An *Arabidopsis* Fip1 Homolog Interacts with RNA and Provides Conceptual Links with a Number of Other Polyadenylation Factor Subunits*5*

Received for publication, October 7, 2005, and in revised form, November 2, 2005. Published, JBC Papers in Press, November 10, 2005, DOI 10.1074/jbc.M510964200

Kevin P. Forbes, Balasubrahmanyam Addepalli, and Arthur G. Hunt1

From the Plant Physiology, Biochemistry, and Molecular Biology Program, Department of Plant and Soil Sciences, University of Kentucky, Lexington, Kentucky 40546-0312

The protein Fip1 is an important subunit of the eukaryotic polyadenylation apparatus, since it provides a bridge of sorts between poly(A) polymerase, other subunits of the polyadenylation apparatus, and the substrate RNA. In this study, a previously unreported *Arabidopsis* Fip1 homolog is characterized. The gene for this protein resides on chromosome V and encodes a 1196-amino acid polypeptide. Yeast two-hybrid and *in vitro* assays indicate that the N-terminal 137 amino acids of the *Arabidopsis* Fip1 protein interact with poly(A) polymerase (PAP). This domain also stimulates the activity of the PAP. Interestingly, this part of the *Arabidopsis* Fip1 interacts with *Arabidopsis* homologs of CstF77, CPSF30, CFIm-25, and PabN1. The interactions with CstF77, CPSF30, and CFIm-25 are reminiscent in various respects of similar interactions seen in yeast and mammals, although the part of the *Arabidopsis* Fip1 protein that participates in these interactions has no apparent counterpart in other eukaryotic Fip1 proteins. Interactions between Fip1 and PabN1 have not been reported in other systems; this may represent plant-specific associations. The C-terminal 789 amino acids of the *Arabidopsis* Fip1 protein were found to contain an RNA-binding domain; this domain correlated with an intact arginine-rich region and had a marked preference for poly(G) among the four homopolymers studied. These results indicate that the *Arabidopsis* Fip1, like its human counterpart, is an RNA-binding protein. Moreover, they provide conceptual links between PAP and several other *Arabidopsis* polyadenylation factor subunit homologs.

The polyadenylation of messenger RNAs in the nucleus is an important step in the biogenesis of mRNAs in eukaryotes. This RNA processing reaction adds an essential cis element, the poly(A) tail, to the 3′-end of a processed pre-mRNA. This process is also coupled with many other steps in mRNA biogenesis (1). Thus, some polyadenylation factors are associated with transcription factors and recruit parts of the polyadenylation apparatus to the transcription initiation complex (2). Polyadenylation is linked to pre-mRNA splicing in a number of ways. For example, interactions between the polyadenylation and splicing machineries are important for the definition of 3′-terminal exons in animal cells (3, 4). Other interactions help to modulate different processing fates for pre-mRNAs, thus contributing to the scope of alternative splicing and polyadenylation in eukaryotes. The polyadenylation apparatus interacts with the C-terminal domain of the large subunit of RNA polymerase II (5–9) and with factors that play roles in transcription termination (10); these interactions suggest a central role for 3′-end processing in the termination of transcription by RNA polymerase II and subsequent recycling of polymerase II for new rounds of initiation.

Polyadenylation is mediated by a multifactor complex in yeast and mammals. 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) to the processed pre-mRNA. 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),2 cleavage-stimulatory factor (CstF), and cleavage factors I and II (CFIm and CFIml, respectively) (1). Each of these factors in turn consists of several distinct subunits. With the exception of CFIm1 (the two subunits of which are not obviously apparent in the yeast proteome), yeast possesses a similar array of polyadenylation factor subunits that form a somewhat different set of chromatographically distinct factors, namely cleavage and polyadenylation factor and cleavage factor I (1). Interestingly, the enzyme that adds poly(A) (poly(A) polymerase, or PAP) is part of the cleavage and polyadenylation factor in yeast nuclear extracts but fractionates largely as a separate protein in mammalian extracts. Whereas there are differences in the chromatographic behaviors of the complexes in mammals and yeast, most of the functions of the individual subunits seem to be similar. Besides the PAPs, this includes RNA binding by CPSF160, CPSF30, and CstF64 and their yeast counterparts (Yhh1p, Yth1p, and Rna15p, respectively) (11–17) and bridging between factors (CstF77 and its yeast counterpart RNA14p, hFip1p and the yeast counterpart Fip1p) (18–22). Of particular interest are the protein Fip1p and its human counterpart, hFip1. In yeast, Fip1p appears to be the principal means by which PAP is linked with the rest of the cleavage and polyadenylation factor. Fip1p is the only polyadenylation factor subunit that has been shown to interact with PAP (23). Fip1p also interacts with Yhh1p, Yth1p, P62p, and RNA14, components of the two major polyadenylation complexes (cleavage and polyadenylation factor and cleavage factor I) in yeast (13, 22). The human homolog, hFip1, interacts with PAP and CPSF160 (the mammalian counterpart of Yhh1p) and has been recently recognized as an authentic subunit of...
CPSF (18). The yeast and human Fip (factor interacting with golv(A) polymerase) proteins have somewhat contrasting properties; the yeast protein lacks an RNA-binding domain and inhibits the nonspecific activity (24) (e.g. activity on RNA substrates that do not possess authentic polyadenylation signals) of PAP, whereas the human Fip1 can bind RNA and stimulates PAP activity (18). Kaufmann et al. (18) 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 cleavage and polyadenylation factor, may stimulate PAP much as does the human Fip1. In this light, the functioning of Fip in the two systems may be relatively conserved, serving to promote PAP activity via some sort of tethering to the RNA substrate.

Plant polyadenylation signals have been well characterized and found to be distinct in many ways from their mammalian and fungal counterparts (25, 26). However, the properties of the plant polyadenylation apparatus are less well understood. Bioinformatic analysis of the Arabidopsis genome indicates that plants possess genes that encode most of the subunits of the mammalian polyadenylation complex.1 Insertions in two of these (encoding homologs of CPSF100 and CPSF73, respectively) lead to embryo lethality (27, 28). The Arabidopsis CPSF100 protein interacts with at least one of the four PAPs (29), an interaction that seems to be unique to the plant polyadenylation machinery. There is a degree of novelty in the properties of the Arabidopsis homologs of the CstF subunits, in that one of the three proteins (AtCstF50) does not interact with AtCstF77 (30), in contrast to what has been shown in the mammalian complex (31). The CstF64-CstF77 interaction does seem to be evolutionarily conserved (30). Arabidopsis possesses four PAP-encoding genes (32). Three of the corresponding PAP isosforms are similar in size to each other, whereas the fourth is much smaller, lacking an obvious nuclear localization signal. An Arabidopsis homolog of the yeast polyadenylation factor subunit Psf2p (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 CCA-encoding RNAs and consequently to regulate flower timing (33).

As mentioned above, the yeast and mammalian Fip1 proteins are important bridging factors in the polyadenylation complex, providing links between PAP, RNA, and other multisubunit complexes. These links presumably recruit PAP or stabilize the association of PAP with the apparatus and may contribute to the differential recognition of various RNAs by the 3′-processing machinery. In this report, we present a characterization of an Arabidopsis Fip1 isoform (geneid At5g58040, termed AtFip1(V)). We find that this protein binds RNA; interacts with the Arabidopsis polyadenylation factor subunits AtPAP, AtCstF77, AtCPSF30, AtCFI-25m, and AtPabN1; and stimulates nonspecific PAP activity. The abilities to bind RNA; interact with AtPAP, AtCstF77, and AtCPSF30; and stimulate PAP activity are properties that the AtFip1(V) shares with its human counterpart. The interaction with AtCFI-m-25 may also be analogous to a recently reported CFI-m-hFip1 interaction (47). However, the interaction with AtPabN1 has not been reported in other systems and may reflect a unique aspect of the plant polyadenylation machinery. Taken together, these results indicate that AtFip1(V) coordinates a number of polyadenylation factor subunits with PAP and with RNA.

EXPERIMENTAL PROCEDURES

Plant Materials—Arabidopsis thaliana seed was obtained from Lehle Seeds. Seeds were germinated and plants were cultivated in the greenhouse or growth room 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 medium (500 mg of sucrose, 215.5 mg of Murashige and Skoog Basal Medium (Sigma), and 25 mg of MES was inoculated with 30–40 sterilized seeds and grown for 2–3 weeks at room temperature under a 12-h light, 12-h dark cycle. For scoring T-DNA insertion plants, this medium was supplemented with kanamycin (50 μg/ml).

PCR Genotyping of Salk T-DNA Lines—The T-DNA insertion line, SALK_087117, was generated by the Salk Institute (available on the World Wide Web at signal.salk.edu/cgi-bin/tdnexpress), and seed was obtained from the Arabidopsis Biological Resource Center. Seeds from the stock center were germinated, and the plants were allowed to self-fertilize so as to generate a bulk stock of seed. Seeds from this bulked population were germinated, and plants were cultivated in the greenhouse or growth room. DNA from single leaves was extracted using a rapid homogenization plant leaf DNA amplification kit (Cartagen). Extracted DNA was used in a typical PCR (see supplemental materials) with AtFip1(V)-specific primers 5′-GIFIP and 3′-INT (Table 1), yielding a 1-kb genomic DNA fragment, and with a combination of AtFip1(V)-and T-DNA- specific primers (AtFip1(V)-specific primer 5′-INT and a T-DNA-specific primer Lbbl, yielding a ~500-bp fragment) for verifying the presence and location of the T-DNA insertion.

RNA Isolation from Arabidopsis and Generation of First Strand cDNA—Total RNA was isolated from Arabidopsis leaves using either an SV Total RNA Isolation Kit (Promega), RNaseasy® Plant Mini Kit (Qiagen), or Trizol (Invitrogen), following the manufacturers’ 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 AtFip1(V) cDNAs—cDNAs derived from the Arabidopsis AtFip1(V) gene were isolated by PCR and RT-PCR. Potential Fip1-encoding genes were identified in the Arabidopsis genome (available on the World Wide Web at www.Arabidopsis.org/home.html) with TBLASTN and BLASTP (34) using the human and yeast Fip1 amino acid sequences as search queries. Based on the results, primers were designed to amplify the cDNA coding region of the AtFip1(V) gene (Table 1). Various combinations of these primers were used in PCRs with first strand cDNA or with a 3–6-kb cDNA expression library for Arabidopsis (CD4-16; ABRC-DNA Stock Center (35)) as templates. PCR products were subcloned into pGEM-T Easy vector and sequenced by automated sequencing (ABI Prism 310 Genetic Analyzer; PerkinElmer Life Sciences) using the BigDye Terminator Cycle Sequencing Ready Reaction kit (ABI prism) and T7 and SP6 primers.

Three pGEM clones that represent the 5′-end (bases 1–1478), middle region (bases 1220–2692), and 3′-end (bases 2672–3588) of the full-length coding region, respectively, were generated. Full-length clones were assembled in pGEM using common restriction enzyme sites; the C-terminal domain of AtFip1(V) was amplified and subcloned using full-length cDNAs. One clone containing the 5′-end of AtFip1(V) contained a premature stop codon after residue 137 of the protein, presumably due to a PCR error. This stop codon conveniently delimits the highly divergent part of the N terminus of AtFip1(V) (Fig. 1A); for this reason, it was selected to produce yeast two-hybrid and protein expression clones. Cloning details are provided in supplemental materials.

For expression analysis of AtFip1 genes in different Arabidopsis tissues, PCR amplification was done with 1.5 μl of first strand cDNA
Cloning of Arabidopsis cDNAs Encoding Arabidopsis Polyadenylation Factor Subunits—Data base searches of the Arabidopsis genome using TBLASTN and BLASTP with the yeast Pfs2 and human CstF50, -64, and -77 subunits as well as mammalian CFIm-25 and PabN1 as search queries identified potential homologs for each subunit. Based on the sequence information, primers were designed to amplify the cDNA coding regions of these genes (Table 1). Clones were generated by PCR or RT-PCR, subcloned in pGEM, and sequenced. In cases where partial clones were produced, full-length cDNAs were assembled using com-

| Gene or use | Primer designation | Sequence (5’→3’) |
|-------------|-------------------|-----------------|
| **AtFip1(V)** | 5’FL | CCGCATGGAAGGAAGCAGATGAGTTCGGA |
| | 5’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 5’INT1 | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’INT1 | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’INT2 | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’FL | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW2 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW3 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |

**Table 1**

| Gene or use | Primer designation | Sequence (5’→3’) |
|-------------|-------------------|-----------------|
| **AtCstF50** | 5’FL | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 5’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’FL | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW2 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW3 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |

**Table 1**

| Gene or use | Primer designation | Sequence (5’→3’) |
|-------------|-------------------|-----------------|
| **AtCstF64** | 5’FL | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 5’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’FL | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW2 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW3 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |

**Table 1**

| Gene or use | Primer designation | Sequence (5’→3’) |
|-------------|-------------------|-----------------|
| **AtCFI-25** | 5’FL | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 5’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’INT | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’FL | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 5’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW1 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW2 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |
| | 3’GW3 | TTAATACGACTCACTATAGGGAAACACGCTGAAATCACCAGTCTCTCT |

**Table 1**

| Gene or use | Primer designation | Sequence (5’→3’) |
|-------------|-------------------|-----------------|
| **AtPABN1** | 5’FL | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |
| | 3’FL | CCGGATCCCGAGTGTACGAGCAAGAAGTGGGGA |

**Table 1**

| Gene or use | Primer designation | Sequence (5’→3’) |
|-------------|-------------------|-----------------|
| **CaMV RNAs** | 1, 2, 3 | T7-STS |
| 4 | T7-NUE |
| 1, 4 | CaMV61/80R |
| 3 | CaMV20/1R |
| 2 | CaMV 50/70R |
| rbcS-E9 RNA | T7-E9 |
| E9 61–80 | E9 61–80 |

(ProSTAR; Stratagene) added to 100 ng of each primer, 0.8 mM dNTPs, 5.0 μl of Ultra HF PCR buffer (Stratagene), and 2.5 units of Pfu Turbo DNA polymerase (Stratagene) in a 50-μl reaction.
Interactions Involving Arabidopsis Fip1

To produce the GST-FipC fusion protein, 10-ml overnight cultures of recombinant BL21-SI cells (Invitrogen) were used to inoculate 200 ml of media, and cells were grown at 37 °C until an A600 of 0.8. Expression of the fusion protein genes was then induced by the addition of NaCl to a final concentration of 0.3 M. After additional growth for 3 h at 37 °C, cells were harvested and resuspended in 5 ml of lysis buffer. Cells were disrupted by sonication (three bursts, 30 s each), and debris was removed by centrifugation. To purify GST fusion proteins, lysates were then incubated for 1 h with glutathione-Sepharose beads (that had been 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.

For the poly(G) affinity purification of GST-FipC, 100 μl of poly(G)-agarose (Sigma) was pelleted, washed three times with NEB, and briefly pelleted by centrifugation. The pellet was then incubated with −30 μg of GST-FipC (in 100 μl of NEB) at 30 °C for 20 min. The agarose was then briefly pelleted and washed three times with NEB. Thirty μl of SDS-sample buffer was added to the agarose, boiled for 10 min, and then briefly pelleted. Twenty μl of the sample was separated by SDS-PAGE.

In Vitro Determination of Protein-Protein Interactions—To measure protein-protein interactions, 20-μl aliquots of E. coli cell extracts containing maltose-binding protein (MBP) or CBD fusion proteins were mixed with 10 μl of cell extract containing the GST-tagged target fusion protein; reactions were brought to a final volume of 100 μl with control E. coli lysate. All reactions had 0.1% Nonidet P-40. After 30 min at 30 °C, these mixtures were added to amyllose beads (New England Biolabs) or calmodulin affinity resin (Stratagene) that had been pretreated (for 30 min) with control cell extract, 100 μg/ml bovine serum albumin, and 0.1% Nonidet P-40. After 5 min, the matrix was collected and washed five times with lysis buffer containing 0.1% Nonidet P-40, and proteins that were retained were analyzed by SDS-PAGE and immunoblotting. GST fusion proteins were detected using alkaline phosphatase-conjugated anti-GST antibodies (Sigma).

Electrophoretic Mobility Shift Assays—For electrophoretic mobility shift assays, RNA and proteins were incubated at 30 °C in NEB supple-

JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 281 • NUMBER 1

JANUARY 6, 2006

179

would enable the production of GST- or histidine-tagged proteins, respectively. In addition, AtCstF77, AtCPSF30, and AtPabN1 were subcloned into pMAL-C2x (New England Biolabs) so as to produce maltose-binding protein fusions. AtCstF1m-25 was cloned into a Gateway-converted form of pCgi-72 (In Vitro Technologies) that had been pretreated (for 30 min) with control cell extract, 100 μg/ml bovine serum albumin, and 0.1% Nonidet P-40. After 5 min, the matrix was collected and washed three times with lysis buffer containing 0.1% Nonidet P-40, and proteins that were retained were analyzed by SDS-PAGE and immunoblotting. GST fusion proteins were detected using alkaline phosphatase-conjugated anti-GST antibodies (Sigma).
ment with KCl (final concentration 60 mM) and MgCl₂ (final concentration 1.2 mM). After 20 min, 0.1 volume of gel loading buffer II (AMBION) diluted in NEB (1:40 ratio of GLBII to NEB) was added to the reactions, and samples were loaded onto nondenaturing gels (4% acrylamide, 0.08% bisacrylamide, cast in TBE). After electrophoresis, gels were transferred to Whatmann paper and dried under vacuum. Dried gels were developed on a storage phosphor screen and visualized by ImageQuant.

RESULTS

To identify possible Arabidopsis counterparts of Fip1, a BLAST search of the Arabidopsis genome was conducted using the human Fip1 amino acid sequence as a query. This process yielded two candidate genes located on chromosomes 5 (At5g58040) and 3 (At3g66652). The chromosome 3 gene had been identified earlier as a possible Fip1 homolog (18), but the chromosome 5 gene had not been commented on in the earlier study. Amino acid sequence alignments (not shown) and consideration of the general putative domain organization of the two Arabidopsis proteins and hFip1 (Fig. 1A) suggested that the chromosome 5 gene encoded a likely Fip1 homolog (AtFip1(V)). Accordingly, this gene was selected for the studies that follow. As a prelude, cDNAs that span the open reading frame encoded by this gene were cloned and sequenced; this process allowed revision of the predicted exon structure (Fig. 2). Briefly, recombinant Arabidopsis poly(A) polymerase (the product of At4g32850 (32)) and the histidine-tagged N-terminal 137 amino acids of AtFip1(V) were assayed for incorporation of label from [α-32P]ATP into poly(A). Reactions were performed in a volume of 30 μl at 30 °C for the times indicated. Reactions were terminated by incubating with 2.5 μl of STOP solution (2.5% SDS, 135 mM EDTA, 5 mg/ml Proteinase K) for 10 min at 37 °C, and the labeled RNAs were recovered and separated on sequencing gels. Dried gels were exposed to a phosphorimaging screen, autoradiographs were developed, and the quantities of labeled polymer were determined with ImageQuant software (Amersham Biosciences). The results were plotted as arbitrary values against time.
Fip1 protein (Fig. 1C). The sequence similarity outside of this domain is much more modest (not shown). However, several identifiable trends in other parts of the protein could be identified, and these trends correspond to some degree to domains that are present in hFip1 and its yeast counterpart (Fig. 1A).

RT-PCR analysis of RNA isolated from leaves, stems, flower tissues, and roots indicated that this gene is expressed in all of these tissues (Fig. 2B). Retrieval and inspection of the expression data of this gene from Genevestigator (available on the World Wide Web www.genevestigator.ethz.ch (42)) corroborated the RT-PCR data; specifically, there was little difference in the expression of this gene in the different tissues analyzed in the microarray data set that is available (not shown). These data indicate that At5g58040 is expressed ubiquitously in Arabidopsis.

A T-DNA insertion within the AtFip1(V) locus, SALK_087117, could be found in the Salk Institute T-DNA insertion library data base (available on the World Wide Web at signal.salk.edu/cgi-bin/tdnaexpress (43)). According to the data base, the T-DNA insertion lies within the sixth exon (Fig. 2A); this was confirmed using primers (illustrated in Fig. 2A) situated within the T-DNA and in the expected flanking sequences of the AtFip1(V) gene (not shown). PCR genotyping of 42 progeny derived from self-crosses of kanamycin-resistant individuals from the ABRC, using pairs of primers specific for AtFip1(V) that flank the T-DNA insert revealed that all 42 plants had at least one of the wild-type alleles. Further screening was performed by plating seeds from several of these plants on kanamycin-containing media; this would identify individuals containing the complete T-DNA and serve as a phenotypic confirmation of the PCR genotyping. Media containing the selectable marker associated with the T-DNA. Seeds from 23 lines were germinated on the selectable media. Of these 23 lines, 15 were heterozygous for the selectable marker and eight lines showed wild-type (kanamycin sensitive) background. These results indicate that AtFip1(V) is an essential gene.

In mammals and yeast, Fip1 interacts with a subset of other polyadenylation factor subunits (18, 23). To examine this aspect of Fip1 function in Arabidopsis, the ability of AtFip1(V) to interact with a battery of other Arabidopsis polyadenylation factor subunits was measured. For these assays, two parts of AtFip1(V) were cloned into yeast two-hybrid vectors; one part consisted of the N-terminal 137 amino acids of the protein, whereas the other consisted of the C-terminal 789 amino acids. These portions were tested for interactions with the Arabidopsis homologs of CstF77, CstF64, CstF50, CPSF160, CPSF100, CPSF73 (both homologs) (28), CPSF30, FY, PabN, and PAP (specifically, the isoform encoded by the Arabidopsis PAP gene situated on chromosome IV (32)). In these assays, no interactions involving the C-terminal portion of AtFip1(V) could be ascertained (not shown). However, these assays also suggested that the first 137 amino acids of AtFip1(V) interact with the Arabidopsis homologs of PAP, CPSF30, CstF77, CFIm-25, and PabN1 (summarized in Table 2). In contrast, no discernible interaction was observed with the Arabidopsis homologs of CPSF160, CPSF100, CPSF73 (neither of the two distinct isofoms), CstF64, CstF50, and FY (Table 2). To account for the possibility that these negative results might be due to the omission of the central part of AtFip1(V), the assays with the Arabidopsis homologs of CPSF160, CPSF100, CPSF73 (both isofoms), CstF64, CstF50, and FY were repeated with another AtFip1(V) construct containing the N-terminal 492 amino acids of AtFip1(V). Again, no discernible interactions were observed. With the exception of FY, all of the partners for which negative results were obtained with AtFip1(V) have yielded positive results in other tests, indicating that the two-hybrid constructs enable the production of functional proteins in yeast. In addition, all of the interacting partners identified in this screen (e.g. CPSF30, CstF77, CFIm-25, and PabN1) have yielded negative results when tested in other combinations, indicating that these various proteins have specificity with respect to the proteins with which they interact in yeast cells.

The positive results of the two-hybrid assays were corroborated with a second measure, namely the measurement of copurification of AtFip1(V) with other polyadenylation factor subunits. Initial focus was placed on PAP, since this interaction is central to the functioning of Fip1 in yeast and mammals (18, 23, 44). For these assays, E. coli extracts containing a fusion protein (GST-PAP in Fig. 3A) that contains amino acids 130–503 of the Arabidopsis chromosome IV-encoded PAP were mixed with extracts containing an MBP fusion with the N-terminal part of AtFip1(V) that was implicated in the two-hybrid assays (MBP-FipN; Fig. 3A). (This part of the PAP is the most highly conserved among the

![FIGURE 2. A, an exon-exon map of the AtFip1 gene that resides on chromosome V. The thin lines represent introns, and the thick lines indicate exons. The arrows indicate the locations of oligonucleotide primers used in B. The location of the T-DNA insert within the gene is located above, B, RT-PCR analysis of AtFip1-V expression. RNA isolated from the indicated tissue (labeled above; flower (f), leaf (l), stem (s), and root (r)) was analyzed by RT-PCR, with primers (5INT1 and 3INT2 in A) that flank the last intron of chromosome V Fip1. For comparison, primers specific for the Arabidopsis tubulin gene (At5g62690) were used. When reverse transcription was omitted from these reactions, no amplification products were made (data not shown).]

**TABLE 2**

| Polyadenylation factor subunit | A. thaliana gene designation* | Interaction with Fip1** |
|-------------------------------|-------------------------------|------------------------|
| CPSF160                       | At5g51660                     | –                      |
| CPSF100                       | At5g23880                     | –                      |
| CPSF73(I)                     | At1g61010                     | –                      |
| CPSF73(II)                    | At2g01730                     | –                      |
| CPSF30                        | At1g30460                     | +                      |
| CstF77                        | At1g17760                     | +                      |
| CstF64                        | At1g1800                      | –                      |
| CstF50                        | At5g69040                     | +                      |
| FY                            | At5g13480                     | +                      |
| CFIm-25                       | At4g29820                     | +                      |
| PabN1                         | At5g51120                     | +                      |
| PAP(V)                        | At4g32850                     | +                      |

* Arabidopsis gene designation for the protein denoted in the leftmost column.
** Interaction in the yeast two-hybrid assay using the N-terminal 137 amino acids of AtFip1(V). –, no interaction (e.g. the number of colonies growing after 5 days on ALW-selective medium was less than 10% of those seen on LW medium after 3 days), +, positive for the interaction (e.g. the number of colonies growing after 5 days on ALW-selective medium was between 50 and 200% of those seen on LW medium after 3 days).
Interactions Involving Arabidopsis Fip1

FIGURE 3. AtFip1(V) interacts with PAP in vitro. A, illustration of the structure of the GST-PAP and MBP-FipN fusion proteins. B, results of in vitro interaction assays. The top and middle panels show immunoblots developed with anti-GST antiserum; the bottom panels show stained gels showing the quantities of MBP and MBP-FipN present in the experiment. Roughly 70% of the input GST sample for each experiment is shown under input. The panels under baits show the GST-tagged proteins that copurify with MBP or MBP-FipN, respectively. Copurification of GST and MBP was not tested, since GST did not bind to MBP-FipN. In the bottom panel for the MBP-FipN samples, bands denoted with an asterisk indicate the major full-sized MBP-FipN polypeptides, as determined by immunoblotting with anti-MBP antiserum (not shown), present after the affinity purification. nd, experiment was not performed.

four Arabidopsis PAP isoforms (32).) Control reactions contained MBP. After purification using amylose resin, the copurification of AtPAP[IV] with the MBP "baits" was evaluated by immunoblotting, using anti-GST antibodies. The results of this experiment corroborated those of the two-hybrid analysis. Specifically, GST-AtPAP[IV] copurified with MBP-FipN but not MBP (Fig. 3B, row of samples designated GST-PAP). The GST-AtPAP[IV] protein was not retained in experiments using just MBP, indicating that the copurification was not due to interactions involving the MBP portion of the baits. GST itself did not copurify with MBP-FipN (Fig. 3B, row of samples designated GST), ruling out the possibility that the copurification of GST-AtPAP[IV] and MBP-FipN was due to interactions involving the GST part of the GST-AtPAP[IV] fusion protein. The absence of copurifying GST-AtPAP[IV] in the MBP sample was not due to a low quantity of MBP in the experiments, since ample quantities of the bait proteins could be detected after SDS-PAGE and staining (Fig. 3B, row of samples designated stained proteins).

For other assays, extracts containing a GST-AtFip1(V) fusion protein (GST-FipN in Fig. 4A) that contained the N-terminal domain implicated in the two-hybrid tests were mixed with extracts containing MBP fusions with CPSF30, CstF77, or PabN1 (MBP-CstF77, MBP-CPSF30, and MBP-PABN1; Fig. 4A) or extracts containing a calmodulin-binding protein (CBD)-CFIm-25 fusion (CFI-25-CBD in Fig. 4A). Control reactions contained MBP or a CBD-chloramphenicol acetyltransferase fusion (CAT-CBD in Fig. 4A). After purification using amylose resin, the copurification of AtFip1(V) with this battery of "baits" was evaluated by immunoblotting, using anti-GST antibodies. The results of these experiments corroborated those of the two-hybrid analysis. Thus, GST-FipN copurified with MBP-CstF77, MBP-CPSF30, MBP-PabN1, and CFI25-CBD after purification of the latter fusions on their respective affinity matrices (Fig. 4B, row of samples designated GST-FipN). The GST-FipN protein was not retained in experiments using just MBP or CAT-CBD, indicating that the copurification with the other "baits" was not due to interactions involving the MBP or CBD portions of the baits. GST itself did not copurify with MBP-CstF77, MBP-CPSF30, MBP-PabN1, and CFI25-CBD (Fig. 4B, row of samples designated GST), ruling out the possibility that the copurification of GST-FipN was due to interactions involving the GST part of the GST-FipN fusion protein. The absence of copurifying GST-FipN in the MBP and CAT-CBD samples was not due to a low quantity of the bait proteins in the experiments, since ample quantities of the various bait proteins could be detected after SDS-PAGE and staining (Fig. 4B, row of samples designated stained proteins).

The C-terminal part of the human Fip1 protein contains an arginine-rich region that is suggestive of an RNA-binding protein; consistent with this, the human Fip1 protein does bind RNA and shows a preference for U-rich sequences (18). The C-terminal part of AtFip1(V) also contains an arginine-rich region, raising the possibility that the Arabidopsis protein also binds RNA. To test this, the C-terminal portion of AtFip1(V) (Fig. 5A) was produced as a GST fusion protein and assayed for RNA binding activity. The protein that was produced in E. coli was heterogeneous, with some full-sized polypeptide as well as several GST-containing polypeptides that were apparently truncated at the C terminus of the AtFip1(V) part of the fusion protein (based on the mobilities of these forms and their cross-reactivities with anti-GST antibodies (Fig. 5A)). This heterogeneity could not be reduced beyond a point with any of a number of strategies (not shown), so the preparations shown in Fig. 5A were analyzed for RNA binding. For this, uniformly labeled RNA
was incubated with purified protein preparations for 20 min at 30 °C and aliquots of the binding mixtures separated on nondenaturing acrylamide gels. The RNA used in these studies contains a functional poly(A) site derived from the pea rbcS-E9 gene. As shown in Fig. 5B, purified GST-AtFip1(V) was able to bind this RNA; in contrast, purified GST lacked this activity. To assess possible sequence preferences of FipC for RNA binding, a range of excesses of each of the four homopolymers (poly(A), poly(C), poly(G), and poly(U)) was added to binding reactions. As shown in Fig. 6, RNA binding was inhibited by even modest excesses of poly(G). Slight inhibition was seen with large excesses of poly(A), but poly(U) and poly(C) did not inhibit binding to the labeled RNA.

The array of GST-AtFip1(V) polypeptides that can be detected in E. coli extracts represent a “deletion” series of a sort, in that they define a set of polypeptides that share a common N terminus (the GST part of the fusion protein) and differ at their C termini. This afforded an opportunity to map, to a first approximation, the amino acid sequences in the C terminus that are required for RNA binding. Accordingly, the mixture of GST-AtFip1(V) polypeptides was incubated with poly(G)-Sepharose, and proteins that bound were analyzed by SDS-PAGE and immunoblotting. The results showed that a subset of the GST-AtFip1(V) proteins were able to bind poly(G)-Sepharose (Fig. 7). Based on the mobilities of those polypeptides that bound as well as those that did not, it is concluded that RNA binding requires the arginine-rich domain of AtFip1(V).

Plant polyadenylation signals consist of a distinctive set of cis elements, elements situated relatively far 5' downstream of the poly(A) site (FUEs), elements situated within 30 nt of the poly(A) site (NUEs), and the cleavage/polyadenylation site itself (CS). To examine the possible preference of AtFip1(V) for any of these classes of elements, different parts of the cauliflower mosaic virus (CaMV) polyadenylation signal were tested for the ability to compete with the labeled rbcS-E9 RNA for binding to AtFip1(V). The CaMV poly(A) signal was chosen, because, unlike most plant 3'-untranslated regions, the CaMV signal directs mRNA 3'-end formation at a single site and thus consists of a relatively simple array of cis elements (45, 46). Four RNAs were tested as competitors for binding of AtFip1(V) to the labeled rbcS-E9 RNA. One of the competitors (labeled A in Fig. 8A) contained all sequences extending from 181 nt 5' to 80 nt 3' of the CaMV polyadenylation site. A second site (labeled B) contained sequences from 181 to 50 nt 5' of the poly(A) site; this portion contains the FUE but lacks the NUE and poly(A) site itself. A third RNA (labeled C) extended from 181 nt 5' of the poly(A) site to the poly(A) site itself. A fourth RNA (labeled D) included sequences from 30 nt 5' to 80 nt 3' of the poly(A) site. As seen in Fig. 8A and B, a 50-fold molar excesses of RNAs 1, 2, and 3 all reduced binding of AtFip1(V) to the labeled RNA (the reduction is apparent as a dramatic increase in the quantity of free RNA in the experiment). In contrast, a similar excess of RNA 4 had no effect on the binding of the labeled RNA. This experiment indicates that AtFip1(V) has a decided preference for CaMV-derived RNAs that contain the FUE of the polyadenylation signal.

The human Fip1 protein is able to stimulate the nonspecific activity of PAP, and both the N-terminal (PAP-binding) and C-terminal (RNA-binding) portions of the protein are required for this stimulation (18). Attempts to recapitulate this aspect of the AtFip1(V)-AtPAP interaction were not successful, because full-length AtFip1(V) consistently copurified with E. coli nuclease to an extent that precluded the assay of PAP activity. However, in light of the observation that amino acids 1–137 of AtFip1(V) interacted with PAP, the effects of this domain on PAP activity were examined. For this, recombinant FipN was produced as a histidine-tagged protein (Fig. 9A). The Arabidopsis PAP (IV) polyadenylates the RNA template, and the quantity of product increases over time (Fig. 9B). In the presence of FipN, the overall quantities of product were greater at all times, when compared with the PAP alone (Fig. 9B). This indicates that the N-terminal domain of AtFip1(V) that interacts with PAP can also stimulate the activity of PAP.

**DISCUSSION**

The results described in this paper indicate that an Arabidopsis Fip1 homolog (AtFip1(V)) possesses two distinct and separable domains. One of these, situated within the N-terminal 137 amino acids of the protein, is involved in interactions with a number of other polyadenylation factor subunits, the Arabidopsis counterparts of PAP, CstF77, CPSF30, CstFm25, and PabN1. The other, located roughly in the C-terminal 789 amino acids of the protein, possesses an RNA-binding domain that has a preference for poly(G) among the four homopolymers. This RNA-binding domain includes an arginine-rich region. Some of these properties are reminiscent of those of the human Fip1 protein (hFip1). The latter interacts with PAP (through as yet unidentified domains) and binds to RNA through an arginine-rich C-terminal domain (much as does AtFip1(V)). It has been proposed that these two properties are manifest in a singular function of hFip1, to tether PAP to a cleaved polyadenylation substrate, thereby overcoming the inherently low affinity of PAP for RNA. It is tempting to suggest a similar action for AtFip1(V).

The various protein-protein interactions involving the N terminus of AtFip1(V) are of interest for a number of reasons. First and foremost, these interactions provide a conceptual link between PAP and several other Arabidopsis homologs of eukaryotic polyadenylation factor subunits. Three of these interactions (between AtFip1(V) and PAP, CstF77, and CPSF30) have been reported in mammals and/or yeast (18, 23, 44) and are likely to be diagnostic of evolutionarily conserved functions. Interestingly, no interactions between AtCPSF160 and AtFip1(V) were apparent in the assays performed here. The significance of this observation is unclear at this time; it is possible that this particular interaction cannot be recapitulated in yeast cells, or it may be that, in contrast to what has been reported in mammals (18), these two proteins do not interact in plants.

An interaction analogous to that described here between AtFip1(V) and AtCstFm-25 has not been noted in mammalian systems. However, it has recently been reported that the Cfim polyadenylation factor can recruit CPSF to RNAs that contain sequence motifs recognized by...
CFIm (47). Moreover, CFIm and hFip1 can act together with PAP to promote sequence-specific polyadenylation of RNAs containing preferred CFIm binding sites (UGUAN) (47). Thus, there is precedence of a sort for the interaction between AtFip1(V) and AtCFIm-25. Whether this precedent extends to functions analogous to those reported for CFIm is unclear. It is possible that, as in mammals, CFIm in plants recognizes motifs related to UGUAN, such elements occur often in FUEs, and it has been proposed that CFIm in mammals recognizes positionally analogous sequences upstream of the canonical polyadenylation signal AAUAAA (47). This would be interesting, since AtFip1(V) itself has a modest preference for the FUE (Fig. 8); the presence of multiple FUE-recognizing proteins in a single complex would be consistent with the redundant nature of FUEs in plant polyadenylation signals (39, 45). However, as discussed below, there are other possibilities that must be considered. Resolution of these various scenarios will require further study.

The interactions involving the Arabidopsis PabN1 is at this time unprecedented, having not been reported in other systems. The significance of this is unclear, since it is possible that analogous interactions have not been assayed as has been done in this study. It is possible, however, that plant Fip1 proteins may engage in interactions that do not occur in other eukaryotes. Regardless, at this time, it is hard to ascertain the functional significance of the interaction between AtPabN1 and Fip1[V] in Arabidopsis. In mammals, PabN1 interacts directly with PAP and controls both the processivity of the enzyme and length of the added poly(A) tail (48, 49); neither of these biochemical properties require Fip1. Whether the Arabidopsis PabN1 is tethered to PAP through AtFip1(V) or has unanticipated functions in addition to (or besides) poly(A) length control is an issue for future investigation.

A surprising aspect of the set of interactions reported here is that they all involve a relatively small part of AtFip1(V). It seems unlikely that all five of the interacting proteins identified in this study can bind to the N-terminal 137 amino acids of AtFip1(V) at the same time. Rather, two alternative explanations for these interactions seem more plausible. It may be that all six proteins can and do reside in a single multimeric complex, albeit one that can be assembled through more than one combination of protein–protein interactions. Thus, for example, PabN1 might be held in such a hypothetical complex via interactions with AtFip1(V) or PAP. Alternatively, there may be more than one configuration of a complex that involves AtFip1(V), none of which would include all of the proteins identified in this study as AtFip1(V)-interacting partners. The different possible configurations might reflect a progression of sorts through the cleavage and polyadenylation process, with different proteins being recruited, through AtFip1(V), to the reaction at different steps. It is also possible that there exists different stable complexes that include AtFip1(V). Different complexes might be responsible for recognition of different sets of polyadenylation signals or for linking 3′-end formation with other processes within the nucleus.

The RNA binding properties of AtFip1(V) are interesting. This protein displays a marked preference for poly(G) among the four RNA homopolymers. Of the three components of a plant polyadenylation signal, only the so-called FUE has a bias toward G content (those FUEs that have been experimentally determined are rich in UG-containing motifs (25, 26)). Thus, one interpretation of the homopolymer competition studies is that AtFip1(V) binds to the FUE in a plant polyadenylation signal. This hypothesis is supported by the different abilities of the CaMV-derived RNAs used in this study to bind to AtFip1(V). Specifically, RNAs that contain the FUE of the CaMV poly(A) signal were effective competitors with the labeled rbcS-E9 RNA for binding, but an RNA that lacked the FUE was unable to compete. This observation suggests that AtFip1(V) is an FUE-binding protein and that it may be the FUE recognition factor for polyadenylation in plants. The possibility that AtFip1(V) is an FUE-binding factor in turn suggests that one or more of its interacting protein partners may be involved in recognition of the NUE and/or cleavage site, the other two known cis elements for polyadenylation of plant mRNAs. Two of the interacting proteins identified in this study, AtCPSF30 and AtPabN1, are probably RNA-binding proteins. Additionally, AtCstF77 interacts with AtCstF64 (30), which is expected to bind RNA. In mammals, CFIm, which contains the mammalian homolog of AtCFI-25, is an RNA-binding factor, and the 25-kDa
subunit contacts the RNA (50, 51). None of the corresponding plant factors have been correlated with any of the three poly(A) signal cis elements; however, the convergence of so many possible RNA-binding activities on AtFip1(V) raises the possibility that one or more of these interacting partners may be involved in recognition of NUEs and cleavage sites.

Neither of the two functional domains identified in this study includes the conserved sequence signature that is found in Fip1 in mammals and yeast. This raises a question as to the function of this conserved motif. In yeast, the Fip1 signature is part of, or adjacent to, the part of Fip1p that is involved in interactions with Yth1p, the yeast counterpart of CPSF30 (44). The Fip1 signature is not essential for the interaction between AtFip1(V) and AtCPSF30, but our results do not address the possibility that additional contacts may be made between these two proteins via the Fip1 motif. There is no obvious sequence similarity between the N-terminal 137 amino acids of AtFip1(V) and the yeast Fip1p, but it remains a possibility that Yth1p also might engage in multiple contact with Fip1p.

Interestingly, the N-terminal domain of AtFip1(V) is, by itself, capable of stimulating the activity of the Arabidopsis PAP (Fig. 9). Although this characteristic is similar in some respects to the stimulatory effects of the human Fip1 protein, it is different in the sense that the stimulation of PAP by the human Fip1 requires both the N-terminal PAP-interacting part and the C-terminal RNA-binding domain (18). Technical difficulties preclude an analysis of the effects of the complete AtFip1(V) protein on PAP activity. However, if one aspect of the functioning of Fip1 pro-
Interactions Involving Arabidopsis Fip1

teins is to tether PAP to the RNA substrate (as has been proposed by Kaufmann et al. (18)), then our results suggest that PAP stimulation may be a multifaceted property, depending on RNA binding by Fip1 and on direct Fip1-PAP contacts for a modification of the inherent activity of PAP.

In summary, the AtFip1(V) protein possesses two functional domains that are distinct from the conserved Fip1 motif that it shares with other eukaryotic Fip1 proteins. One of these is involved in a number of interactions with other Arabidopsis polyadenylation factor subunits and provides conceptual links between these subunits and PAP. The other is an RNA-binding domain that binds with a preference for RNAs that contain functional FUEs. The properties of these two domains lend themselves to a model whereby the AtFip1(V) protein interacts with the FUE in the primary transcript and acts, through either a succession of different complexes or one of several distinct complexes, to effect the cleavage and polyadenylation of RNAs in the nucleus.

Acknowledgments—We are grateful for the technical assistance of Carol Von Lanken and thank Dr. Quinn Li (Miami University, Oxford, OH) for the generous gifts of plasmids with genes for Arabidopsis CPSF subunit homologs.

REFERENCES

1. Proudfoot, N., and O’Sullivan, J. (2002) Curr. Biol. 12, R855–R857
2. Dantonel, J. C., Murthy, K. G., Manley, J. L., and Tora, L. (1997) Nature 389, 399–402
3. Lutz, C. S., and Alwine, J. C. (1994) Genes Dev. 8, 576–586
4. Takagaki, Y., and Manley, J. L. (2000) Mol. Cell. Biol. 20, 585–594
5. Lanken and thank Dr. Quinn Li (Miami University, Oxford, OH) for the generous gifts of plasmids with genes for Arabidopsis CPSF subunit homologs.

REFERENCES

1. Proudfoot, N., and O’Sullivan, J. (2002) Curr. Biol. 12, R855–R857
2. Dantonel, J. C., Murthy, K. G., Manley, J. L., and Tora, L. (1997) Nature 389, 399–402
3. Lutz, C. S., and Alwine, J. C. (1994) Genes Dev. 8, 576–586
4. Takagaki, Y., and Manley, J. L. (2000) Mol. Cell. Biol. 20, 585–594
5. Lanken and thank Dr. Quinn Li (Miami University, Oxford, OH) for the generous gifts of plasmids with genes for Arabidopsis CPSF subunit homologs.