals (for review; Zhao et al., 1999a) and subsequently found in cellular genes (Moreira et al., 1995; Phillips and Virtanen, 1997; Brackenridge and Proudfoot, 2000; Aissouni et al., 2002).

Polyadenylation of transcripts for the cellular genes encoding complement factor 2 and lamin B2 is activated by their respective USEs *in vivo* (Moreira et al., 1995; Brackenridge and Proudfoot, 2000). Legrende and Gautheret (2003) have detected the accumulation of uridines upstream of the AAUAAA signal using a computational survey in a large number of human genes. This suggests that USEs are a common feature of polyadenylation signals.

Yeast polyadenylation signals also consist of at least three sequence elements (Figure 1.1): a UA-rich efficiency element and related sequences, an A-rich positioning element, and the site of cleavage and polyadenylation (Graber et al., 1999a, b). The UA-rich efficiency element (EE) and related sequences are located upstream of the cleavage site at a variable distance and functions to activate the positioning element (Guo and Sherman, 1996a, b; Zhao et al., 1999a; Dichtl and Keller, 2001). The sequence UAUUA is the most predominant EE used by yeast genes with other related sequences to UAUUA functioning as EEs (Graber et al., 1999a, b; Zhao et al., 1999a). The deletion of this element in the RNA precursors of the *GAL7* gene, prevents 3' end processing *in vitro* and *in vivo* (Hyman and Moore, 1993). However, Dichtl and Keller (2001) reported that the deletion of the EE in the gene *sCYC1* did not influence poly(A) site cleavage. This suggests that EE may be dispensable or may require interaction with other sequences for function.

The positioning element (PE; Figure 1.1), an A-rich sequence, directs cleavage approximately 10-30 nucleotides downstream of this sequence (Guo and Sherman, 1996a, b; Zhao et al., 1999a). The sequences AAUAAA and AAAAAA are the predominant PEs found in yeast genes (Graber et al., 1999a, b). Deletion or mutations in PE sequences change the location of the cleavage site, but not the level of mRNA processing (Guo and Sherman, 1996a, b; Zhao et al., 1999a; Dichtl and Keller, 2001). The PE contributes to poly(A) site cleavage only when an intact EE is present (Dichtl and Keller, 2001). This suggests that PE and EE elements work in concert for proper 3' end processing.

The poly(A) site in yeast is defined by the dinucleotide sequence UA or CA similar to mammalian poly(A) sites (Graber et al., 1999a). The poly(A) site is flanked by U-rich sequences upstream and downstream of the cleavage site which are presumed to be part of the cleavage site (Graber et al., 1999a, b; Dichtl and Keller, 2001). Mutations

within these U-rich elements of the cleavage site inhibit cleavage and the U-content at the poly(A) site reflects the strength of the poly(A) signal (Dichtl and Keller, 2001). This suggests that these U-rich elements represent a genuine poly(A) signal that works in concert with the poly(A) site.

As observed with mammalian and yeast polyadenylation signals, plant polyadenylation signals also consist of at least three sequence motifs (Figure 1.1): a U-rich far upstream element, an A-rich near upstream element, and the site of cleavage and polyadenylation (Rothnie, 1996; Li and Hunt, 1997). Far upstream elements (FUEs) are found at a variable distance from the near upstream element (NUE), ranging from 13-100 nucleotides upstream of the NUE. FUEs are U-rich (Hunt, 1994) and the hexanucleotides UUGUAU and UUGUAA are suggested to be the predominant consensus sequence as determined by computational analysis of plant ESTs (Graber et al., 1999a). Initial studies on plant polyadenylation signals identified FUEs in the cauliflower mosaic virus (CaMV) 19S/35S transcription unit, a T-DNA octopine synthase gene, and two cellular genes, the pea ribulose-1,5-bisphosphate carboxylase small subunit gene and the maize 27 kDa zein mRNA (summarized in Hunt, 1994). Deletion of these FUEs decreases the efficiency of their poly(A) sites. This suggests that FUEs are a general feature of plant poly(A) signals.

Plant polyadenylation signals may contain one or more NUEs as well as multiple cleavage sites. This may be due to potential for alternative polyadenylation. In genes with multiple cleavage sites, each site is controlled by a different NUE (MacDonald et al., 1991; Mogen et al., 1992). NUEs have been considered to be analogous to the mammalian polyadenylation signal AAUAAA, based on location and similarities in sequence (Figure 1.1). NUEs are found 6-40 nucleotides from their corresponding cleavage site and A-rich in sequence (Hunt, 1994). Computational analysis of plant ESTs, suggests a consensus sequence of AAUAAA and AAUGAA (Graber et al., 1999a). Mutational analysis has revealed that NUEs are essential for 3' end formation and can be as big as 10 nucleotides (Li and Hunt, 1995). NUEs are considered a distinct element and may function in concert with FUEs.

The cleavage and polyadenylation site [poly(A) site] in plants is usually situated in a U-rich region (Figure 1.1; Hunt, 1994). Computational analysis by Graber et al. (1999a) identified a dinucleotide sequence of UA and CA at the site of cleavage flanked

by downstream and upstream U-rich sequences (yA). This suggests that plant poly(A) sites resemble the yeast poly(A) site and are independent poly(A) signals.

In eukaryotes, a common minimal polyadenylation signal consists of a tripartite signal composed of an A-rich sequence, a U-rich element and a poly(A) cleavage site. When comparing the mammalian, yeast and plant poly(A) signals (Figure 1.1), it suggests that plant poly(A) signals appear to be more closely related to those of yeast than those of animals. In contrast with mammalian poly(A) signals, yeast and plant poly(A) signals lack the involvement of a downstream element. The U-rich sequences flanking the poly(A) site in yeast and plants may function as a DSE. A U-rich element has been identified in a few mammalian genes that surrounds the poly(A) site (Moreira et al., 1995; Brackenridge and Proudfoot, 2000). This suggests that U-rich poly(A) sites may play a more important role in 3' end formation.

mRNA Cleavage and Polyadenylation Apparatus

Multiple-subunit complexes in yeast and mammals mediate the cleavage and polyadenylation reaction (Figures 1.2 and 1.3). This complex recognizes the polyadenylation signal in the pre-mRNA, cleaves the pre-mRNA at a site that is defined by the *cis* elements, and adds a defined tract of poly(A). The subunits that comprise the cleavage and polyadenylation apparatus in these two organisms exhibit conservation.

In mammals, the factors involved in this process have been classified according to chromatographic and biochemical behaviors, and termed Cleavage and Polyadenylation Specificity Factor (CPSF), Cleavage stimulatory Factor (CstF), and Cleavage Factors I and II (CFIm and CFII_m, respectively). Each of these factors in turn consists of several distinct subunits. Early preparations of CPSF from calf thymus or HeLa cell extracts identified CPSF as large protein complex containing four subunits of 160, 100, 73, and 30 kDa, referred to as CPSF160, CPSF100, CPSF73, and CPSF30, respectively (Bienroth et al., 1991; Murthy and Manley, 1992). CPSF is involved in cleavage as well as polyadenylation (Bienroth et al., 1991). Recently, a fifth integral subunit of CPSF has been identified, hFip1 (66 kDa), the human homolog to the yeast Fip1p (Kaufmann et al., 2004). Interactions between the CPSF subunits, has been demonstrated by co-purification and co-immunoprecipitation with antibodies (Bienroth et al., 1991; Murthy and Manley,

1992; Jenny et al., 1994; Kaufmann et al., 2004). CPSF160 interacts with all four of the other subunits (Figure 1.2) and has been suggested to preferentially bind the poly(A) signal AAUAAA based on UV cross-linking studies (Keller et al., 1991; Jenny and Keller, 1995). In addition, CPSF160 has been shown to interact with PAP and the CTD of RNA polymerase II (Pol II; see below) (Murthy and Manley, 1995; McCracken et al., 1997). CPSF30 is another protein that has also been shown to bind RNA with a preference for poly(U) sequences (Barabino et al., 1997). hFip1, which has been reported to interact with CPSF30, CPSF160, CstF77 (see below), and PAP (see below), and also binds RNA with a preference for U-rich sequence elements (Kaufmann et al., 2004). The CPSF73 subunit has been suggested to function as the 3' processing endonuclease (Ryan et al., 2004) and the function of the CPSF100 subunit remains largely undefined.

The binding of AAUAAA by CPSF, through its 160 kDa subunit, is strengthened by the interaction with CstF (Bienroth et al., 1993). CstF is necessary for cleavage but not polyadenylation (Gilmartin and Nevins, 1989; Takagaki et al., 1989) and consists of three polypeptides of 77, 64, and 50 kDa, referred to as CstF77, CstF64, and CstF50. CstF64 binds to the downstream U/GU-rich element in pre-mRNA (Takagaki and Manley, 1997; Perez Canadillas and Varani, 2003). CstF77 is a bridging factor between CstF64 and CstF50 and also mediates the interaction between CstF and CPSF by interacting with its 160 kDa subunit. This suggests that the binding of poly(A) signals by the complex is strengthened by the interaction between CPSF160 and CstF77. CstF50 is one of the two subunits that bridge Pol II to the cleavage and polyadenylation apparatus through an interaction with the CTD of Pol II, the other subunit being CPSF160 (McCracken et al., 1997). In addition, the CstF50 subunit is suggested to be a link to other cellular processes (see below). Another subunit that has been identified, referred to as symplekin, interacts with CstF64 and its proposed function is to facilitate or maintain properly assembled CstF and to possibly help hold together the complete polyadenylation complex (Takagaki and Manley, 2000).

Cleavage factors I and II (CFIm and CFII_m; m=mammalian) are required only for the cleavage reaction. CFIm consists of a mixture of four polypeptides of 72, 68, 59, and 25 kDa and has been demonstrated to facilitate complex formation, enhance the overall efficiency of poly(A) site cleavage, and regulate pre-mRNA 3' processing (Ruesegger et

al., 1996; Ruegsegger et al., 1998; Brown and Gilmartin, 2003). The three larger polypeptides appear to be highly related in their amino acid sequence. The 25 kDa subunit can interact with any one of the larger polypeptides to form a heterodimer. The 25 and 68 kDa subunits are sufficient to reconstitute CFIm activity *in vitro* (Ruegsegger et al., 1998). CFIm has been shown to bind RNA, which is dependent on the 25 and 68 kDa heterodimer (Dettwiler et al., 2004). CFIIIm consists of two components CFIIAm and CFIIIm, based on their activities during purification. CFIIAm is an essential fraction and CFIIIm is a non-essential, but stimulatory component of the cleavage reaction *in vitro* (de Vries et al., 2000). Little is known about the composition of CFIIIm and several polypeptides have been identified in CFIIAm. CFIIAm contains the human homologs to two yeast 3' end processing factors, Pcf11p and Clp1p (see below), as well as CFIm and several splicing and transcription factors (de Vries et al., 2000). de Vries et al. (2000) also observed that hClp1 interacts with CFIm and CPSF. This suggests that CFIIIm bridges these two factors within the cleavage complex.

Poly(A) polymerase (PAP; 82/77 kDa), is the enzyme that synthesizes the poly(A) tail, and it is well established that this activity is essential for proper 3' end formation of mRNAs in eukaryotic cells (Zhao et al., 1999a). The function and regulation of PAPs are discussed in the following section. In the mammalian complex, PAP has been reported to interact with CPSF160, hFip1, the nuclear poly(A) binding protein (PABN1), and the 25 kDa subunit of CFIm (Murthy and Manley, 1995; Kim and Lee, 2001; Kerwitz et al., 2003; Kaufmann et al., 2004). PAP has also been suggested to be required for cleavage of most pre-mRNAs (Murthy and Manley, 1995).

Two factors mentioned above, but not described, Pol II and PABN1 are essential for 3' end formation. Pol II (200 kDa) interacts with CPSF and CstF. It has been suggested that Pol II recruits polyadenylation factors to poly(A) signals and is required, in the absence of transcription, for 3' end formation (Hirose and Manley, 1998). PABN1 (33 kDa) binds the elongating poly(A) tail and serves to control poly(A) tail length (Bienroth et al., 1993) and has recently been reported to be necessary for the stimulation of PAP activity (Kerwitz et al., 2003). Studies by Chen et al. (1999) suggest that PABN1 can interact with CPSF30 *in vitro*. Taken together, these results suggest that Pol II and PABN1 are bona fide factors of the mammalian cleavage and polyadenylation complex.

Protein factors involved in cleavage and polyadenylation of precursor mRNAs have been identified in yeast based on biochemical fractionation studies and are required for one or both of the steps involved in 3' end formation. These factors are termed cleavage factor IA (CFIA), cleavage factor IB (CFIB), cleavage factor II (CFII), polyadenylation factor I (PFI), and poly(A) polymerase (Pap1p) (Chen and Moore, 1992; Kessler et al., 1996). CFIA, IB, and II have been suggested to be sufficient for cleavage and CFIA and CFIB, Pap1p, PFI, and poly(A) binding protein (Pab1p) is required for specific poly(A) addition (Zhao et al., 1999a). PFI, CFII, and Pap1p can also be found in a larger complex, cleavage and polyadenylation factor (CPF) (Preker et al., 1997). For simplicity, the model in Figure 1.3 was assembled according to CPF. All of the genes that have been cloned for the yeast cleavage and polyadenylation subunits, and the subsequent mutational analysis of these genes, suggests that they are essential for cell viability (Zhao et al., 1999a). In the subsequent description of these factors, mammalian homologs will be in parentheses following the listed subunit where applicable.

CFIA consists of four protein subunits of 76, 72, 50, and 38 kDa termed Rna14 (CstF77), Pcf11p (hPcf11), Clp1p (hClp1), and Rna15 (CstF64) respectively (Minvielle-Sebastia et al., 1994; Kessler et al., 1996; Amrani et al., 1997). Rna14 tightly associates with Rna 15 and interacts with Pcf11p, which also simultaneously binds Clp1p and Rna15 (Zhao et al., 1999a; Gross and Moore, 2001a). CFIA binds to the A-rich positioning element (PE) via its Rna15 subunit (Gross and Moore, 2001b). This suggests that CFIA directs the cleavage site via Rna15 and interactions with other cleavage factors.

CFIB consists of one subunit of 73 kDa termed Hrp1p (Kessler et al., 1996). Hrp1p interacts with Rna14 and Rna15 from CFIA and binds to the UA-rich upstream efficiency element (Zhao et al., 1999a; Gross and Moore, 2001a; Gross and Moore, 2001b). The five proteins of CFIA/B (Rna14, Pcf11p, Clp1p, Rna15 and Hrp1p) in combination are necessary and efficient for reconstitution of cleavage activity with CFII (described below) and polyadenylation with PFI, Pap1p, and Pab1p (described below; Gross and Moore, 2001a).

CFII consists of four subunits of 150, 105, 100, and 90 kDa termed Yhh1p (CPSF160), Ydh1p (CPSF100), Ysh1p (CPSF73), and Pta1p (symplekin) respectively

(Stumpf and Domdey, 1996; Zhao et al., 1997; Zhao et al., 1999a; Zhao et al., 1999b). *In vitro* pull-down assays with recombinant proteins have shown that Ydh1p interacts with Yhh1p, Ysh1p, Pta1p, Pcf11p, and Pfs2p (described below) (Kyburz et al., 2003). Yhh1p interacts with Rna14, to a lesser extent with Pcf11p, and weakly to Clp1p (Kyburz et al., 2003). Pta1p binds Pcf11p weakly and Ysh1p interacts with Clp1p, Pcf11p, and Yth1p (see below) (Barabino et al., 2000; Kyburz et al., 2003). Dichtl and Keller (2001; 2002a) have suggested that CFII binds the U-rich cleavage/poly(A) site. Recognition of the poly(A) site by CFII has been suggested to be through the Yhh1p and Ydh1p subunits, since they have a preference for U-rich sequences (Dichtl and Keller, 2001; 2002a). This suggests that the site of cleavage is directed by CFIA binding of the positioning element and protein-protein interactions between CFIA and CFII subunits, and as a result CFII recognizing the cleavage site.

PFI was first identified as an activity involved in poly(A) addition but not cleavage (Chen and Moore, 1992). PFI consists of four polypeptides of 58, 55, 53, and 24 kDa termed Pfs1p or Mpe1p, Fip1p (hFip1), Pfs2p, and Yth1p (CPSF30) respectively (Preker et al., 1995; Preker et al., 1997; Ohnacker et al., 2000; Vo et al., 2001). Fip1p interacts with Yth1p, Pfs2p, Rna14, and Pap1p (Preker et al., 1995; Barabino et al., 1997). The gene for Pfs1p has not been characterized and a protein similar to Pfs1p, Mpe1p, has been suggested to interact functionally with Pcf11p (Vo et al., 2001). Like Yhh1p and Ydh1p, Yth1p has been reported to bind U-rich sequences, which suggests Yth1p may be directly involved in the recognition of poly(A) sites (Barabino et al., 2000).

Recently, another sub-complex termed APT (Associated with Pta1p) has been identified using a TAP (tandem affinity purification)-tagging approach to identify components of the yeast cleavage and polyadenylation machinery and factors that link 3' end formation to transcription (Nedea et al., 2003). TAP purification identified APT and core cleavage and polyadenylation factor components CFII, PFI, and Pap1p. APT contains Pta1p and six other polypeptides, with a few of them being involved in transcription elongation and termination. Nedea et al. (2003) have suggested that APT is initially recruited to promoters and therefore suggest that the yeast cleavage and polyadenylation complex is recruited to the transcription unit.

Poly(A) polymerase (Pap1p; 64 kDa), as with the mammalian PAP, is the enzyme that synthesizes the poly(A) tail (Zhao et al., 1999a). The function and regulation of PAPs are discussed in the following section. Fip1p is the only subunit reported to interact with Pap1p, and unlike the mammalian PAP, it is not required for efficient cleavage of pre-mRNA *in vitro* (Lingner et al., 1991a; Lingner et al., 1991b; Preker et al., 1995).

Poly(A) binding proteins in yeast have also been identified. It has been suggested that there is a requirement for two poly(A) binding proteins, Nab2p (55 kDa) and Pab1p (70 kDa), for proper control of polyadenylation (Figure 1.3). Nab2p is the predominant nuclear poly(A) binding protein and Pab1p the predominant cytoplasmic form, although found in the nucleus at very low levels (Anderson et al., 1993a; Hector et al., 2002). Nab2p binds poly(A) and has been suggested to play a role in mRNA export (Anderson et al., 1993b; Green et al., 2002). Hector et al. (2002) have speculated that Nab2p regulates poly(A) length by functioning as a polyadenylation termination factor (terminates poly(A) synthesis). Pab1p has also been suggested to function in limiting the length of poly(A) and interacts with the Rna15 subunit of CFIA, but is not required for cleavage (for review see Zhao et al., 1999a). Two roles have been suggested for Pab1p regulating tail length: suppression of Pap1p activity by limiting its access to RNA substrate (Zhelkovsky et al., 1998) and recruitment of poly(A)-specific nuclease (PAN; Figure 1.3) (Lowell et al., 1992; Deardorff and Sachs, 1997). Pan consists of two subunits of 127 and 76 kDa termed Pan2 and Pan3 respectively (Zhao et al. 1999a), which are encoded by genes that are not essential for cell growth (Boeck et al., 1996; Brown et al., 1996). Pan2 interacts with Pan3 and Pab1p, and in combination, is suggested to shorten the poly(A) tail via nuclease activity (Brown and Sachs, 1998). Hector et al. (2002) have proposed a hypothesis for Nab2p and Pab1p function in the regulation of poly(A) tail length. They have suggested that Nab2p is recruited initially to the mRNA 3' end by the growing poly(A) tail and then Pab1p recruits Pan activity for trimming of the poly(A) tail. The importance of poly(A) binding protein function on mRNA 3' end formation in yeast should be strengthened by further experimentation.

The yeast RNA polymerase (Pol II; Rpb1) has not been reported to have an affect of cleavage and polyadenylation. However, mutations in poly(A) signals (Russo and

Sherman, 1989) and genes encoding cleavage factors (Birse et al., 1998; Greger et al., 2000) result in defects in termination of transcription by Pol II. In addition, the CTD of Pol II in yeast has been reported to interact with components of CFIA (Barilla et al., 2001) and CPF (Rodriguez et al., 2000; Dichtl et al., 2002b; Dichtl et al., 2002a). Furthermore, the Rna15 subunit of CFIA has been found to interact with two proteins involved in the initiation of transcription by Pol II, Sub1 and Mbp1 (Aranda and Proudfoot, 2001; Calvo and Manley, 2001). These observations suggest that Pol II and 3' end processing factors mediate a coupling of polyadenylation and transcription.

Plant polyadenylation signals have been well characterized and require a minimal tripartite signal as observed in mammalian and yeast signals (described above). Although similar in structure, there are some unique aspects of plant poly(A) signals, suggesting differences in 3' end formation in higher eukaryotes. The properties of the plant cleavage and polyadenylation apparatus is poorly understood when compared to data available for plant poly(A) signals. Genes that encode products with significant similarity to mammalian and yeast polyadenylation factor subunits have been identified in various plant genomes (Q. Q. Li and A. G. Hunt, unpublished observations). Protein-protein interactions between some of these cloned homologs have recently been reported. *Arabidopsis* possesses four PAP-encoding genes (Addepalli et al., 2004) and at least one of the four PAPs interacts with the *Arabidopsis* homolog to CPSF100 (Elliott et al., 2003). Also, the *Arabidopsis* CstF subunits (AtCstF50, AtCstF64, AtCstF77) have a conserved CstF interaction, an interaction between CstF77 and CstF64 (Yao et al., 2002). In contrast to what has been shown in the mammalian CstF factor (Takagaki and Manley, 2000), AtCstF77 does not interact with AtCstF50. The *Arabidopsis* poly(A) binding protein (PAB3) has also been reported to have an evolutionarily conserved function in mRNA biogenesis and export (Chekanova and Belostotsky, 2003). Recent reports have suggested links of mRNA 3' end processing machinery in modulating plant development. Gene disruptions, by T-DNA insertions or Dissociation (*Ds*) insertion, in *Arabidopsis* genes encoding homologues of CPSF100 and CPSF73 lead to embryo lethality respectively (Meinke et al., 2003; Xu et al., 2004). An *Arabidopsis* homologue of the yeast polyadenylation factor subunit Pfs2p (the *Arabidopsis* protein has been termed FY) has been shown to act in concert with the flower-timing regulatory protein FCA to

promote alternative polyadenylation of FCA-encoding RNAs and consequently to regulate flower timing (Simpson et al., 2003). Analysis of the *Arabidopsis* genome has identified a more complex set of RNA-binding proteins than found in *C. elegans* and *D. melanogaster*, suggesting that RNA processing may be more diverse in plants (Lorkovic and Barta, 2002). These reports suggest that the analysis of mRNA 3' end formation in plants is feasible. While many of the other potential *Arabidopsis* cleavage and polyadenylation factors have yet to be characterized, it suggests that polyadenylation is likely to be important in regulated as well as constitutive gene expression.

Function and Regulation of Poly(A) Polymerases

Poly(A) polymerase, the catalytic unit of the polyadenylation machinery, is a template-independent RNA polymerase that incorporates ATP at the 3' end of mRNA into a poly(A) tail. This template independent activity is referred to as non-specific PAP activity (for review see Zhao et al., 1999a). PAPs belong to a large superfamily of nucleotidyl transferases (Holm and Sander, 1995; Martin and Keller, 1996; Aravind and Koonin, 1999; Martin et al., 1999). PAPs possess a catalytic domain near their N-terminus and an RNA-binding domain that overlaps a nuclear localization signal (NLS) near the C-terminus (Zhelkovsky et al., 1995; Martin and Keller, 1996; Martin et al., 1999).

Martin et al. (2000) have solved the crystal structure of the mammalian PAP and as expected, the catalytic domain shares structural homology with other nucleotidyl transferases. In the absence of other poly(A) factors, mammalian PAP activity is low and distributive due to its weak and non-specific binding to RNA (Wahle and Keller, 1992). However, in the presence of CPSF, PAP specifically and efficiently polyadenylates RNAs containing poly(A) signals (Keller et al., 1991; Wahle, 1991). The C-terminal domain of PAP in mammals has been shown to be a regulating domain of PAP activity. The C-terminus contains an S/T rich region, which contains multiple cyclin-dependent kinase (cdk) sites, which are phosphorylated *in vitro* and *in vivo* by p34cdc/cyclinB (Colgan et al., 1998). PAP activity is regulated by phosphorylation during the cell cycle and p34cdc/cyclinB hyperphosphorylates PAP during mitosis and meiosis (M phase), resulting in repression of PAP activity (Colgan et al., 1996; Colgan et al., 1998). PAP

activity is also repressed upon interaction of its C-terminus with the U1 snRNP A protein (U1A) (Gunderson et al., 1997). The C-terminus of PAP is also involved in protein-protein interactions with the splicing factor U2AF65 (Vagner et al., 2000), the 25 kDa subunit of CFI (Kim and Lee, 2001), and 14-3-3 ζ a member of the 14-3-3 protein family (Kim et al., 2003). 14-3-3 ζ inhibits PAP activity *in vitro* and overexpression of 14-3-3 ζ leads to shorter poly(A) tails *in vivo* (Kim et al., 2003). These reports demonstrate that PAP function is regulated by several factors/components found within its C-terminus.

The yeast poly(A) polymerase, Pap1p, is similar to its mammalian counterpart based on domain organization and crystal structure (Bard et al., 2000). However, Pap1p lacks the regulatory C-terminal domain found in mammalian PAPs. Unlike the mammalian PAP, Pap1p has a high affinity for and rapidly elongates non-specific RNA, in a processive manner (Zhelkovsky et al., 1998). Fip1p is the only known protein to interact with Pap1p from the cleavage and polyadenylation complex, and regulates the activity of PAP through multiple interactions (Helmling et al., 2001). It has been reported that Pap1p is phosphorylated and ubiquitinated at the S/G₂ stage of the cell cycle and phosphorylation of Pap1p appears to inactivate Pap1p (Mizrahi and Moore, 2000). In addition, Pap1p interacts with Uba2, a homolog of ubiquitin-activating (E1) enzyme, and Ufd1, whose function is probably linked to the ubiquitin-mediated protein degradation pathway (del Olmo et al., 1997). These proteins are suggested to play a role in regulating levels of poly(A) polymerase in the cell, to eliminate defective Pap1p or Pap1p independent of polyadenylation specificity factors (ie. Pap not found in the 3' end complex). Overexpression of PAP in chicken cells has been shown to be detrimental to cell growth (Zhao and Manley, 1998). This supports the hypothesis that Pap, not in the complex, could be lethal to the cell and regulation by protein degradation is essential.

Poly(A) polymerases that correspond to chloroplast and nuclear enzymes have been characterized to some extent from several plant species (Hunt and Messing, 1998). Polyadenylation has been characterized in the chloroplast and polyadenylated RNAs are degraded rapidly (Kudla et al., 1996; Lisitsky et al., 1996). This suggests that chloroplasts exhibit an analogous function seen in bacteria, a role for polyadenylation in accelerated RNA turnover (Ingle and Kushner, 1996). Little is known about how this functions in the regulation of gene expression in the chloroplast. The *Arabidopsis* genome possesses four

nuclear PAP-encoding genes, and recombinant proteins of all four possess PAP enzyme activity (Hunt et al., 2000; Addepalli et al., 2004). All four share the conserved N-terminal catalytic core with mammals and yeast, but differ substantially at their C-terminus. Alternative splicing may be a potential means of regulating these genes and their gene products. All four share novel alternative splicing events, which differ in the alternative splicing of PAP-related mRNAs documented in mammals (Addepalli et al., 2004). In addition, these *Arabidopsis* PAPs are expressed at the level of mRNA in tissue specific ways, suggestive of differential roles in polyadenylation. Some of which may relate to mRNA specific modifications if these genes are expressed. The function of these PAPs and how they are involved in cleavage and polyadenylation is unknown and will require further experimentation.

Importance of the Fip1 Subunit

Of all the subunits involved in cleavage and polyadenylation in mammals and yeast, hFip1/Fip1p is of particular interest. Fip1 appears to be an important bridge between several polyadenylation factors (see above; Figure 1.2 and 1.3). In yeast, Fip1p is the only polyadenylation factor subunit that has been shown to interact with PAP. Fip1p also interacts with Yhh1p (CPSF160), Yth1p (CPSF30), Pfs2p, and RNA14 (CstF77), components of the two major polyadenylation complexes (CPF and CFII) in yeast. The human homologue, hFip1, interacts with PAP, CPSF160, CPSF30, and CstF77 and has been recently recognized as an authentic subunit of CPSF. The yeast and human Fip1 proteins have somewhat contrasting properties; the yeast protein lacks an RNA-binding domain and inhibits the non-specific activity of PAP, while the human Fip1 can bind RNA and stimulates PAP activity. Kaufmann et al. (2004) have suggested that these contrasting properties may reflect the differing RNA-binding abilities of the two proteins, and that the yeast protein, in concert with other components of CPF, may stimulate PAP much as does the human Fip1. In support of this possibility, purified CPF stimulates PAP activity *in vitro* (Preker et al., 1997). In this regard, the functioning of Fip1 in the two systems may be relatively conserved, serving to promote PAP activity via some sort of tethering to the RNA substrate and regulating poly(A) tail synthesis.

The Interplay of 3' End Formation and Other Cellular Processes

A number of recent studies indicate that messenger RNA 3' end formation is coupled with other processes that impact the production and/or stability of an mRNA (Proudfoot and O'Sullivan, 2002). The most obvious and established linkage with 3' end formation is that with transcription initiation and termination. Much of this linkage involves the transcribing enzyme RNA polymerase II and specifically the C-terminal domain of its large subunit. Recent evidence implicates the CTD in enhancing capping, splicing and polyadenylation (Hirose and Manley, 2000; Bentley, 2002). In mammals, the CTD binds subunits of CPSF and CstF (McCracken et al., 1997). Likewise in yeast, the CTD of RNA polymerase II interacts with subunits of the yeast equivalents of CstF and CPSF (Barilla et al., 2001; Dichtl et al., 2002a). The 64 kDa subunit of CstF associates with PC4, a transcriptional co-activator (Calvo and Manley, 2001), and its yeast counterpart (RNA15p) interacts with Sub1, which functions as an antiterminator (Aranda and Proudfoot, 2001), thus implicating 3' end processing factors in transcription termination. This suggests that coupling of transcription and polyadenylation is conserved throughout eukaryotes. A general transcription factor TFIID, also co-purifies with CPSF, and as mentioned above CPSF interacts with the CTD (Dantanel et al., 1997). CPSF transfers from TFIID to pol II with the initiation of transcription (Dantanel et al., 1997). These observations suggest that CPSF is associated with the transcription apparatus from the earliest steps at the promoter, first via TFIID and then via the CTD of RNA polymerase II.

The polyadenylation apparatus interacts in specific ways with various splicing factors. Lutz et al. (1996) have shown an interaction between the U1 snRNP-A protein and the 160 kDa subunit of CPSF that increases polyadenylation efficiency *in vitro*. This stimulation has been suggested to be due to U1A increasing the RNA-binding affinity of CPSF160 for the AAUAAA polyadenylation signal. Polyadenylation is thought to play an active role in the removal of the last intron and the definition of 3' terminal exons (Wassarman and Steitz, 1993; Lutz and Alwine, 1994). Mutations of the 3' splice site of the last exon reduces polyadenylation and mutations of the AAUAAA and the downstream elements of the poly(A) signal decrease splicing efficiency (Cooke et al., 1999), thus suggesting that these elements are involved in the coupling of splicing and

polyadenylation. Mutation of the 5' splice site in the last intron has no effect on polyadenylation, however a site for U1 snRNP binding (a 5' splice site) within the last exon can negatively effect both polyadenylation and splicing (Cooke et al., 1999). In summary, the definition of the last exon, whether by protein-protein interactions with snRNPs and the 3' end polyadenylation apparatus, or through sequence elements, may couple splicing and polyadenylation processes. These reports suggest that splicing plays a more global role in RNA processing through effects on polyadenylation.

Other evidence suggests that polyadenylation plays an important role in stable RNA turnover and regulation. The exosome, a complex of 3' to 5' exonucleases, processes rRNAs, snoRNAs, and snRNAs in the nucleolus and also participates in the mRNA degradation in the cytoplasm in both mammalian and yeast cells (Mitchell et al., 1997; Anderson and Parker, 1998; Butler, 2002; van Hoof et al., 2002; van Hoof and Parker, 2002). Recently, it has been reported that the exosome also functions in mRNA degradation in the nucleus (Bousquet-Antonelli et al., 2000; Libri et al., 2002; Torchet et al., 2002; Zenklusen et al., 2002). Rrp6p interacts genetically and physically with Pap1p in yeast and has been proposed to degrade mRNAs in competition with Pap1p (Briggs et al., 1998; Burkard and Butler, 2000). The protein Rrp6p is a subunit of the nuclear exosome, which plays a role in the precise 3' end formation of many mRNAs and the degradation of excess rRNA and processing of their intermediates (Briggs et al., 1998). This suggests a link between the nuclear exosome and the 3' processing system. In yeast, rRNAs can be polyadenylated to a certain level and polyadenylated rRNAs accumulate in an Rrp6p mutant strain of yeast (Kuai et al., 2004). The poly(A) polymerase (Pap1p) appears to be responsible for the polyadenylation of these rRNAs and it has been suggested that Rrp6p is competing with the polyadenylation machinery to shorten the poly(A) tail of rRNAs (Kuai et al., 2004).

The 50 kDa subunit of CstF interacts with a protein (BARD1) that plays a role in the repair of DNA damage in mammalian cells (Kleiman and Manley, 1999, 2001). BARD1 has also been implicated in tumor suppression through its association with the BRCA1 protein. BRCA1, a breast and ovarian cancer tumor suppressor protein, is suggested to maintain genomic stability in response to DNA damage. BRCA1 has also been shown to interact with several components of the transcription complex. These

interactions are suggestive of a link between the 3' end processing machinery and chromosome dynamics or DNA repair.

Each of the above mentioned cellular processes of transcription, splicing, RNA turnover, and DNA damage repair have been linked to polyadenylation. This suggests that the networking of all these processes is a conserved regulating step in gene expression and that mRNA maturation is not dependent on one process alone. No one process working independently and no one reaction more important than the other for proper mRNA processing/maturation. Overall, proper mRNA maturation is an essential step in maintaining cell viability.

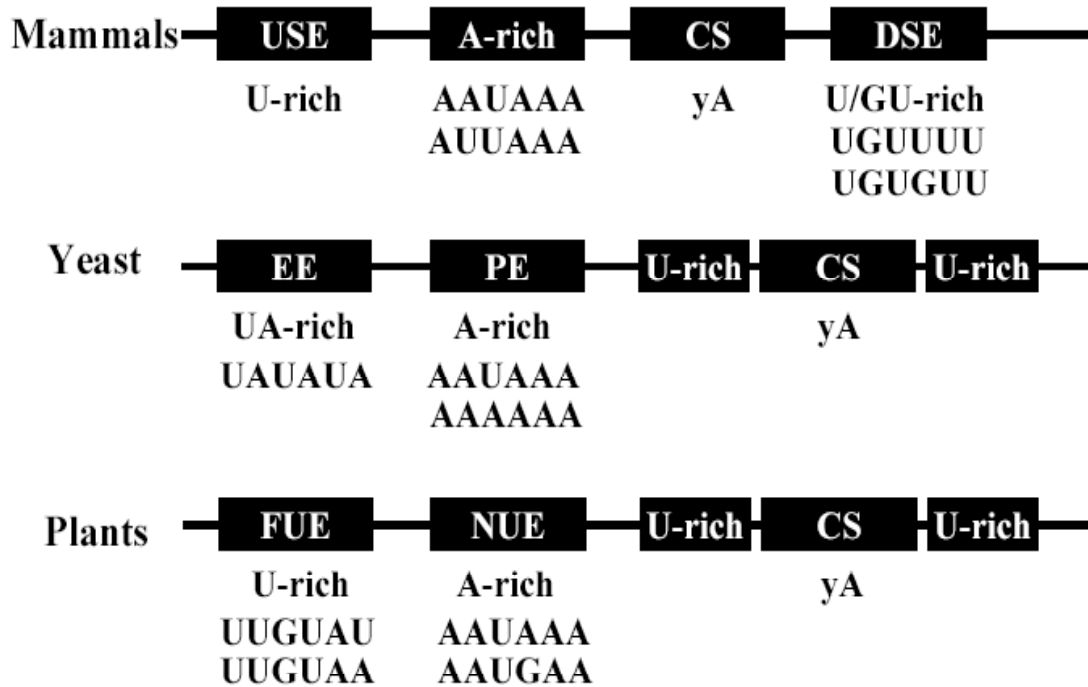


Figure 1.1. Eukaryotic Poly(A) Signals.

Graphical representation of the polyadenylation signals for mammals, yeast, and plants. Black boxes represent the elements found in each signal (see text for description). Sequences below each element represent the predominant consensus sequence. CS represents the cleavage poly(A) site (y = pyrimidine).

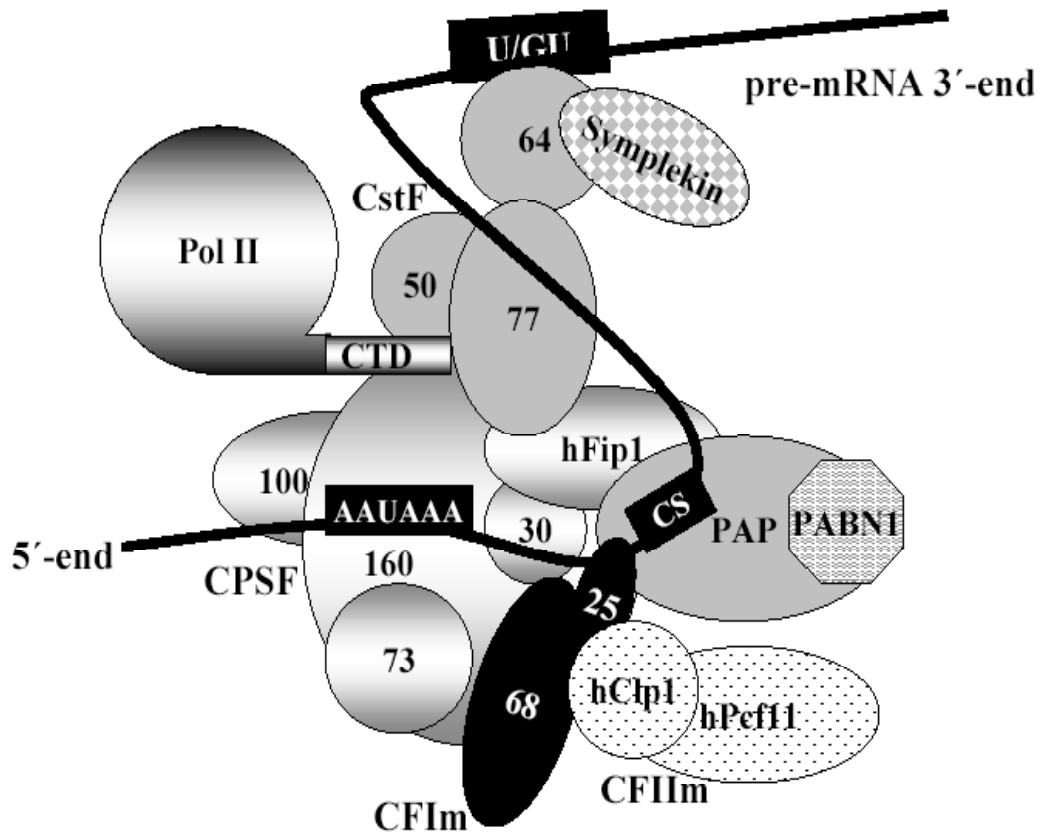


Figure 1.2. Mammalian Cleavage and Polyadenylation Complex.

A schematic model of the mammalian cleavage and polyadenylation apparatus (see text for description). Patterns and shading represent individual subunits or subunits within the same factor, for example, CstF factors are all gray (see text for function).

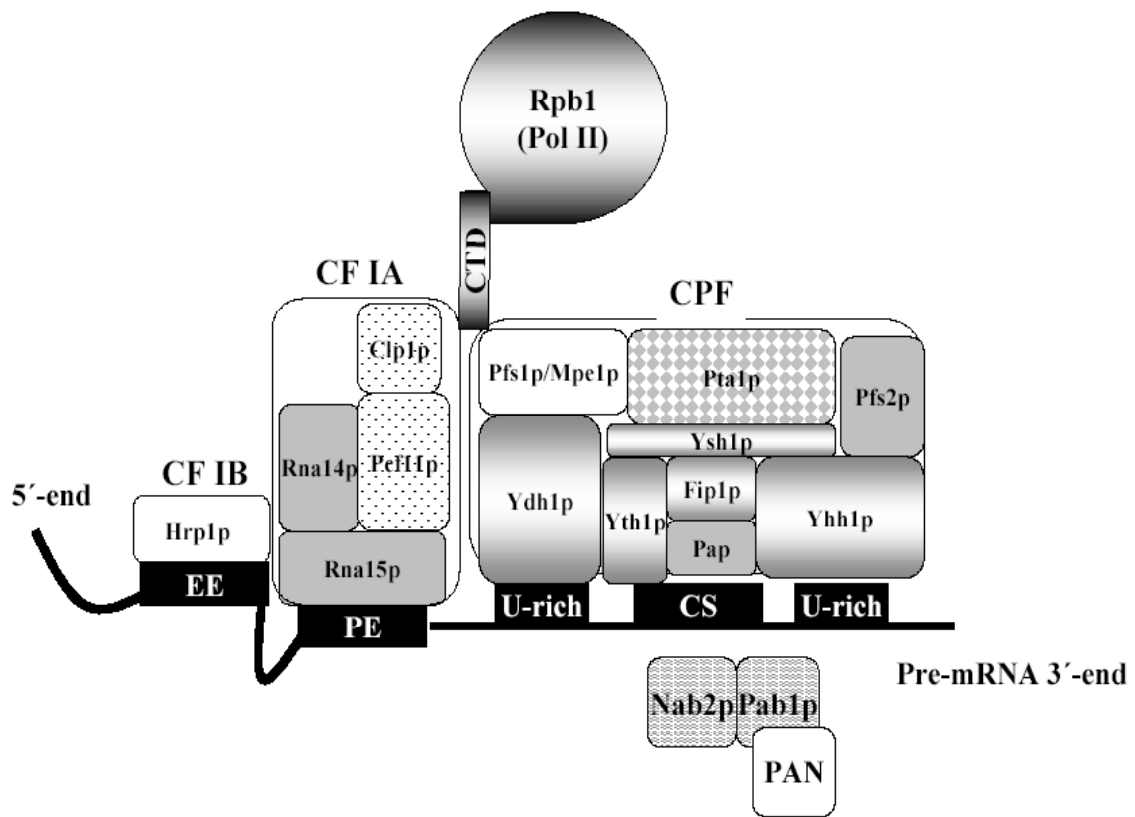


Figure 1.3. Yeast Cleavage and Polyadenylation Complex.

A schematic model for the yeast cleavage and polyadenylation apparatus (see text for description). Patterns and shading represent the homologs identified in the mammalian complex in Figure 1.2. Not all protein-protein interactions are represented for simplicity (see text for function).

CHAPTER TWO

Characterization of Putative Polyadenylation Factor-B (PPF-B), an Inhibitor of Non-Specific Poly(A) Polymerase Activity

INTRODUCTION

Poly (A) tails are a feature of most mature eukaryotic mRNAs (Proudfoot, 1991). Polyadenylated mRNAs have been identified as well in mitochondria (Ojala et al., 1981), bacteria (Sarkar, 1997), and chloroplasts (Kudla et al., 1996; Lisitsky et al., 1996). In eukaryotes, it is now assumed that poly (A) tails play essential roles in conferring mRNA stability, promoting translation, and facilitating the transport of processed mRNAs from the nucleus to the cytoplasm (Manley and Proudfoot, 1994). In contrast, in bacteria, it appears that poly (A) tails serve to increase the rate at which RNAs are degraded *in vivo* (Xu et al., 1993; Hajnsdorf et al., 1995; O'Hara et al., 1995) and *in vitro* (Xu and Cohen, 1995).

Compared to other eukaryotic systems, nuclear RNA metabolism in plants is poorly understood. This is especially true for mRNA 3' end formation and polyadenylation. Thus, while a good understanding of the RNA sequence requirements for mRNA 3' end formation has been obtained (Rothnie, 1996; Li and Hunt, 1997; Graber et al., 1999a), little is known about the factors that recognize plant polyadenylation signals or polyadenylate mRNAs in plants. Plant genes that encode products with significant similarity to mammalian and yeast polyadenylation factor subunits can be identified in various genomes (Q. Q. Li and A. G. Hunt, unpublished observations), but the functioning of most of these in the process of mRNA 3' end formation in plants remains to be demonstrated or understood. In addition, plants possess at least one identifiable enzymatic activity that is associated with mRNA polyadenylation, namely a poly(A) polymerase (PAP). Probable nuclear PAPs (as opposed to the chloroplast PAP; (Das Gupta et al., 1998; Li et al., 1998) were described as long as 26 years ago (Mans and Huff, 1975) and have been reported in a number of plant species (summarized in Hunt and Messing, 1998).

An alternative to sequence comparisons for the identification and characterization of polyadenylation factors is to identify factors that modify the activity of poly(A) polymerases. For example, the mammalian factor CPSF has the property that it can stimulate the activity of the PAP when using RNAs that possess authentic polyadenylation signals (Christofori and Keller, 1988; Keller et al., 1991). Using this as a

rationale, we set out to identify potential polyadenylation factors from plants that change the activity of the plant PAP. In so doing, we have discovered a novel factor (termed herein as PPF-B) that effectively inhibits the non-specific activity of plant PAPs. We describe in this chapter the identification, purification, and characterization of this from pea nuclear extracts.

RESULTS

Purification of an Inhibitor of Poly(A) Polymerase

To identify factors that inhibit PAP activity, pea nuclear extracts were first fractionated on DEAE-Sepharose much as is done to isolate the nuclear PAP (nPAP). Initial immunoblot studies indicated that these nuclear extracts were free of detectable ribulose-1,5-bisphosphate carboxylase/oxygenase (not shown; see also Das Gupta et al., 1998; Li et al., 1998), indicating a lack of chloroplast contamination of the nuclei used to prepare the extracts. Column fractions were then assayed in the presence of nPAP. For this, the non-specific assay used to purify and characterize the pea chloroplast PAP was used (Das Gupta et al., 1995). In this assay, enzyme, labeled ATP, and poly(A) (average length about 300 nts) is incubated at 30°C for given periods of time. Subsequently, the reactions are terminated by the addition of phenol + chloroform and incorporation of label into polynucleotide assessed by determining the label that is retained on DEAE ion-exchange paper in high phosphate. This assay has the advantage that it is sensitive and can be used to follow the purification of the plant PAPs through stages in which small amounts of nuclease activity are present. Moreover, the assay is non-specific in the sense that it does not require a polyadenylation signal.

As seen in Figure 2.1A, a broad range of DEAE column fractions (fractions 19-29) inhibited PAP activity. The inhibitory activity eluted before the PAP (which was present in fractions higher than 32). Fractions showing this inhibition were pooled, dialyzed, and further purified by chromatography on CM-Sepharose (Figure 2.1B). CM fractions showing this inhibition (fractions 24-30) were pooled, dialyzed and further purified by Mono-Q FPLC (Figure 2.1C). Mono-Q fractions 8-14 were pooled and dialyzed and this factor was termed putative polyadenylation factor-B (PPF-B). Unless otherwise noted, for the studies that follow, the MonoQ-purified factor was used.

PPF-B inhibits Plant PAP Activities, but Stimulates the Yeast PAP

PPF-B was first identified as an inhibitor of the pea nuclear PAP. Given the sequence similarity between the plant and other eukaryotic nPAPs, we hypothesized that PPF-B might act on other eukaryotic nPAPs in a similar manner. Accordingly, the affects

of PPF-B on the yeast PAP were determined (Figure 2.2). In contrast to what is observed with the plant nPAP, PPF-B stimulated yPAP activity approximately 4-fold. This was observed under conditions used to assay plant PAPs (as shown) or under conditions more optimal for the yeast PAP (not shown). This result indicates that PPF-B has a distinct PAP preference, perhaps suggestive of specific protein-protein interactions. In addition, these observations suggest that PPF-B does not inhibit PAP activity indiscriminately because of inherent ATPase, nuclease or protease activities.

In parallel to the studies with the yeast enzyme, the effects of PPF-B on the cpPAP were also tested. The results of this experiment are also shown in Figure 2.2. As can be seen, PPF-B proved to be as effective an inhibitor of the cpPAP as it was of the nPAP. Because the cpPAP is inherently more active (on a specific activity basis) than the pea nPAP, thus affording a more sensitive measure of inhibitory activity by PPF-B, most of the subsequent characterizations that follow were carried out using the cpPAP as a source of PAP.

The nPAP preparations we used in screening column fractions are themselves derived from nuclear extracts, and are only partially pure with respect to the nPAP. To test whether inhibition of this activity by PPF-B was due to an effect on the nPAP, as opposed to a secondary effect due to other components that copurify with the nPAP, the effects of PPF-B on recombinant nPAP (Hunt et al., 2000) were tested. As shown in Figure 2.2, PPF-B was as effective with the recombinant nPAP as with the partially purified preparation. Thus, we conclude that PPF-B acts directly on either the cpPAP or nPAP.

PPF- Inhibitory Activity is Not Due to PAP, Substrate, or Product Degradation

Among the possible means by which PPF-B inhibits PAP activity are a number of somewhat trivial ones, such as prior hydrolysis of the ATP used by PAP, degradation of the RNA substrates or products of PAP, and proteolytic breakdown of PAP itself. While the stimulatory effect of PPF-B on the yeast PAP argues against these as a means of action of PPF-B, it remained possible that PPF-B might have activities that are restricted to one or both of the plant PAPs (and thus not be manifest when assayed using the yeast enzyme). Accordingly, studies were conducted to further address these possibilities.

To test the hypothesis that inhibition of PAP activity by PPF-B is due to hydrolysis of this substrate, the ATPase activity of various combinations of PPF-B and cpPAP was directly measured. For this, we used conditions identical to those used for assay of PAP, and assayed the labeled nucleotides by thin layer chromatography. As shown in Figure 2.3, there was no difference in the profile of labeled compounds in reactions in the presence or absence of PPF-B. As expected, with the negative control with just reaction buffer, ATP was the predominant species (Figure 2.3; -PAP, -PPF-B). In reactions with PAP alone (Figure 2.3; +PAP, -PPF-B), some breakdown of ATP to ADP and AMP was observed, but considerable quantities of ATP remained at the end of the assay. The profile of labeled compounds in samples containing PAP and PPF-B (Figure 2.3; +PAP, +PPF-B) was identical to that seen with PAP alone. Thus, inhibition of PAP by PPF-B is not due to hydrolysis of ATP prior to the action of PAP. PPF-B alone had little or no detectable ATPase activity (Figure 2.3, +PPF-B, -PAP).

To test the proteolytic activity of PPF-B, different plant PAPs (nPAP and the two components of the cpPAP, respectively) were incubated with PPF-B under standard PAP reaction conditions, but without labeled nucleotide. After 30 minutes, the samples were dissolved in SDS-PAGE sample buffer and analyzed by immunoblotting, using the appropriate antisera. As is seen in Figure 2.4, inclusion of PPF-B in PAP reactions had no effect on the mobility of polypeptides recognized by antisera specific for the nPAP, the cpPAP, and PNP (the RNA binding cofactor for the cpPAP). These results indicate that inhibition of PAP activity by PPF-B is not due to proteolytic breakdown of the PAP.

Other studies indicated that PPF-B was free from detectable nuclease activity (not shown), and this factor did not inhibit the yPAP (Figure 2.2). These observations show that inhibition of PAP by PPF-B cannot be attributed to the breakdown of the RNA substrate for the PAP reaction. However, it remained a possibility that inhibition might be a manifestation of specific breakdown of the polyadenylate product of the reaction. To test this, the effects of PPF-B on "pre-formed" labeled poly(A) (the product of the PAP reaction itself) was assessed. Using the standard assay, PAP reactions were initiated and PPF-B was subsequently added at 15 minute intervals. Each reaction was incubated for a total of 120 minutes and PAP activity determined as usual. The rationale behind this experiment was that, should PPF-B inhibit PAP activity by selective degradation of the

poly(A) product of the reaction, then addition of PPF-B after allowing PAP to react for shorter periods of time would result in a degradation of the pre-formed poly(A). The results of this experiment are shown in Figure 2.5. Addition of PPF-B after short periods of time resulted in decreases in total incorporation of label into poly(A), as expected. However, these low levels of incorporation were still significantly greater than the levels seen when the inhibitor was added at the onset of the reaction. Moreover, when PPF-B was added at times more than one hour after initiation of the reaction, virtually no difference in the quantity of labeled product formed was observed when compared with reactions carried out in the absence of PPF-B. This result, along with those mentioned at the beginning of this paragraph, indicates that PPF-B does not affect the poly(A) product and is not a nuclease. Based on the experiments summarized in Figures 2.3-2.5, it suggests that the products, substrates, and proteins involved in the PAP reaction are not affected by PPF-B.

A Distinct RNA Co-purifies with PPF-B

To understand the composition of PPF-B, aliquots of the purified factor were treated with proteinase K and micrococcal nuclease. This was used to determine if PPF-B inhibitory activity is dependent on a protein or nucleotide factor or a combination of both. PPF-B was sensitive to both treatments (data not shown). Therefore, this suggests PPF-B consists of a protein and nucleotide factor. Polypeptides were not identified when PPF-B extracts were resolved by SDS-PAGE followed by Coomassie brilliant blue and silver staining (data not shown).

In mammals, the U1 snRNP is an inhibitor of the poly(A) polymerase, owing to the interaction of either the U1A or 70kD subunit of U1 with PAP (Gunderson et al., 1997; Gunderson et al., 1998). Could PPF-B be U1 or an analog of U1? To test this, RNAs isolated from PPF-B were assessed by northern blot analysis using a probe prepared from cloned pea U1 snRNA. This probe hybridized efficiently to the corresponding RNA in preparations from whole nuclei (Figure 2.6A), but failed to hybridize to any nucleic acids in the sample from PPF-B (Figure 2.6A). This suggests that PPF-B is not the pea U1 snRNP.

The results of PPF-B sensitivity to nuclease treatment and northern blot analysis, suggest that PPF-B contains an RNA component that is not from the pea U1 snRNP. To further explore this hypothesis, PPF-B-containing column fractions were extracted with phenol+chloroform and possible nucleic acids in the aqueous phase recovered by ethanol precipitation using glycogen as a bulk carrier. Nucleic acids in the resulting preparations were then end-labeled using poly(A) polymerase (commercial preparations, from yeast) and [γ -P³²]-Cordycepin-5' Triphosphate and analyzed by electrophoresis on agarose gels. For comparison, total RNA isolated from pea nuclear extracts was also end-labeled and analyzed in a similar manner. The results, shown in Figure 2.6B, indicate the presence of a distinct labeled molecule, probably an RNA (owing to the specificity of the yeast PAP) in PPF-B. This putative RNA was somewhat larger than the pea U1 RNA. This result supports the conclusion drawn from the nuclease sensitivity studies, that PPF-B contains an RNA component that is needed for full activity. Attempts to clone this RNA by reverse transcriptase (RT) and nested PCR were unsuccessful.

PPF-B Co-Purifies with a Linker Histone

Our initial preparations of PPF-B revealed no obvious polypeptides. Therefore, we scaled up our process (more plant material) to acquire visible polypeptides for analysis. Analysis of these new preparations of PPF-B indicated the presence of a single predominant polypeptide of about 30 kDa (Figure 2.7A). Further immunoblot analysis, using antibodies that recognize other possible plant polyadenylation-related proteins (PABP, nPAP, Fip1, and plant homologues of the 100 kDa subunit of CPSF and 50 kDa subunit of CstF) indicated that none of these proteins were present in the purest preparations of PPF-B (Figure 2.7B). Thus, PPF-B does not contain polypeptides related to PAP, nor does it co-purify with four other possible polyadenylation factors.

The predominant polypeptide in PPF-B, of approximately 30 kDa, was isolated and excised for proteolytic digestion with endo-LysC and NH₂-terminal sequencing to acquire peptide sequences (see Materials and Methods). Two peptide sequences were obtained from this polypeptide and this information was used to search available databases. The database search yielded an exact match with a protein from garden pea (*Pisum sativum*), Ribosome-Sedimenting Protein (RSP; Figure 2.8A). RSP displays

significant homology to a number of histone H1 proteins in various databases. The most prominent among these were several histone H1s from a variety of plant species (Figure 2.8B). For example, the amino acid identity of RSP and the tomato, tobacco, pea, and Arabidopsis histone H1 was 62%, 52%, 48%, and 43% respectively over the entire length of the RSP protein. Interestingly, both of the peptide sequences that were obtained were derived from RSP, as opposed to other linker histones (such as the pea histone H1 whose sequence is shown in Figure 2.8B). This suggests that PPF-B may consist solely of a particular linker histone.

Several attempts were made to isolate and characterize GST-RSP fusion proteins from *E. coli*. However, the fusion protein produced was very susceptible to bacterial proteolysis, thereby precluding a direct assay of the fusion protein (data not shown). Thus, as an alternative, co-purification experiments using DNA-cellulose were performed. PPF-B extracts were first incubated with DNA-cellulose and then resolved by SDS-PAGE. As seen in Figure 2.8C, RSP was removed from the extract and was bound to the DNA-cellulose (compare lanes 2 and 3). This suggests that RSP has affinity for DNA, which is not surprising, given that it is a linker histone. The supernatant collected from the DNA-cellulose incubation was tested for inhibitory activity of PAP. While the preparation before DNA-cellulose treatment was able to inhibit 97% of the PAP activity, the supernatant after treatment only inhibited 8% of the PAP activity (Figure 2.8C). Thus, like RSP itself, the inhibitory activity of PAP by PPF-B bound to DNA-cellulose. This suggests, at the very least, that inhibitory activity of PAP and RSP might be closely associated.

It has been shown that human linker histones affect the ability of the SWI/SNF ATP-dependent remodeling complex to remodel chromatin *in vitro* (Hill and Imbalzano, 2000; Ramachandran et al., 2003). Human forms of the SWF/SNF complex have been shown to be inactivated by phosphorylation in mitosis (Muchardt et al., 1996; Sif et al., 1998). With that in mind, we set out to determine if PPF-B activity could be regulated by phosphorylation activities. PPF-B was treated with calf intestinal alkaline phosphatase (CIAP; Promega) and then tested with PAP. However, buffer control (buffer treated with CIAP) affected PAP activity and thus phosphorylation regulation of PPF-B activity could not be determined (data not shown).

RSP Localizes to the Nucleus and Cytoplasm

RSP was identified as a novel component of ribosome-sedimenting activity and the plant cytoskeleton (Abe and Davies, 1995; Davies et al., 2001). RSP is an atypical linker histone and its sub-cellular localization was previously unknown. A novel histone H1 variant has been shown to be localized in nucleoli of higher plant cells (Tanaka et al., 1999). Therefore, we set out to determine the localization of RSP in tobacco leaves by generating autofluorescent protein fusions. Co-localization of RSP with an *Arabidopsis* nuclear PAP (AtPAPIV; At4g32850) was also tested.

In order to evaluate localization, plasmids (pGD) carrying autofluorescent protein fusions were introduced into tobacco leaves by particle bombardment (see Materials and Methods). The pGD binary vector enables high level of transient expression in plant cells (Goodin et al., 2002). RSP was fused to a red fluorescent protein (DsRed; pGDR vector) and AtPAPIV was fused to a green fluorescent protein (GFP; pGDG vector). Observations with the green/red filters to reveal the fluorescence patterns of GFP (Figure 2.9A) and DsRed (Figure 2.9B) indicated that both proteins accumulate throughout the epidermal cell (Figure 2.9D). The GFP-PAPIV fusion was co-expressed with DsRed in the same epidermal cell (Figure 2.9 E-H). It suggests that GFP-PAPIV (Figure 2.9 E) localizes to the nucleus when compared to GFP alone (Figure 2.9A). When images E-G were the overlaid (Figure 2.9H), it also suggested that GFP-PAPIV was nuclear localized, because DsRed saturates the nucleus. The DsRed-RSP fusion was co-expressed with the GFP-PAPIV fusion in the same epidermal cell (Figure 2.9 I-L). DsRed-RSP appears to be targeted to the periphery of the epidermal cell and potentially the nucleus (Figure 2.9 J), as evidenced by the red speckles seen where GFP-PAPIV localizes (comparing Figure 2.9 J to I). Fluorescence of GFP and DsRed proteins and fusion proteins could easily be detected in epidermal cells without chlorophyll interference as shown in panels C, G, and K (Figure 2.9). DAPI staining of the nuclei in these tissues was unsuccessful, thus definitive results indicating localization to the nucleus in these experiments is premature. However, epidermal cells lack chloroplasts and GFP-PAPIV localizes to a single organelle in the cell, presumably the nucleus. For these reasons, we

tentatively conclude that PAP-IV is nuclear localized. Also, we tentatively conclude that RSP may co-localize with PAP-IV and be nuclear localized.

Highly Pure PPF-B Contains an AMP Forming Activity

As described above, our purification process of PPF-B was scaled up and a polypeptide target linked to inhibitory activity was identified. With this new preparation, assays for substrate (ATP) breakdown were performed. As seen in Figure 2.10, concentrated PPF-B co-purifies with an activity that converts ATP to AMP. This activity was observed in the presence as well as in the absence of PAP. Given that the inhibitory activity of PAP by PPF-B was bound by DNA-cellulose (Figure 2.8C), we wanted to determine if inhibitory and AMP-forming activities could be eluted from the DNA-cellulose. PPF-B extracts were first incubated with DNA-cellulose and then washed with buffer containing different concentrations of KCl. As seen in Figure 2.11A, inhibitory activity eluted from the DNA-cellulose in buffer containing 500 mM KCl. Eluted proteins from each wash were also analyzed. Different amounts of RSP eluted in each wash and it appears that more RSP eluted in 500 mM KCl than any other wash. This suggests that the inhibitory activity and RSP are closely associated. However, supernatant from PPF-B treated with DNA-cellulose (Figure 2.11A; lane 2) appears to have equal amounts of RSP and one-tenth the inhibitory activity of the 500 mM wash (Figure 2.11A; lane 6). This suggests that RSP is not the inhibitory factor found in PPF-B.

These above mentioned washes were also analyzed for substrate (ATP) breakdown. As seen in Figure 2.11B, AMP-forming activity eluted from the DNA-cellulose. PAP assays were performed with washes in the presence or absence of PAP. AMP-forming activity eluted in the 500 mM wash, which contains the PPF-B inhibitory activity. This suggests that inhibitory and AMP-forming activities are closely associated. Thus, characterization of this AMP-forming activity may lead to the identification of PPF-B.

DISCUSSION

We have characterized a plant polyadenylation factor (termed PPF-B) that inhibits the non-specific activity of PAPs from the nucleus and chloroplast. This factor can be resolved from other putative plant polyadenylation factors (PAP, putative subunits of CstF and CPSF, Fip1 and a nuclear poly[A] binding protein) for which probes or assays are available, and also from at least one snRNP (U1) that has been shown to inhibit PAP in mammalian systems. The inhibition of the plant PAPs is not due to depletion of PAP polypeptides, substrates, or products, as shown by direct assay and by the lack of effect on yPAP. Inhibitory activity is associated with a novel linker histone (Davies et al., 2001). RSP appears to be localized to the outer membrane and potentially the nucleus of the cell. Highly pure PPF-B also has a co-purifying AMP forming activity that is distinct from its inhibitory activity. Taken together, these properties indicate that PPF-B is a novel factor, unlike other polyadenylation-related factors that have been described to date.

In some ways, the activity of PPF-B is similar to that of the Fip1p subunit of the yeast polyadenylation factor CPF. By itself, freed from the other subunits of CPF, Fip1p inhibits the non-specific activity of the yeast PAP by blocking access of the RNA primer to an RNA-binding domain of PAP (Zhelkovsky et al., 1998). In contrast, intact CPF (which contains Fip1p) stimulates the non-specific activity of the yeast PAP (Preker et al., 1995; Helmling et al., 2001). The polypeptides recognized in pea nuclear extracts by anti-sera developed against an *Arabidopsis* Fip1, were absent in PPF-B extracts. This suggests that PPF-B does not contain Fip1. The lack of other immunologically related subunits of the polyadenylation complex (PAP, PAB, CstF50, CPSF30) in PPF-B suggests that PPF-B does not contain the cleavage and polyadenylation complex and inhibition is linked to another cellular process.

Kleiman and Manley (1999) have reported that BARD1 interacts with the polyadenylation factor CstF50 and inhibits 3' end processing of mRNA precursors. BARD1 is associated with BRCA1 and several lines of evidence support the role of BRCA1 in maintaining genomic stability in response to DNA damage. BRCA1 has also been shown to interact with several components of the transcription complex. The

BARD1-CstF50 interaction which inhibits 3' end formation has been linked to DNA damage and tumor suppression (Kleiman and Manley, 2001). These authors have proposed that the inhibition of polyadenylation by BARD1-BRCA1 is a means of preventing the polyadenylation (and subsequent transport, presumably) of aberrantly-terminated RNAs associated with RNA polymerase situated (paused) at sites of DNA damage. There are no apparent homologues in the *Arabidopsis* database for BARD1 or BRCA1 (data not shown), which raises the question as to the existence in plants of an analogous mechanism for limiting the polyadenylation of aberrant RNAs. The inhibition of PAP by PPF-B may be analogous to the BARD1-Cstf50 association in the prevention of deleterious polyadenylation of mRNAs transcribed from damaged DNA in plants. Linker histones can be modified in response to DNA damage (e.g., de Murcia et al., 1988), thus providing precedent for a link between linker histones and DNA damage. Therefore, RSP may be the link of repair complexes in response to DNA damage and directly or indirectly, serve to limit or prevent polyadenylation of aberrant RNAs.

Among the subcellular locations that linker histones may be found is the nucleolus; for example, Tanaka et al. (1999) have reported that a distinctive plant linker histone localizes to the nucleolus. This raises the possibility that PPF-B may as well be associated with the nucleolus. This in turn suggests that PPF-B may act to inhibit nPAP in the nucleolus. This inhibition may serve to prevent inadvertent polyadenylation of rRNA precursors (an eventuality that one might expect to lead to breakdown of the rRNA, given the parallels between exosomes and degradosomes). Inherent in this model is the suggestion that the nPAP may play a role in nuclear RNA turnover in plants, much as the cpPAP and bacterial PAP do in their respective compartments. It is thus of interest to note that recent studies have shown that the yeast PAP interacts with Rrp6p, an exosome subunit (Burkard and Butler, 2000). Moreover, polyadenylated forms of various small nuclear and nucleolar RNAs accumulate in yeast exosome mutants (van Hoof et al., 2000), much as do polyadenylated RNAs in nuclease mutants of *E. coli* (Sarkar, 1997). Thus, an alternative to the model that PPF-B links DNA damage with polyadenylation holds that PPF-B may facilitate RNA ribosomal RNA maturation (or that of other stable RNAs that are processed by the exosome) by limiting or preventing their polyadenylation.

The inhibition of the cpPAP by PPF-B, and the association of PPF-B with a linker histone, suggests an alternative model for the role of PPF-B in RNA metabolism in the plant nucleus. The chloroplast PAP requires PNP as an RNA binding cofactor *in vitro* (Li et al., 1996). We have found that partially purified nPAP is also able to utilize PNP (a chloroplast enzyme) as an RNA-binding cofactor (Hunt et al., 2000). This commonality suggests possible sites of action for PPF-B, these being steps or reactions that involve PNP. In bacteria, PNP is a component of the so-called degradosome, a complex responsible for RNA turnover (Carpousis et al., 1999). Eukaryotic cells possess analogous complexes (termed by others as exosomes) that contain subunits that are related to subunits of the bacterial degradosome (Decker, 1998; van Hoof and Parker, 1999). Among the functions that exosomes serve in eukaryotic cells are processes involved in ribosomal RNA, snoRNA, and snRNA maturation (Allmang et al., 1999).

Recombinant *Arabidopsis* PAP and purified chloroplast PAP can polyadenylate RNAs associated with a purified exosome-like activity from pea nuclear extracts (Dattaroy and Hunt, 2002). The ability of PAP to use RNAs associated with a degradative process as template RNA further strengthens the hypothesis of polyadenylation linked to turnover. In this context, PPF-B may be a bridge between polyadenylation and RNA turnover (or ribosomal RNA maturation, or both). This factor may perhaps serve to partition the nPAP between biosynthetic and degradative functions.

The linker histone that is associated with PPF-B has previously been identified as a protein that co-purifies with cytoskeleton-associated polysomes (Davies et al., 2001). The significance of this is unclear. This would imply a cytoplasmic role for PPF-B in plant cells, perhaps one related to interactions with cytoplasmic exosomes. Alternatively, cytoskeleton-associated RSP may not be capable of inhibiting PAPs. However, PAP functions in translational control of certain mRNAs by cytoplasmic polyadenylation. Control of PAP level is essential for cytoplasmic polyadenylation and early development in *Drosophila* (Juge et al., 2002). This may be the link RSP has in the cytoplasm, by regulation of cytoplasmic PAP involved in increasing translation efficiency.

BRCA1/ BARD1 can assemble a ubiquitin ligase (E3) activity and E1-mediated monoubiquitination *in vitro* (Chen et al., 2002). Poly-ubiquitin chains are believed to play a role in signaling for DNA repair (Spence et al., 1995; Hofmann and Pickart, 1999).

Thus suggesting a linkage of BRCA1/BARD1-ubiquitin in tumor suppression. The reaction of ubiquitination requires the sequential actions of three enzymes, an activating enzyme (E1), a conjugating enzyme (E2) and a ligase (E3). E1-ubiquitin activating enzymes require ATP for activation. Ubiquitin is linked to E1 through AMP. Thus, E1s generate AMP during the activation reaction. Concentrated PPF-B contains an AMP forming activity and the inhibitory activity is linked to RSP. This raises the possibility that PPF-B contains an E1-activating enzyme, drawing a link to chromatin surveillance and polyubiquitination as mentioned above. The yeast PAP interacts with Uba2, an E1 homolog, and Ufd1, a protein whose function is linked to the ubiquitin-mediated protein degradation pathway (Olmo et al., 1997). It is possible that the inhibitory activity of PPF-B is linked to an E1 enzyme by way of a linker histone (RSP) to ubiquitinate its target, PAP. The hypothesis of linking RSP to DNA damage complexes adds credence to this proposed function, by preventing polyadenylation through regulating subunits of the cleavage and polyadenylation complex by ubiquitination. *In vitro* E1-ubiquitin activating assays were performed with PPF-B and commercially available ubiquitin and rabbit E1 as a control (data not shown). In these experiments no activation was observed. However, the expected activity may require plant ubiquitin or ubiquitin like protein. Immunoblots with plant specific ubiquitin and E1-3 antibodies and E1-activating assays with recombinant plant ubiquitins would aide in determining if E1 is present in PPF-B.

Further evidence for the potential of post-translational modification of PAP is provided by the identification of potential sites of modification in its amino acid sequence. As seen in Figure 2.12, all four of the recently reported *Arabidopsis* PAPs (Addepalli et al., 2004), contain a potential binding site for sumolation. The consensus binding motif for SUMO is \square KXE, where \square is a hydrophobic amino acid, K is lysine, X is any amino acid, and E is glutamic acid (Melchior, 2000). SUMO, small ubiquitin-like modifier, is a post-translational modifier related to ubiquitin. SUMO attachment employs an E1-E2-E3 conjugation similar to ubiquitination (see above). SUMO has not been shown to be a targeting modifier for proteolysis like ubiquitin. SUMO has been shown to function in cellular localization, activation, and protection from ubiquitin-mediated proteolysis (Muller et al., 2001). This suggests that *Arabidopsis* PAPs may be modified by SUMO. Since SUMO requires an E1-activating enzyme for attachment, it may be that

PPF-B contains an E1-SUMO activating enzyme. This is consistent with the observation that PPF-B lacks proteolytic activity against PAPs.

To summarize, we have identified a novel factor (PPF-B) that is unlike other polyadenylation-related factors that have been described to date. PPF-B inhibits poly(A) polymerase activity and it does not contain known plant poly(A) factors (including Fip1). PPF-B has allowed us to hypothesize conceptual links of polyadenylation to several processes including DNA damage, RNA turnover and protein regulation by post-translational modifiers. Links with these processes have been reported already in other eukaryotic systems. Further experiments will help us understand more about PPF-B function and how other cellular processes are linked and affected by polyadenylation.

MATERIALS AND METHODS

Plant Material

Garden pea (*Pisum sativum* cv. Laxton Progress) seed was obtained from Kentucky Garden Supply (Lexington, KY). Seeds were sown in the greenhouse and harvested after 2-3 weeks of growth, when plants had two to three extended open trifolates. Tobacco plants (*Nicotiana tabacum*) used for particle bombardment were supplied by Dr. Randy Dinkins (PGEL; University of Kentucky, Lexington, KY).

Protein Purification

Pea nuclei and nuclear extracts were prepared as described by Yang and Hunt (Yang and Hunt, 1994). Nuclear extracts were examined by immunoblot analysis to verify that they were free of chloroplast contaminants; these characterizations have been described in detail elsewhere (Das Gupta et al., 1998; Li et al., 1998). Nuclear extracts were diluted 2-fold with NEB (40 mM KCl, 25 mM HEPES-KOH, pH 7.9, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF, 10% glycerol and 5 mg/ml of each of leupeptin, chymostatin, and antipain) prior to DEAE chromatography. Protein contents were determined using the BioRad dye reagent and bovine serum albumin as a standard. Ten mg of nuclear protein was loaded onto a 20 ml DEAE-Sepharose column (Sigma) that had been equilibrated with NEB. The column was washed with 40 ml of NEB at a rate of 0.5ml/min and proteins retained in the column were eluted with a 80 ml gradient (40-500mM) of KCl in NEB. The column was then washed with 40 ml of NEB containing 500 mM KCl and 40 ml of 1M KCl in NEB. Five ml fractions were collected and 20 μ l of each fraction were assayed for the inhibition of non-specific PAP activity.

Fractions containing PPF-B activity were pooled, dialyzed against NEB, and loaded onto a 15 ml CM-Sepharose (Sigma) column at a flow rate of 0.5ml/min. The column was washed with 40 ml NEB and bound proteins eluted with a 60 ml gradient of 40-790 mM KCl in NEB. The column was then washed with 40 ml of NEB containing 1M KCl. One ml fractions were collected after the first wash and 20 μ l assayed for inhibitory activity. Active fractions were pooled, dialyzed against NEB, and loaded at 0.5ml/min onto a MonoQ HR5/5 column (Pharmacia) that had been equilibrated with

NEB. The column was developed with a linear gradient (10 ml) of 40 mM to 500 mM KCl in NEB. Fractions of 0.5 ml were collected, and 20 μ l of each fraction were assayed for PPF-B activity. Active fractions were pooled and used for the characterization. Typical yields were about 2 ml of purified factor containing 10-20 mg/ml of protein.

The polypeptide composition of PPF-B was analyzed by separating peak fractions by SDS-PAGE (12.5% gels), transferring to PVDF membranes, and staining the membranes with Coomassie Brilliant Blue.

The nuclear PAP was purified from nuclear extracts as described by Hunt et al. (2000). Likewise, recombinant nPAP was isolated from *E. coli* carrying a pGEX2T-nPAP construct (Hunt et al., 2000). The cpPAP was purified from extracts prepared from the leaves of young pea seedlings as described previously (Das Gupta et al., 1995). In this study, the enzyme purified through the heparin-Sepharose step was used; this preparation contains both subunits of the active chloroplast poly(A) polymerase complex (the PAP and polynucleotide phosphorylase; (Li et al., 1996, 1998). The yeast PAP was purchased from US Biochemicals; the activity of this enzyme in the buffers used for the assay of the chloroplast and plant nuclear PAPs was confirmed prior to the studies described here.

For purification of PPF-B with DNA-cellulose, 50 mg of DNA-cellulose was washed 3 times with 1 ml of NEB, and resuspended in 300 μ l of the same buffer. 100 μ l this suspension was pelleted, the supernatant removed, and 70 μ l of Mono-Q pure PPF-B was added to the pellet. The pellet was suspended thoroughly, and the mixture incubated for 30 minutes at 25°C. The mixture was then centrifuged for 10 seconds (10,000 rpm) and the supernatant collected and used for assays as described above or resolved by SDS-PAGE along with the DNA-cellulose pellet in 2X SDS-PAGE sample buffer. The gel was stained with Coomassie Brilliant Blue (Wong et al., 2000), and then with silver (using a kit supplied by Bio-Rad, as recommended by this manufacturer). Stained polypeptides were then visualized on a light box. For elution of proteins, DNA-cellulose was incubated with NEB containing 125 mM KCl at room temperature for 30 minutes. The DNA-cellulose was pelleted and supernatant collected for analysis. The DNA-cellulose pellet was then incubated again with NEB containing 250 mM KCl and supernatant collected and analyzed. This was repeated three more times with NEB containing 500 mM, 1M, and 2 M KCl and supernatants collected and analyzed.

Immunoblot Analyses

For immunoblots, proteins were separated by SDS-PAGE and transferred to a nitrocellulose membrane using a Trans-Blot Cell (Bio-Rad Laboratories) following the manufactures recommendations. Filters were then washed and probed with antibodies (at dilutions of 1/500 to 1/1000) as described elsewhere (Yang and Hunt, 1994). The antibody preparations used for detection of the cpPAP, polynucleotide phosphorylase, CPSF100, and poly(A) binding protein have been described (Yang and Hunt, 1994; Das Gupta et al., 1998; Li et al., 1998; Elliott et al., 2003). Antibodies that recognize the plant nuclear PAP and *Arabidopsis* homologues of the 100 kDa subunit of CPSF and 50 kDa subunit of CstF were prepared using glutathione-S-transferase fusion proteins as immunogens. For antibodies against the plant nPAP, the amino acid coding sequence of a rice EST whose predicted product is significantly related to the bovine and yeast PAPs (accession C26207) was amplified by PCR and cloned onto pGEX-2T (Pharmacia). To raise antibodies against the CstF50 homologue, the coding region corresponding to the C-terminal 228 amino acids of an *Arabidopsis* gene (accession BAB10643) whose product is significantly similar to the mammalian protein was amplified by RT/PCR and cloned onto pGEX-2T (Pharmacia). *E. coli* carrying the appropriate recombinant plasmid were grown, induced with IPTG, and protein was purified by affinity chromatography using glutathione-Sepharose 4B, as recommended by the manufacturer (Pharmacia). The purified fusion proteins were used without further purification. These preparations were used to inject rabbits as described by Yang and Hunt (Yang and Hunt, 1994). The production of antibodies against AtFip1 is described in chapter 3.

Enzyme Assays

The assay for non-specific plant PAP activity has been described elsewhere (Das Gupta et al., 1995; Li et al., 1996). Unless noted otherwise all reactions were incubated at 30°C for 2 hours. Bovine serum albumin and poly (A) were omitted from the assays for protease characterization of PPF-B.

For ATPase assays, PAP reactions were carried out and 1 μ l aliquots of each reaction were applied to polyethyleneimine-cellulose thin layer chromatography plates (Fischer Scientific). The plates were developed in 0.8 M LiCl and visualized by autoradiography. One μ l each of 20 mM AMP, ADP, and ATP was applied as a standard (visualized by irradiating plates directly with ultraviolet light).

To estimate possible proteinase activity, proteins from the PAP reactions were solubilized by the addition of an equal volume of 2X SDS-PAGE sample buffer (1X = 0.1 M Tris HCl, pH 6.8, 2.9 M β -mercaptoethanol, 4% SDS, 0.2% bromophenol blue, 20% glycerol), separated by SDS-PAGE, and transferred to a nitrocellulose membrane using a Trans-BlotCell (Bio-Rad) following the recommendations of the manufacturer. Filters were then washed and probed with antibodies (at dilutions of 1/500 to 1/1000) as described elsewhere (Sambrook et al., 1989; Yang and Hunt, 1994).

Northern Blot Analysis of RNA Isolated from PPF-B and Crude Nuclear Extracts

Total RNA from whole nuclei and MonoQ purified PPF-B fractions were extracted with phenol + chloroform and recovered by ethanol precipitation using 20 μ g of glycogen as a bulk carrier. RNA samples were separated on a 2% agarose gel cast in formaldehyde (Sambrook et al., 1989) and transferred to a nylon membrane. These filters were hybridized at 40° C for 16 hours with a probe prepared using a U1 cDNA (see the following paragraph), washed three times in 2X SSC + 0.1% SDS at 40° C, and visualized by autoradiography.

To prepare the U1 snRNA cDNA, 0.5 mg of total RNA from pea nuclear extracts was used in a one-step RT-PCR reaction with specific oligonucleotides for U1 snRNA (5':ATACTTACCTGGATGGGGTCGAT and 3':GGGCCGCGCGAACGCAGGCCCC TCAGGCAA) (Hanley and Schuler, 1991). The RT-PCR product of 155 bp was used to generate a radiolabeled probe (Prime-It II/Stratagene) for Northern Analysis (see below).

3' End Labeling and Northern Analysis of RNA Extracts from PPF-B

Total RNA from whole nuclei and MonoQ purified PPF-B fractions were extracted with phenol + chloroform and recovered by ethanol precipitation using 20 μ g of glycogen as a bulk carrier. These RNA preps were end-labeled with purified yeast PAP

(US Biochemicals), briefly described in Das Gupta et al. (1995), with the exception of using 10 μ Ci of [α -P³²]-Cordycepin 5'-Triphosphate per reaction (Dupont-NEN). These labeled RNAs were separated on a 2% formaldehyde/agarose gel, transferred to nylon membrane and visualized by autoradiography

Peptide Sequencing of PPF-B

PPF-B extract, purified through Mono-Q, was separated by SDS-PAGE. Proteins were transferred to Immobilon-P membranes (Millipore) by electroblotting and the resolved polypeptides visualized by staining briefly with Coomassie Brilliant Blue. The 30 kDa polypeptide from PPF-B was excised and submitted to the University of Kentucky Macromolecular Structure Analysis Facility for proteolytic digestion with endo-LysC and NH₂-terminal peptide sequencing. Picomole yields of sequence were recovered per input protein. The resulting sequences were used to search various databases.

GFP and DsRed Fusion Constructs and Particle Bombardment

For transient GFP and DsRed assays, the pGDG (GFP) and pGDR (DsRed) fusion plasmids were used (Goodin et al., 2002). Briefly, RSP was PCR amplified with oligonucleotide primers (CCAAGATCTATGGCGTCTGAAGAGCCCACCACCGT and CCAAGATCTTCACTTCTTCTCTTAGCTGCTG; in the 5' to 3' direction) using a full-length cDNA as template (a gift from Dr. Shunnosuke Abe; Ehime University, Matsuyama, Japan). PCR products of predicted size (891 bp; all coding region) were subcloned into the EcoRV site of pGEM-T Easy (Promega), per manufacturers instructions, and resulting clones were sequenced with T7 and SP6 oligonucleotide primers. pGEM-RSP clones were then digested with BglII and the resulting fragment (903 bp) was ligated into BamHI digested pGDR plasmid by T4 DNA ligase (Invitrogen). Recombinants were sequenced with oligonucleotide primer (in the 5' to 3' direction) TTTCTTGGCCTTGGGCTCCTTGGG to determine if the DsRed fusion is in frame. AtPAPIV was PCR amplified with oligonucleotide primers (CCGAGATCTATGGTGGG TACTCAAATTTAGGTGGT and TGAAGTGAAGTTCGCGGG; in the 5' to 3' direction) using the pGBD-IV clone (see Materials and Methods in chapter 3) as

template. PCR products of expected size (2403 bp) were subcloned into pGEM-T and sequenced as described above. The pGEM-AtPAPIV clone was then digested with BglII and SalI. The resulting fragment (approximately 2400 bp) was then ligated into BglII/SalI digested pGDG by T4 DNA ligase and digested again with HindIII. The resulting clone was then sequenced with oligonucleotide primer (in the 5' to 3' direction) CTCAGCCAA TATATCATGCAAGATGATGAA to determine if the GFP fusion is in frame.

These resulting fusion plasmids were introduced into tobacco abaxial leaf tissues by particle bombardment using a PDS1000 DuPont Bio-Rad Microprojectile delivery system (Bio-Rad Laboratories) and leaf pieces were analyzed by microscopy 24 hours after bombardment.

Laser Scanning Confocal Microscopy

Confocal microscopy was done using an Olympus BX61WI confocal laser scanning microscope equipped with helium/neon lasers and multi-tracking. GFP was excited at 488 nm, and the resulting fluorescence filtered through a primary dichroic (UV/488/543/633), 570 nm secondary dichroic, and BP505-550 nm emission filters to the photomultiplier tube (PMT) detector. DsRed was excited at 543 nm, and the emission was passed through the same primary and secondary dichroic mirrors, and through a LP570 nm emission filter to the PMT detector. Images were captured using Olympus software and converted to TIFF files. At least three transformed cells were examined for each gene of interest.

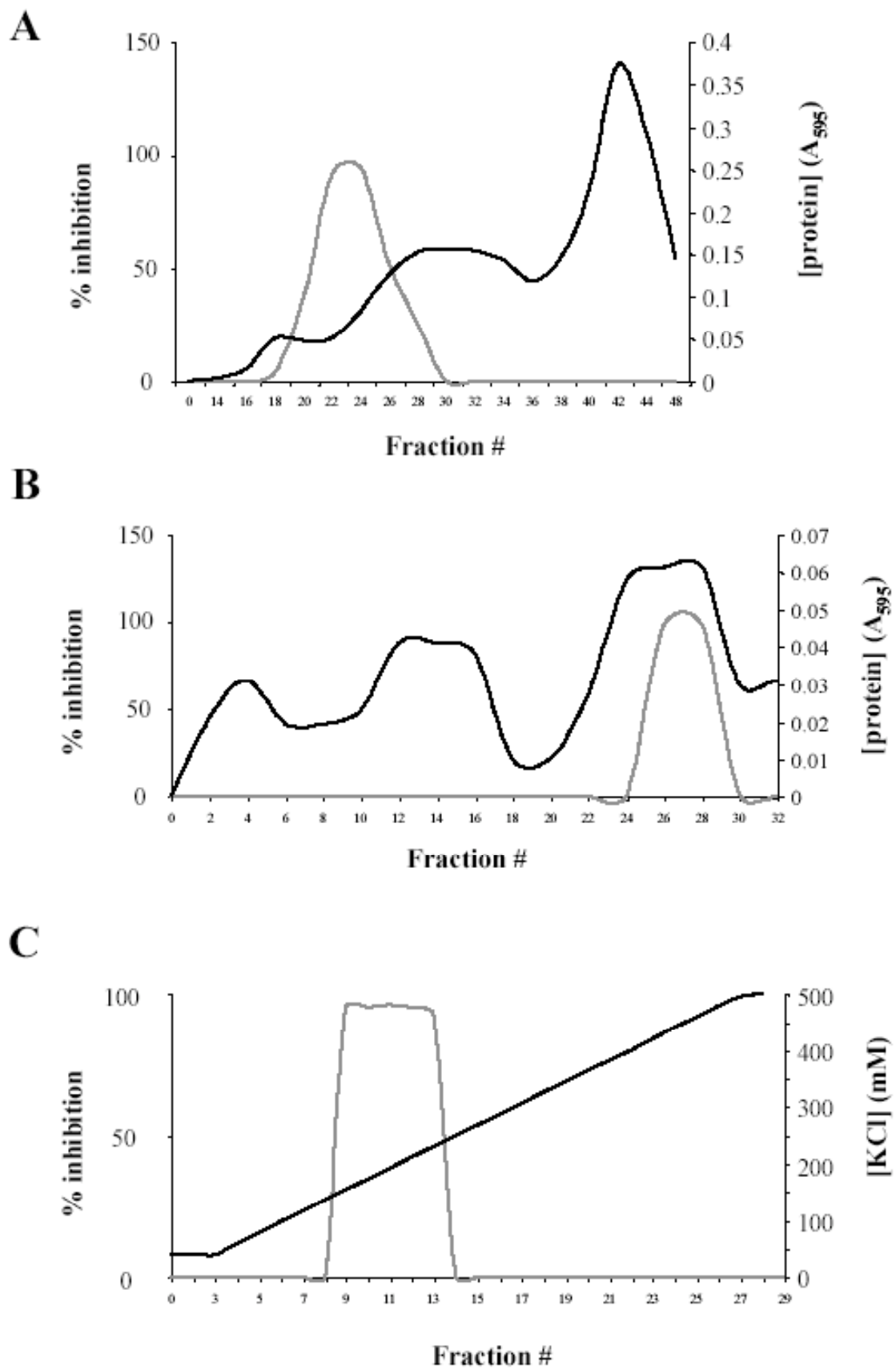


Figure 2.1. Purification and Activity of PPF-B.

Summary of a typical PPF-B purification: Pea nuclear extracts were dialyzed and then fractionated on DEAE-Sepharose(A). DEAE fractions (19-29) containing PPF-B were then pooled and dialyzed and further purified by CM-Sepharose(B). CM fractions containing PPF-B (24-30) were pooled and dialyzed and further purified by Mono-Q FPLC (C). (See Materials and Methods)

- (A) Twenty μ l of each DEAE-Sepharose fraction were screened for PAP-inhibitory activity in polyadenylation reactions. Each reaction contained 20 μ l (ca. 800 ng of total protein) of partially-purified nPAP. Inhibitory activity (gray line) and total protein (black line) were plotted as shown.
- (B) Twenty μ l of each CM-Sepharose fraction were screened for PAP-inhibitory activity in polyadenylation reactions. Each reaction contained 20 μ l of partially-purified nPAP. Inhibitory activity (gray line) and total protein (black line) were plotted as shown.
- (C) Twenty μ l of Mono-Q fractions were screened for PAP-inhibitory activity in polyadenylation reactions. Each reaction contained 20 μ l of partially-purified nPAP. Inhibitory activity (gray line) and [KCl] (black line) were plotted as shown (Total protein concentrations were too low to be plotted). In A-C, the percent inhibition represents the activity of the fractions compared to the non-specific activity of nPAP in the absence of added factors.

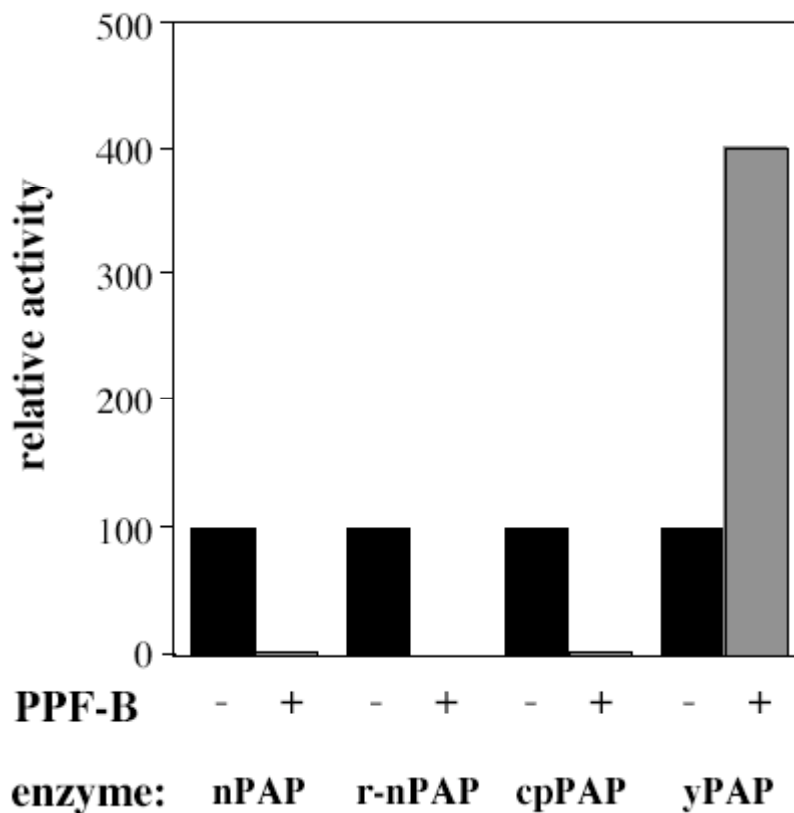


Figure 2.2. Effects of PPF-B on Different Poly(A) Polymerases

PAP reactions were performed with the chloroplast PAP (cpPAP; 3.6 μ g), the pea nPAP (400 ng), recombinant *Arabidopsis* nPAP (220 ng), and the yPAP (1 μ l of a commercial preparation) as described in Materials and Methods. Dark bars represent the relative activity of each enzyme cpPAP (372U/mg protein), pea nPAP (129U/mg protein), recombinant *Arabidopsis* nPAP (1500U/mg protein), and yPAP (7.32U/ 1 μ l of a commercial preparation) in the absence of PPF-B; these activities have been set at 100% to simplify presentation. Light bars denote the relative activities of each enzyme in the presence of 5 μ l (ca. 40 ng) of MonoQ-purified PPF-B.

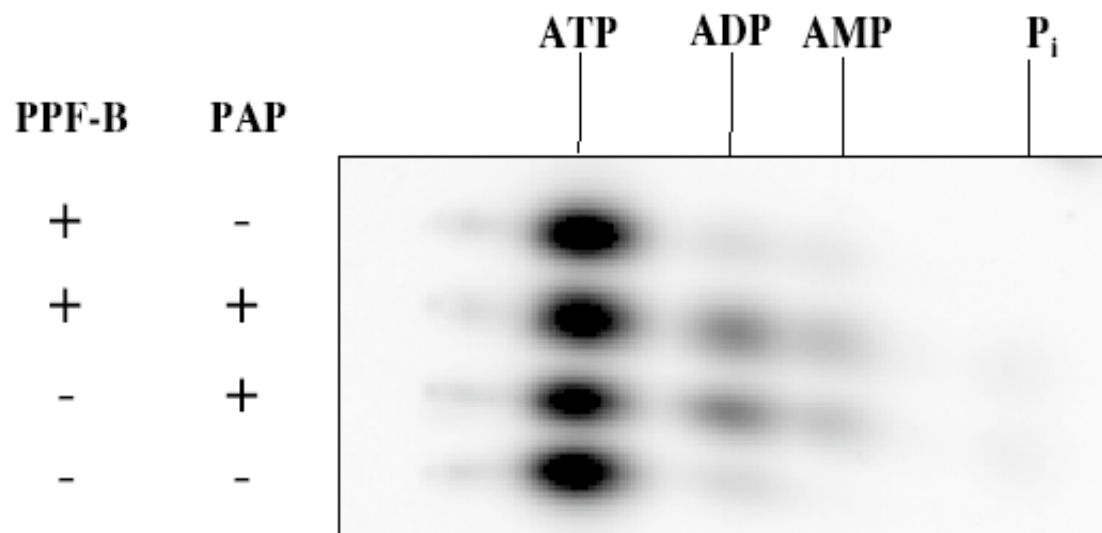


Figure 2.3. PPF-B does not Promote the Hydrolysis of ATP.

PAP reactions containing the indicated enzymes were conducted as described in Methods. After stopping each reaction, 1 μ l of the aqueous phase of each reaction was analyzed by TLC. The locations of ATP, ADP, AMP, and inorganic phosphate (Pi) are indicated on top. For these reactions, 3 μ g of cpPAP and 40 ng of PPF-B were used, as indicated.

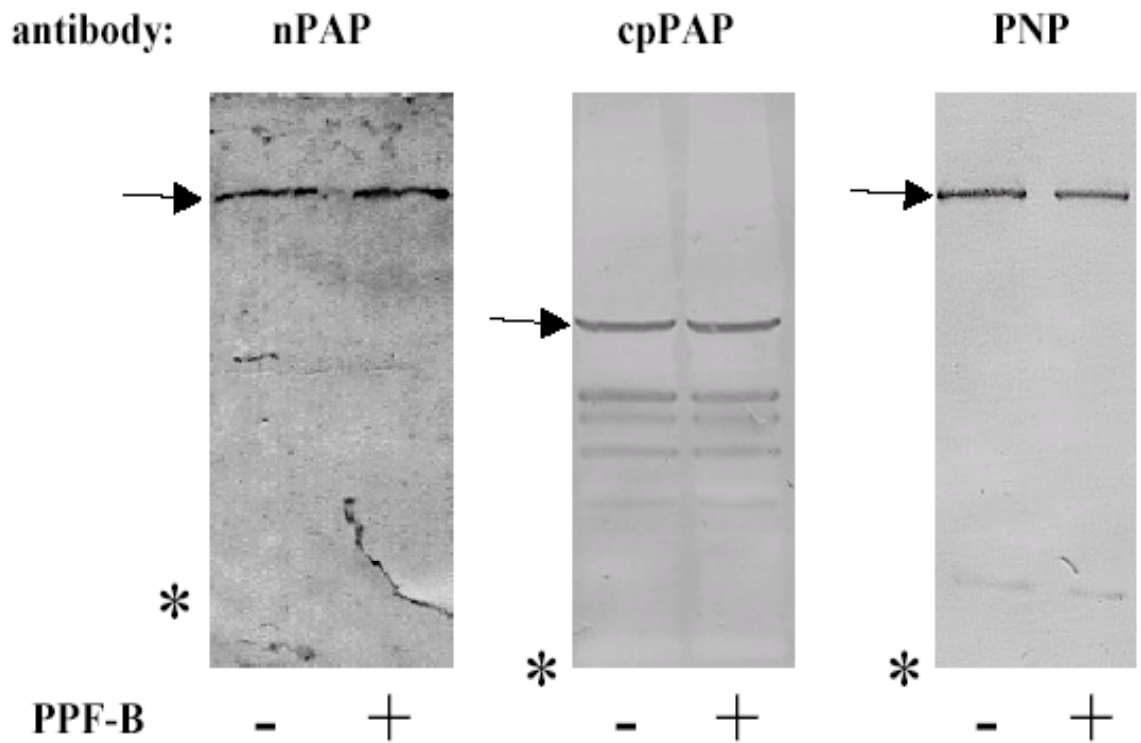


Figure 2.4. PPF-B does not act Proteolytically on Plant PAPs

nPAP (800 ng), cpPAP (30 μ g) or polynucleotide phosphorylase (PNP; 6 μ g) were incubated in the presence or absence (as indicated) of PPF-B (160 ng), incubated for 30 min in PAP reaction conditions, and analyzed by immunoblotting, using the antibody indicated above the blot. The mobilities of "authentic" PNP, nPAP, and cpPAP, respectively, are indicated with arrows to the left of each blot. The mobilities of a 28 kDa size standard are denoted with a star near the bottom of each blot.

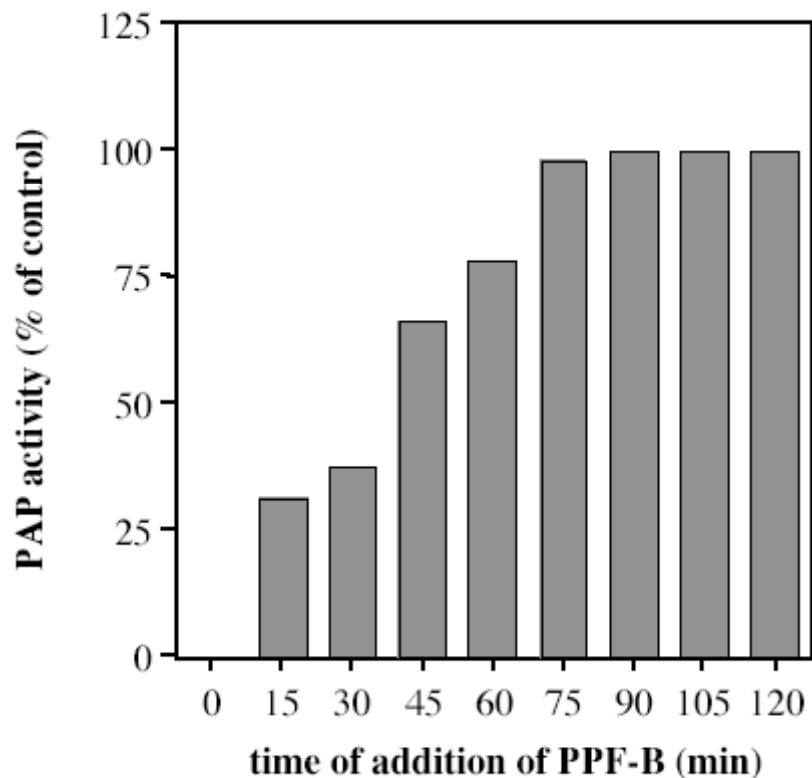
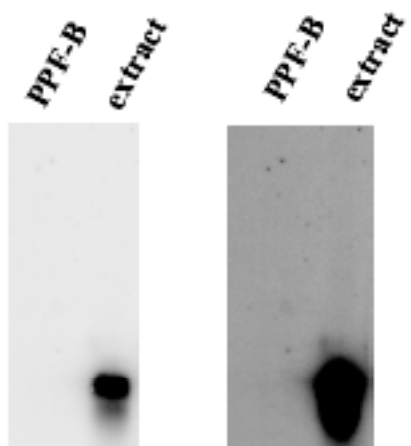


Figure 2.5. PPF-B does not Degrade Pre-formed Products of the PAP Reaction.

PAP reactions were conducted using 7.2 μ g of the cpPAP as described in Methods. At the times indicated, 160 ng of PPF-B was added to each reaction. All reactions were incubated for a total of 120 min and the label incorporated into poly(A) determined. The quantity for each sample was plotted as a percent of activity in the absence of PPF-B.

A



B

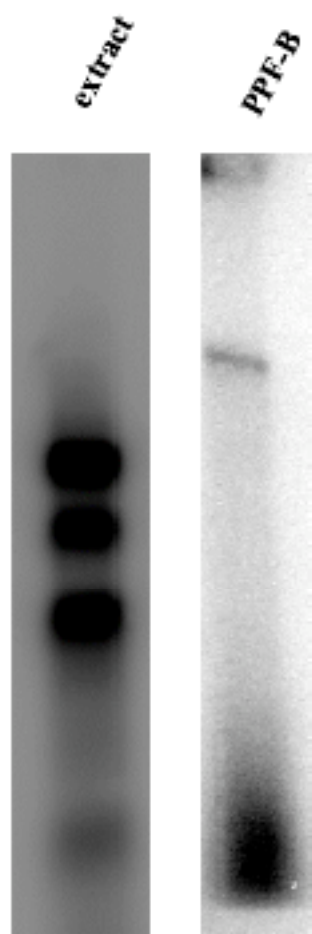


Figure 2.6. PPF-B RNA is Distinct from U1 snRNP.

- (A) PPF-B is distinct from U1 snRNP. RNAs from pea nuclear extracts and PPF-B were separated on a 2% formaldehyde/agarose gel and hybridized with a probe prepared using a pea U1 snRNA PCR product. Two exposures are shown; the lighter (left panel) indicates the quality of the RNA isolated from crude extracts, and the darker one (right panel) shows an absence of U1-related RNAs in the PPF-B sample. The RNAs analyzed, represent 40% of the purified RNAs from 100 μ l of PPF-B and crude extract, respectively.
- (B) Total RNAs extracted from pea nuclear extracts and MonoQ PPF-B were end-labeled with yeast PAP via [α -P³²]-Cordycepin 5' Triphosphate and analyzed as described under Materials and Methods. The RNA extracts were separated on a 2% formaldehyde/agarose gel.

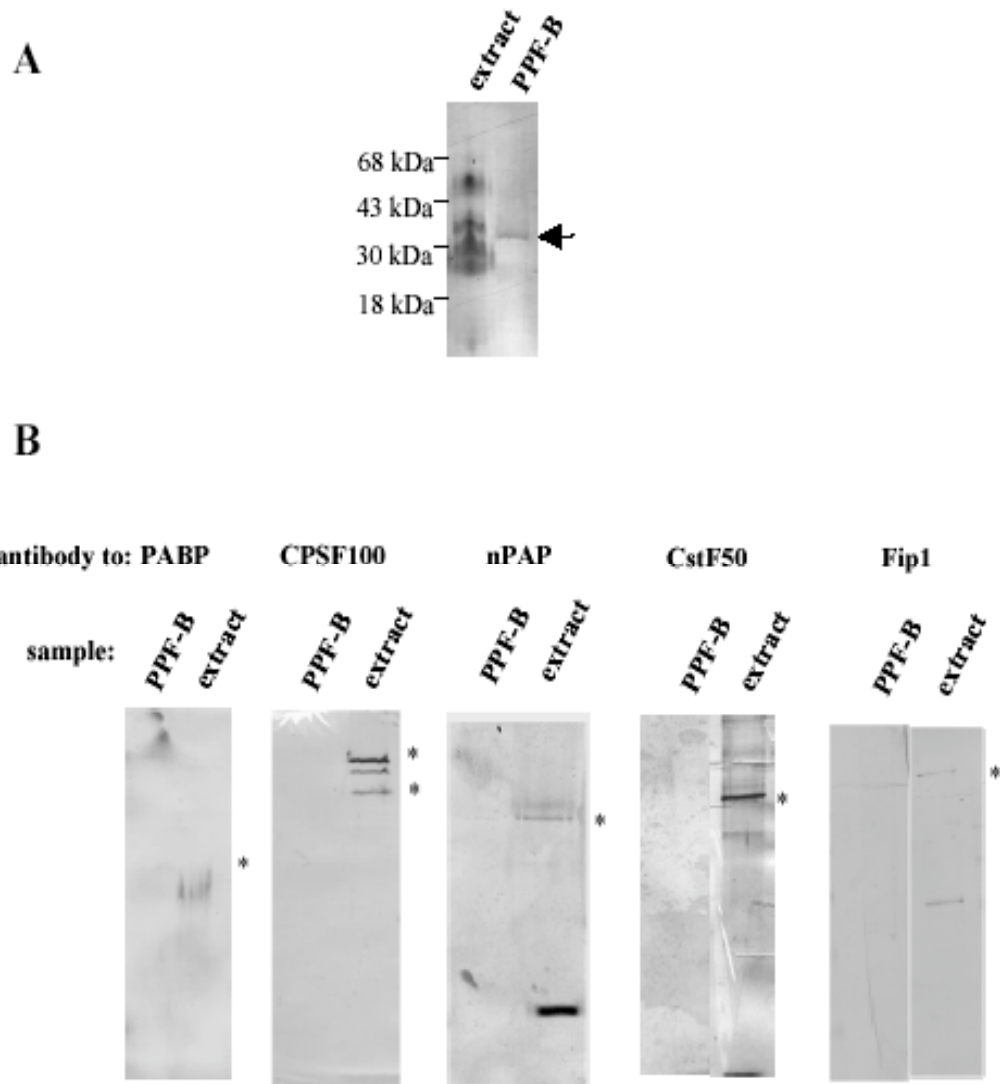


Figure 2.7. PPF-B does not appear to be Immunologically Related to Other Plant Poly(A) Factors.

(A) Thirty-five μ l (ca. 280 ng) of Mono-Q pure PPF-B ("PPF-B") was analyzed by SDS-PAGE, transferred to nitrocellulose, followed by Coomassie brilliant blue staining. The "extract" sample corresponds to pooled CM-Sepharose PPF-B fractions. The location of prestained size standards is shown on the left.

(B) Immunoblot analysis of PPF-B. Thirty-five μ l of PPF-B that had been purified through CM-Sepharose were analyzed by immunoblotting, using antibodies that recognize pea poly(A) binding proteins, a putative plant homologue of the 100 kDa subunit of CPSF, 50 kDa subunit of CstF, Fip1, and the nPAP. For comparison, 25 μ l (ca. 16 μ g) of crude nuclear extract was also analyzed. The positions of the PABP, CPSF100, CstF 50, Fip1 and nPAP polypeptides are indicated with stars next to the appropriate blot. The anti-CPSF100 antibodies consistently recognize two polypeptides, of about 100 and 70 kDa in size. The additional polypeptide seen in the blot shown here is not consistently seen.

A

MASEPTTVAVEQPIVEEPEAVDTFPPVVNESEEPTAKPKKAPKEPKAKKAPAKPRTHIPTYEEMVK
AIVALKERNGSSQYAIAKFIEEKHKQLPSNFKLLLVQIRKLVASGKLVKVKASYKLPKSSAAKPA
 KKPAAAKSKAKPKAKAATKSKAKPAAKAKPAAKAKPAAKAKPAAKAKPAAKAKPAAKAKPAAKI
 VKTAAAKPAAKAKPAAKPKAAAKAKPAAKAKPAAKAKPAAKAKPAAKAKPAAKAARTSTRTSPAAAAAAPKPA
 KKAAPVKKTPVKA~~AKAKAK~~TAKSPAKKAAAKRGKK

B

```

RSP      1:MAEEPTTtaveq...pVveEgeAvdGfppvvneseetAkpkkapK.....SpKAKKpApkpr.....pHPpEEMvKeAIVaLNEnGSSQAIaKFI: 87
TomH1   1:MAEEEPviva.....evVeEgLApevkvdean...PpAAsgkaKK.....StKAKKpAAPRHrsatp...pHPpYfEMIDAIvGLNEnGSSQhAITKFI: 85
TobH1   1:MAEEEPVvat.....eVtEPLvveeepakEd...nppAvesdepKk...ekStKAKKpAAPRHrnp...pHPpYfEMIDAIvGLNEnGSSQhAITKFI: 88
PeaH1   1:MAEEEPVavetvpepVtEpttitepevpkE...ePkAeveEtKa...k.gsKpKpGskpRn.p..ashPPYfEMIDAIvSLNEnGSSQAIaKFI: 91
AtH1    1:MsIEEenupt.....tVdsgAdtVvkspgk...kPaAsggksKktttakatKpvrAAAFcKkkttsHPpEEMIDAIvGLNEnGSSQAIaKFI: 91

RSP      88:EKkKqLPSfFrKLLLVQAIrIVASCKLVKVKsFrKLPKks...saMkEeKN.PAAAsKakpJaRAtkaKAKPAKkEpaakAaEAKKApAIAKPA:183
TomH1   86:EKkKqLPSfFrKLLLVQAIrIVASCKLVKVKsFrKLPKks...saMkEeKN.PAAAsKakpJaRAtkaKAKPAKkEpaakAaEAKKApAIAKPA:185
TobH1   89:EdKkKqLPSfFrKLLLVQAIrIVASCKLVKVKsFrKLPKks...saMkEeKN.PAAAsKakpJaRAtkaKAKPAKkEpaakAaEAKKApAIAKPA:178
PeaH1   92:EKkKqLPeLFrKLLLVQAIrIVASCKLVKVKsFrKLPKks...saMkEeKN.PAAAsKakpJaRAtkaKAKPAKkEpaakAaEAKKApAIAKPA:172
AtH1    92:EKkKqLPSfFrKLLLVQAIrIVASCKLVKVKsFrKLPKks...saMkEeKN.PAAAsKakpJaRAtkaKAKPAKkEpaakAaEAKKApAIAKPA:175

RSP      184:AKkKpAAKpAAKpVktAkkpaakaPaakPAaakaPaKkKaKaMaPaakAKAARSTTnsEaaaaaPpAAKkApvKktPvAaakAita:283
TomH1   186:AKkKpAAKpAAKpVktAkkpaakAaakpAavKPAAP.aakKlavKniKaM..ttAKVktA...Es....rkApkktPaKkPvKkap.AinV:273
TobH1   179:pKaKivKAKpAAKpVktAkkpaakAa.....KPAaekpAtpvktKaaKpA...ekAKVAKT...Esrk.aapVvAKKpVvKaaPaaksvkAita:265
PeaH1   173:AKkKpAAKpAAKpVktAkkpaakAa.....KPAaekpAtpvktKaaKpA...ekAKVAKT...Esrk.aapVvAKKpVvKaaPaaksvkAita:250
AtH1    176:paaKvVAAKpAAKpVktAkkpaakAa.....KPAaekpAtpvktKaaKpA...ekAKVAKT...Esrk.aapVvAKKpVvKaaPaaksvk...V:259

RSP      284:KSPAKKAAaKRC.kK:297
TomH1   274:KSPAKKAtpKRC.KK:287
TobH1   266:KSPAKKAsarkg.KK:279
PeaH1   251:KSPvKvVAKKqgK:265
AtH1    260:KSPAKKAsrka.kK:273

```

C

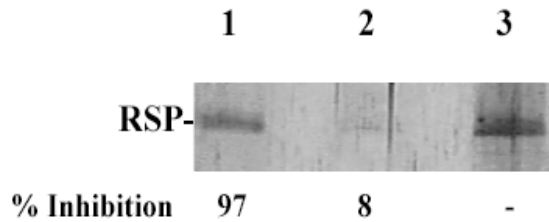


Figure 2.8. Association of PPF-B with a Linker Histone.

- (A) Amino acid sequence of Ribosome-Sedimenting Protein (RSP; accession BAA78535), a linker histone that is identical to the major polypeptide present in PPF-B. The two peptides whose sequences were obtained by direct sequencing are underlined.
- (B) Alignment of RSP with other plant linker histones. Databases were searched by BLAST and the most significant similarities tabulated among plant species are shown. The similarity of RSP with histone H1s from tomato *Lycopersicon esculentum* (TomH1; Accession P37218), tobacco *Nicotiana tabacum* (TobH1; Accession BAA88671.1), pea *Pisum sativum* (PeaH1; Accession P08283), and *Arabidopsis thaliana* (AtH1; v P26569) is shown. Homology was determined with ClustalW and the results were formatted using MacBoxshade. Conserved residues are in black boxes with white capital letters, residues of identity are boxed in gray with black capital letters, and other amino acids with black lower case letters.
- (C) Co-purification of RSP and PPF-B on DNA-Cellulose. Seventy μ l of PPF-B extract was incubated with DNA-cellulose for 30 minutes at room temperature. The DNA-cellulose was pelleted and 40 μ l of the supernatant analyzed by SDS-PAGE followed by Coomassie brilliant blue and Silver staining (lane 2). For comparison, 40 μ l of untreated PPF-B (lane 1) and the proteins that adhered to the DNA-cellulose were also analyzed (lane 3). The location of RSP is shown to the left of the gel. Untreated and treated PPF-B were also assayed for PAP inhibitory activity using 5 μ l of the cpPAP as the PAP source. The results, expressed as percent of PAP activity (compared with the cpPAP control), are shown beneath the gel.

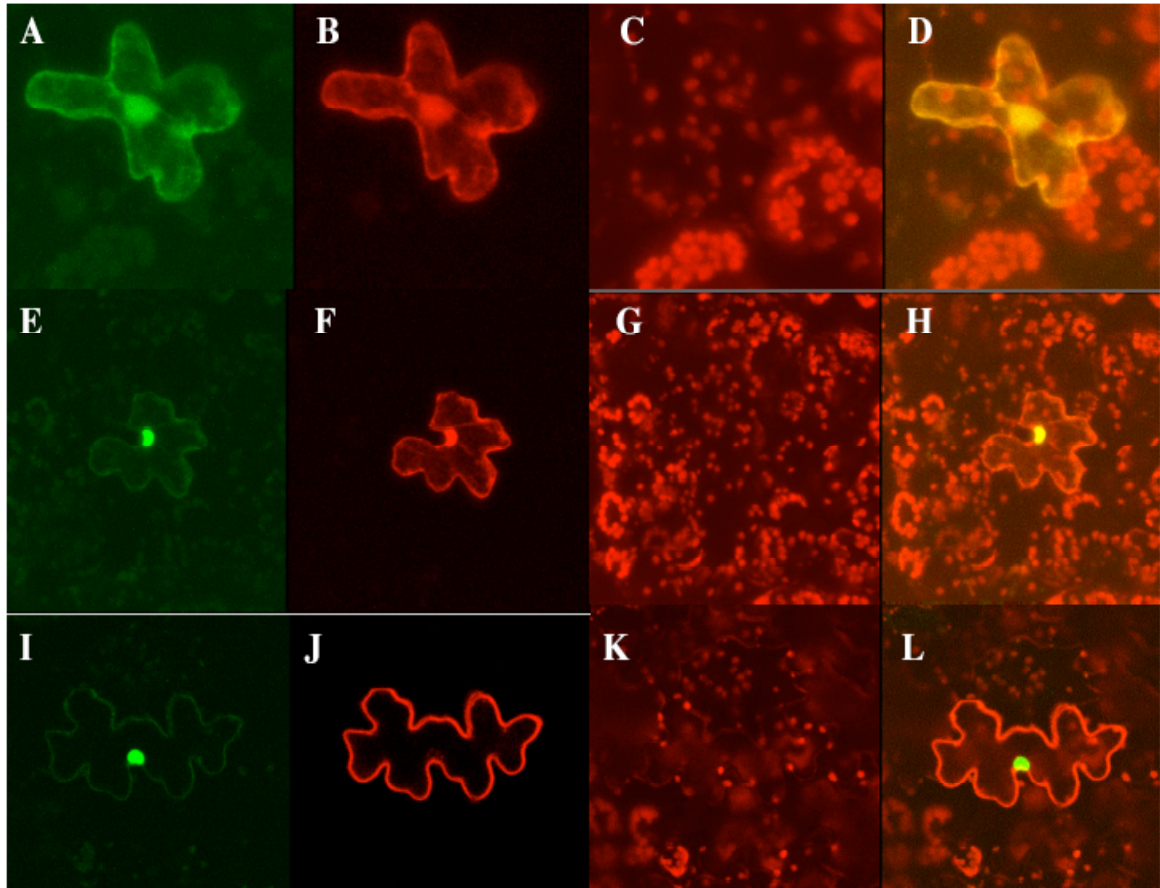


Figure 2.9. Sub-Cellular Localization of RSP with an *Arabidopsis* PAP.

Laser-scanning confocal micrographs showing fluorescence of leaf cells following particle bombardment carrying pGDG or pGDR plasmids expressing the autofluorescent proteins GFP, DsRED, PAP and RSP proteins and their autofluorescent protein fusions. Panels A to L depict micrographs showing expression patterns in selected whole cells. Panels (A) and (B) Co-expression of GFP expressed from pGDG and DsRed expressed from pGDR respectively. Panels (E) and (F) Co-expression of GFP:PAP and DsRed respectively. Panels (I) and (J) Co-expression of GFP:PAP and DsRed:RSP respectively. Panels C, G, and K represent background excitation of chlorophyll. Panel D overlay of images A-C. Panel H overlay of images E-G. Panel L overlay of images I-K. Green and Red channels used as described under Materials and Methods.

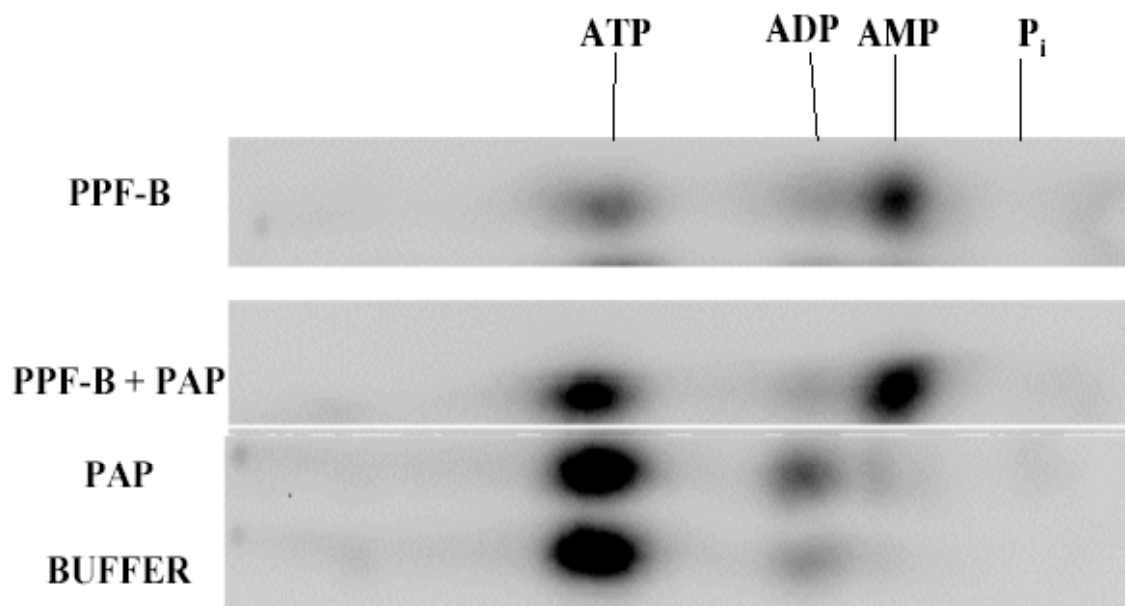
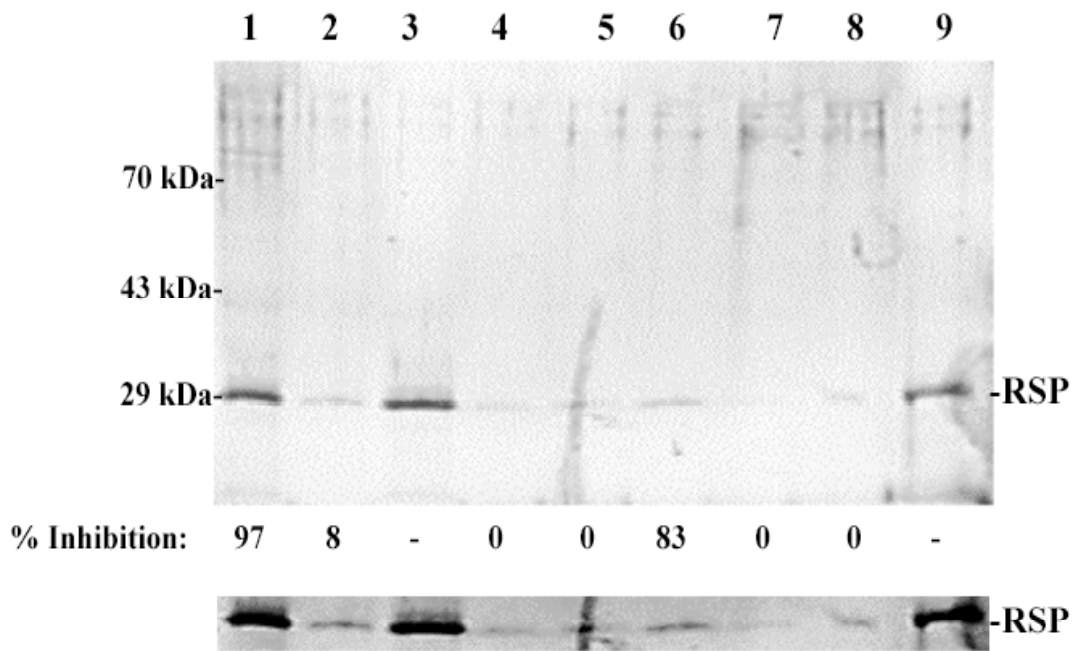


Figure 2.10. Pure PPF-B Contains an AMP-Forming Activity.

PAP reactions containing the indicated enzymes were conducted as described in Methods. After stopping each reaction, 1 μ l of the aqueous phase of each reaction was analyzed by TLC. The locations of ATP, ADP, AMP, and inorganic phosphate (Pi) are indicated on top. NEB was used as a buffer control. For these reactions, 3 μ g of cpPAP and 160 ng of PPF-B were used, as indicated

A



B

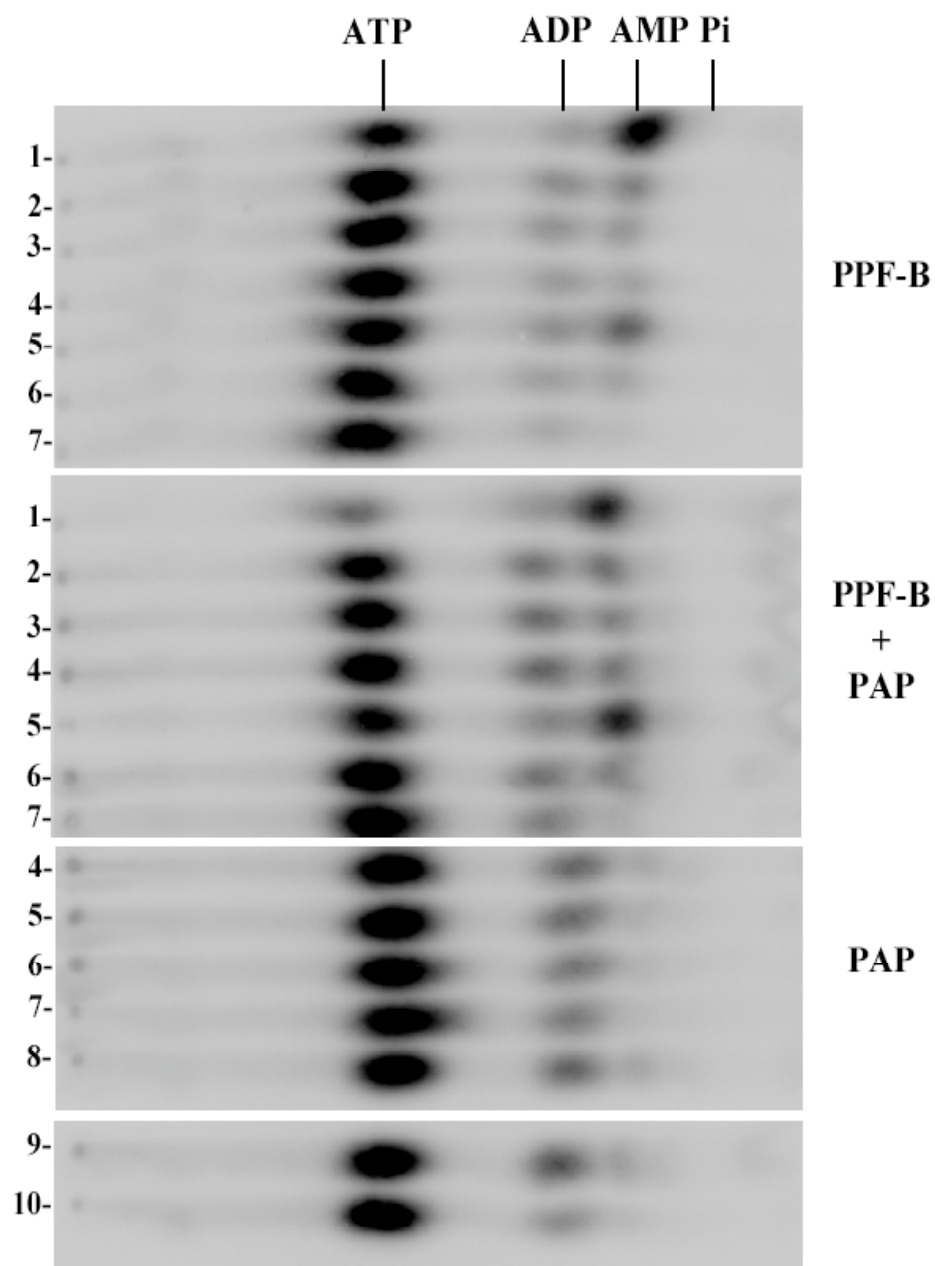


Figure 2.11. AMP-Forming Activity can be Eluted From DNA-Cellulose.

- (A)** Purification of AMP-forming activity with DNA-cellulose. Seventy μ l of PPF-B extract was incubated with DNA-cellulose for 30 minutes at room temperature. The DNA-cellulose was pelleted and supernatant collected. The DNA-cellulose was then incubated with 40 μ l of NEB containing 125 mM (lane 4), 250 mM (lane 5), 500 mM (lane 6), 1 M (lane 7), 2 M KCl for 30 minutes at room temperature (see Materials and Methods). Supernatants were analyzed by SDS-PAGE followed by Coomassie brilliant blue and Silver staining. For comparison, 40 μ l of untreated PPF-B (lane 1) and the proteins that adhered to the DNA-cellulose were also analyzed (lane 3). Lane 9 represents the proteins that adhered to the DNA-cellulose after the NEB-KCl washes. The location of RSP is shown to the right of the gel. Elutions were assayed for PAP inhibitory activity using 5 μ l of the cpPAP as the PAP source. The results, expressed as percent of PAP activity (compared with the cpPAP control), are shown beneath the gel. The location of prestained size standards is shown on the left. The lower panel is an enhanced picture of the gel region containing RSP, to help visualize the amounts of RSP present.
- (B)** NEB- KCl washes [125mM (lane 3), 250 mM (lane 4), 500 mM (lane 5). 1 M (lane 6), 2 M (lane 7)] from A were incubated in PAP reactions containing the indicated enzymes as described in Methods. Untreated (lane 1) and treated (lane 2) PPF-B were also assayed with 3 μ g of cpPAP (lane 9). After stopping each reaction, 1 μ l of the aqueous phase of each reaction was analyzed by TLC. The locations of ATP, ADP, AMP, and inorganic phosphate (Pi) are indicated on top. NEB was treated with DNA-cellulose and incubated with PAP (lane 8). NEB was also used as a buffer control (lane 10).



Figure 2.12. *Arabidopsis* PAPIs Contain Potential Sites for Sumolation.

Sequence alignment of nuclear *Arabidopsis* PAPIs regions that contain potential sumolation sites; AtPAPI (At1g17980; aa 386-413), AtPAPII (At2g25850; aa 393-421), AtPAPIII (At3g06560; aa 338-363), AtPAPIV (At4g32850; aa 395-423) (Addepalli et al., 2004). Homology was determined with ClustalW and the results were formatted using MacBoxshade. Amino acids that are identical in all four are shaded black, with white lettering. Positions that are similar are shaded gray, with white lettering. Black bars above and below the sequence represent the potential sumolation sites.

CHAPTER THREE

Characterization of *Arabidopsis* Genes Encoding Fip1 Homologs

INTRODUCTION

Compared to other eukaryotic systems, nuclear RNA metabolism in plants is poorly understood. This is especially true for mRNA 3' end formation and polyadenylation. Thus, while a good understanding of the RNA sequence requirements for mRNA 3' end formation has been obtained (Rothnie, 1996; Li and Hunt, 1997; Graber et al., 1999a), little is known about the factors that recognize plant polyadenylation signals or polyadenylate mRNAs in plants. Plant genomes possess genes encoding probable homologs of most of the polyadenylation factor subunits that have been identified in mammalian and yeast (Q.Q Li and A.G. Hunt, unpublished observations), but the functioning of most of these in the process of mRNA 3' end formation in plants remains to be demonstrated or understood.

In mammals and yeast, Fip1 appears to be an important bridge between several polyadenylation factors. Fip1p interacts with PAP, Yhh1p, Yth1p, Pfs2p, and RNA14 in yeast (Preker et al., 1995). Fip1p is the only polyadenylation factor subunit known to date to interact directly with PAP. The human homolog, hFip1, interacts with PAP, CPSF30, CPSF160, and CstF77 (the mammalian counterparts of Yth1p, Yhh1p, RNA14 respectively) and has been recently recognized as an integral subunit of CPSF (Kaufmann et al., 2004). The yeast and human Fip1 proteins have different structural make-ups and biochemical properties. The yeast protein (Fip1p) lacks an RNA-binding domain and inhibits the non-specific activity of PAP (the ability to extend any RNA template lacking a polyadenylation signal), while the human Fip1 can bind RNA and stimulates PAP activity. In both systems, Fip1 is a possible bridging factor, providing links between PAP, RNA, and other multi-subunit complexes. These links serve to promote PAP activity via some sort of tethering function to the RNA substrate. Since we have identified a novel factor from pea nuclear extracts that inhibits the non-specific activity of PAP, PPF-B, we set out to identify genes from *Arabidopsis* that encode a Fip1 homolog.

In this chapter, we present the characterization of *Arabidopsis* Fip1 homologs encoding genes on chromosome III (AtFip1-III) and V (AtFip1-V). We find that AtFip1-V is related to hFip1 by way of protein sequence motifs and AtFip1-III appears to be divergent from other Fip1 homologs. AtFip1-V contains domains that stimulate the non-

specific activity of at least one *Arabidopsis* nuclear PAP, bind RNA, and interact with other cleavage and polyadenylation homologs from *Arabidopsis*. The abilities of AtFip1-V to bind RNA and stimulate PAP activity, along with the interactions with AtCstF77 and AtCPSF30, are properties that AtFip1-V shares with its human counterpart. Taken together, these results suggest that AtFip1-V may coordinate several polyadenylation factor subunits and may play a role in regulating the activity of PAP in plants.

RESULTS

Identification and Characterization of Possible Arabidopsis Fip1 Genes

Database searches of the *Arabidopsis* genome using the human Fip1 (hFip1) and *Saccharomyces cerevisiae* Fip1p (ScFip1p) protein as search queries, revealed two potential Fip1 genes, residing on chromosomes III and V (gene designations: At3g66652 and At5g58040 respectively). At the time, no EST data was available for either gene in the databases. Therefore, cDNA sequences encompassing the predicted coding regions were generated by RT/PCR (using total RNA as a template for reverse transcription) and PCR (using cDNA libraries) with specific oligonucleotide primers (see Materials and Methods) and sequencing. A cDNA of 3588 bp, all coding region, was obtained for chromosome V Fip1 (AtFip1-V) that encodes a protein of 1196 amino acids (Figure 3.1A) with a predicted molecular weight of 133 kDa. A partial cDNA of 1980 bp of the predicted 2943 bp coding region derived from chromosome III Fip1 (AtFip1-III) was obtained. This partial clone lacks 963 bp of the predicted 3' end of the coding region. RT/PCR with specific primers for the 3' end generated a fragment which contains predicted introns. This suggests genomic DNA contamination, but the first strand cDNA was generated from total RNA. Thus, the 3' end of the chromosome III Fip1 gene is unknown. The partial cDNA encodes a protein of 660 amino acids (Figure 3.1B) with a predicted molecular weight of 75 kDa.

Human and yeast Fip1 proteins share a similar domain organization representative of an acidic N-terminus, a conserved FIP1 domain followed by a proline-rich domain. The hFip1 also has a C-terminal extension with a domain rich in arginine and aspartate (RD domain), followed by an arginine-rich region, which includes a predicted bipartite nuclear localization signal (Kaufmann et al., 2004). AtFip1 peptide sequences were analyzed by Motif-Scan:PROSITE (Falquet et al., 2002) to identify potential domains. A schematic alignment of the domains identified in these proteins is shown in Figure 3.2A. The hFip1 and ScFip1p domain organization is adapted from Kaufmann et al. (2004). As seen with other eukaryotic Fip1 homologs, AtFip1-III and V contain an acidic N-terminus and the conserved FIP1 domain. AtFip1-V also contains a C-terminal arginine-rich domain with two predicted bipartite nuclear localization signals, and an N-

terminal glycine-rich region not seen in the hFip1 and ScFip1p. This architecture suggests that AtFip1-V is related to hFip1. In contrast, AtFip1-III seems to be more divergent from other Fip1 homologs based on. Although, the C-terminus of AtFip1-III has not been determined, it may be more like other Fip1s.

In our database searches for the *Arabidopsis* Fip1 genes, the highly conserved FIP1 domain was the only domain identified in both genes. The Fip1 domain is approximately 70 amino acids in length. Multiple sequence alignments of the Fip1 domain from the two AtFip1 proteins, the human Fip1 and the yeast Fip1p showed a high degree of conservation (Figure 3.2B). AtFip1-V had about 26%, 38%, 41% identity and 40%, 56%, 50% similarity to hFip1, ScFip1p, and AtFip1-III, respectively. AtFip1-III is 23%, 34% identical and 38%, 40% similar to hFip1 and ScFip1p, respectively. The human and yeast Fip1 domains have 40% identity and 53% similarity.

To determine the expression profiles of AtFip1-V and AtFip1-III, total RNA was isolated from four different tissues of *Arabidopsis*: flower, leaves, stems and roots. cDNA was generated from this RNA by RT using oligo-d(T) and PCR was performed with gene-specific primers spanning a region including introns. PCR products of the predicted spliced gene fragments were observed in all four tissues for chromosome III and V AtFip1 (Figure 3.3). Analysis of public micro-array data using GENEVESTIGATOR (Zimmermann et al., 2004), verifies expression of AtFip1-V and AtFip1-III in all four tissues used in the above experiment (data not shown). The public micro-array data compiled at GENEVESTIGATOR also indicates that chromosome V Fip1 is expressed at higher levels than chromosome III Fip1 in different organs or tissues (data not shown). In either case, these results suggest that both AtFip1 genes are universally expressed in the four *Arabidopsis* tissues.

In comparison to AtFip1-III, AtFip1-V has moderately higher identity and similarity scores, contains more structurally related peptide domains found in Fip1 homologs, a clone of the full-length coding region available, and is expressed at higher levels in the plant. Therefore, the AtFip1-V gene was selected for further characterization.

Polypeptides Related to AtFip1 are Present in Pea Nuclear and Arabidopsis Extracts

The predicted product of the AtFip1-V gene is predicted to be nuclear localized. To test this hypothesis, pea nuclear extracts and *Arabidopsis* extracts were probed with antibodies raised against the N-terminal 137 amino acids of the AtFip1-V protein. These antibodies were generated with a GST-tagged fusion protein and therefore the antibodies were pre-incubated with GST to reduce background. These antibodies consistently recognize two polypeptides in pea nuclear extracts and multiple polypeptides in *Arabidopsis* extracts (Figure 3.4). Importantly, these polypeptides are not recognized when the antibody is pre-incubated with GST-Fip1 fusion proteins. Therefore, this suggests that these polypeptides are authentic and contain at least part of the first 137 amino acids of AtFip1-V. At this time, probing of *Arabidopsis* nuclear extracts is limited, due to low amounts of total protein extracted from *Arabidopsis* nuclei. In any case, these results support the hypothesis that the products of the AtFip1-V encoding gene are present in the nuclei of plants.

Interactions Between AtFip1-V and Other Arabidopsis Cleavage and Polyadenylation Homologs in Yeast

In the human and yeast cleavage and polyadenylation apparatus, Fip1 is an important bridge between many factors (see introduction). With the use of a yeast two-hybrid system, we tested AtFip1-V for protein-protein interactions with other plant cleavage and polyadenylation homologs. The full-length cDNA of AtFip1-V was split into two fragments and fused to the activation domain (AD). These cDNA fragments, FipN and FipC, span amino acids 1-137 and 407-1192 respectively. *Arabidopsis thaliana* (At) homologs to the mammalian CstF (50, 64 and 77) and CPSF (30, 73, 100, and 160) subunits and the yeast Pfs2 were fused to the DNA binding domain (BD) and then introduced into the yeast strain PJ69-4 with FipN or FipC AD fusions. As seen in figure 3.5, transformants from most of these fusion genes with empty BD or AD vector did not grow on selection plates minus adenine, indicating no self-activation. Growth on selection plates minus adenine was observed with the BD-AtCPSF30/empty-AD transformant, suggesting self-activation activity of AtCPSF30. Thereafter, AtFip1-V (FipN) and AtCPSF30 were fused to BD and AD respectively and retested. In addition,

an *Arabidopsis* PAP gene AtPAPIV, which is a nuclear poly(A) polymerase located on chromosome IV, was also tested with FipN in the experiment.

As summarized in figure 3.5, FipN showed interactions with AtCstF77, AtCPSF30 and AtPAPIV. No interactions were detected with FipN and the other *Arabidopsis* homologs. The FipC region of the AtFip1-V protein appears to not interact with the above mentioned *Arabidopsis* homologs, thus suggesting a different function other than protein-protein interactions. However, we cannot rule out further interactions, possibly involving other *Arabidopsis* cleavage and polyadenylation homologs or the full length AtFip1 protein.

Effects of AtFip1-V on PAP Activity

In eukaryotes, PAP functions in the cleavage and polyadenylation complex with other proteins to specifically polyadenylate cleaved mRNA precursor. This activity is referred to as specific PAP activity. PAP enzymatic activity can be studied independently of its association with other factors, due to the ability to extend any RNA template, an activity known as non-specific PAP activity. Mammalian and yeast PAP proteins both have very diverse non-specific activities. The activity of the mammalian PAP is low and distributive due to its weak and non-specific binding to RNA (Wahle, 1991). Yeast PAP has a high affinity for and rapidly elongates non-specific RNA, in a processive manner (Zhelkovsky et al., 1998).

In yeast the only protein known to date to interact with PAP directly is Fip1p (Preker et al., 1995). In contrast to yeast Pap1, mammalian PAP interacts with CPSF160 and the recently identified human homolog to the ScFip1p, hFip1 (Kaufmann et al., 2004). Both Fip1 proteins have been shown to affect PAP activity. ScFip1p inhibits PAP activity by shifting it from a processive to a distributive enzyme. In contrast, the hFip1 stimulates the non-specific activity of PAP.

From the interaction studies, AtFip1-V interacted with AtPAPIV by way of the FipN fragment. Recombinant FipN (aa 1-137) was produced and assayed with recombinant AtPAP(IV) (Figure 3.6A). Although recombinant FipN is predicted to be approximately 15 kDa, it resolved in denaturing gels with an apparent molecular weight of 24 kDa. The *Arabidopsis* PAP(IV) polyadenylates the RNA substrate into a large

extension poly(A) product and the quantity of products increase over time, analogous to the processive activity of the yeast PAP (Figure 3.6B). In the presence of FipN, large extension products were observed at earlier time points, when compared to the PAP alone (Figure 3.6A). Moreover, the overall quantities of product were greater at all times in the presence of FipN (3.6 C). This suggests that AtFip1-V contains a domain that stimulates poly(A) synthesis, analogous to hFip1.

RNA Binding Characterization of AtFip1-V

The C-terminal arginine-rich domain of hFip1 binds to RNA with a preference for U-rich sequence elements (Kaufmann et al., 2004). Like hFip1, AtFip1-V contains an arginine-rich C-terminus. This prompted us to investigate the RNA binding characteristics of AtFip1. A GST-tagged FipC (GST-FipC; aa 407-1192) recombinant protein was produced and purified (Figure 3.7A). The predicted molecular weight of the GST-FipC fusion is 115 kDa. As seen in Figure 3.7A, GST-FipC made in *E. coli* generates breakdown products as well as full-length GST-FipC fusions that are recognized by GST anti-serum.

To determine if GST-FipC binds RNA, we generated internally labelled RNAs containing poly(A) signals for this purpose. We used RNA containing the poly(A) signal from the pea *rbcS-E9* gene because it contains an FUE, three NUES, and three sites of cleavage and polyadenylation (Figure 3.8A; Mogen et al., 1992). This RNA contained sequences upstream (-145) and downstream (+81) of the second cleavage and polyadenylation site found in the *rbcS-E9* poly(A) signal. The GST-FipC mixture was able to bind RNA containing this poly(A) signal as seen in electrophoretic mobility shift assays (EMSAs) (Figure 3.7B). With EMSAs, the shift of RNA template into a larger complex (bound RNA), based on mobility, suggests RNA-binding. GST purified from *E. coli* had no RNA-binding activities as seen in Figure 3.7B.

To further characterize the RNA binding properties of FipC, we performed competition EMSA binding experiments with internally labeled *rbcS-E9* and increasing amounts of RNA homopolymers as competitors. Homopolymers were used to determine if FipC has a preference for sequences rich in A, C, G or U. FipC binding of RNAs containing the *rbcS-E9* polyadenylation signal was most efficiently competed by poly(G)

(Figure 3.7B). In contrast, poly(A) and poly(C) were less efficient and poly(U) did not compete (Figure 3.7B). The ratio of bound to unbound RNA was plotted against the fold excess amount of each competitor, which further supported these results (Figure 3.7C). Thus, the FipC preparation appears to have a preference for poly(G).

In order to narrow down the RNA binding domain of FipC, GST-FipC was purified with poly(G) agarose. Moderate amounts of the full-length fusion protein were produced, as well as some breakdown products (Figure 3.7A). Three of these GST-FipC proteins were pulled down when incubated with poly(G) agarose (Figure 3.7A). The extent of FipC present in the proteins was estimated by comparison with size standards. All three of these fusion proteins contain the bulk of the arginine-rich domain. Therefore, the RNA binding domain of AtFip1-V is within the arginine-rich C-terminus.

Further EMSA competition experiments were performed to determine the binding site preference within the *rbcS-E9* polyadenylation signal. For this, yeast tRNA and different RNAs derived from the *rbcS-E9* (Mogen et al., 1992) and cauliflower mosaic virus (CaMV) polyadenylation signals (Mogen et al., 1990) were used as competitors (Figure 3.8A). CaMV poly(A) signal also contains an FUE, NUE (AAUAAA) and a cleavage and polyadenylation site. As seen in Figure 3.8B, the wild type *rbcS-E9* that contains sequences +81 (81 bases) downstream of the second polyadenylation site was the only RNA that could reduce RNA binding. This suggests sequences found from +6 to +81 in the *rbcS-E9* 3' region, are bound by FipC. RNAs that end at +6 and +81 in the *rbcS-E9* were used in increasing amounts as competitors to determine this. Excess RNA containing +6 to +81 sequences efficiently competed for FipC RNA binding (Figure 3.8C). The ratio of bound to unbound RNA was plotted against the fold excess amount of each competitor, which further supported these results (Figure 3.7C). In addition, the plot suggests that RNAs with just sequences +6 downstream of the second polyadenylation site could compete but not to the same degree as RNAs ending at +81 of the RNA (Figure 3.8C). Since +81 RNA contains three cleavage sites while +6 contains one, this result suggests that AtFip1-V may have a preference for cleavage and polyadenylation sites found in plant genes.

Finally, further EMSA experiments were performed with GST-FipC and RNA containing the *rbcS-E9* polyadenylation signal (-145 to +81) to determine RNA affinity

for FipC. As seen in Figure 3.9A, as we dilute GST-FipC (from 2.14 μ M to 30nM), the amount of complex formed (bound RNA), decreases. The ratio of bound to unbound RNA was plotted against the molar amount of protein used in each experiment (Figure 3.9B). In this experiment we were not able to reach saturation. For this, we roughly estimated the apparent K_D values for FipC to be $>1 \mu$ M. Kaufmann et al. (2004) have reported that the hFip1 has an apparent K_D value of 2-4nM. This is significantly higher than FipC, and suggests that FipC has a low affinity for RNA.

DISCUSSION

Fip1 is an interesting subunit of the cleavage and polyadenylation apparatus and genes encoding predicted Fip1 subunits can be found in the genomic databases of several eukaryotes. Fip1 is considered an important subunit involved in cleavage and polyadenylation since it serves at least two functions, one being the bridge between subunits within the polyadenylation complex, and second regulating the activity of PAP. We have identified two *Arabidopsis* Fip1 homologs encoded by genes residing on chromosomes III (AtFip1-III) and V (AtFip1-V). These genes are universally expressed in different organs and tissues. AtFip1-V, which can be detected in nuclear extracts, interacts with AtPAPIV, AtCstF77, and AtCPSF30 and stimulates non-specific polyadenylation by PAP. AtFip1-V also binds preferentially to poly(G) as well as RNAs that possess polyadenylation signals.

The human and yeast Fip1s share protein-protein interactions with three other subunits besides PAP; CPSF160/Yhh1p, CPSF30/Yth1p, and CstF77/RNA14. In addition, the yeast protein interacts with Pfs2p. The yeast two-hybrid interactions reported here show that the N-terminus of AtFip1-V (FipN; aa 1-137) interacts with the *Arabidopsis* homologs of CstF77, CPSF30 and PAPIV. This suggests a conservation of these protein-protein interactions in eukaryotes. FipN has also been shown to interact by yeast two-hybrid and *in vitro* pull-down assays with the *Arabidopsis* homologs of CFI-25 and PABN1 (Balasubrahmanyam Addepalli; personal communication). The interactions with PABN1 and CFI-25 have not been reported in other systems, and may reflect a unique aspect of the plant polyadenylation machinery. AtCstF77 has also been shown to interact with AtCstF64 (Kevin Forbes and Arthur Hunt, unpublished observations; Yao et al., 2002). Therefore, this suggests a strong linkage of AtFip1-V to the plant cleavage and polyadenylation apparatus and suggests that AtFip1-V coordinates a number of polyadenylation factor subunits.

Stimulation of PAP activity by Fip1 has been reported with the human Fip1. This stimulation is dependent on the C-terminus of hFip1 that binds a U-rich sequence element. Therefore, hFip1 mediated stimulation of PAP may result from tethering PAP to the RNA. However, we have seen that an N-terminal domain of AtFip1-V (FipN), which

lacks the RNA binding domain, stimulates PAP activity. This suggests that AtFip1-V may be involved in at least the poly(A) synthesis step and does not require the RNA-binding domain.

The C-terminus of AtFip1-V has not been reported to interact with other polyadenylation factor subunits. This suggests that it functions specifically in RNA binding. While the subunits of the mammalian and yeast polyadenylation complexes that bind polyadenylation signals have been identified, there have been no reports of an *Arabidopsis* protein binding to specific polyadenylation signals. Plant genes possess polyadenylation signals that are distinctly different from their mammalian and yeast counterparts and consist of three different classes of *cis* elements: NUEs, FUEs, and cleavage and polyadenylation sites themselves (Hunt, 1994; Rothnie, 1996; Graber et al., 1999a). AtFip1-V preferentially binds poly(G) and RNA containing sequences +6 to +81 of the *rbcS-E9* polyadenylation signal. This suggests that sequences +6 to +81 contains a sequence element that AtFip1-V preferentially binds. From this study, we were not able to definitively identify where in these sequences binding occurs. Sequences from +38 to +48 (GUUGUGUGUUG) might be the preferred binding site of AtFip1-V given the affinity for poly(G). RNA containing sequences up to +6 also compete for AtFip1-V RNA-binding, but to a lesser degree. This RNA contains only one cleavage and polyadenylation site and no stretch of G-rich sequences downstream of this site. Therefore, it is possible that AtFip1-V may have a preference for cleavage and polyadenylation sites and not downstream sequences.

The RNA-binding activity of AtFip1-V lies in the arginine-rich C-terminus of the protein. Arginine-rich motifs (ARM) are found in a variety of RNA-binding proteins and often recognize a particular structure rather than a sequence (Burd and Dreyfuss, 1994). Structural elements have also been proposed to be determinants of poly(A) site recognition (Graveley et al., 1996; Klasens et al., 1999; Hans and Alwine, 2000). Accordingly, the RNA target of AtFip1-V could require a specific structure and not a specific sequence. Two human immunodeficiency virus (HIV) proteins that contain ARMs, Tat and Rev, that are required for HIV replication, have a K_D of 5nM and 1nM respectively (Burd and Dreyfuss, 1994). As previously mentioned, hFip1 has a K_D of 2-4nM. In contrast AtFip1-V, as an ARM containing protein, has a low affinity for RNA

($K_D = >1 \mu\text{M}$). Some other RNA-binding proteins involved in cleavage and polyadenylation have a low affinity for RNA. For example, CstF64 has a reported K_D of 1.5 μM to 0.1 μM , and PABN1 a K_D of 8 μM to 0.1 μM , which is dependent on RNA substrate length (Takagaki and Manley, 1997; Dichtl et al., 2002a; Kuhn et al., 2003).

The site or structure to which AtFip1-V binds may require other subunits to determine specificity. For example, in mammals CPSF confers specificity by binding to the poly(A) signal (AAUAAA) through its 160 kDa subunit CPSF160. The downstream element (G/U) is recognized by CstF through its 64 kDa subunit CstF64 (Beyer et al., 1997; Takagaki and Manley, 1997). Both CPSF and CstF bind RNA weakly, but form a strong cooperative complex when bound together to the same pre-mRNA (Bienroth et al., 1993). Therefore, multiple subunit interactions may confer specificity. *Arabidopsis* homologs of CstF64, CPSF30, CPSF100, CPSF160, and CFI-25 are potential candidates for binding FUE or NUES in plant polyadenylation signals. These subunits may determine the specificity of AtFip1-V by direct protein-protein or multiple protein interactions in the complex. Further experiments are required to determine if other *Arabidopsis* subunits bind RNA and what sequences they preferentially bind.

The Fip1 gene found on chromosome III in *Arabidopsis* (AtFip1-III) does contain at least the conserved acidic N-terminus and Fip1 domain found in eukaryotic Fip1s. There is no evidence so far of *Arabidopsis* polyadenylation factors interacting with AtFip1-III. Since it does contain conserved domains that are indicated to be a hub for poly(A) factors, it may function in cleavage and polyadenylation. Further experiments will be needed to determine if AtFip1-III is involved in polyadenylation.

Based on the data described in this chapter, we propose a model for the organization of the plant cleavage and polyadenylation apparatus around Fip1 (Figure 3.10A). This model is based on individual protein-protein interactions between subunits. In Figure 3.10B we propose a model for poly(A) signal recognition by the complex in Figure 3.10A. For simplicity, we have only included potential RNA binding proteins and the subunits that may link them together. Further analysis of protein-protein interactions with multiple subunits (more than two) will be required to determine if all these interactions occur simultaneously.

To summarize, the studies described in this chapter showed that several homologs to mammalian and yeast polyadenylation subunits have been identified in *Arabidopsis*. The *Arabidopsis* Fip1 subunit has analogous functions to those found with the human Fip1 and yeast Fip1p. Further characterization of AtFip1 should allow us to better understand the mechanism by which PAP activity is regulated, and how it functions in mRNA 3' end formation in plants.

MATERIALS AND METHODS

Plant Materials

Arabidopsis thaliana seed was obtained from Lehle Seeds. Seed were germinated and plants cultivated in the greenhouse for 3-4 weeks. Plants were harvested before as well as after the flowering stage. Leaves, stems and flowers were used for total RNA isolation (see below). Root material was gathered from seedlings that were grown in liquid culture under lights with shaking. 50 ml of germination media (500mg Sucrose, 215.5mg Murashige and Skoog Basal Medium (Sigma), 25mg MES) was inoculated with 30-40 sterilized seeds and grown for 2-3 weeks at room temperature under 12 hour lights.

RNA Isolation from Arabidopsis and Generation of First Strand cDNAs

Total RNA was isolated from *Arabidopsis* leaves using either an SV Total RNA Isolation Kit (Promega), RNeasy[®] Plant Mini Kit (Qiagen), or Trizol (Invitrogen), per the manufactures instructions. Reverse transcription experiments were conducted using the total RNA, oligo-dT and Superscript RT II (Invitrogen), oligo-dT with the ProSTAR[®] Ultra HF RT-PCR system (Stratagene) or random primers using a RETROscript[®] First Strand Synthesis Kit (Ambion).

Isolation and Characterization of Arabidopsis Fip1 cDNAs

cDNAs derived from Fip1-encoding mRNAs were isolated from a number of sources by PCR and RT/PCR. Potential Fip1-encoding genes were identified in the *Arabidopsis* genome (<http://www.arabidopsis.org/home.html>), with TBLASTN and BLASTP using the human and yeast Fip1 amino acids sequences as search queries. Based on the results, primers were designed to amplify the cDNA coding regions of AtFip1-III and V listed in Table I. AtFip1-III primers 5'FL/3'INT2 and 5'INT/3'FL were used in a continuous (1-step) reverse transcriptase PCR (RT-PCR) reaction. The reaction mixture contained: 350 mM Tris-HCl (pH 8.8), 250 mM KCl, 2.5 mM DTT, 6mM MgSO₄, 0.5% Triton X-100, 1 mM deoxynucleotide triphosphates (dNTPs), 30 U RNase inhibitor (Eppendorf), 100 U SuperScript RT II (Invitrogen), 1 U Taq DNA polymerase (Invitrogen), 100 ng of each primer, and 200 ng *Arabidopsis* RNA. PCR reactions were

run at 55°C for 30 min, 95°C for 2 min, then 35 cycles of 95°C for 30 s, 55°C for 45 s, 72°C for 90s, and one step at 72°C for 7 min. RT-PCR products of expected sizes (1970 bp and 1295 base pairs) were subcloned into the EcoRV site of pGEM-T Easy (Promega), per manufacture instructions, clones were sequenced using the listed primers and T7 and SP6 primers (). DNA sequencing was done by automated sequencing (Perkin & Elmer, ABI Prism 310 Genetic Analyzer) using a BigDye Terminator Cycle Sequencing Ready Reaction kit (ABI prism).

Primer sets for AtFip1-V (5'FL/3'INT, 5'INT/3'INT1, 5'INT1/3'FL) were used in PCR reactions with first strand cDNA or with a 3-6 kb cDNA expression library for *Arabidopsis* (CD4-16; ABRC-DNA Stock Center; (Kieber et al., 1993) as templates. PCR reactions contained a final concentration of 1X PCR buffer (Invitrogen), 5mM MgCl₂, 125µM dNTPs, and 1U of Taq DNA polymerase (Invitrogen), with 100 ng of each primer and template. PCR reactions were run at 95°C for 2 min, then 35 cycles of 95°C for 30 s, 55°C for 60 s, 72°C for 90s, and one step at 72°C for 7 min. PCR products of expected size (1478bp, 1460bp, and 899bp) were subcloned into pGEM-T Easy vector and sequenced as described above. These three pGEM clones represent the 5'end (N; bases 1-1478), middle region (M; bases 1220-2692) and 3'end (C; bases 2672-3588) of the full-length coding region, respectively. All three clones are in the 5' to 3' direction from the T7 promoter in the vector. The pGEM-N and pGEM-M clones were digested with *Ava*I and *Nsi*II (American Allied Biochemical; AAB). All further digests were performed with AAB restriction enzymes. The resulting fragments M fragment (approx. 1200 bp) was ligated into pGEM-N with T4 DNA ligase (Invitrogen). This clone pGEM-N+M (bases 1–2692) and the pGEM-C clone (bases 2672-3588) were transformed into a *dam*-/*dcm*- strain of *E. coli* (Strain ER2925-New England Biolabs; NEB) to eliminate DNA methylation of a *Cla*I restriction site in both clones. Plasmid DNA was purified and digested with *Nco*I and *Cla*I. The resulting N+M fragment (2672 bp) was ligated into pGEM-C. The resulting clone yields a full-length protein coding region of 3588 bp, termed pGEM-AtFip1-V(FL). One clone was isolated for the very 5'end of AtFip1-V (N) that contained a premature stop codon. This recombinant can only encode amino acids 1-137. This recombinant interacted with other subunits and was used for subsequent characterizations.

For expression analysis of AtFip1 genes in different *Arabidopsis* tissues, PCR amplification was done with 1.5 μ l of first strand cDNA (ProSTAR; Stratagene) added to 100ng of each primer, 0.8mM dNTPs, 5.0 μ l Ultra HF PCR buffer (Stratagene), and 2.5 units of Pfu Turbo DNA Polymerase (Stratagene) in a 50 μ l reaction. Primers used for AtFip1-III and AtFip1-V amplification were 5'FL/3'INT and 5'INT/3'INT2 respectively.

Cloning of Arabidopsis Cleavage and Polyadenylation Homologs cDNAs

Database searches of the *Arabidopsis* genome with the yeast Pfs2 and human CstF 50, 64 and 77 as search queries using TBLASTN and BLASTP at the previously mentioned web pages, identified potential homologs to each subunit. Based on the sequence information, primers were designed to amplify the cDNA coding regions of these genes using primers listed in Table I. All of the following clones were generated by PCR under the same reaction conditions used for AtFip1-V.

For the cloning of the *Arabidopsis* Pfs2 homolog, AtPfs2 primer sets of 5'FL/3'INT and 5'INT/3'FL were used in a PCR reaction with first strand cDNA as template. PCR products of approx. sizes (1212bp and 897bp) were subcloned into pGEM-T Easy vector and sequenced as mentioned above. Both clones were in the 5' to 3' direction from the T7 promoter in the vector. Both clones were then digested with EcoRV. The pGEM-1212 and pGEM-897 clones were further digested with Sall and EcoRI, respectively, generating a fragment from pGEM-897 of approx. 900bp. The fragments were then treated with DNA Polymerase I, Large (Klenow) Fragment (Invitrogen) and purified. The 900bp fragment was ligated into pGEM-1212 and digested again with Sall. The resulting clone represents a protein-coding region that is 1945bp.

For the cloning of the *Arabidopsis* CstF50 homolog, AtCstF50 primers 5'FL/3'INT were used in a PCR reaction with first strand cDNA as template. A PCR product of expected size (1190 bp) was subcloned into pGEM-T easy vector and sequenced as mentioned above. This clone is base 1 to 1190 of the predicted 1290bp coding region. The 3' end was amplified in a continuous (1-step) reverse transcriptase RCR (RT-PCR) reaction as mentioned above using primers 5'INT/3'FL. A RT-PCR product of expected size (685 bp) was digested with BglII and EcoRI and then subcloned into BamHI and EcoRI digested pBLUESCRIPT-(KS) (Stratagene) using T4 DNA ligase

(Invitrogen) and resulting clone sequenced. The pGEM clone was then digested with SpeI and NotI restriction enzymes and the resulting fragment (approx. 1200 bp) from the digest was ligated into the pBLUE clone that had also been digested with SpeI and NotI. The resulting clone was then digested with BamHI and then religated together to yield a full-length protein coding region of 1290 bp.

For the cloning of the *Arabidopsis* CstF64 homolog, AtCstF64 primers 5'FL/3'FL were used in a PCR reaction with a 2-3 kb cDNA expression library for Arabidopsis (CD4-15; ABRC-DNA Stock Center; (Kieber et al., 1993) as template. The resulting PCR product of expected size (1386 bp) was subcloned into pGEM-T Easy vector and sequenced as previously mentioned above. This represents the full-length protein coding region of the cDNA for AtCstF64.

For the cloning of the *Arabidopsis* CstF77 homolog, AtCstF77 primers 5'FL/3'FL were used in a PCR reaction with first strand cDNA as template. PCR product of expected size (2320 bp) was subcloned into pGEM-T Easy vector and sequenced as previously mentioned above. This represents the full-length protein coding region of the cDNA for AtCstF77.

A cDNA for an *Arabidopsis* PAP from chromosome IV was also generated. The protein coding region was amplified and cloned into three different fragments, the NTD which spans bases 1 to 910, MR that spans bases 390 to 1500, and CTD which spans bases 1440 to 2403. Clones of MR and CTD fragments were gifts in the form of pGEM-T Easy vector clones from Lisa Meeks. The NTD fragment was amplified from first strand cDNA using AtPAPIV primers 5'FL/3'NT in a PCR reaction. PCR product of expected size (910 bp) was subcloned into SmaI digested pBLUESCRIPT-(KS) and sequenced. pBLUE-NTD (pB-N) and pGEM-MR clones were then digested with NheI and NotI restriction enzymes. A fragment of approx. 1 kB from pGEM-MR was ligated into the pB-N clone. The resulting pB-N-MR clone and pGEM-CTD were then digested with StuI and NotI restriction enzymes. A fragment of approx. 900 bp from the pGEM-CTD digest was then ligated into pB-N-MR yielding a full-length clone of 2403 bp.

Yeast Two-hybrid Assay

A Gal4-based two-hybrid system was used as described previously (James et al., 1996). The yeast strain used was PJ69-4 (*MATa, ade2, trp1- Δ 901, leu2-3, 112, his3-200, gal4901, leu2-3, 112, his3-200, gal4 Δ , gal80 Δ , ura3-52, met2::GAL7-lacZ, ADE2::Gal2-ADE2, LYS2::GAL1-HIS3*) and the expression vectors were pGAD-C(1) and pGBD-C(1) for activation domain (AD) and binding domain (BD), respectively.

AtFip1-V and the previously mentioned *Arabidopsis* polyadenylation homologs were cloned into AD and BD expression vectors using Gateway Δ cloning technology (Invitrogen). For this, the entire coding sequences of AtPfs2, AtCStf50, AtCstF64, AtCstf77 and AtPAPIV were amplified by PCR using 5'GW/3'GW primers listed in Table I. PCR products were mobilized into pDONR 201 vector to generate Entry clones by a recombination reaction via BP clonase (Invitrogen) per manufacture instructions. Also, an EST for AtCstF77 (EST Id: 297902; bases 1290-2320, aa 491-793) was amplified with 5'GW1/3'GW primers and an entry clone generated as described above. The AtFip1-V gene was also amplified with 5'GW/3'GW and 5'GW1/3'GW1 primers to generate two entry clones that span amino acids 1-137 (FipN) and 407-1192 (FipC) respectively. The yeast expression vectors, pGAD-C(1) and pGBD-C(1), were converted to Gateway compatible vectors (provided by Lisa Meeks and Dr. A.G. Hunt). The entry clones were then subcloned into the converted pGAD-C(1) and pGBD-C(1) vectors by a recombination reaction using LR clonase (Invitrogen). The resulting expression clones were sequenced to ensure that the gene fusions were in correct reading frame. The pGAD-C(1) and pGBD-C(1) clones of AtCPSF factors used in this experiment were supplied from Drs. Ruqiang Xu and Qingshun Quinn Li (Miami University, Oxford, OH).

Yeast cells were transformed with plasmid DNA using the PEG/Li Acetate method (Gietz et al., 1992). Two-hybrid analysis was carried out by plating yeast transformants on defined media containing sucrose as a carbon source, and lacking the nutritional supplements suited for selection of transformants (leucine and tryptophan) and for identification of interactions (adenine).

Production of Recombinant Proteins

The coding region of AtFip1-V (FipFL) was amplified by PCR using 5'GW/3'GW1 primers in Table I and subcloned into a series of Gateway vectors. As mentioned previously, the PCR fragment was subcloned into pDONR 201 to generate an entry clone and then into pDEST-17 (Invitrogen) using Gateway recombination technology. The AtPAPIV and AtFip1-V-FipFL (which is actually FipN) entry clones were also subcloned into pDEST-17. pDEST-17 is used to generate an N-terminal 6X-His-tagged fusion protein. The resulting recombinant plasmids were introduced into Rosetta(D3) cells (Novagen) for the production of protein.

Extracts containing the appropriate fusion protein were prepared after induction of Rosetta(D3) cells (Novagen) containing recombinant pDEST-17 constructs. Briefly, an overnight 10 ml culture of LB(+) (LB plus 100 µg/ml ampicillin and 25 µg/ml chloramphenicol) was used to inoculate 200 ml of LB(+) media and cells were grown at 37°C until an OD₆₀₀ of 1.0 to 1.2. Expression of the fusion protein genes was then induced by addition of 200 µl 1 M IPTG. After an additional growth of 2 hours at 37°C, cells were harvested and resuspended in 5 ml of binding buffer (20mM Tris-HCl, pH 7.9; 500mM NaCl, and 5mM imidazole). Cells were disrupted by sonication (3 bursts, 30 s each), debris was removed by centrifugation, and lysates passed through a nitrocellulose filter (0.45 µm pore size). The filtrate was then loaded onto a His-Bind resin column (Novagen) equilibrated with binding buffer. The matrix was washed with binding buffer containing 40 mM imidazole. Histidine-tagged proteins were then eluted with binding buffer containing 125mM imidazole. The eluted proteins were dialyzed against NEB (40 mM KCl, 25 mM HEPES-KOH, pH 7.9, 0.1 mM EDTA, 1 mM DTT, 1 mM PMSF, and 10% glycerol). Protein quantities were estimated by 10% SDS-PAGE and staining with Coomassie brilliant blue using bovine serum albumin as a standard.

The first 483 amino acids of the N-terminus of AtFip1-V (FipN; actually amino acids 1-137) and amino acids 407-1196 (FipC), were subcloned into pDEST-15 by recombination using GATEWAY technology. pDEST-15 is used to generate a N-terminal GST-tagged fusion protein. These recombinant plasmids were introduced into BL21-SI cells (Invitrogen) for production of protein.

Overnight cultures (10 ml LB media plus ampicillin; LB+A) of BL21-SI cells

containing recombinant pDEST-15 plasmids was used to inoculate 200 ml of LB+A media and cells were grown at 37°C until an OD₆₀₀ of 0.8. Expression of the fusion protein genes was then induced by addition of NaCl to a final concentration of 0.3 M. After an additional growth of 3 hours at 37°C, cells were harvested and resuspended in 5 ml of lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM PMSF). Cells were disrupted by sonication (3 bursts, 30 s each) and debris was removed by centrifugation. For subsequent stages, lysates were kept on ice. The lysate was then incubated for 1 hour with Glutathione-Sepharose beads equilibrated with lysis buffer with gentle agitation. After incubation, the glutathione-sepharose beads were pelleted by brief centrifugation, and washed twice with lysis buffer containing 2 M NaCl and finally two more times with lysis buffer alone. Proteins bound to the beads were eluted with glutathione elution buffer (20 mM reduced glutathione, 50 mM Tris-HCl, pH 8.0) and then dialyzed overnight with NEB.

FipC was further purified with poly(G) agarose. One hundred μ g of poly(G) agarose (Sigma) was washed three times with NEB and briefly pelleted by centrifugation. The pellet was then incubated with approximately 32 μ g of FipC at 30° C for twenty minutes. The agarose was then briefly pelleted and washed three times with NEB. Thirty μ l of SDS-sample buffer was added to the agarose, boiled for ten minutes, and then briefly pelleted. Twenty μ l of the sample was separated by SDS-PAGE.

Antibodies and Immunoblotting

Fip1 polyclonal antibody was prepared against GST-FipN fusion protein. At least 1 mg of fusion protein was sent to Cocalico Biologics, Inc. to inject rabbits and collect bleeds. Protein concentration was determined using the Bradford reagent (Bio-Rad Laboratories). In short, two rabbits (UKY 50 and UKY 51) were injected with antigen (GST-FipN). Each rabbit was given three booster injections at day 14, 21, and 49. FipN-specific antibodies were purified using an affinity-purification procedure as described (Sambrook et al., 1989). Briefly, antigen (GST-FipN) was resolved by 10 % SDS-PAGE, transferred to nitrocellulose (Millipore), GST-FipN strips incubated with bleeds, and then specific antibodies eluted off of nitrocellulose strips.

For immunoblots, proteins were separated by SDS-PAGE and transferred to nitrocellulose membrane using a Trans-Blot Cell (Bio-Rad Laboratories) following the manufactures recommendations. Filters were washed and probed with antibodies at dilutions of 1/2000 unless otherwise noted. Anti-GST alkaline phosphatase conjugated antibody was purchased from Sigma and Anti-His antibody was purchased from Invitrogen.

Poly(A) polymerase Assay

Recombinant AtPAPIV and AtFip1 were assayed in NEB at a volume of 30 μ l, plus 4 μ l of 15 mM MgCl₂ and 6 μ l of reaction mix (15 μ l polyvinyl alcohol, 15 μ l PAP RXN Mix (167 mM Tris-HCl, pH 8.0 267 mM KCL, 0.33 mM EDTA, 3.33 mM DTT, 0.67% nonidet-P40) 6 μ l 20 mM ATP, and 60 units of RNase Inhibitor (Eppendorf), final volume 50 μ l with NEB). When needed RNA substrate, poly(A) of a length of 14 bases (Dharmacon RNA Technologies), was added as indicated, 10 μ Ci of [α -P³²]-ATP (3000Ci/mmol) were added, to give a final ATP concentration of 0.36 mM and 0.69 μ Ci/nmol. Reactions were incubated at 30°C for the times indicated. Amount of protein and RNA used in each reaction is mentioned in the figure legends. Reactions were terminated by incubating with 2.5 μ l of STOP solution (2.5% SDS, 135mM EDTA, 5mg/ml Proteinase K), for 10 min at 37°C. Fifteen μ l of Tris-HCl, pH 8.0 was then added to the terminated reaction. Equal volume of Phenol:Chloroform:IAA (AMBION) was then added and centrifuged for 5 min. The aqueous phase was then precipitated in 340 μ l of 100% ethanol with 40 μ g glycogen as a bulk carrier at. RNA was pelleted by centrifugation at full speed for 10 min (Eppendorf; Centrifuge Model 5415C). The pellet was washed with 70% ethanol and dried under vacuum for 10 min. The pellet was resuspended in 2-4 μ l Gel Loading Buffer II (AMBION) and denatured at 70°C for 10 min and then stored on ice until loading on a gel. Samples were loaded onto a 7.6% Acrylamide; 8.3 M Urea gel cast in 1X TBE. Gels were pre-run at constant current of 15 mAMPs for 45 min. Samples were resolved at a constant current of 15 mAMPs until the dye front reached the bottom of the gel. Gels were transferred to Whatmann paper (Fisher Scientific) and dried under vacuum at 80°C for 2 hours. Dried gels were developed on a storage phosphor screen and visualized by ImageQuant (Amersham Biosciences). With

the use of ImageQuant, the top half of each lane was boxed and analyzed. ImageQuant generates volume values, which we termed arbitrary values, subtracting time zero as background from each value and the results plotted.

Electrophoretic Mobility Shift Assays (EMSAs)

For EMSAs, RNA and proteins were incubated at 30°C for 20 min in NEB supplemented with KCL (final concentration 60 mM) and MgCl₂ (final concentration 1.2 mM). Gel Loading Buffer II (AMBION) diluted in NEB (1:40; GLBII to NEB) was added to the reactions at one tenth the reaction volume. Samples were analyzed on 4% non-denaturing gels (2 g Acrylamide, 0.04 g Bis-Acrylamide, 50 ml 1X TBE, 800 μ l 15% APS, 100 μ l TEMED). Gels were transferred to Whatmann paper and dried under vacuum at 80°C for 2 hours. Dried gels were developed on a storage phosphor screen and visualized by ImageQuant. Bound and unbound RNA substrate in each lane were boxed and analyzed by ImageQuant. The volume values generated were corrected by subtracting background, and the results were plotted accordingly.

Wild type and mutant derivatives of the rbcS-E9 and CaMV polyadenylation signals were made from pBLUESCRIPT derivatives (Mogen et al., 1990; Mogen et al., 1992). The length and position of the mutant RNAs is described in the text. DNA templates for these wild type and mutant RNAs were used in an *in vitro* transcription reaction. Uniformly labeled RNA substrates for EMSAs, were generated with T7 RNA polymerase of the AMPLISCRIBE or RIBOSCRIBE transcription kit (Epicentre) with [γ -P³²]-ATP. Cold competitors were generated with T7 RNA polymerase of the AMPLISCRIBE transcription kit. RNA was quantitated as mentioned previously.

RNA homopolymers (poly(A), poly(C), poly(G), poly(U)) were purchased from Sigma and yeast tRNA from Invitrogen.

Pea Nuclei Isolation and Nuclear Protein Extraction and Arabidopsis plant extracts

Pea nuclei and nuclear extracts were prepared as described in chapter 2. *Arabidopsis* plant material was ground up in liquid nitrogen and then boiled in SDS-Sample buffer.

Table I. List of oligonucleotides used for PCR and sequencing

Gene	Primer designation	Sequence (5'-3')
AtFip1-V	5'FL 5'INT 5'INT 3'INT 3'INT1 3'INT2 3'FL 5'GW 5'GW1 3'GW 3'GW1	CCGCATGGAAGAGGACGATGAGTTCGGA CCCGGATCCGAGTTAGCTGCAGCAACAGGGGCA CCCAGATCTGGTCCGAAGATCGATCATCAAGG GCGAATTCACCCGAGGGTTCATCCTCATG CCCGAATTCTTGATGATCGATCTTCGGAACCTC CTAGTTTTGAGGAAATGGATGATG TTATGCGTATCCCTCCCTATTCTTACACA GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATG GAAGAGGACGATGAGTTC GGGGACAAGTTTGTACAAAAAAGCAGGCTTCGAG TTAGCTGCAGCA GGGGACCACTTTGTACAAGAAAGCTGGGTATCAAC CCGAGGGTTCATCCTC GGGGACCACTTTGTACAAGAAAGCTGGGTATTATG CGTATCCCTCCCTATTCTTACACA
AtFip1-III	5'FL 5'INT 3'INT1 3'INT2 3'FL	ATGGATTCTACTGATGATGATTTCGGT AAAGAAGAGGATGATCGATAT TATGCATCTTGCTTTCTCAGGTGTCCT TTCATATCTATCACGATGCAA TTAACTTGCACACCATCTCCT
AtCstF50	5'FL 5'INT 3'INT 3'FL 5'GW 3'GW	CGCGAATTCATGGGGAATAGTGGAGATTTG CCAGATCTTTCTTCGACTTCTCCAAAACCACGGCT GTTATGGTTAGAAGGCCACTTTGCCACTTT CCGGAATTCCTAAACGGATTCCTTCCAGAACCGAAT GGGGACAAGTTTGTACAAAAAAGCAGGCTGGATGG GGAATAGTGGAGATT GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAA CGGATTCCTTCCAGAA
AtCstF64	5'FL 3'FL 5'GW 3'GW	GGAGATCTGCCATGGCCATGGCTTCATCATCATCC CCAGATCTATCGATTGAAGGCTGCATCATGTGG GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGC TTCATCATCATCCCA GGGGACCACTTTGTACAAGAAAGCTGGGTAGTGAAG GCTGCATCATGTGGT
AtCstF77	5'FL 3'INT 3'FL 5'GW 5'GW1 3'GW	ATGGCTGATAAGTACATCGTCGAG TTCCAGAAAGTGCTTCTTTCATTC TTAGCCAGTGCTACCAGAAAGCTCGCCAGA GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGC TGATAAGTACATCGTC GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTTAAG CACGTTACCAGTTGA GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGCC AGTGCTACCAGAAAG
AtPfs2	5'FL 5'INT 3'INT 3'FL 5'GW	ATGTACGCCGCGGCGATATGCACAGG AGTGTTTGGGATCTTGCATGGCATCCT AAGAACATCTCGGGGATTATCTGCAGG CTACTGATGTTGCTGATIGTTGTTTGG GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGTA CGCCGGCGGCGATATG

Table I (continued).

AtPfs2 (continued)	3'GW	GGGGACCACTTTGTACAAGAAAGCTGGGTACTACTG ATGTTGCTGATTGTT
AtPAP	5'FL	CCGAGATCTATGGTGGGTACTCAAAAATTTAGGTGGT
	3'INT	CCGAGATCTACGGTCCCAGACAGGAAATCCAAGCTC
	5'GW	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGT GGTACTCAAAAATTT AGGTGGT
	3'GW	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGCT CTGTCTCCGACTTCTCCATC

A

MEEDDEFGLYSDVLQPFQPPVVLPPPPPLPHRSIDLNLRSQDQDVSEPNSAPISRVSNDNAVKLSTQDA 70
TRQAIVDGGGDDKMSFDIEEPDADSTPTIPGLFVTGALPGLATDRGVSQVTTTRIEQQVGGGGDGGYGGQ 140
GEGDDWDSSEDDLQIVLNDSSRNVMIGGADRRSRMGDNEDDDDEDEDPLVIVADTDPNQPMEEQMWGE 210
DGLQGIIEGDGKDGGEAGKGSPPGGATGPPKAGYSSHGYPFHSSQFKYVRPGAAPIPGGAASVGGPSSGQV 280
RPPANLGP MAGRGRGDWRPLGMRNASAAQKGFHQPWGSNTAGRGLDFTLPSHKTIFEV³³⁶IDS³³⁷FEEK³³⁸PWRY³³⁹ 350
PGVEMTDYFNFG³⁴⁰LNEES³⁴¹WKDY³⁴²CKQLD³⁴³QHRIQ³⁴⁴TTMQ³⁴⁵SRIRVYESGR³⁴⁶TDQ³⁴⁷GYD³⁴⁸PD³⁴⁹LPPELAAATGAQGV³⁵⁰PVD 420
SSNLVKPDSVQGDSAKVPANVRPTLPPGRPIPVETGSGERLPSIDTRAPMRDLDAIIEIVCQDSHEDEP 490
SGENGTDAQDSSSLPGENVPVETSIVNNKRPDTESAEHSPAQDEPHKNLLKKQDDEISRSTDSGQSFSSS 560
PVGDRGTRSSSV⁵⁶¹DREDV⁵⁶²GGEAG⁵⁶³KDAEM⁵⁶⁴GEEL⁵⁶⁵KMS⁵⁶⁶FTSP⁵⁶⁷QSAV⁵⁶⁸QEDD⁵⁶⁹GGES⁵⁷⁰KTERS⁵⁷¹SESS⁵⁷²KARS⁵⁷³GS⁵⁷⁴HR⁵⁷⁵DF⁵⁷⁶Q 630
QEEDVIQDKHSSRPANNRKQYDNNAPHQSRKNQDRGKEMERTRAASKGGRENSNPHMELDSTYIYSIASR 700
EDFDKRKERDVDGAVWRRKEDDPYSRRGGDEGSRKRDREDDPGFRORGKMRNEIRSKDDQVPSRKHMDD 770
AGMRNIYEPDDHINKRRKDEEYLRRSRPEKNEISYGQRESMSRVKREDDRLEHQKRDVQHKIRDDFDH 840
GSLRQRDDIYMQRDGNERLRERDVLDKLKLPHEDGISARGRERQVAVRGHRGSEDRSSRMKDEYKASDKE 910
HVTKDTLRHAKQTKRRDYPGESSSHRHGHEDFSARTDNIVNNEKKPRQERTGAKIDKFIDTLDGQRLQD 980
RKHKDSRRKIKEQREGTESLSKQGEQNGSSVVTGSKGTNDARNCRSEIPHQPNTAKRHKENASSGDEIHD 1050
SKRGRTKLERWASHKEREDAVSAKSSSIS¹⁰⁵¹SKLEEK¹⁰⁵²ENNT¹⁰⁵³NGRL¹⁰⁵⁴SEPV¹⁰⁵⁵HGS¹⁰⁵⁶IGKSR¹⁰⁵⁷DV¹⁰⁵⁸TEEK¹⁰⁵⁹IGH¹⁰⁶⁰DLADTK 1120
DGSEKGPGRHLDTVEK¹¹²¹LK¹¹²²RSER¹¹²³FK¹¹²⁴LP¹¹²⁵M¹¹²⁶TEK¹¹²⁷DT¹¹²⁸GV¹¹²⁹KK¹¹³⁰ME¹¹³¹SET¹¹³²LPSA¹¹³³KIE¹¹³⁴GP¹¹³⁵VD¹¹³⁶SE¹¹³⁷GE¹¹³⁸YV¹¹³⁹W¹¹⁴⁰DERS¹¹⁴¹CVR 1190
IGREYA 1196

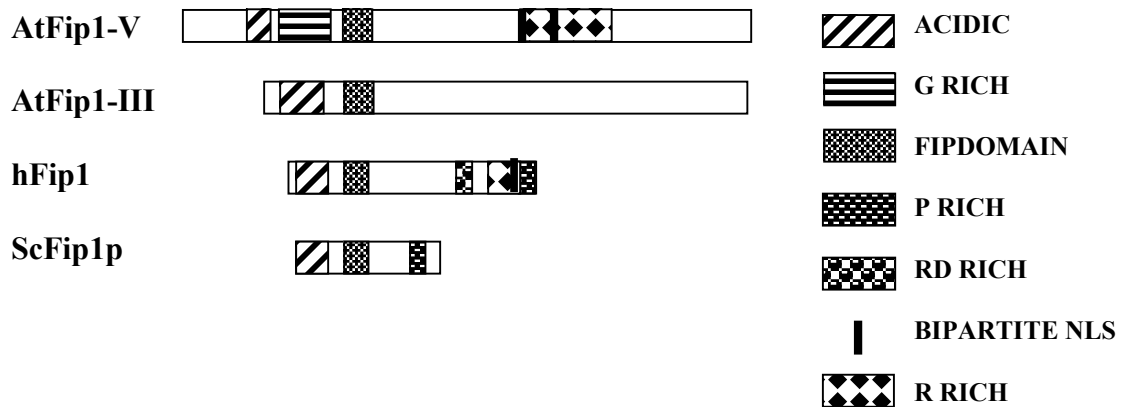
B

MDSTDDDFGLYVDDAKFQATDAFESECATNSGEDKGFEE¹⁴⁰TVKSDSEGEVKKFDV¹⁴¹VAKDSSPCDDDDCAM 70
NLTEADEESE¹⁴²FS¹⁴³SDS¹⁴⁴DDL¹⁴⁵NIVL¹⁴⁶KDD¹⁴⁷SKAL¹⁴⁸PASC¹⁴⁹V¹⁵⁰NTN¹⁵¹FG¹⁵²GYE¹⁵³ASK¹⁵⁴ASS¹⁵⁵F¹⁵⁶QRR¹⁵⁷WTR¹⁵⁸NAS¹⁵⁹ANN¹⁶⁰ACID¹⁶¹PSL 140
GMSQYRYSF¹⁶²PN¹⁶³WS¹⁶⁴RTP¹⁶⁵FD¹⁶⁶VN¹⁶⁷LD¹⁶⁸VLE¹⁶⁹KK¹⁷⁰PWR¹⁷¹DP¹⁷²GT¹⁷³DT¹⁷⁴SD¹⁷⁵FF¹⁷⁶NF¹⁷⁷GL¹⁷⁸NE¹⁷⁹Q¹⁸⁰SW¹⁸¹KDY¹⁸²CK¹⁸³PL¹⁸⁴GRA¹⁸⁵IE¹⁸⁶VR¹⁸⁷GG¹⁸⁸TLE¹⁸⁹R 210
IP²¹⁰SAD²¹¹LRR²¹²PR²¹³DP²¹⁴PG²¹⁵VVI²¹⁶QIP²¹⁷VT²¹⁸ND²¹⁹VE²²⁰EL²²¹PV²²²RT²²³PE²²⁴KARC²²⁵ITS²²⁶NE²²⁷AS²²⁸RS²²⁹DV²³⁰SH²³¹SY²³²GK²³³DL²³⁴NS²³⁵VY²³⁶GS²³⁷PK²³⁸DEA 280
FV²⁸⁰GQ²⁸¹EE²⁸²NAG²⁸³SF²⁸⁴S²⁸⁵GE²⁸⁶KS²⁸⁷LPT²⁸⁸EN²⁸⁹CC²⁹⁰S²⁹¹REAT²⁹²PS²⁹³DK²⁹⁴EM²⁹⁵LE²⁹⁶KE²⁹⁷KE²⁹⁸ES²⁹⁹VC³⁰⁰NS³⁰¹DE³⁰²TD³⁰³PSS³⁰⁴VER³⁰⁵ESS³⁰⁶LG³⁰⁷DR³⁰⁸IR³⁰⁹LS³¹⁰P³¹¹T 350
SS³¹⁰SV³¹¹GINE³¹²ES³¹³DD³¹⁴YET³¹⁵ES³¹⁶LK³¹⁷DS³¹⁸AT³¹⁹DD³²⁰Q³²¹REV³²²ST³²³PP³²⁴Q³²⁵EAR³²⁶LAE³²⁷HEA³²⁸IS³²⁹IK³³⁰REG³³¹DS³³²GM³³³HS³³⁴RR³³⁵SH³³⁶ED³³⁷SS³³⁸KR 420
HC⁴²⁰GR⁴²¹AG⁴²²YAR⁴²³YV⁴²⁴KD⁴²⁵AS⁴²⁶PT⁴²⁷PD⁴²⁸GR⁴²⁹GK⁴³⁰VG⁴³¹SL⁴³²QGL⁴³³YR⁴³⁴DS⁴³⁵NKN⁴³⁶W⁴³⁷Q⁴³⁸NG⁴³⁹PP⁴⁴⁰IT⁴⁴¹LER⁴⁴²DE⁴⁴³TE⁴⁴⁴GK⁴⁴⁵GV⁴⁴⁶HYY⁴⁴⁷RE⁴⁴⁸KSH⁴⁴⁹GR⁴⁵⁰LN 490
SS⁴⁹⁰VD⁴⁹¹HD⁴⁹²HR⁴⁹³HR⁴⁹⁴HF⁴⁹⁵GWR⁴⁹⁶NN⁴⁹⁷KES⁴⁹⁸SL⁴⁹⁹GR⁵⁰⁰GF⁵⁰¹DHS⁵⁰²NS⁵⁰³YK⁵⁰⁴CG⁵⁰⁵TH⁵⁰⁶LKEY⁵⁰⁷TS⁵⁰⁸RS⁵⁰⁹SS⁵¹⁰FD⁵¹¹LN⁵¹²QR⁵¹³NS⁵¹⁴RS⁵¹⁵SS⁵¹⁶FK⁵¹⁷EED⁵¹⁸DRY⁵¹⁹G⁵²⁰WHH 560
RER⁵⁶⁰KY⁵⁶¹VHER⁵⁶²SPI⁵⁶³RAY⁵⁶⁴ENY⁵⁶⁵KERN⁵⁶⁶GCD⁵⁶⁷WL⁵⁶⁸RE⁵⁶⁹YY⁵⁷⁰ED⁵⁷¹CIP⁵⁷²IT⁵⁷³DM⁵⁷⁴YR⁵⁷⁵YR⁵⁷⁶SENS⁵⁷⁷SA⁵⁷⁸HAI⁵⁷⁹HN⁵⁸⁰LKH⁵⁸¹SP⁵⁸²END⁵⁸³LY⁵⁸⁴CRR 630
RGGYDYNLHRDRYEDGVHRVESRIPFELAY 660

Figure 3.1. Amino acid sequences of *Arabidopsis thaliana* Fip1 proteins.

- (A) Amino acid sequence of AtFip1 that resides on chromosome V (AtFip1-V; At5g58040). The conserved Fip1 domain (amino acids 336-398) is framed and predicted bipartite nuclear localization signals (nls) are underlined. Sequence deduced by RT-PCR.
- (B) Amino acid sequence of AtFip1 that resides on chromosome III (AtFip1-III; At3g66652). The conserved Fip1 domain (amino acids 140-209) is framed. Sequence deduced by RT-PCR. This sequence represents 67% of the predicted coding region for AtFip1 that resides on chromosome III.

A



B

```

S. cerevisiae ...mtgsDldkegifdsvglttIdpavlkeKPKWRGPGanLSDYFNYGfNEfWnEYLlrqeAlqqdynprri
H. sapiens ...kvkgVDldaPgsingvp_lEVDLdsfEDKPKWRGPGaDLSDYFNYGfNEdWkaYcekqkRirmglevipv
A. thaliana(III) lgmsqyrysfpnFwsrt...pfDVNLDVLEKPKWRGPGGdLSDYFNYGfNEqSWKDYCNplgRaievrggtle
A. thaliana(V) sntagrgDftlPshkt...lfeVDLdsfEEKPKWRYPGVEMTDYFNfGfNEeSWKDYCNqldqhriqttmqsr
  
```

Figure 3.2. Alignment of Peptide Domains Found in Fip1 Proteins.

(A) Schematic alignment of AtFip1-V, AtFip1-III, human Fip1 (*H. sapiens*; hFip1), and *S. cerevisiae* Fip1p (Swissprot or Trembl accession numbers: tr|Q9H077 and sp|P45976 respectively). All four were aligned according to the conserved Fip1 domain. The patterns used for acidic, the glycine-rich domain, the conserved Fip1 domain, the proline-rich domain, the mixed charged domain (RD), the predicted bipartite nls, and the arginine-rich regions are depicted on the right. The AtFip1 motifs were determined by PROSITE (Motif-Scan; Falquet et al., 2002) and the hFip1 and ScFip1 were adapted from Kaufmann et al. (2004).

(B) Alignment of the conserved Fip1 domain of the two *Arabidopsis* Fip1s with the human Fip1 and the yeast Fip1p. Homology was determined with ClustalW and the results were formatted using MacBoxshade. Amino acids that are identical in all four are shaded black, with white lettering. Positions that are similar are shaded gray, with white lettering.

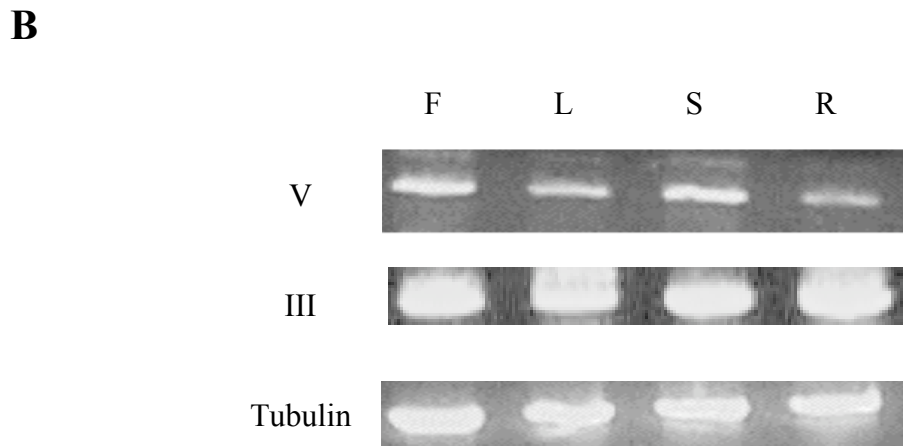
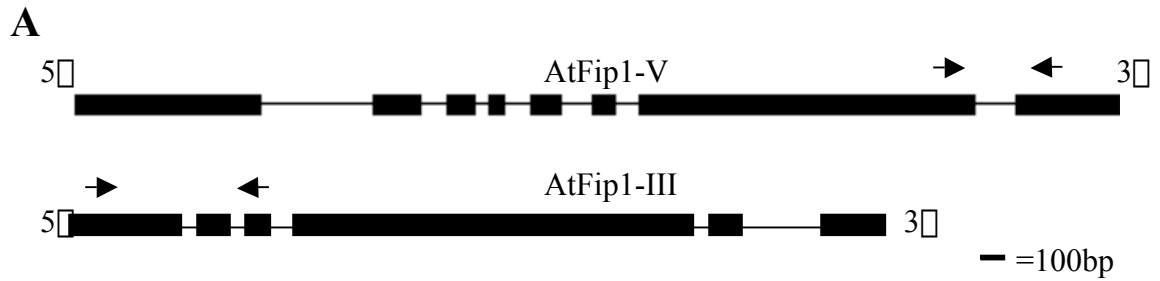


Figure 3.3. AtFip1 Genes are Expressed in all *Arabidopsis* Tissues.

(A) Intron-exon maps of the AtFip1 gene that resides on chromosome V and III. Lines represent introns and bold lines indicate exons. Arrows indicate location of oligonucleotide primers used in (B). III is the predicted intron and exon map

(B) RT-PCR analysis of AtFip1-V expression. RNA isolated from the indicated tissue (labeled above: F-flower, L-leaf, S-stem, R-root) was analyzed by RT-PCR, with primers shown in (A) that flank the last intron of chromosome V Fip1 and the first two introns of chromosome III Fip1. For comparison, primers specific for *Arabidopsis* tubulin gene (At5g62690) were used. When reverse transcriptase was omitted from these reactions, no amplification products were made (data not shown).

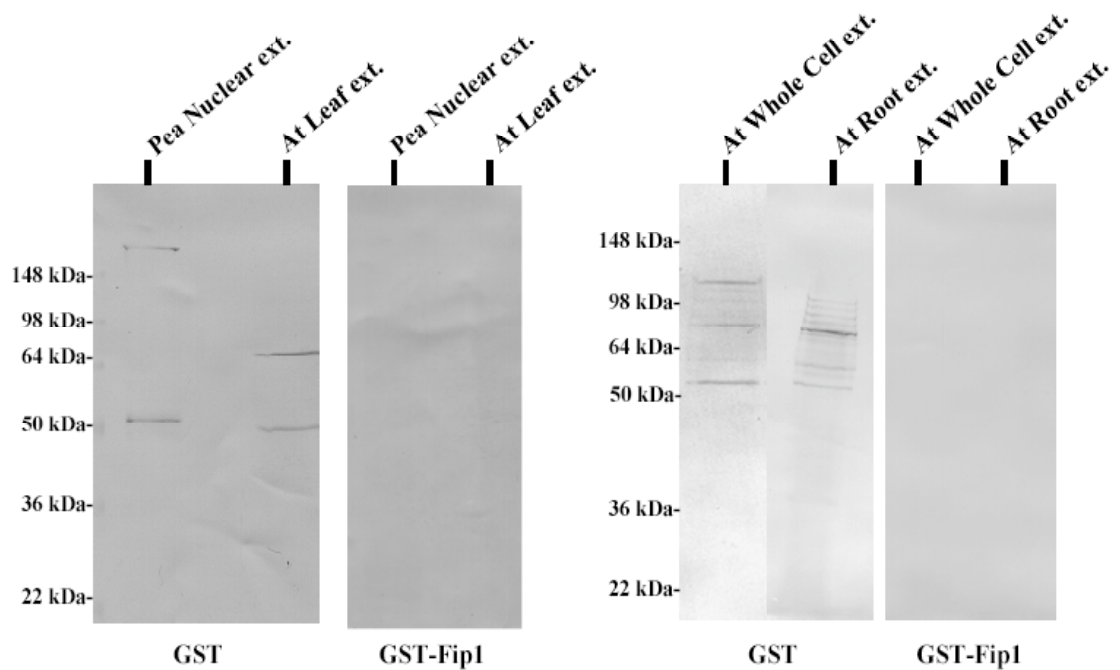


Figure 3.4. Polypeptides Related to the *Arabidopsis* Fip1 are Present in Plant Extracts.

Pea nuclear extracts, *Arabidopsis* whole cell (plant), leaf and root extracts were separated by SDS-PAGE, transferred to nitrocellulose, and probed with anti-AtFip1 (1:2000 dilution). Extracts are listed above each panel. At whole cell extract equals whole plant (all tissues mixed together). Antibodies were pre-incubated with recombinant GST or GST-Fip1 fusion protein, listed below each panel. The location of prestained size standards are shown to the left of the panels.

		Binding domain fusion (pGBD)									
		empty	CstF 50	CstF 64	CstF 77C	Pfs2 FY	CPSF 30	CPSF 73(I)	CPSF 73(II)	CPSF 100	CPSF 160
Activation domain fusion (pGAD)	empty	-	-	-	-	-	+	-	-	-	-
	FipN	-	-	-	+	-	+	-	-	-	-
	FipC	-	-	-	-	-	ND	-	-	-	-
	CstF 77FL	-	-	ND	ND	-	ND	ND	ND	ND	-
	CstF 64	-	-	ND	+	-					

		Binding domain fusion (pGBD)	
		empty	FipN
Activation domain fusion (pGAD)	empty	-	-
	CPSF 30	-	+
	PAP (IV)	-	+

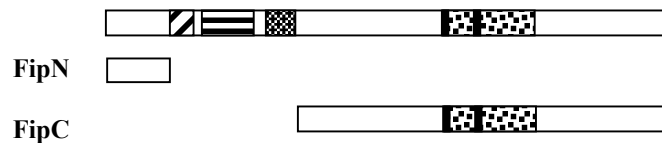
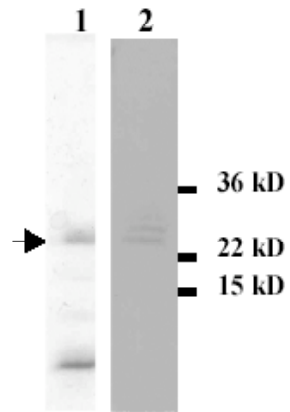


Figure 3.5. Qualitative Interactions Between AtFip1-V and Other *Arabidopsis* Processing Homologs in yeast.

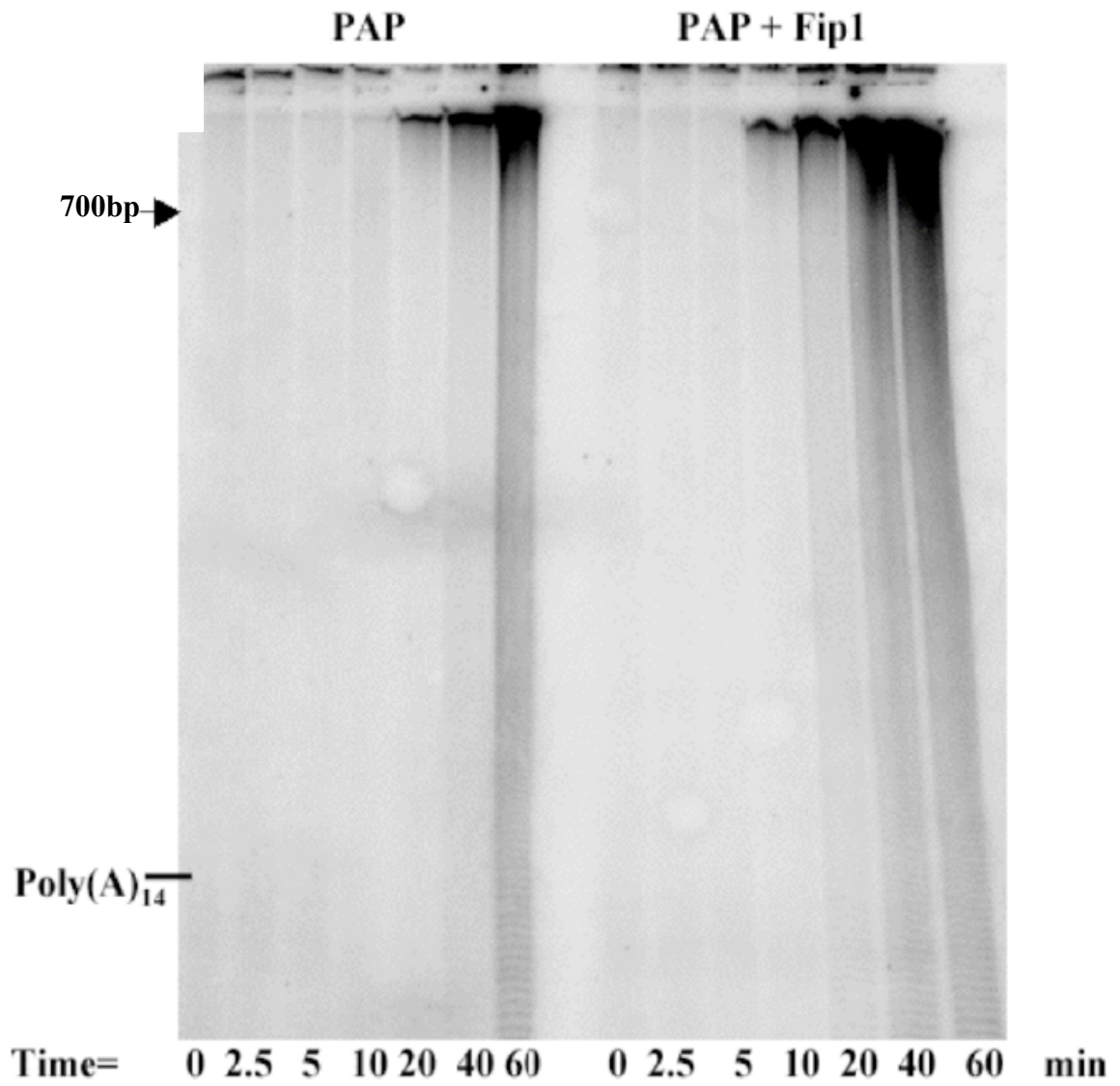
Yeast two-hybrid assays of AtFip1-V with *Arabidopsis* homologs to CstF50 (At5g60940), CstF64 (At1g71800), CstF77C (At1g17760; amino acids 491-793), Pfs2 (At5g13480), CPSF30 (At1g30460), CPSF73(I) (At1g61010), CPSF73(II) (At2g01730), CPSF100 (At5g23880), CPSF 160 (At5g51660), mCFI-25 (At4g29820), PAB2 (At5g51120), and PAP(IV) (At4g32850). Detection of pair-wise protein-protein interactions was determined by growth of the yeast transformed with the indicated plasmids on selection plates, minus adenine: +, growth; - no growth; ND, not determined. Positive (+) for the interaction, means that the number of colonies growing after five days on ALW selective media were between 50 and 200% of those seen on LW media after 3 days. Negative (-) for the interaction, means that the number of colonies growing after

five days on ALW selective media were less than 10% of those seen on LW media after 3 days. Empty vectors were used to rule out self-activation of reporter gene. The AtFip1 fragments used in this experiment are depicted schematically below the full-length protein. FipN spans amino acids 1-137 and FipC amino acids 407–1196. To simplify, the prefix At was omitted.

A



B



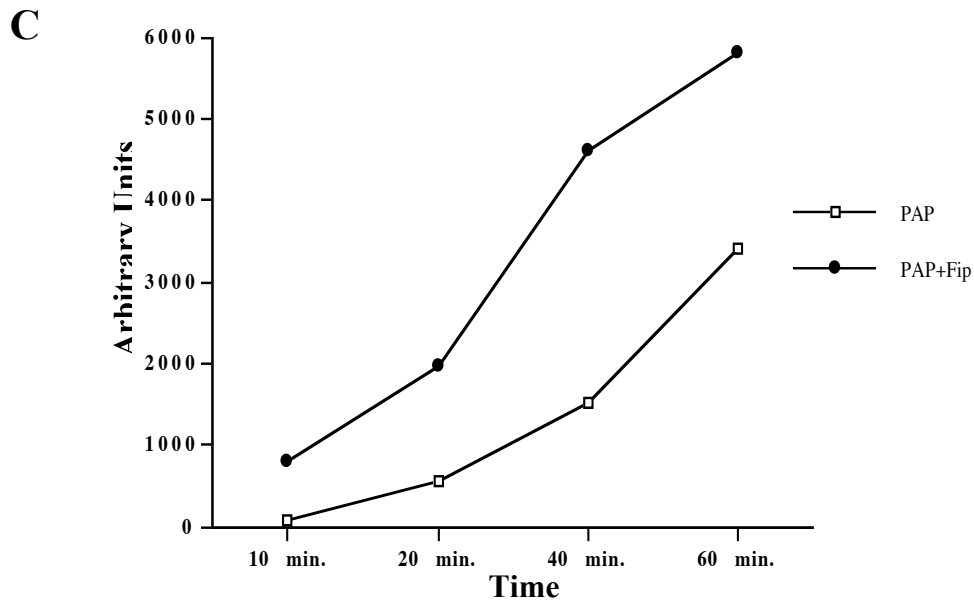
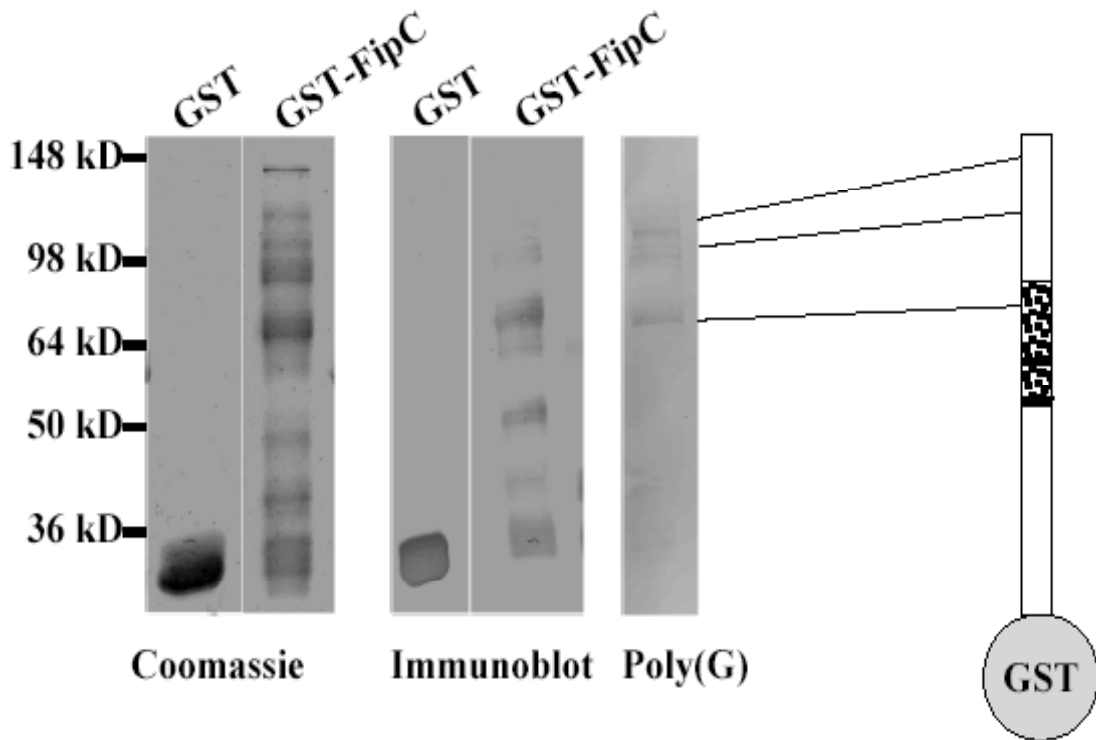


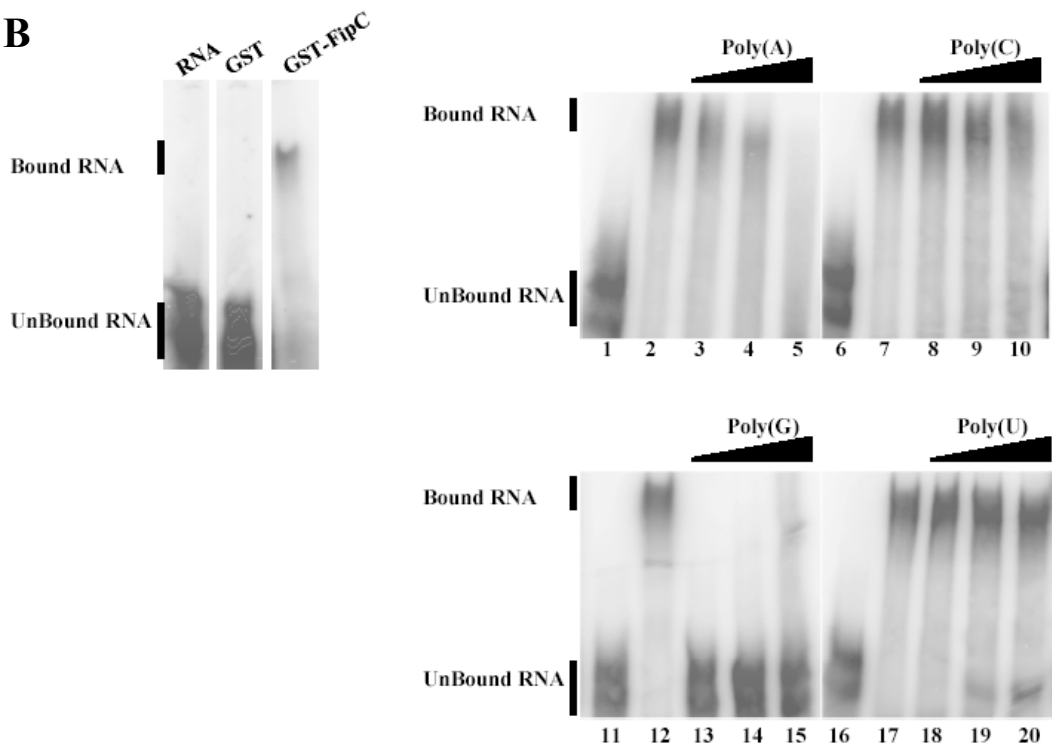
Figure 3.6. AtFip1 Stimulates the Non-Specific Activity of AtPAP(IV).

- (A)** Recombinant His₆-tagged AtFip1-V (FipN; amino acids 1-137) was separated by SDS-PAGE, stained with Coomassie brilliant blue (lane 1) and probed with anti-His after transfer to nitrocellulose (lane 2). Arrow on the left indicates the recombinant His₆-AtFip1-V (FipN). Protein size standards are indicated on the right.
- (B)** Poly(A), of a length of 14 bases (220pmol), was incubated with 10 pmol of His₆-AtPAP(IV) either alone or together with 6 pmol of His₆-AtFip1-V (FipN) in PAP assays (see Materials and Methods). At the time points indicated below the panels, reactions were stopped and RNA precipitated (see Materials and Methods). The RNA was analyzed on a denaturing 7.6% polyacrylamide gel and visualized by autoradiography. The poly(A) template is indicated on the left, as well as the approximate location of the 700bp DNA marker.
- (C)** The autoradiograph was quantitated for the reaction products of the 0, 10, 20, 40, and 60 minute time point for PAP and PAP plus Fip lanes by ImageQuant Software. The results were plotted as arbitrary values against time.

A



B



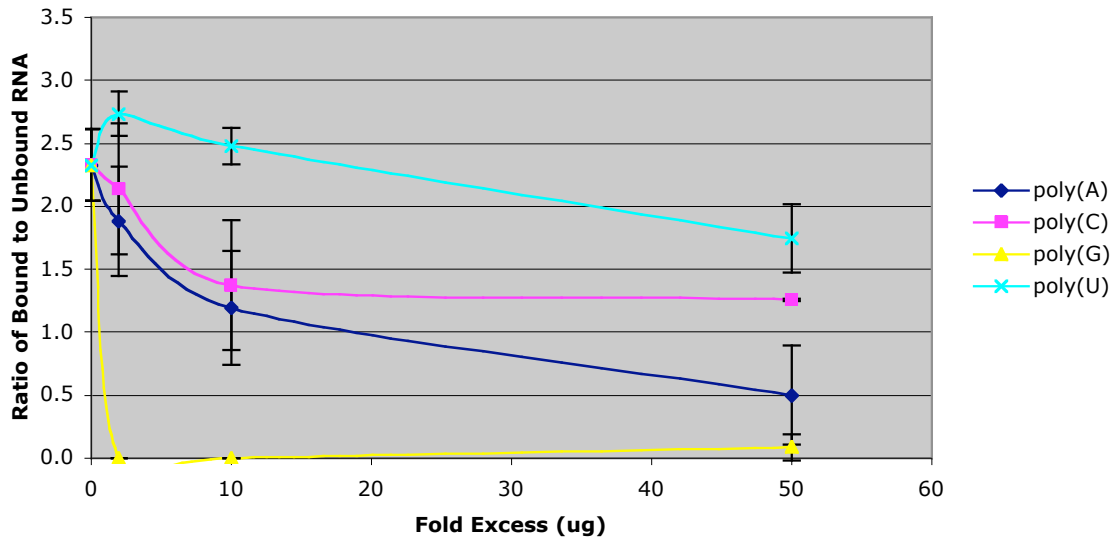
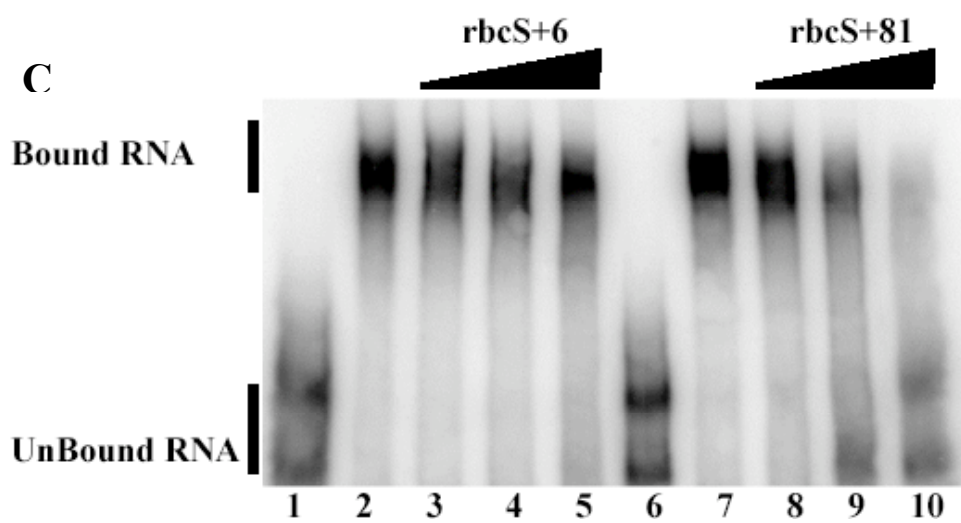
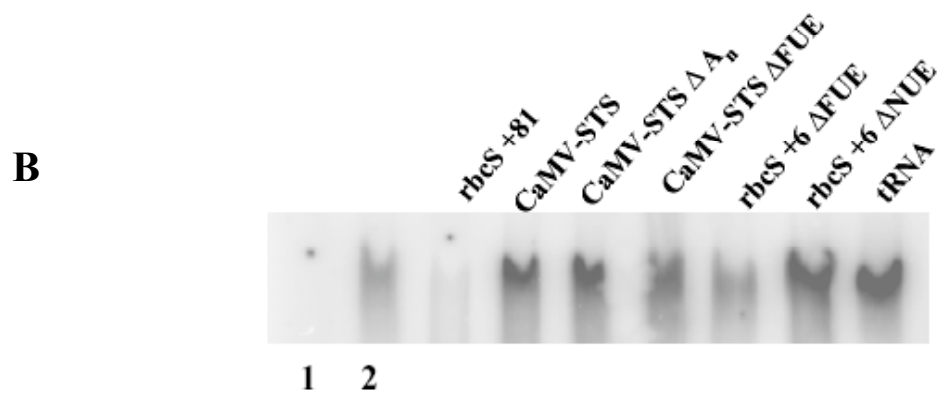
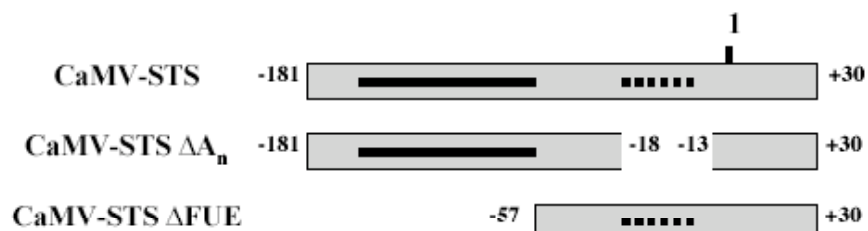
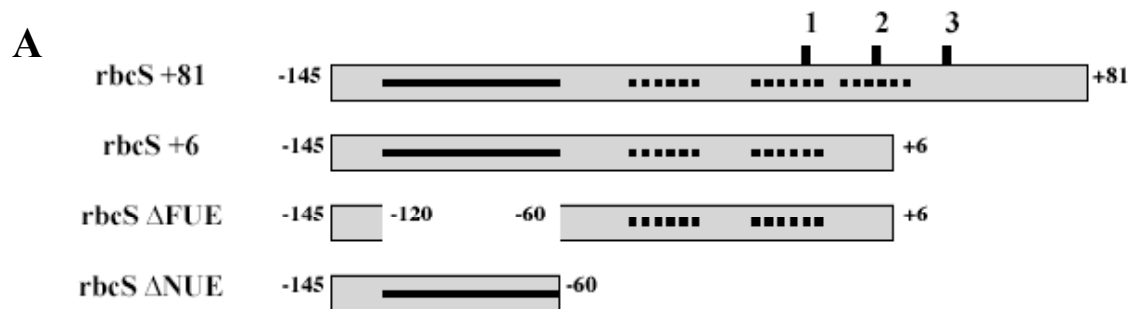
C

Figure 3.7. The C-terminal Domain of AtFip1 (FipC) Binds RNA.

- (A)** GST and N-terminal GST-tagged FipC fusion proteins (aa 407–1196) were separated by SDS-PAGE, stained with Coomassie brilliant blue (left panel) and probed with anti-GST after transfer to nitrocellulose (middle panel). GST-FipC was incubated with poly(G) agarose, extensively washed, boiled in SDS-sample buffer, and separated by SDS-PAGE and probed with anti-GST after transfer to nitrocellulose (right panel). Protein size standards are indicated on the left. The schematic on the right represents GST-FipC. The lines indicate approximately how much of FipC these fragments contain.
- (B)** Ninety-seven ng of RNA substrate (761.9 μ Ci/nmol) containing the polyadenylation signal of the *rbcS-E9* gene (-145 to +81), uniformly labeled, was incubated with GST (4 μ g) and GST-FipC (4 μ g). The reactions were analyzed on a native 4% polyacrylamide gel (left panels). Competition EMSA experiments with 1.25 μ g GST-FipC, 3.92 μ g (lanes 3, 8, 13, and 18), 19.6 μ g (lanes 4, 9, 14, and 19), or 98 μ g (lanes 5, 10, 15, and 20) of competitor RNA (right panels). Lanes 2, 7, 12, and 17

contained no competitor. RNA substrate alone (1.96 μ g; 11.2 μ Ci/nmol; lanes 1, 6, 11, and 16). Bound and unbound RNA substrate is indicated on the left.

(C) Bound and unbound RNA substrate in the autoradiographs were quantitated by ImageQuant Software. The results were plotted as a ratio of bound to unbound RNA substrate against the fold excess (by weight) of competitor RNA.



D

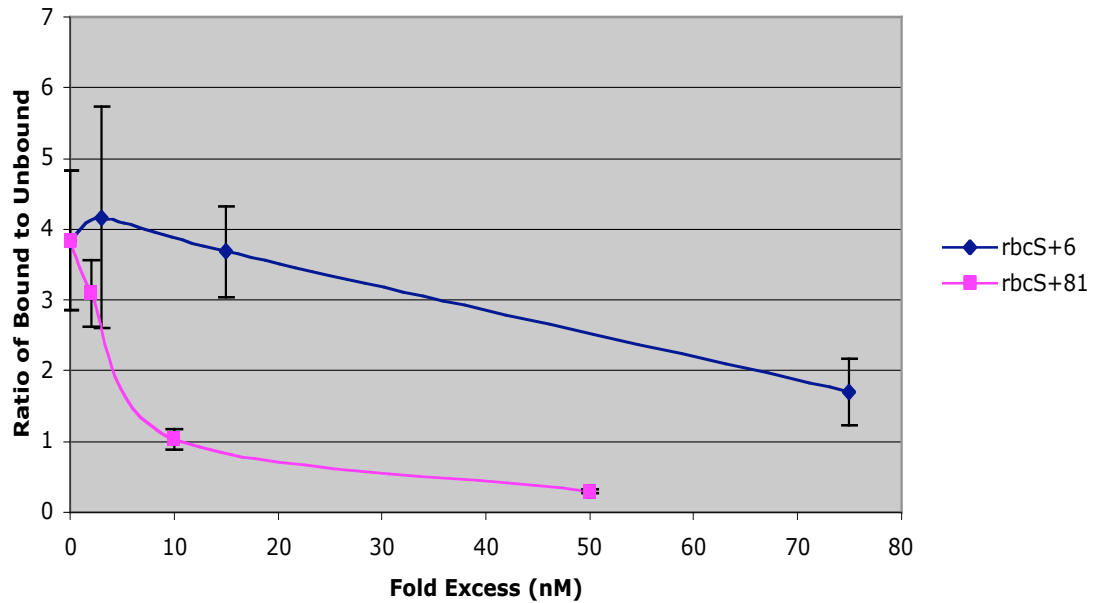


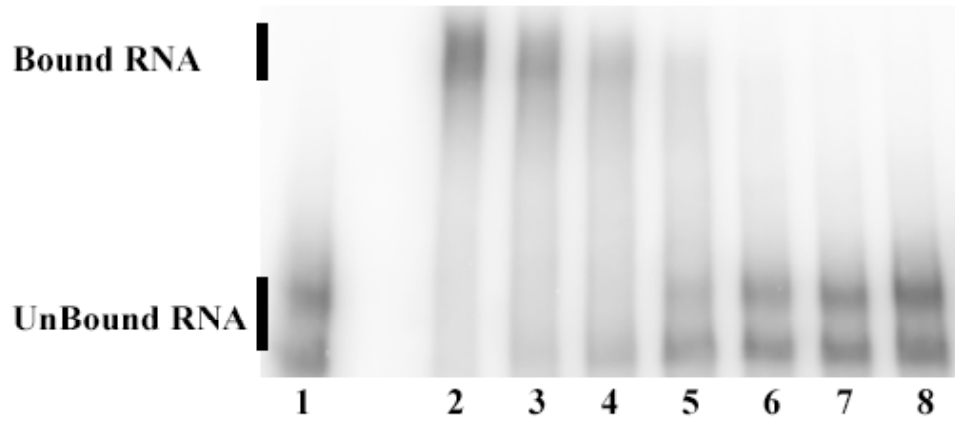
Figure 3.8. The C-terminal Domain of AtFip1 (FipC) Binds Polyadenylation Sites.

- (A) The structure of the polyadenylation signals of the *rbcS*-E9 gene and cauliflower mosaic virus (CaMV) and the mutant variants used in competition EMSA experiments in B. The relative positions of far upstream elements (FUEs; solid line), near upstream elements (NUEs; dashed line), and polyadenylation sites (1,2,3). Negative values are nucleotide distances upstream from polyadenylation sites and positive values are distances downstream from polyadenylation sites. All *rbcS*-E9 signals start -145 upstream from the second polyadenylation site. CaMV polyadenylation signal is not to scale.
- (B) Competition EMSA experiments with 4 μ g GST-Fip1C (lane 2), and competitor wild type or mutant *rbcS*-E9 and CaMV RNA at a fifty-fold excess (4.85 μ g) as indicated above panel. Lane 1 is RNA alone (97 ng; 761.9 μ Ci/nmol). The reactions were analyzed on a native 4% polyacrylamide gel. Bound RNA is shown.
- (C) Competition EMSA experiments with 1.25 μ g GST-Fip1C (lane 2 and 7), incubated with RNA (490ng; 11.2 μ Ci/nmol; lane 1), and 980ng (lane 3 and 8), 4.9 μ g (lane 4 and 9), or 24.5 μ g (lane 5 and 10) of the indicated *rbcS*-E9 RNA as competitor

(above gel). The reactions were analyzed on a native 4% polyacrylamide gel. Bound and unbound RNA substrate is indicated on the left.

(D) Bound and unbound RNA substrate in the autoradiographs were quantitated by ImageQuant Software. The results were plotted as a ratio of bound to unbound RNA substrate against the fold excess (molar) of competitor RNA.

A



B

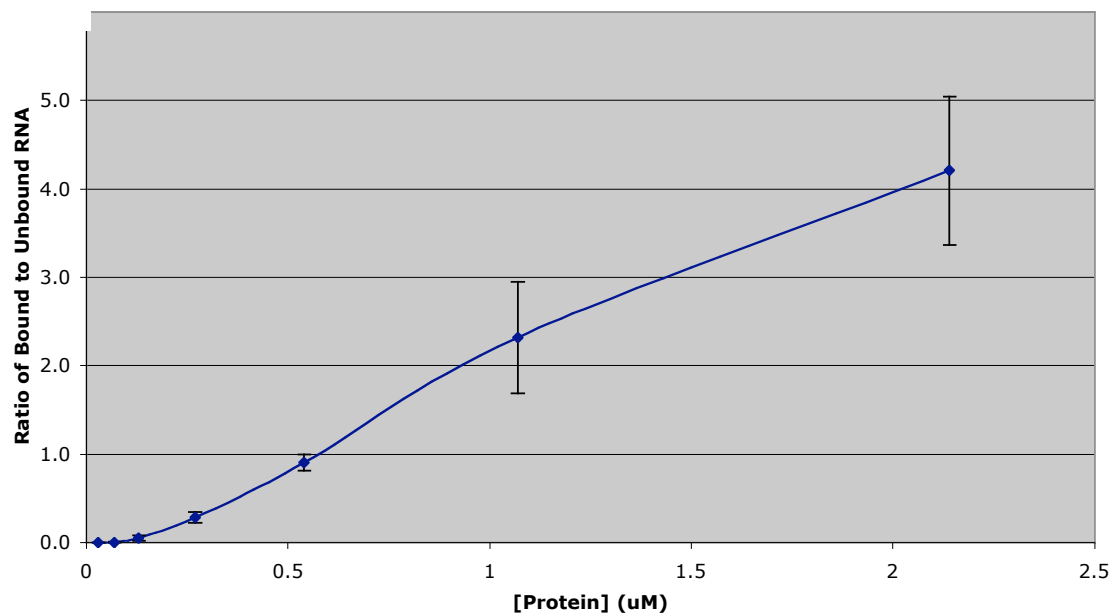
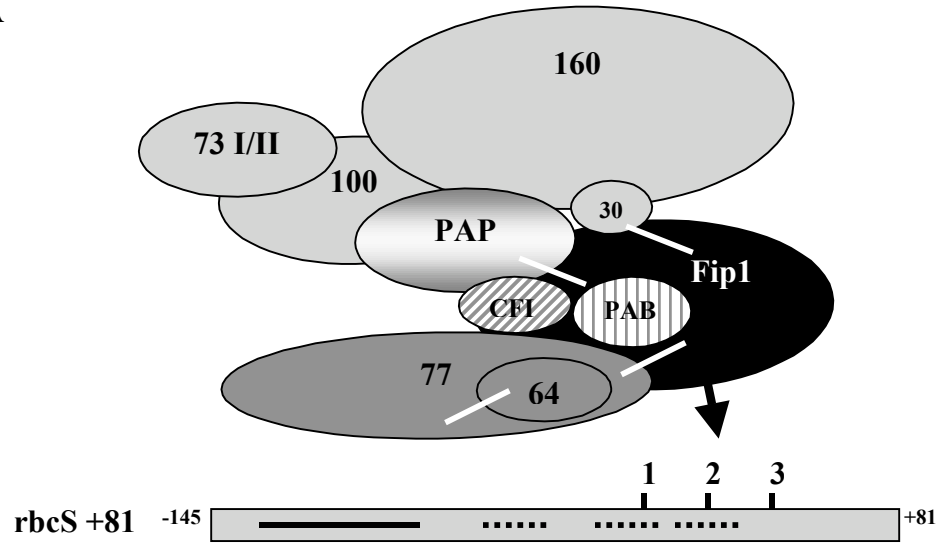


Figure 3.9. The C-terminal Domain of AtFip1 (FipC) has a low Affinity for RNA

(A) RNA substrate (rbcS+81; 1.96 μg; 11.2 Ci/nmol; lane 1) was incubated with 2.14 μM (lane 2), 1.07 μM (lane 3), 0.54 μM (lane 4), 0.27 μM (lane 5), 0.13 μM (lane 6), 0.07 μM (lane 7), 0.03 μM (lane 8) of GST-FipC. The reactions were analyzed on a native 4% polyacrylamide gel. Bound and unbound RNA substrate is indicated on the left.

(B) Bound and unbound RNA substrate in the autoradiographs were quantitated by ImageQuant Software. The results were plotted as a ratio of bound to unbound RNA substrate against the protein concentration of GST-FipC.

A



B

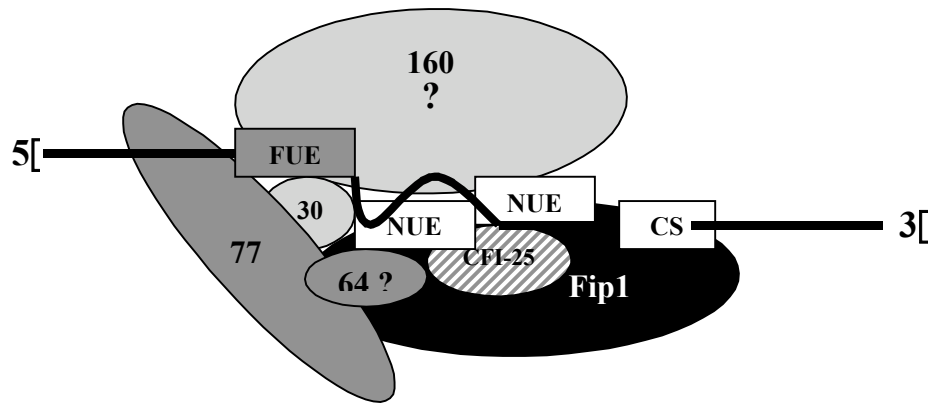


Figure 3.10. Model for the Organization of the Plant Cleavage and Polyadenylation Apparatus Around Fip1.

- (A) Multiple interactions between Fip1 and other components of the polyadenylation machinery. The mapping of protein-protein interactions between the *Arabidopsis* CPSF subunits (in light grey) has also been established and included in this model (Xu et al., manuscript in preparation). AtCstF subunits are dark gray. To simplify, the prefix At was omitted. The rbcS-E9 RNA from -145 to +81 is below the model (see figure 3.8 for description). The white lines indicate the interactions demonstrated in this thesis.
- (B) Model of potential RNA-binding proteins in the plant cleavage and polyadenylation apparatus and how they may be linked together by Fip1 and CstF77. CstF77 interactions between CPSF30 and CPSF 160 has been established by Drs. Ruqiang Xu and Qingshun Quinn Li (Miami University, Oxford, OH). Preliminary RNA-binding experiments have suggested CPSF30 binding of FUE (Kim Delaney and Arthur Hunt, unpublished observations).

CHAPTER FOUR

Future Directions

Gene expression in eukaryotes can be regulated at several levels within the cell. The proper processing of precursor mRNAs into mature forms is a conserved regulated step in gene expression. Mature mRNAs are generated from larger precursors generated by RNA Pol II in transcription, and are processed by splicing, 5' capping and 3' polyadenylation. All four of these processes can take place independently *in vitro*. However, recent discoveries in the field of gene expression indicate that pre-mRNA processing reactions are coupled *in vivo* (for review see (Wahle and Ruegsegger, 1999; Zhao et al., 1999a; Calvo and Manley, 2003). This suggests that defects in polyadenylation may have a profound affect on many of the processing reactions involved in mature mRNA formation.

With the identification and characterization of PPF-B from pea nuclear extracts, presented in chapter 2, we have attempted to conceptually link the regulation of PAP activity to other cellular processes. PPF-B may be involved in any number of cellular processes or may be a link between two different processing reactions. It is premature for us to say that PPF-B is involved in pre-mRNA processing and further experimentation is needed to address this possibility. The question still remains, what is the PAP inhibitory factor in PPF-B extracts? We were able to visualize other polypeptides found in PPF-B extracts stained with Coomassie in SDS-PAGE gels, but have not pursued these as potential targets. There are potentially 10 other polypeptides present in PPF-B extracts besides RSP, which is the major polypeptide (data not shown). Further purification or concentrating methods may allow us to generate enough material for protein identification by Mass Spectrometry or N-terminal sequencing. However, a problem that arises with identifying each of these polypeptides is the lack of genomic sequences available for garden pea *Pisium sativum* when compared to other plant genomes. A solution to this dilemma would be to attempt the purification of PPF-B from *Arabidopsis* nuclear extracts. The *Arabidopsis* genome has been fully sequenced and annotated which would aid in the identification of polypeptides found in *Arabidopsis* PPF-B extracts, possibly linking one or more of these proteins to inhibition of PAP activity. PPF-B does inhibit a recombinant *Arabidopsis* PAP suggesting that this activity may be conserved across different plant species. Currently our lab is working out different conditions for

nuclei isolation and protein nuclear extractions from *Arabidopsis*, which hopefully will make the purification of PPF-B from *Arabidopsis* a possibility.

Little is known about the regulation of poly(A) polymerases in plants, except that the genes encoding poly(A) polymerases exhibit novel alternative splicing events which may generate several isoforms of PAP (Addepalli et al., 2004). There is no evidence supporting post-translational modifications of PAPs in plants and what effects this may have on their function. In mammals and yeast, poly(A) polymerases have been reported to be cell cycle regulated by different post-translational modifications such as phosphorylation and ubiquitination (Colgan et al., 1996; Colgan et al., 1998; Mizrahi and Moore, 2000). From our characterization of PPF-B, we have suggested that the AMP-forming activity may be from an Ubiquitin/SUMO (E1)-activating enzyme. Analysis of peptide sequences found in *Arabidopsis* poly(A) polymerase, identified a potential site for SUMO attachment. Several E1, E2-conjugating enzymes, E3-ligases and Ubiquitin/Ubiquitin-like genes have been identified and characterized in *Arabidopsis* (for review see (Bachmair et al., 2001) and SUMO modifying proteins have also been identified (Kurepa et al., 2003). SUMO has been implicated to be essential for cell cycle progression in yeast (Seufert et al., 1995; Johnson and Blobel, 1997; Johnson et al., 1997) and suggested to function in stress protection and/or repair in *Arabidopsis* (Kurepa et al., 2003). Since *Arabidopsis* possesses all the enzymes involved in sumolation, *Arabidopsis* PAPs may be regulated in some way by SUMO attachment. Purification and analysis of *Arabidopsis* PAPs from nuclear extracts may allow us to identify post-translational modifications like sumolation or phosphorylation. The use of recombinant PAPs incubated *in vitro* with plant nuclear extracts may also identify modifications made to the enzyme. In general, this may allow us to identify post-translational modifications of *Arabidopsis* PAPs and determine what effects these modifications have, if any, on the enzymes activity. Further experiments on post-translational modifications of PAPs, as well as other plant cleavage and polyadenylation subunits would shed some light on how the complex is regulated during the cell cycle and under native and stress conditions as well.

There are two Fip1 genes found in the *Arabidopsis* genome, AtFip1-III and AtFip1-V. AtFip1-V appears to be analogous to the human Fip1 (hFip1), based on

protein domain organization, ability to bind RNA and the ability to stimulate non-specific PAP activity. These characterizations have only been done with fragments of the AtFip1-V protein. We plan on generating the full-length protein to test if it exhibits the same properties *in vitro* as its individual domains. We cannot rule out the possibility that the full-length protein may act differently than the individual domains themselves. While we have identified one domain in AtFip1-V that interacts with PAP, we do not know where AtFip1-V binds to on PAP. With the use of truncated forms of PAP, we should be able to identify the Fip1 interacting domain on PAP by yeast two-hybrid and *in vitro* pull-down assays. This would allow us to better understand the biochemistry of PAP activity and how Fip1 affects this activity, based on where the PAP-Fip1 interaction is occurring. Once identified, it raises the following question: is this Fip1 interacting domain found in other *Arabidopsis* PAPs? The characterization of AtFip1-V was only done with one of the four *Arabidopsis* PAPs. Further experiments will need to be done to identify the Fip1 interacting domain and if Fip1 stimulation of PAP is universal.

Little is known about the function of AtFip1-III, which raises an important question: why does *Arabidopsis* possess two Fip1 genes? AtFip1-III does contain the conserved acidic N-terminus and Fip1 domain found in eukaryotic Fip1s, but lacks the RNA-binding domain found in the hFip1 and AtFip1-V. Two-hybrid assays will need to be done with AtFip1-III to determine if it interacts with other polyadenylation homologs (as well as what affects it may have on PAP activity *in vitro*). The Fip1 genes may have evolved separate functions in *Arabidopsis* one analogous to hFip1 and the other to the yeast Fip1p. Further experiments will need to be done to address the function of both of these genes and if they both are required for cleavage and polyadenylation.

AtFip1-V binds RNA and appears to have a preference for poly(G) and RNAs possessing cleavage and polyadenylation sites found in the pea rbcS-E9 poly(A) signal. At first glance, one interpretation of the homopolymer competition studies is that AtFip1-V binds to the FUE in the plant polyadenylation signal. FUEs have been experimentally determined to be rich in U, G, or UG sequence motifs. However, the competition experiments indicate that RNAs that contain FUEs from two plant polyadenylation signals are no more effective in reducing RNA binding than competitors that lack the FUEs. Poly(G) assumes distinctive structures in solution, raising the possibility that an

RNA structure, rather than specific sequence, may be the target for AtFip1-V binding. There appears to be no obvious G-rich sequence found in the *rbcS-E9* poly(A) signal between +6 and +81. This raises the question, does AtFip1 bind structural elements or specific sequences in the poly(A) signal? RNase H footprinting experiments using RNAs generated from the 3' end of the pea *rbcS-E9* gene and protecting oligos should allow us to identify where on the RNA AtFip1-V binds. This may represent a specific sequence that AtFip1-V binds or a structural element in the RNA. Further analysis could be done using selection-amplification (SELEX) of RNAs to identify a consensus sequences that AtFip1-V binds. In either case, this information could be used to probe the *Arabidopsis* genome to identify genes that contain these structural elements or consensus sequences at their 3' ends. This may allow us to identify sites of cleavage and polyadenylation and class sets of genes accordingly.

Why does *Arabidopsis* possess a Fip1 protein that appears to be analogous to the hFip1 when AtPAP enzymes are processive? Besides AtPAP-IV, the other three *Arabidopsis* PAPs are also processive polymerases (unpublished observations). The answer may be that AtFip1-V plays a role in both the cleavage and polyadenylation reactions, possessing yeast and human Fip1 functions. The hFip1 subunit is suggested to stimulate PAP by increasing its affinity for RNA, thus making it processive. hFip1 requires both the N-terminus and C-terminal RNA-binding domains for stimulation (Kaufmann et al., 2004). However, AtFip1-V has an N-terminal domain that stimulates PAP independent of its C-terminal RNA-binding domain, which suggests that AtFip1-V at least plays a role in the polyadenylation reaction. This N-terminal domain is also suggested to be a hub for other poly(A) factors. We do not know if all of these interactions occur at the same time or independent of each other, which opens up the possibility for AtFip1-V having another function. AtPAP has been shown to interact with other subunits, which suggests that Fip1 may not be required for PAP recruitment to the complex. Therefore, AtFip1-V may play a role in complex formation by its N-terminal hub and C-terminal RNA-Binding to cleavage sites, independent of its interaction with PAP. This also suggests that Fip1 is at the center of a network of protein-protein and protein-RNA interactions. Once cleavage occurs, the complex may dissociate, and the N-terminus may be free to bind and direct the PAP to the free 3' end, thus kick starting the

enzyme by this positioning, resulting in stimulation. Further experiments need to be done to test for the effects other subunits have on AtFip1-V in the presence of PAP and whether these interactions are exclusive. This may allow us to link the other subunits to the polyadenylation reaction and/or the cleavage reaction.

Do other plant polyadenylation factors affect AtFip1-V binding? The subunits that interact with AtFip1-V (CstF77, CPSF30, PAP, CFI-25m, PABN1; see Chapter 3) may have an effect on the affinity AtFip1-V has for RNA. Preliminary results suggest that AtCPSF30 binds FUEs (Kim Delaney and Arthur Hunt, unpublished observations). The binding of FUE by AtCPSF30 and the protein-protein interaction with AtFip1-V may direct or specify what cleavage and polyadenylation site AtFip1-V binds. This may prove functionally analogous to what is suggested with CFIA directing CFII recognition of the cleavage site in yeast. Also, this dimer of Fip1-CPSF30 may form a more stable RNP complex. Since the reported interaction between AtCPSF30 and AtFip1-V is via the N-terminus of AtFip1-V, the full length AtFip1-V protein will be required for this experiment, given that the RNA-binding domain is in the C-terminus of AtFip1-V.

The idea of adding several subunits at a time *in vitro* to test for changes in function or activity of individual subunits is one of the aims of our laboratory, which is the *in vitro* reconstitution of the plant cleavage and polyadenylation apparatus. *In vitro* reconstitution of the cleavage and polyadenylation reaction has been accomplished with purified factors from mammalian and yeast extracts. Cleavage and polyadenylation activities in yeast can also be reconstituted with recombinant proteins of CFIA and purified CFII, PFI, and Pap1p (Gross and Moore, 2001a). This has allowed for the identification of the subunits required for one or both of the steps involved in 3' end formation. We are taking two approaches to identifying the factors involved in one or both of the processes in plants: one is the cloning of the Arabidopsis homologs to the mammalian and yeast poly(A) factors and generating recombinant proteins and the other is purification of whole or parts of the complex from nuclear extracts by tandem affinity purification (TAP)-tagging of individual subunits. Currently, we have one activity we can test for *in vitro* with recombinant proteins, that is the non-specific poly(A) polymerase activity of PAP. By including AtFip1-V in this reaction we have observed stimulation of the non-specific activity of PAP. What happens when we add subunits that interact with

AtFip1-V, the other reported subunits that interact with PAP, or a combination of both? When we have all of the *Arabidopsis* homologs cloned and recombinant proteins generated, we can test for the effects of multiple-subunit interactions and what affects they might have on non-specific PAP activity. Full reconstitution of specific cleavage and polyadenylation in the mammalian and yeast systems, with recombinant proteins, has not been very successful. Most of the mammalian and yeast polyadenylation factors can be identified in various plant genomes, but other essential factors required for specific cleavage and polyadenylation may go unnoticed. For example, there are several potential homologs to the yeast polyadenylation factor Hrp1p in the *Arabidopsis* genome, making it difficult to identify the *Arabidopsis* Hrp1p. Hrp1p in yeast is required for the cleavage reaction. Other methods will need to be employed to identify subunits like Hrp1p in the plant polyadenylation apparatus. Therefore, reconstitution of the *Arabidopsis* cleavage and polyadenylation apparatus may not be feasible. The second method mentioned above, TAP-tagging, may solve this problem. The use of TAP-tagging of yeast polyadenylation subunits has proven to be a useful tool in identifying subunits of the polyadenylation complex and other proteins associated with the complex that function in transcription initiation and termination (Vo et al., 2001; Nedea et al., 2003). TAP-tagging in plants has proven to be very effective at isolating and purifying heteromeric protein complexes (Rohila et al., 2004). What subunit of the plant polyadenylation apparatus would be a good candidate for TAP-tagging? We have suggested that AtFip1-V coordinates a number of polyadenylation subunits and therefore it could be inferred that TAP-tagging of AtFip1-V would increase our chance of pulling down the whole or parts of the polyadenylation complex when compared to using other subunits. This method will not only allow us to identify subunits involved in cleavage and polyadenylation, but also proteins suggested to be involved in other cellular processes, thus potentially linking mRNA 3' end formation in plants to these processes. Purification of the complex by this method may also allow us to test for specific cleavage and polyadenylation with RNAs containing poly(A) signals.

REFERENCES

- Abe, S., and Davies, E.** (1995). Methods for isolation and analysis of the cytoskeleton. *Methods Cell Biol* **50**, 223-236.
- Addepalli, B., Meeks, L.R., Forbes, K.P., and Hunt, A.G.** (2004). Novel alternative splicing of mRNAs encoding poly(A) polymerases in Arabidopsis. *Biochim Biophys Acta* **1679**, 117-128.
- Aissouni, Y., Perez, C., Calmels, B., and Benech, P.D.** (2002). The cleavage/polyadenylation activity triggered by a U-rich motif sequence is differently required depending on the poly(A) site location at either the first or last 3'-terminal exon of the 2'-5' oligo(A) synthetase gene. *J Biol Chem* **277**, 35808-35814.
- Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D.** (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *Embo J* **18**, 5399-5410.
- 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.S., and Parker, R.P.** (1998). The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *Embo J* **17**, 1497-1506.
- Anderson, J.T., Paddy, M.R., and Swanson, M.S.** (1993a). PUB1 is a major nuclear and cytoplasmic polyadenylated RNA-binding protein in *Saccharomyces cerevisiae*. *Mol Cell Biol* **13**, 6102-6113.
- Anderson, J.T., Wilson, S.M., Datar, K.V., and Swanson, M.S.** (1993b). NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. *Mol Cell Biol* **13**, 2730-2741.
- Aranda, A., and Proudfoot, N.** (2001). Transcriptional termination factors for RNA polymerase II in yeast. *Mol Cell* **7**, 1003-1011.

- 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.
- Bachmair, A., Novatchkova, M., Potuschak, T., and Eisenhaber, F.** (2001). Ubiquitylation in plants: a post-genomic look at a post-translational modification. *Trends Plant Sci* **6**, 463-470.
- 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.
- Barabino, S.M., Hubner, W., Jenny, A., Minvielle-Sebastia, L., and Keller, W.** (1997). The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev* **11**, 1703-1716.
- Bard, J., Zhelkovsky, A.M., Helmling, S., Earnest, T.N., Moore, C.L., and Bohm, A.** (2000). Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. *Science* **289**, 1346-1349.
- 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*. *Proc Natl Acad Sci U S A* **98**, 445-450.
- Beelman, C.A., and Parker, R.** (1995). Degradation of mRNA in eukaryotes. *Cell* **81**, 179-183.
- Bentley, D.** (2002). The mRNA assembly line: transcription and processing machines in the same factory. *Curr Opin Cell Biol* **14**, 336-342.
- Beyer, K., Dandekar, T., and Keller, W.** (1997). RNA ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3'-end processing of pre-mRNA. *J Biol Chem* **272**, 26769-26779.
- Bienroth, S., Keller, W., and Wahle, E.** (1993). Assembly of a processive messenger RNA polyadenylation complex. *Embo J* **12**, 585-594.
- 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.

- Birse, C.E., Minvielle-Sebastia, L., Lee, B.A., Keller, W., and Proudfoot, N.J.** (1998). Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280**, 298-301.
- Boeck, R., Tarun, S., Jr., Rieger, M., Deardorff, J.A., Muller-Auer, S., and Sachs, A.B.** (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. *J Biol Chem* **271**, 432-438.
- Bousquet-Antonelli, C., Presutti, C., and Tollervey, D.** (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102**, 765-775.
- Brackenridge, S., and 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.
- Briggs, M.W., Burkard, K.T., and Butler, J.S.** (1998). Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. *J Biol Chem* **273**, 13255-13263.
- Brodsky, A.S., and Silver, P.A.** (2000). Pre-mRNA processing factors are required for nuclear export. *Rna* **6**, 1737-1749.
- Brown, C.E., and Sachs, A.B.** (1998). Poly(A) tail length control in *Saccharomyces cerevisiae* occurs by message-specific deadenylation. *Mol Cell Biol* **18**, 6548-6559.
- Brown, C.E., Tarun, S.Z., Jr., Boeck, R., and Sachs, A.B.** (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.
- Burd, C.G., and Dreyfuss, G.** (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265**, 615-621.
- Burkard, K.T., and Butler, J.S.** (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. *Mol Cell Biol* **20**, 604-616.
- Butler, J.S.** (2002). The yin and yang of the exosome. *Trends Cell Biol* **12**, 90-96.

- 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.** (2003). Strange bedfellows: polyadenylation factors at the promoter. *Genes Dev* **17**, 1321-1327.
- Caponigro, G., and Parker, R.** (1996). Mechanisms and control of mRNA turnover in *Saccharomyces cerevisiae*. *Microbiol Rev* **60**, 233-249.
- Carpousis, A.J., Vanzo, N.F., and Raynal, L.C.** (1999). mRNA degradation. A tale of poly(A) and multiprotein machines. *Trends Genet* **15**, 24-28.
- Chekanova, J.A., and Belostotsky, D.A.** (2003). Evidence that poly(A) binding protein has an evolutionarily conserved function in facilitating mRNA biogenesis and export. *Rna* **9**, 1476-1490.
- Chen, A., Kleiman, F.E., Manley, J.L., Ouchi, T., and Pan, Z.Q.** (2002). Autoubiquitination of the BRCA1*BARD1 RING ubiquitin ligase. *J Biol Chem* **277**, 22085-22092.
- Chen, F., MacDonald, C.C., and Wilusz, J.** (1995). Cleavage site determinants in the mammalian polyadenylation signal. *Nucleic Acids Res* **23**, 2614-2620.
- 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, 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.
- 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, D.F., and Manley, J.L.** (1997). Mechanism and regulation of mRNA polyadenylation. *Genes Dev* **11**, 2755-2766.
- Colgan, D.F., Murthy, K.G., Prives, C., and Manley, J.L.** (1996). Cell-cycle related regulation of poly(A) polymerase by phosphorylation. *Nature* **384**, 282-285.

- Colgan, D.F., Murthy, K.G., Zhao, W., Prives, C., and Manley, J.L.** (1998). Inhibition of poly(A) polymerase requires p34cdc2/cyclin B phosphorylation of multiple consensus and non-consensus sites. *Embo J* **17**, 1053-1062.
- Cooke, C., Hans, H., and Alwine, J.C.** (1999). Utilization of splicing elements and polyadenylation signal elements in the coupling of polyadenylation and last-intron removal. *Mol Cell Biol* **19**, 4971-4979.
- Dantoni, J.C., Murthy, K.G., Manley, J.L., and Tora, L.** (1997). Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. *Nature* **389**, 399-402.
- Das Gupta, J., Li, Q., Thomson, A.B., and Hunt, A.G.** (1995). Characterization of a novel plant poly(A) polymerase. *Plant Science* **110**, 215-226.
- Das Gupta, J., Li, Q.S., Thomson, A.B., and Hunt, A.G.** (1998). Characterization of a cDNA encoding a novel plant poly(A) polymerase. *Plant Mol Biol* **37**, 729-734.
- Dattaroy, T., and Hunt, A.G.** (2002). Polyadenylation of RNAs Associated with a nucleus-localized Phosphorolytic Nuclease. *J Plant Biochemistry and Biotechnology* **11**, 21-25.
- Davies, E., Stankovic, B., Azama, K., Shibata, K., and Abe, S.** (2001). Novel components of the plant cytoskeleton: a beginning to plant 'cytomics'. *Plant Sci* **160**, 185-196.
- 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.
- 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.** (1998). The exosome: a versatile RNA processing machine. *Curr Biol* **8**, R238-240.
- del Olmo, M., Mizrahi, N., Gross, S., and Moore, C.L.** (1997). The Uba2 and Ufd1 proteins of *Saccharomyces cerevisiae* interact with poly(A) polymerase and affect the polyadenylation activity of cell extracts. *Mol Gen Genet* **255**, 209-218.

- 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.
- Dichtl, B., and Keller, W.** (2001). Recognition of polyadenylation sites in yeast pre-mRNAs 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.
- Duvel, K., Valerius, O., Mangus, D.A., Jacobson, A., and Braus, G.H.** (2002). Replacement of the yeast TRP4 3' untranslated region by a hammerhead ribozyme results in a stable and efficiently exported mRNA that lacks a poly(A) tail. *Rna* **8**, 336-344.
- Eckner, R., Ellmeier, W., and Birnstiel, M.L.** (1991). Mature mRNA 3' end formation stimulates RNA export from the nucleus. *Embo J* **10**, 3513-3522.
- Elliott, B.J., Dattaroy, T., Meeks-Midkiff, L.R., Forbes, K.P., and Hunt, A.G.** (2003). An interaction between an Arabidopsis poly(A) polymerase and a homologue of the 100 kDa subunit of CPSF. *Plant Mol Biol* **51**, 373-384.
- Falquet, L., Pagni, M., Bucher, P., Hulo, N., Sigrist, C.J., Hofmann, K., and Bairoch, A.** (2002). The PROSITE database, its status in 2002. *Nucleic Acids Res* **30**, 235-238.
- Fortes, P., Beloso, A., and Ortin, J.** (1994). Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport. *Embo J* **13**, 704-712.
- Gallie, D.R.** (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev* **5**, 2108-2116.
- Gietz, D., St Jean, A., Woods, R.A., and Schiestl, R.H.** (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* **20**, 1425.

- Gilmartin, G.M., and Nevins, J.R.** (1989). An ordered pathway of assembly of components required for polyadenylation site recognition and processing. *Genes Dev* **3**, 2180-2190.
- Goodin, M.M., Dietzgen, R.G., Schichnes, D., Ruzin, S., and Jackson, A.O.** (2002). pGD vectors: versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves. *Plant J* **31**, 375-383.
- 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. *Proc Natl Acad Sci U S A* **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.
- Graveley, B.R., Fleming, E.S., and Gilmartin, G.M.** (1996). RNA structure is a critical determinant of poly(A) site recognition by cleavage and polyadenylation specificity factor. *Mol Cell Biol* **16**, 4942-4951.
- Green, D.M., Marfatia, K.A., Crafton, E.B., Zhang, X., Cheng, X., and Corbett, A.H.** (2002). Nab2p is required for poly(A) RNA export in *Saccharomyces cerevisiae* and is regulated by arginine methylation via Hmt1p. *J Biol Chem* **277**, 7752-7760.
- Greger, I.H., Aranda, A., and Proudfoot, N.** (2000). Balancing transcriptional interference and initiation on the GAL7 promoter of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **97**, 8415-8420.
- Gross, S., and Moore, C.** (2001a). Five subunits are required for reconstitution of the cleavage and polyadenylation activities of *Saccharomyces cerevisiae* cleavage factor I. *Proc Natl Acad Sci U S A* **98**, 6080-6085.
- Gross, S., and Moore, C.L.** (2001b). Rna15 interaction with the A-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. *Mol Cell Biol* **21**, 8045-8055.
- Gunderson, S.I., Polycarpou-Schwarz, M., and Mattaj, I.W.** (1998). U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell* **1**, 255-264.

- Gunderson, S.I., Vagner, S., Polycarpou-Schwarz, M., and Mattaj, I.W.** (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**, 761-773.
- Guo, Z., and Sherman, F.** (1996a). 3'-end-forming signals of yeast mRNA. *Trends Biochem Sci* **21**, 477-481.
- Guo, Z., and Sherman, F.** (1996b). Signals sufficient for 3'-end formation of yeast mRNA. *Mol Cell Biol* **16**, 2772-2776.
- Hajnsdorf, E., Braun, F., Haugel-Nielsen, J., and Regnier, P.** (1995). Polyadenylation destabilizes the rpsO mRNA of Escherichia coli. *Proc Natl Acad Sci U S A* **92**, 3973-3977.
- 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.
- Hanley, B.A., and Schuler, M.A.** (1991). cDNA cloning of U1, U2, U4 and U5 snRNA families expressed in pea nuclei. *Nucleic Acids Res* **19**, 1861-1869.
- Hans, H., and Alwine, J.C.** (2000). Functionally significant secondary structure of the simian virus 40 late polyadenylation signal. *Mol Cell Biol* **20**, 2926-2932.
- Hector, R.E., Nykamp, K.R., Dheur, S., Anderson, J.T., Non, P.J., Urbinati, C.R., Wilson, S.M., Minvielle-Sebastia, L., and Swanson, M.S.** (2002). Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. *Embo J* **21**, 1800-1810.
- Helmling, S., Zhelkovsky, A., and Moore, C.L.** (2001). Fip1 regulates the activity of Poly(A) polymerase through multiple interactions. *Mol Cell Biol* **21**, 2026-2037.
- Hill, D.A., and Imbalzano, A.N.** (2000). Human SWI/SNF nucleosome remodeling activity is partially inhibited by linker histone H1. *Biochemistry* **39**, 11649-11656.
- Hilleren, P., and Parker, R.** (1999). mRNA surveillance in eukaryotes: kinetic proofreading of proper translation termination as assessed by mRNP domain organization? *Rna* **5**, 711-719.
- Hirose, Y., and Manley, J.L.** (1998). RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395**, 93-96.

- Hirose, Y., and Manley, J.L.** (2000). RNA polymerase II and the integration of nuclear events. *Genes Dev* **14**, 1415-1429.
- Hofmann, R.M., and Pickart, C.M.** (1999). Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**, 645-653.
- 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.C.** (1996). Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol Cell Biol* **16**, 1534-1542.
- Hunt, A.** (1994). Messenger RNA 3' End Formation in Plants. *standard.dtl Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 47-60.
- Hunt, A.G., and Messing, J.** (1998). in *A Look Beyond Transcription: Mechanisms Determining mRNA Stability and Translation in Plants. Proceedings of the 19th UC-Riverside Annual Symposium on Plant Physiology. Riverside CA 1997.*,pp. 29-39.
- Hunt, A.G., Meeks, L.R., Forbes, K.P., Das Gupta, J., and Mogen, B.D.** (2000). Nuclear and chloroplast poly(A) polymerases from plants share a novel biochemical property. *Biochem Biophys Res Commun* **272**, 174-181.
- Hyman, L.E., and Moore, C.L.** (1993). Termination and pausing of RNA polymerase II downstream of yeast polyadenylation sites. *Mol Cell Biol* **13**, 5159-5167.
- Ingle, C.A., and Kushner, S.R.** (1996). Development of an in vitro mRNA decay system for *Escherichia coli*: poly(A) polymerase I is necessary to trigger degradation. *Proc Natl Acad Sci U S A* **93**, 12926-12931.
- James, P., Halladay, J., and Craig, E.A.** (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.
- Jenny, A., and Keller, W.** (1995). Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyadenylation specificity factor. *Nucleic Acids Res* **23**, 2629-2635.

- 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.
- Johnson, E.S., and Blobel, G.** (1997). Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *J Biol Chem* **272**, 26799-26802.
- Johnson, E.S., Schwienhorst, I., Dohmen, R.J., and Blobel, G.** (1997). The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. *Embo J* **16**, 5509-5519.
- Juge, F., Zaessinger, S., Temme, C., Wahle, E., and Simonelig, M.** (2002). Control of poly(A) polymerase level is essential to cytoplasmic polyadenylation and early development in *Drosophila*. *Embo J* **21**, 6603-6613.
- Kaufmann, I., Martin, G., Friedlein, A., Langen, H., and 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**, 616-626.
- 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.
- 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.
- Kessler, M.M., Zhao, J., and Moore, C.L.** (1996). Purification of the *Saccharomyces cerevisiae* cleavage/polyadenylation factor I. Separation into two components that are required for both cleavage and polyadenylation of mRNA 3' ends. *J Biol Chem* **271**, 27167-27175.
- Kieber, J.J., Rothenberg, M., Roman, G., Feldmann, K.A., and Ecker, J.R.** (1993). CTR1, a negative regulator of the ethylene response pathway in *Arabidopsis*, encodes a member of the raf family of protein kinases. *Cell* **72**, 427-441.
- 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, J.H., and Lee, Y.** (2003). Regulation of poly(A) polymerase by 14-3-3epsilon. *Embo J* **22**, 5208-5219.
- Klasens, B.I., Thiesen, M., Virtanen, A., and Berkhout, B.** (1999). The ability of the HIV-1 AAUAAA signal to bind polyadenylation factors is controlled by local RNA structure. *Nucleic Acids Res* **27**, 446-454.
- Kleiman, F.E., and Manley, J.L.** (1999). Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. *Science* **285**, 1576-1579.
- Kleiman, F.E., and Manley, J.L.** (2001). The BARD1-CstF-50 interaction links mRNA 3' end formation to DNA damage and tumor suppression. *Cell* **104**, 743-753.
- Kuai, L., Fang, F., Butler, J.S., and Sherman, F.** (2004). Polyadenylation of rRNA in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **101**, 8581-8586.
- Kudla, J., Hayes, R., and Grusissem, W.** (1996). Polyadenylation accelerates degradation of chloroplast mRNA. *Embo J* **15**, 7137-7146.
- Kuhn, U., Nemeth, A., Meyer, S., and Wahle, E.** (2003). The RNA binding domains of the nuclear poly(A)-binding protein. *J Biol Chem* **278**, 16916-16925.
- Kurepa, J., Walker, J.M., Smalle, J., Gosink, M.M., Davis, S.J., Durham, T.L., Sung, D.Y., and Vierstra, R.D.** (2003). The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. *J Biol Chem* **278**, 6862-6872.
- 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.
- Legendre, M., and Gautheret, D.** (2003). Sequence determinants in human polyadenylation site selection. *BMC Genomics* **4**, 7.
- Lei, E.P., and Silver, P.A.** (2002). Protein and RNA export from the nucleus. *Dev Cell* **2**, 261-272.
- Li, Q., and Hunt, A.G.** (1995). A near-upstream element in a plant polyadenylation signal consists of more than six nucleotides. *Plant Mol Biol* **28**, 927-934.
- Li, Q., and Hunt, A.G.** (1997). The Polyadenylation of RNA in Plants. *Plant Physiol* **115**, 321-325.

- Li, Q.S., Gupta, J.D., and Hunt, A.G.** (1996). A plant poly(A) polymerase requires a novel RNA-binding protein for activity. *J Biol Chem* **271**, 19831-19835.
- Li, Q.S., Gupta, J.D., and Hunt, A.G.** (1998). Polynucleotide phosphorylase is a component of a novel plant poly(A) polymerase. *J Biol Chem* **273**, 17539-17543.
- Libri, D., Dower, K., Boulay, J., Thomsen, R., Rosbash, M., and Jensen, T.H.** (2002). Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol* **22**, 8254-8266.
- 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.
- Lisitsky, I., Klaff, P., and Schuster, G.** (1996). Addition of destabilizing poly (A)-rich sequences to endonuclease cleavage sites during the degradation of chloroplast mRNA. *Proc Natl Acad Sci U S A* **93**, 13398-13403.
- Lorkovic, Z.J., and Barta, A.** (2002). Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Res* **30**, 623-635.
- Lowell, J.E., Rudner, D.Z., and Sachs, A.B.** (1992). 3'-UTR-dependent deadenylation by the yeast poly(A) nuclease. *Genes Dev* **6**, 2088-2099.
- Lutz, C.S., and Alwine, J.C.** (1994). Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev* **8**, 576-586.
- Lutz, C.S., Murthy, K.G., Schek, N., O'Connor, J.P., Manley, J.L., and Alwine, J.C.** (1996). Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency in vitro. *Genes Dev* **10**, 325-337.
- MacDonald, C.C., and Redondo, J.L.** (2002). Reexamining the polyadenylation signal: were we wrong about AAUAAA? *Mol Cell Endocrinol* **190**, 1-8.

- MacDonald, M.H., Mogen, B.D., and Hunt, A.G.** (1991). Characterization of the polyadenylation signal from the T-DNA-encoded octopine synthase gene. *Nucleic Acids Res* **19**, 5575-5581.
- Manley, J.L.** (1988). Polyadenylation of mRNA precursors. *Biochim Biophys Acta* **950**, 1-12.
- Manley, J.L., and Proudfoot, N.J.** (1994). RNA 3' ends: formation and function--meeting review. *Genes Dev* **8**, 259-264.
- Mans, R.J., and Huff, N.J.** (1975). Utilization of ribonucleic acid and deoxyoligomer primers for polyadenylic acid synthesis by adenosine triphosphate: polynucleotidylexotransferase from maize. *J Biol Chem* **250**, 3672-3678.
- 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., Jenö, P., and Keller, W.** (1999). Mapping of ATP binding regions in poly(A) polymerases by photoaffinity labeling and by mutational analysis identifies a domain conserved in many nucleotidyltransferases. *Protein Sci* **8**, 2380-2391.
- Martin, G., Keller, W., and Doublet, 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., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L.** (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**, 357-361.
- Meinke, D.W., Meinke, L.K., Showalter, T.C., Schissel, A.M., Mueller, L.A., and Tzafrir, I.** (2003). A sequence-based map of Arabidopsis genes with mutant phenotypes. *Plant Physiol* **131**, 409-418.
- Melchior, F.** (2000). SUMO--nonclassical ubiquitin. *Annu Rev Cell Dev Biol* **16**, 591-626.

- 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.
- 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., and 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**, 2794-2802.
- Mogen, B.D., MacDonald, M.H., Graybosch, R., and Hunt, A.G.** (1990). Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. *Plant Cell* **2**, 1261-1272.
- Mogen, B.D., MacDonald, M.H., Leggewie, G., and Hunt, A.G.** (1992). Several distinct types of sequence elements are required for efficient mRNA 3' end formation in a pea *rbcS* gene. *Mol Cell Biol* **12**, 5406-5414.
- Moreira, A., Wollerton, M., Monks, J., and Proudfoot, N.J.** (1995). Upstream sequence elements enhance poly(A) site efficiency of the C2 complement gene and are phylogenetically conserved. *Embo J* **14**, 3809-3819.
- Muchardt, C., Reyes, J.C., Bourachot, B., Leguoy, E., and Yaniv, M.** (1996). The hbrm and BRG-1 proteins, components of the human SNF/SWI complex, are phosphorylated and excluded from the condensed chromosomes during mitosis. *Embo J* **15**, 3394-3402.
- Muller, S., Hoegge, C., Pyrowolakis, G., and Jentsch, S.** (2001). SUMO, ubiquitin's mysterious cousin. *Nat Rev Mol Cell Biol* **2**, 202-210.
- Murthy, K.G., and Manley, J.L.** (1992). Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J Biol Chem* **267**, 14804-14811.
- 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.
- O'Hara, E.B., Chekanova, J.A., Ingle, C.A., Kushner, Z.R., Peters, E., and Kushner, S.R.** (1995). Polyadenylation helps regulate mRNA decay in *Escherichia coli*. *Proc Natl Acad Sci U S A* **92**, 1807-1811.
- Ohnacker, M., Barabino, S.M., Preker, P.J., and Keller, W.** (2000). The WD-repeat protein pfs2p bridges two essential factors within the yeast pre-mRNA 3'-end-processing complex. *Embo J* **19**, 37-47.
- Ojala, D., Crews, S., Montoya, J., Gelfand, R., and Attardi, G.** (1981). A small polyadenylated RNA (7 S RNA), containing a putative ribosome attachment site, maps near the origin of human mitochondrial DNA replication. *J Mol Biol* **150**, 303-314.
- Parker, R., and Song, H.** (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol* **11**, 121-127.
- Perez Canadillas, J.M., and Varani, G.** (2003). Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *Embo J* **22**, 2821-2830.
- Phillips, C., and Virtanen, A.** (1997). The murine IgM secretory poly(A) site contains dual upstream and downstream elements which affect polyadenylation. *Nucleic Acids Res* **25**, 2344-2351.
- Preiss, T., and Hentze, M.W.** (1998). Dual function of the messenger RNA cap structure in poly(A)-tail-promoted translation in yeast. *Nature* **392**, 516-520.
- Preker, P.J., Lingner, J., Minvielle-Sebastia, L., and 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**, 379-389.
- Preker, P.J., Ohnacker, M., Minvielle-Sebastia, L., and 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**, 4727-4737.

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

Proudfoot, N., and O'Sullivan, J. (2002). Polyadenylation: a tail of two complexes. *Curr Biol* **12**, R855-857.

Proweller, A., and Butler, S. (1994). Efficient translation of poly(A)-deficient mRNAs in *Saccharomyces cerevisiae*. *Genes Dev* **8**, 2629-2640.

Proweller, A., and Butler, J.S. (1997). Ribosome concentration contributes to discrimination against poly(A)- mRNA during translation initiation in *Saccharomyces cerevisiae*. *J Biol Chem* **272**, 6004-6010.

Ramachandran, A., Omar, M., Cheslock, P., and Schnitzler, G.R. (2003). Linker histone H1 modulates nucleosome remodeling by human SWI/SNF. *J Biol Chem* **278**, 48590-48601.

Rodriguez, C.R., Cho, E.J., Keogh, M.C., Moore, C.L., Greenleaf, A.L., and Buratowski, S. (2000). Kin28, the TFIIF-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II. *Mol Cell Biol* **20**, 104-112.

Rohila, J.S., Chen, M., Cerny, R., and Fromm, M.E. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J* **38**, 172-181.

Rothnie, H.M. (1996). Plant mRNA 3'-end formation. *Plant Mol Biol* **32**, 43-61.

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.

Ruegsegger, U., Blank, D., and Keller, W. (1998). Human pre-mRNA cleavage factor Im is related to spliceosomal SR proteins and can be reconstituted in vitro from recombinant subunits. *Mol Cell* **1**, 243-253.

Russo, P., and Sherman, F. (1989). Transcription terminates near the poly(A) site in the CYC1 gene of the yeast *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **86**, 8348-8352.

- 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., Sarnow, P., and Hentze, M.W.** (1997). Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* **89**, 831-838.
- Sambrook, J., Fritsch, E.F., and Maniatis, T.** (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ED., Cold Springs Harbor, Cold Springs Harbor, NY.
- Sarkar, N.** (1997). Polyadenylation of mRNA in prokaryotes. *Annu Rev Biochem* **66**, 173-197.
- Searfoss, A.M., and Wickner, R.B.** (2000). 3' poly(A) is dispensable for translation. *Proc Natl Acad Sci U S A* **97**, 9133-9137.
- Seufert, W., Futcher, B., and Jentsch, S.** (1995). Role of a ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature* **373**, 78-81.
- 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.
- Sif, S., Stukenberg, P.T., Kirschner, M.W., and Kingston, R.E.** (1998). Mitotic inactivation of a human SWI/SNF chromatin remodeling complex. *Genes Dev* **12**, 2842-2851.
- Simpson, G.G., Dijkwel, P.P., Quesada, V., Henderson, I., and Dean, C.** (2003). FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. *Cell* **113**, 777-787.
- Spence, J., Sadis, S., Haas, A.L., and Finley, D.** (1995). A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol Cell Biol* **15**, 1265-1273.
- 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**, 1517-1520.
- Takagaki, Y., and 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., Ryner, L.C., and Manley, J.L.** (1989). Four factors are required for 3'-end cleavage of pre-mRNAs. *Genes Dev* **3**, 1711-1724.
- Tanaka, I., Akahori, Y., Gomi, K., Suzuki, T., and Ueda, K.** (1999). A novel histone variant localized in nucleoli of higher plant cells. *Chromosoma* **108**, 190-199.
- Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J., and Tollervey, D.** (2002). Processing of 3'-extended read-through transcripts by the exosome can generate functional mRNAs. *Mol Cell* **9**, 1285-1296.
- Tucker, M., and Parker, R.** (2000). Mechanisms and control of mRNA decapping in *Saccharomyces cerevisiae*. *Annu Rev Biochem* **69**, 571-595.
- Vagner, S., Ruegsegger, U., Gunderson, S.I., Keller, W., and Mattaj, I.W.** (2000). Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *Rna* **6**, 178-188.
- van Hoof, A., and Parker, R.** (1999). The exosome: a proteasome for RNA? *Cell* **99**, 347-350.
- van Hoof, A., and Parker, R.** (2002). Messenger RNA degradation: beginning at the end. *Curr Biol* **12**, R285-287.
- van Hoof, A., Lennertz, P., and Parker, R.** (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol Cell Biol* **20**, 441-452.
- van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R.** (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262-2264.
- 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 Ruegsegger, U.** (1999). 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiol Rev* **23**, 277-295.
- Wallace, A.M., Dass, B., Ravnik, S.E., Tonk, V., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and MacDonald, C.C.** (1999). Two distinct forms of the 64,000 Mr protein of the cleavage stimulation factor are expressed in mouse male germ cells. *Proc Natl Acad Sci U S A* **96**, 6763-6768.
- Wassarman, K.M., and Steitz, J.A.** (1993). Association with terminal exons in pre-mRNAs: a new role for the U1 snRNP? *Genes Dev* **7**, 647-659.
- Wickens, M.** (1990). How the messenger got its tail: addition of poly(A) in the nucleus. *Trends Biochem Sci* **15**, 277-281.
- Wong, C., Sridhara, S., Bardwell, J.C., and Jakob, U.** (2000). Heating greatly speeds Coomassie blue staining and destaining. *Biotechniques* **28**, 426-428, 430, 432.
- Xu, F., and Cohen, S.N.** (1995). RNA degradation in *Escherichia coli* regulated by 3' adenylation and 5' phosphorylation. *Nature* **374**, 180-183.
- Xu, F., Lin-Chao, S., and Cohen, S.N.** (1993). The *Escherichia coli* *pcnB* gene promotes adenylation of antisense RNAI of ColE1-type plasmids in vivo and degradation of RNAI decay intermediates. *Proc Natl Acad Sci U S A* **90**, 6756-6760.
- Xu, R., Ye, X., and Quinn Li, Q.** (2004). *AtCPSF73-II* gene encoding an Arabidopsis homolog of CPSF 73 kDa subunit is critical for early embryo development. *Gene* **324**, 35-45.
- Yang, J., and Hunt, A.G.** (1994). Immunological characterization of plant polyadenylate-binding proteins. *Plant Science* **99**, 161-170.
- Yao, Y., Song, L., Katz, Y., and 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**, 2277-2278.

- Yeung, G., Choi, L.M., Chao, L.C., Park, N.J., Liu, D., Jamil, A., and Martinson, H.G.** (1998). Poly(A)-driven and poly(A)-assisted termination: two different modes of poly(A)-dependent transcription termination. *Mol Cell Biol* **18**, 276-289.
- Zarudnaya, M.I., Kolomiets, I.M., Potyahaylo, A.L., and Hovorun, D.M.** (2003). Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. *Nucleic Acids Res* **31**, 1375-1386.
- Zenklusen, D., Vinciguerra, P., Wyss, J.C., and Stutz, F.** (2002). Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. *Mol Cell Biol* **22**, 8241-8253.
- Zhao, J., Kessler, M.M., and Moore, C.L.** (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**, 10831-10838.
- Zhao, J., Hyman, L., and Moore, C.** (1999a). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **63**, 405-445.
- Zhao, J., Kessler, M., Helmling, S., O'Connor, J.P., and Moore, C.** (1999b). Pta1, a component of yeast CF II, is required for both cleavage and poly(A) addition of mRNA precursor. *Mol Cell Biol* **19**, 7733-7740.
- Zhao, W., and Manley, J.L.** (1998). Deregulation of poly(A) polymerase interferes with cell growth. *Mol Cell Biol* **18**, 5010-5020.
- Zhelkovsky, A., Helmling, S., and Moore, C.** (1998). Processivity of the *Saccharomyces cerevisiae* poly(A) polymerase requires interactions at the carboxyl-terminal RNA binding domain. *Mol Cell Biol* **18**, 5942-5951.
- Zhelkovsky, A.M., Kessler, M.M., and Moore, C.L.** (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**, 26715-26720.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W. (2004).
GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox.
Plant Physiol **136**, 2621-2632.

VITA

Name Kevin Patrick Forbes

Date of Birth February 13th, 1973

Place of Birth San Jose, California

Education

1991-1995: B.S. University of Mary, Bismarck, North Dakota

Major: Biology

Minor: Chemistry

Research Experience

1996-1997: Process Technician

Genencor International

Palo Alto, California

1995-1996: Media Manufacturing Associate II

Baxter Healthcare Corporation-Hyland Division/Biotechnology

Group

Hayward, California

1995: Student Intern

USDA-ARS; Northern Great Plains Research Laboratory

Mandan, North Dakota

1992-1994: Lab Technician

Large Scale Biology, Inc. (Formerly Biosource Technologies)

Morgan Hill, California

Honors

Phillip Morris Fellowship, University of Kentucky, 1998-2000

Dharmacon Award (Honorable Mention-Poster Presentations), Rustbelt RNA Meeting
2000

List of Publications

1. **Forbes K.P., Addepalli B., Hunt A.G.** (2005) An Arabidopsis Fip1p homologue stimulates poly(A) polymerase and provides conceptual links with a number of other polyadenylation factor subunits. Manuscript in preparation.
2. **Addepalli, B., Meeks, L.R., Forbes, K.P., and Hunt, A.G.** (2004). Novel alternative splicing of mRNAs encoding poly(A) polymerases in Arabidopsis. *Biochim Biophys Acta* 1679, 117-128.
3. **Elliott, B.J., Dattaroy, T., Meeks-Midkiff, L.R., Forbes, K.P., and Hunt, A.G.** (2003). An interaction between an Arabidopsis poly(A) polymerase and a homologue of the 100 kDa subunit of CPSF. *Plant Mol Biol* 51, 373-384.
4. **Hunt, A.G., Meeks, L.R., Forbes, K.P., Das Gupta, J., and Mogen, B.D.** (2000). Nuclear and chloroplast poly(A) polymerases from plants share a novel biochemical property. *Biochem Biophys Res Commun* 272, 174-181.