# Calmodulin Interacts with and Regulates the RNA-Binding Activity of an Arabidopsis Polyadenylation Factor Subunit[1][OA]

**Kimberly J. Delaney, Ruqiang Xu[2], Jingxian Zhang[2,3], Q. Quinn Li, Kil-Young Yun, Deane L. Falcone, and Arthur G. Hunt***

Department of Plant and Soil Sciences, University of Kentucky, Lexington, Kentucky 40546–0312 (K.J.D., J.Z., K.-Y.Y., D.L.F., A.G.H.); Department of Botany, Miami University, Oxford, Ohio 45056 (R.X., Q.Q.L.); and Department of Biological Sciences, University of Massachusetts, Lowell, Massachusetts 01845 (K.-Y.Y., D.L.F.)

The Arabidopsis (*Arabidopsis thaliana*) gene that encodes the probable ortholog of the 30-kD subunit of the mammalian cleavage and polyadenylation specificity factor (CPSF) is a complex one, encoding small (approximately 28 kD) and large (approximately 68 kD) polypeptides. The small polypeptide (AtCPSF30) corresponds to CPSF30 and is the focus of this study. Recombinant AtCPSF30 was purified from *Escherichia coli* and found to possess RNA-binding activity. Mutational analysis indicated that an evolutionarily conserved central core of AtCPSF30 is involved in RNA binding, but that RNA binding also requires a short sequence adjacent to the N terminus of the central core. AtCPSF30 was found to bind calmodulin, and calmodulin inhibited the RNA-binding activity of the protein in a calcium-dependent manner. Mutational analysis showed that a small part of the protein, again adjacent to the N terminus of the conserved core, is responsible for calmodulin binding; point mutations in this region abolished both binding to and inhibition of RNA binding by calmodulin. Interestingly, AtCPSF30 was capable of self-interactions. This property also mapped to the central conserved core of the protein. However, calmodulin had no discernible effect on the self-association. These results show that the central portion of AtCPSF30 is involved in a number of important functions, and they raise interesting possibilities for both the interplay between splicing and polyadenylation and the regulation of these processes by stimuli that act through calmodulin.

Alternative processing of pre-mRNAs is an important factor in sculpting the proteome of eukaryotes, including plants. Thus, in Arabidopsis (*Arabidopsis thaliana*), numerous instances of alternative splicing have been reported, a phenomenon that may affect as many as 10% of all genes (Kazan, 2003; Zhu et al., 2003; Iida et al., 2004). Likewise, many reports of regulated alternative processing have been made (Zhou et al., 2003). The scope of stimuli that have been linked to alternative splicing in plants is extensive, and includes light, environmental stress, development, and challenge by pathogens (Marrs and Walbot, 1997; Hayashi and Nishimura, 1999; Mano et al., 1999, 2000; Isshiki et al., 2000; Jordan et al., 2002; Kong et al., 2003; Staiger et al., 2003). The scope of possible alternative polyadenylation in plants is similarly broad; in one report, it is estimated that as many as 25% of all Arabidopsis genes are alternatively polyadenylated (Meyers et al., 2004). The large size of this estimate likely reflects the pervasive microheterogeneity that exists in the 3′-untranslated regions of plant genes (Dean et al., 1986; Manen and Simon, 1993; Hunt, 1994; Klahre et al., 1995; Rothnie, 1996). However, more dramatic instances of alternative polyadenylation in plants have been described. A well-known example in plants is the FY-mediated control of FCA gene expression, an important determinant of flowering time (Amasino, 2003; Quesada et al., 2003; Simpson et al., 2003).

While many instances of regulated alternative RNA processing in plants have been described, the mechanisms that link stimulus and RNA processing are poorly understood. In mammals, regulated as well as constitutive splicing is controlled to a large extent by so-called SR proteins (Sanford et al., 2003; Bourgeois et al., 2004). SR proteins carry domains rich in Ser and Arg that are substrates for cellular protein kinases and phosphatases, which in turn may be linked to numerous stimuli through canonical protein kinase cascades or via regulatory molecules such as calmodulin. Similar SR domain-containing proteins have been implicated

[1] This work was supported by the National Science Foundation (grant no. MCB–0313472 to A.G.H. and Q.Q.L.), the U.S. Department of Agriculture National Research Initiative (grant no. 2001–00911 to D.L.F.), and the Kentucky Tobacco Research and Development Center (D.L.F.).

[2] These authors contributed equally to the paper.

[3] Present address: Department of Surgery, University of Wisconsin-Madison, 600 Highland Ave., Madison, WI 53792.

* Corresponding author; e-mail aghunt00@uky.edu; fax 859–257–7125.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Arthur G. Hunt (aghunt00@uky.edu).

[OA] Open Access articles can be viewed online without a subscription.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.105.070672.

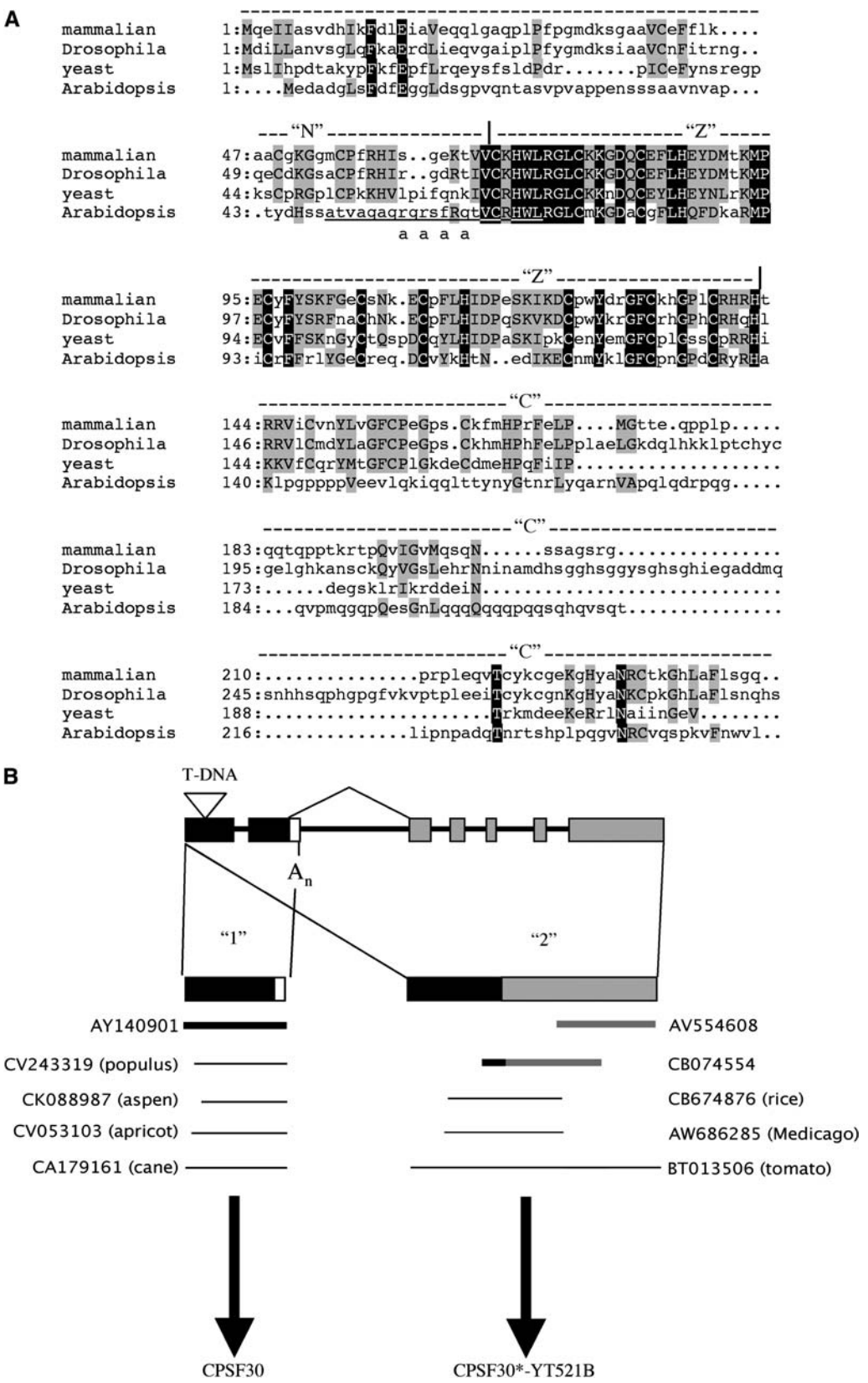


**Figure 1.** Properties of the Arabidopsis gene (At1g30460) that encodes AtCPSF30. A, Sequence alignments of AtCPSF30 with other eukaryotic CPSF30 proteins. Amino acid identity is denoted with white uppercase letters on a black background, similarity with black uppercase letters on a gray background, and unrelated sequences with lowercase lettering. The "N," "Z," and "C"

in alternative splicing in plants (Reddy, 2004), and it is probable that phosphorylation of these proteins is an important aspect of the regulation of the activities of these proteins. For the most part, though, direct links between stimulus and regulated RNA processing in plants have not been made.

Regulated polyadenylation in mammals is affected by a number of different mechanisms. Cell cycle control of polyadenylation is accomplished, at least in part, through phosphorylation of poly(A) polymerase, such that the enzyme is phosphorylated and its activity inhibited as the cell enters M phase (Colgan et al., 1996, 1998; Bond et al., 2000). Alterations in the levels of a male-specific isoform of CstF64 in mouse have been linked with alternative polyadenylation of novel male-specific transcripts (Wallace et al., 1999; Dass et al., 2001a, 2001b; Wallace et al., 2004). Presumably, the efficiency of polyadenylation of male-specific transcripts, which contain noncanonical polyadenylation signals, is determined by overall levels of the male-specific CstF64 isoform. Modulation of the levels of CstF64 (Takagaki and Manley, 1998), and of its RNA-binding activity by hnRNP-F (Veraldi et al., 2001), has been implicated in regulated alternative processing during mouse B cell development; this process is an important control point for immunoglobulin gene expression, as alternative polyadenylation determines the form of immunoglobulin that is produced by the particular cell type.

The Arabidopsis genome possesses genes that are capable of encoding orthologs of virtually the entire suite of known eukaryotic polyadenylation factor subunits (these are summarized at http://www.uky.edu/~aghunt00/polyA2010.html). Among these is a gene that encodes a protein similar to the 30-kD subunit of the mammalian cleavage and polyadenylation specificity factor (CPSF30). CPSF30 and its yeast counterpart, Yth1p, are so-called CCCH-type zinc-finger proteins and bind RNA with characteristics consistent with an involvement in mRNA 3′-end formation (Barabino et al., 1997, 2000; Bai and Tolias, 1998; Tacahashi et al., 2003). Interestingly, CPSF30 is a target for an influenza virus-encoded protein (NS1), a protein that acts to shut down host cell gene in infected cells (Nemeroff et al., 1998). In this report, several biochemical properties of an Arabidopsis ortholog of CPSF30 (AtCPSF30) are described. AtCPSF30 is shown to be nucleus-localized RNA-binding protein that binds calmodulin. Interestingly, its RNA-binding activity is inhibited by calmodulin in a calcium-dependent fashion. AtCPSF30 is also shown to be capable of interacting with itself. All of these properties can be mapped to an evolutionarily conserved core of three CCCH-type zinc-finger motifs and an adjacent domain that is unique to plant CPSF30-like proteins. These characteristics suggest that processes that are mediated by calmodulin signaling may alter mRNA 3′-end formation in plants.

## RESULTS

### Characterization of an Arabidopsis Gene That Encodes a Putative Ortholog of CPSF30

Eukaryotic CPSF30-related proteins contain a distinctive array of CCCH-type zinc-finger motifs (Barabino et al., 1997; Bai and Tolias, 1998). A BLAST (Altschul et al., 1997) search of the Arabidopsis proteome using the bovine CPSF30 amino acid sequence as a query yielded a gene, encoded by At1g30460, capable of encoding a polypeptide that is significantly similar to CPSF30 (Fig. 1A). Interestingly, databases contain contributions that define two different RNAs encoded by At1g30460 (Fig. 1B). The existence of both of these RNAs in Arabidopsis was confirmed by RNA-blot analysis, RT-PCR, and sequencing (Fig. 2A); thus, wild-type plants possess RNAs of the predicted size, transcripts that are absent from plants (termed *oxt6*) that contain a T-DNA insertion within the first exon of At1g30460 (Fig. 2A, oxt6 column). The smaller of the two RNAs (“1” in Fig. 2A) was less abundant than the larger; its presence and correspondence with At1g30460 was further confirmed by RT-PCR analysis, using primers that were specific for the smaller RNA.

The smaller of the two transcripts has a polyadenylation site that lies within an alternatively spliced intron (Fig. 1B) and can encode an approximately 28-kD polypeptide that is similar to eukaryotic CPSF30 proteins. The larger transcript can encode a polypeptide that contains all but the C-terminal 13 amino acids of the 28-kD polypeptide encoded by the smaller transcript, fused to a domain (the so-called YT521-B domain; Stoilov et al., 2002) that has been implicated in pre-mRNA splicing in mammals. Nuclear extracts prepared from wild-type Arabidopsis plants contained

**Figure 1.** (*Continued.*)
domains mentioned in the text are denoted above the Arabidopsis sequence, and are delimited with a “|” symbol above the residues that demarcate the three domains. The domain identified by calmodulin-binding domain prediction programs (http://calcium.uhnres.utoronto.ca/ctdb/ctdb/sequence.html) is underlined, and the four amino acids that were changed to Ala in the 30M mutant are indicated with a lowercase “a” beneath the respective position. GenBank accessions for the eukaryotic sequences are as follows: mammalian (*Bos taurus*), AAC48759; Drosophila (*Drosophila melanogaster*), AAF51453; and yeast (*Saccharomyces cerevisiae*), NP_015432. B, Diagram of the structure of the gene, showing the intron/exon organization (introns are thin lines, exons thick ones), and the two transcripts that encode AtCPSF30 and AtCPSF30-YT521-B, respectively. GenBank accessions for plant-derived expressed sequence tags that correspond to the two transcripts are listed beneath each one. The position of the T-DNA insertion in the *oxt6* mutant is indicated above the respective exon; this site is 147 bp 3′ of the translation initiation codon for this gene.

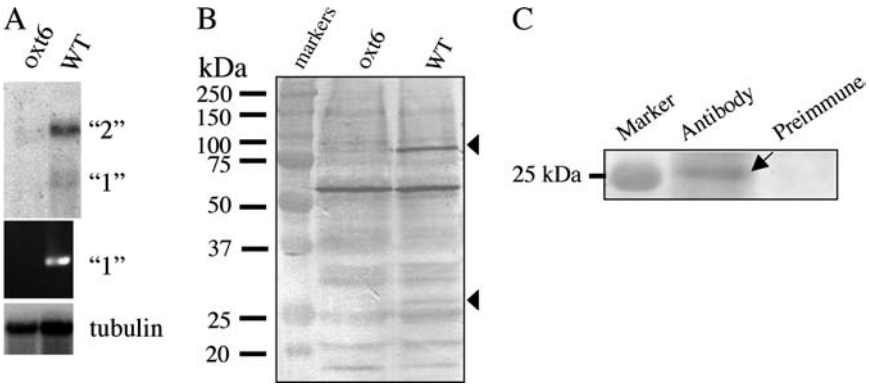


**Figure 2.** At1g30460-encoded proteins are in nuclear complexes. A, Expression of the gene encoding AtCPSF30. Total RNA from wild-type and *oxt6* mutant plants (see Fig. 1 for the location of the T-DNA insert in this mutant) were analyzed by RNA blotting using a probe derived from the first exon (top) or by RT-PCR, using primers specific for the smaller of the two transcripts (middle). At1g30460-derived transcripts are denoted as ''1'' and ''2,'' respectively. The RNA blot was also probed with tubulin sequences (bottom). B, At1g30460-encoded polypeptides can be detected in Arabidopsis nuclear extracts. Immunoblot filters containing separated Arabidopsis nuclear extracts were probed with affinity-purified anti-AtCPSF30 antibody. Protein size markers are on the left lane, and the arrowheads indicate expected bands that correspond to the products of the small and large At1g30460-encoded mRNAs. Nuclear extracts from the *oxt6* mutant (lane 1) and the wild type (lane 2) were used to discriminate two expected proteins encoded by the two mRNAs of At1g30460. The smaller band corresponds to AtCPSF30. C, Coimmunoprecipitation. Equal amounts of Arabidopsis nuclear protein extracts were precipitated by the affinity-purified anti-AtCPSF100 antibody (lane marked ''Antibody'') and preimmune antiserum (lane marked ''Preimmune''), respectively. AtCPSF30 (pointed by arrowhead) was detected from the pellet immunoprecipitated by anti-AtCPSF100 antibody.

a number of polypeptides that were recognized by antisera raised against a peptide present in the two At1g30460-derived proteins (WT lane in Fig. 2B). Importantly, two of these (noted with arrowheads in Fig. 2B) were absent in nuclear extracts prepared from the *oxt6* insertion mutant (Fig. 2B, compare the oxt6 and WT lanes); the mobilities of these are consistent with those predicted from the properties of the two cDNAs. These mutant plants do not express At1g30460-derived RNAs (oxt6 column in Fig. 2A), indicating that the differences seen in the immunoblot represent the two expected polypeptides. The smaller of these two polypeptides could be immunoprecipitated from nuclear extracts by antibodies raised against another Arabidopsis CPSF subunit, CPSF100 (Elliott et al., 2003), but not by the preimmune serum from the same rabbit (Fig. 2C, compare the antibody and preimmune lanes). It was not possible in this experiment to determine the behavior of the larger At1g30460-encoded polypeptide due to comigration with traces of the immunoglobulin that was present in the immunoprecipitates. In any case, these results indicate that both of the At1g30460-encoded proteins are produced and reside in the nucleus of wild-type Arabidopsis plants, and that the smaller protein resides in a complex that includes another Arabidopsis polyadenylation factor subunit.

While the Arabidopsis genome has many other genes that may encode CCCH-like proteins (Fig. 3), only the At1g30460-encoded polypeptide(s) resembles CPSF30 to an extent greater than the characteristic Cys/His fingerprint of CCCH proteins. This is revealed in amino acid alignments, which show that At1g30460 falls within a eukaryotic CPSF30 clade to the exclusion of other Arabidopsis CCCH-like proteins (Fig. 3). Sequences closely related to At1g30460 are present in various plant sequence databases (Figs. 1B and 3); interestingly, several expressed sequence tags exist that include both CPSF30- and YT521-B-encoding regions (Fig. 1B), indicating that this protein domain organization is common to plants.

A more detailed comparison of the smaller At1g30460-derived protein with its mammalian and yeast counterparts revealed a high degree of similarity involving three of the five probable zinc fingers that are in the other eukaryotic CPSF30 proteins (Fig. 1A). This similarity was most striking in the second and fourth zinc fingers (using the mammalian protein as a reference). Outside of these three zinc fingers, the similarity between the Arabidopsis and other eukaryotic proteins was minimal. Interestingly, the Arabidopsis protein lacked the first and fifth zinc fingers found in other eukaryotic CPSF30 proteins, as well as the possible CCHC zinc knuckle motifs found in animal CPSF30s (Fig. 1A).

Alignment of the full-length plant sequences that can be found in databases reveals additional features that are not apparent in the alignments with other eukaryotic CPSF30 proteins (Fig. 4). Thus, while the first zinc-finger motif that is seen in other eukaryotic CPSF30 orthologs is absent in the plant proteins, there is a highly conserved acidic domain at the N termini of the plant proteins (Fig. 4). There is also sequence conservation immediately upstream of the first CCCH motif in the plant proteins, conservation that does not extend to other eukaryotic CPSF30 homologs (compare Figs. 4 and 1A). Downstream from the third

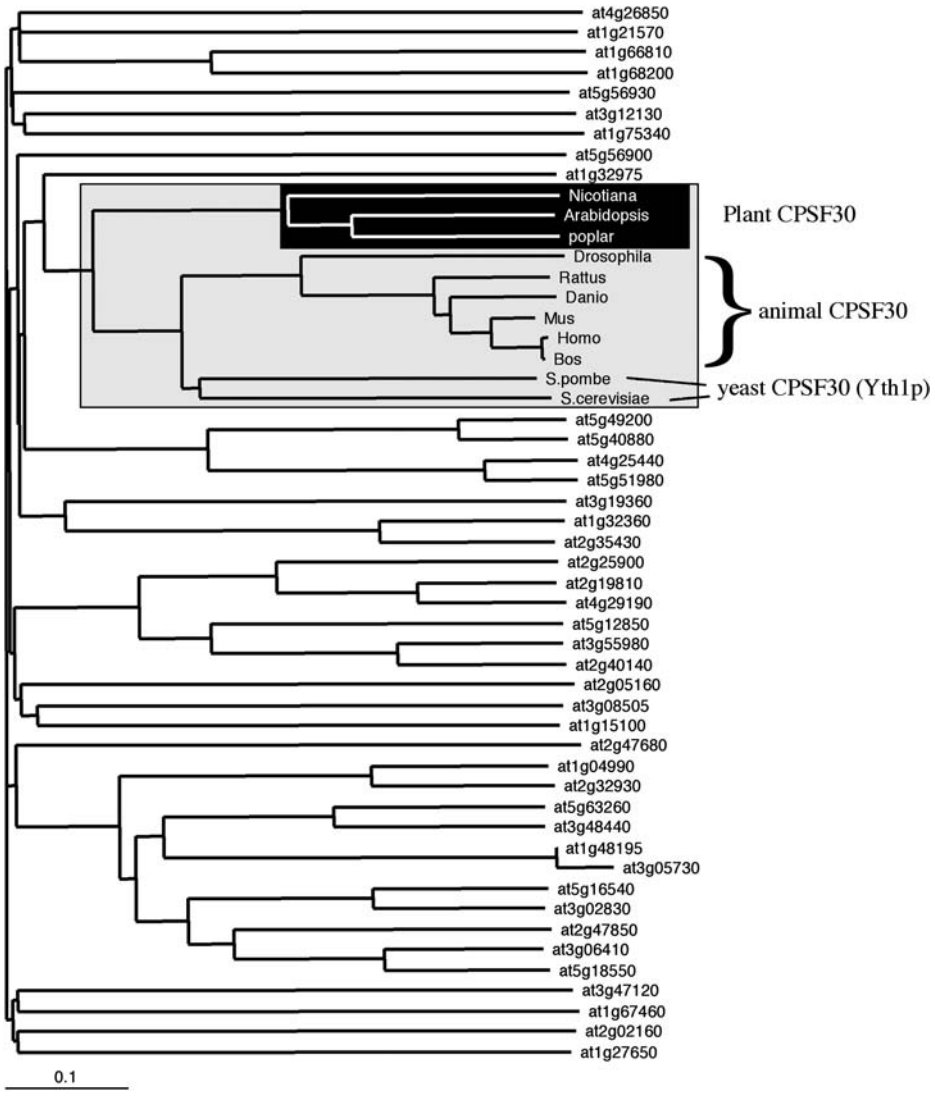


**Figure 3.** Amino acid sequence comparisons of CCCH proteins. A tree diagram is shown that summarizes the amino acid sequences similarities of plant CPSF30-like proteins with other eukaryotic CPSF30s and with the set of Arabidopsis CCCH zinc-finger proteins. The plant and other eukaryotic CPSF30 sequences are highlighted with shaded boxes as indicated. Putative Arabidopsis CCCH motif-containing proteins were identified by BLAST searches; those analyzed are At1g30460 (= AtCPSF30), At1g21570, At5g56930, At2g47680, At3g47120, At1g66810, At3g08505, At3g08505, At5g49200, At4g25440, At1g04990, At1g04990, At5g63260, At2g19810, At1g32360, At2g32930, At5g16540, At5g16540, At5g16540, At2g25900, At5g56900, At2g35430, At5g56900, At5g51980, At3g02830, At5g40880, At1g48195, At1g67460, At1g32975, At2g02160, At2g47850, At1g15100, At4g29190, At3g55980, At3g19360, At2g05160, At1g68200, At5g12850, At2g40140, At3g06410, At2g36040, At3g12130, At3g44785, At4g26850, At3g48440, At5g18550, At1g75340, At1g27650, and At3g05730. GenBank accessions for the other eukaryotic sequences in this analysis were as follows: human, EAL23878; *B. taurus*, AAC48759; Drosophila (*D. melanogaster*), AAF51453; Danio (*Danio rerio*), AAH45289; *S. cerevisiae*, NP_015432; *Schizosaccharomyces pombe*, CAB-61457; poplar (*Populus* spp.), CV-243319; and Nicotiana, CK286112.

(and last) CCCH motif in the plant proteins, there is extensive conservation of a Pro-rich motif; this is the position of the fifth CCCH motif in other eukaryotic CPSF30 proteins. The C termini of the three plant proteins also contain a Gln-rich domain and a highly conserved five-amino acid sequence (PLPQG) near the C terminus (Fig. 4).

### RNA Binding by AtCPSF30 Requires the Central Zinc-Finger Domain and an Adjacent Plant-Specific Motif

CPSF30 and Yth1p are both RNA-binding proteins (Barabino et al., 1997, 2000; Bai and Tolias, 1998; Tacahashi et al., 2003). Consistent with the similarity with other eukaryotic CPSF30 proteins, a maltose-binding protein (MBP)-AtCPSF30 fusion protein was able to bind an RNA that contains the polyadenylation signal from a pea (*Pisum sativum*) *rbcS* gene (Mogen et al., 1992), as indicated in electrophoretic mobility shift assays. MBP fusions containing the full-sized AtCPSF30 polypeptide protein showed a characteristic behavior (Fig. 5A), such that binding was almost undetectable below protein concentrations of 0.5 $\mu$M and almost completely saturating at concentrations above 2 $\mu$M. Purified MBP did not bind RNA (data not shown).

As indicated in Figures 1A and 4, the Arabidopsis CPSF30 contains three conceptual domains: novel plant-specific N and C termini, flanking a conserved central domain defined by the three CCCH-type zinc fingers of the Arabidopsis protein. To better understand the roles of these domains (N, Z, and C, respectively) in the RNA-binding activity of the protein, a series of deletion derivatives was prepared and assayed for RNA binding. Deletion of the entire N terminus (up to the first of the three zinc-finger motifs) dramatically reduced the RNA-binding activity of AtCPSF30 (m5 in Fig. 5B). Thus, at protein concentrations at which maximal binding is seen with the wild-type protein (2–5 $\mu$M), the N-terminal deletion displayed little binding. Some binding could be discerned at concentrations

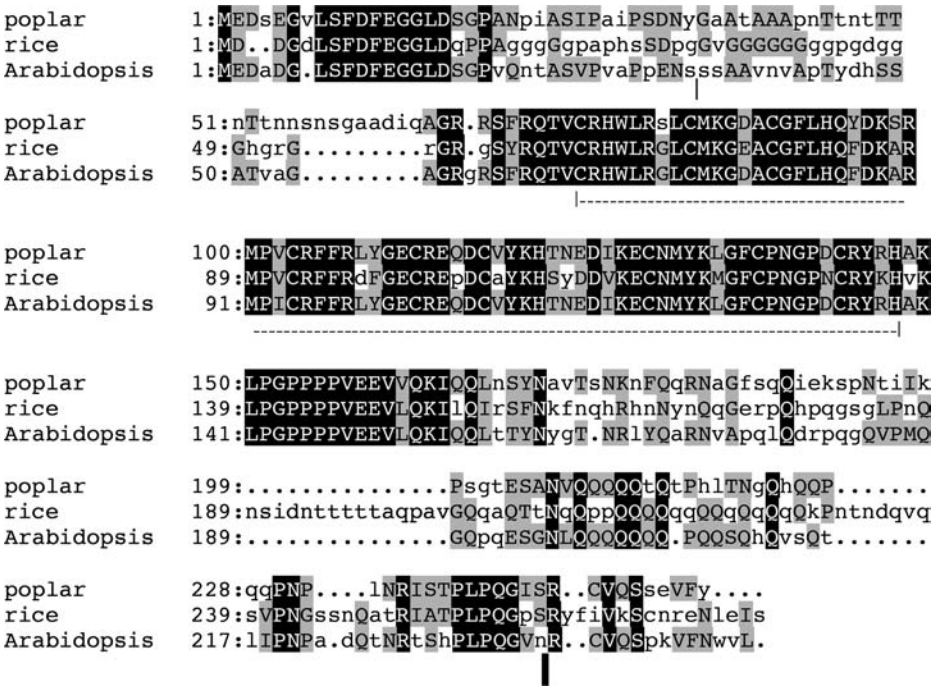


**Figure 4.** Sequence alignments of AtCPSF30 with other plant CPSF30-like proteins. Amino acid identity is denoted with white uppercase letters on a black background, similarity with black uppercase letters on a gray background, and unrelated sequences with lowercase lettering. The central zinc-finger core that is shared with eukaryotic CPSF30 proteins is highlighted with a dashed underline, and the N- and C-terminal limits of this core are denoted with a "|" beneath the sequence. The thin vertical bar beneath the first line of the Arabidopsis sequence denotes the endpoint of the m6 mutation (see Fig. 9). The thick vertical bar beneath the last line of the Arabidopsis sequence shows the location of the splice site that is used in the biogenesis of the larger At1g30460-derived mRNA; amino acid sequences to the C-terminal side of this are absent from the AtCPSF30-YT521-B polypeptide. Accessions are given in the legend for Figure 2.

above 5 $\mu$M; it was not possible to add enough protein to reach saturation in these assays, but the apparent affinity of the RNA for protein in this assay was at least 20-fold less than that seen with the wild-type protein (data not shown). In contrast, deletion of the C terminus had little discernible effect on RNA binding (m4 in Fig. 5B). In particular, there was little difference between the C-terminal truncation and the full-sized protein in terms of the profile of binding activity as a function of protein concentration, with activity decreasing precipitously at concentrations below 1 $\mu$M. As expected (based on the properties of the N-terminal deletion), the central domain by itself did not bind RNA (m2 in Fig. 5B; those mutants that displayed no detectable binding are listed in Fig. 5B). Moreover, the isolated C-terminal domain was also unable to bind RNA (m8 in Fig. 5B). Elimination of the second and third zinc-finger motifs from the m4 mutant eliminated RNA binding as well (m1 in Fig. 5B).

The plant-specific N domain of AtCPSF30 consists of two evolutionarily conserved sequences: an acidic domain at the very N terminus of the protein and a conserved region that abuts the first zinc finger (Fig. 4). The effects on RNA binding of deletion of the N terminus in the m5 mutant might be attributable to elimination of either (or both) of these sequences. To explore the possibilities, the 11 amino acids that are adjacent to the first zinc finger were added to the Z + C and Z proteins (m9 and m10, respectively, in Fig. 5B). In contrast to the m5 protein, RNA binding was readily discernible with the m9 mutant (Fig. 5B). Interestingly, the apparent affinity of RNA for this protein was about 5-fold lower than the wild-type or m4 proteins. While the Z domain by itself (m2) was unable to bind RNA, the m10 mutant displayed considerable RNA-binding activity (Fig. 5B). The apparent affinity of RNA for this protein was some 3- to 5-fold lower than that for the wild-type and m4 proteins, comparable with that of the m9 protein. Based on these results, it can be concluded that RNA binding by AtCPSF30 requires the central zinc-finger domain and the plant-specific sequence immediately adjacent to the first zinc-finger motif.

### An Interaction between Calmodulin and AtCPSF30

Among the computer-assisted analyses of the polypeptide encoded by At1g30460 that were performed was a search for calmodulin-binding motifs. This yielded a possible domain (underlined in Fig. 1A) that abutted the N-terminal zinc-finger motif. Accordingly, the ability of this protein to bind calmodulin was examined using a calmodulin blotting assay. As shown in Figure 6, the full-sized MBP-AtCPSF30 protein as well as a predominant breakdown product displayed calmodulin-binding activity in this assay (lanes marked "WT"). In contrast, comparable quantities of purified MBP lacked calmodulin-binding activity (lanes marked "MBP"). This indicates that the AtCPSF30 portion of the fusion protein is responsible for the observed calmodulin binding.

To test the hypothesis that this calmodulin binding was due to the domain identified in the computational analysis, a mutational approach was taken. Initial studies involved deletion analysis of the protein, examining the abilities of different parts of the protein to bind calmodulin in a far-western assay. For these studies, the three domains of AtCPSF30 were produced in

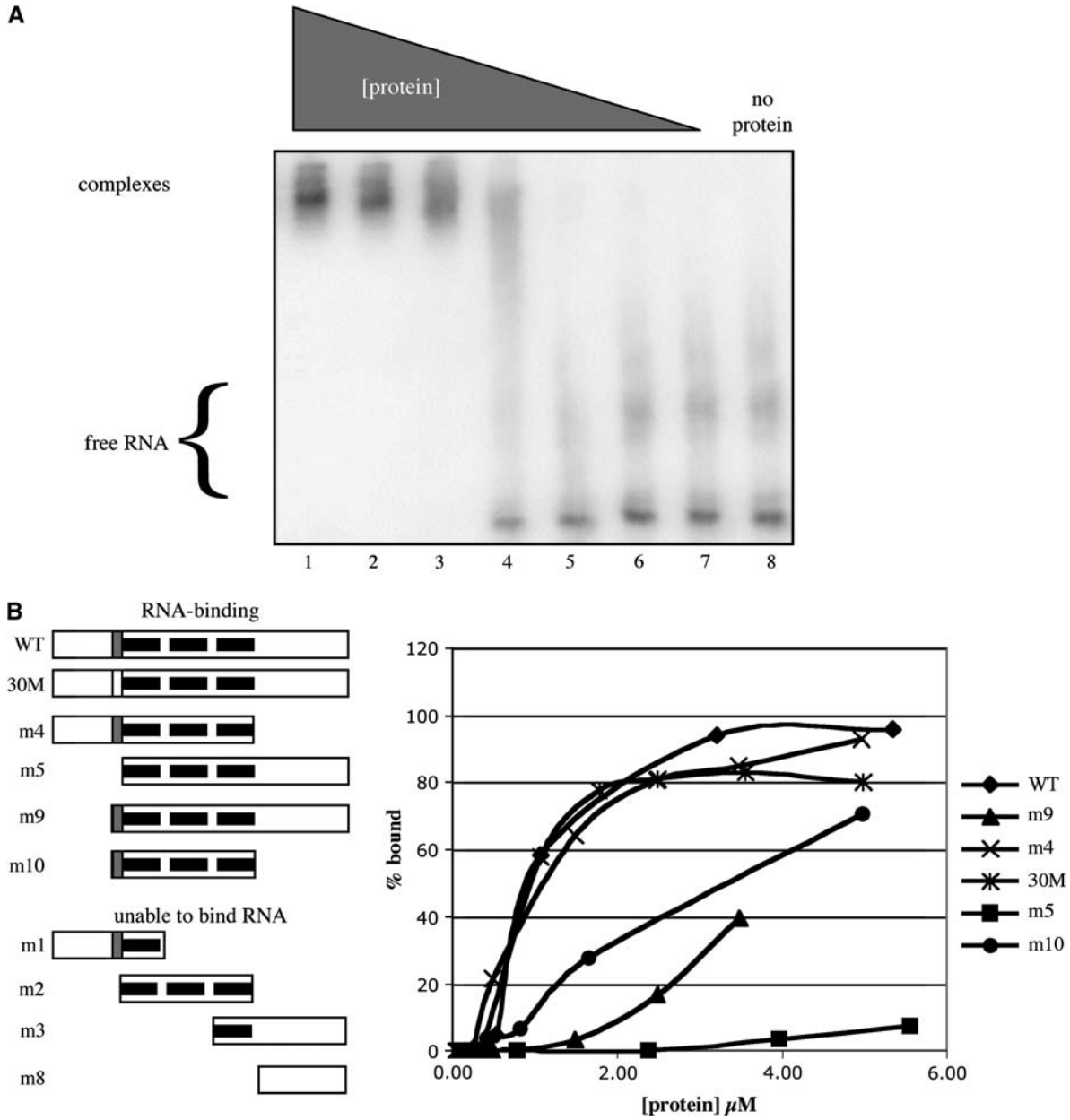


**Figure 5.** Demarcation of the RNA-binding domain of AtCPSF30. A, Binding of RNA by wild-type AtCPSF30 as a function of protein concentration. A total of 3.5 pmol of uniformly labeled RNA containing the polyadenylation signal of the pea *rbcS-E9* gene was incubated with varying quantities of MBP-AtCPSF4 and RNA binding assessed by electrophoresis on native gels and autoradiography. The positions of the RNA-protein complex (complexes) and free RNA are indicated on the left. No protein was added for the sample in lane 8. Lanes 1 to 7 contained 10.5, 7.5, 4.5, 1.5, 0.75, 0.375, and 0.188 $\mu$g of purified protein, respectively. B, RNA binding by mutant forms of AtCPSF30. On the left is a depiction of the different variants of AtCPSF30 that were produced as MBP fusion proteins. The central zinc-finger domain is represented with three black lines within the respective rectangular boxes, and the small domain that is responsible for calmodulin binding shown as a gray box within the larger representation. The clear box in the 30M representation indicates that the calmodulin-binding motif has been changed by mutation to be nonfunctional. Fusion proteins that bound RNA are listed under "RNA-binding," and the plots of binding as a function of protein concentration given on the right. Fusion proteins for which no detectable RNA binding was observed are listed under "unable to bind RNA"; the plots for these proteins are not given, as they would coincide with the *x* axis.

various combinations as MBP fusion proteins. None of the isolated domains (N, Z, and C, respectively) bound calmodulin in this assay (m7 in Fig. 6; data not shown). However, the N + Z construct (m1 in Fig. 6) did bind calmodulin. This suggested that the calmodulin-binding domain might be located near the breakpoints between the N and Z domains, which is consistent with the computer prediction. To confirm this, four point mutations were introduced into the full-sized AtCPSF30 coding sequence, as indicated in Figure 6 ("30M"). In contrast

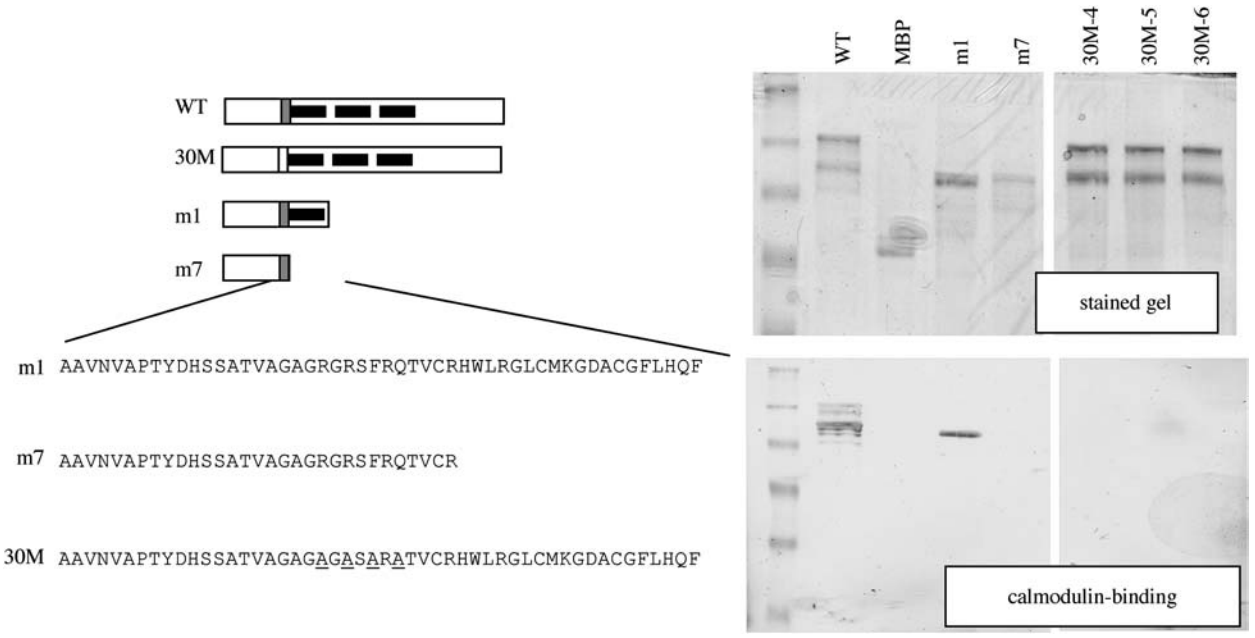


**Figure 6.** Identification of the calmodulin-binding domain of AtCPSF30. Proteins that were analyzed for calmodulin binding are depicted on the left; the explanation of the depiction is as in the legend for Figure 5. The results of SDS-PAGE and staining (top; labeled "stained gel") and far-western blotting using biotinylated calmodulin (bottom; labeled "calmodulin-binding") are shown. 30M-4, 30M-5, and 30M-6 are preparations from three independent mutants in which the putative calmodulin-binding domain was altered as shown on the left (altered amino acids are underlined). The 30M-4 preparation was used in the other studies in this report.

to the full-sized protein, the three independent mutant proteins did not bind calmodulin (Fig. 6, lanes labeled "30M-4," "30M-5," and "30M-6"). From these results, we conclude that the calmodulin-binding domain of AtCPSF30 is situated immediately upstream of the first conserved zinc finger of the protein, in a region that itself is highly conserved in plants (Fig. 4).

The near juxtaposition of the calmodulin-binding and RNA-binding domains in AtCPSF30 suggests that calmodulin may affect the ability of AtCPSF30 to bind RNA. This prediction was tested by examining the effects of calmodulin on RNA binding by AtCPSF30. In the presence of calcium, calmodulin inhibited the RNA-binding activity of AtCPSF30 by approximately 80% ("WT + Ca" sample in Fig. 7). Replacement of calcium chloride with EGTA eliminated the inhibition ("WT + EGTA" sample in Fig. 7).

As further confirmation that calmodulin inhibits RNA binding by AtCPSF30, similar studies were performed with the 30M mutant. As shown in Figure 5B, this mutant displayed an RNA-binding activity that was similar to that of the wild-type protein and the m4 mutant. This indicates that the four Ala substitutions in the 30M mutant do not affect the RNA-binding activity of the protein, even though they are within the region that is needed for RNA-binding activity. To test the hypothesis that the identified calmodulin-binding domain was responsible for the inhibitory effects of calmodulin on RNA binding, the effects of calmodulin on the RNA-binding activity of the mutant were examined. In contrast to the behavior of the wild-type AtCPSF30, the 30M mutant was not affected by calmodulin in the presence of calcium ("MUT + Ca" sample in Fig. 7). RNA binding was not affected by calmodulin when calcium chloride was replaced with EGTA ("MUT + EGTA" sample in Fig. 7). This is consistent with the observation that the 30M mutant does not bind calmodulin and indicates that the conserved sequence immediately preceding the first zinc-finger motif is necessary for the inhibition of RNA binding by calmodulin.

### AtCPSF30 Interacts with Itself

The curve for binding of AtCPSF30 with RNA (Fig. 5B) has characteristics of strong cooperativity. This in turn suggests that AtCPSF30 is capable of interacting with itself. This hypothesis was tested by copurification assays. For this, a series of labeled AtCPSF30 derivatives (illustrated in Fig. 8A) was prepared and used in copurification assays using MBP-AtCPSF30 as bait. As shown in Figure 8B, when AtCPSF30 ("wt") was incubated with purified MBP-CPSF30 and the MBP fusion protein was subsequently purified using amylose resin, there was a significant copurification of the labeled protein (top; lane labeled "wt"). Replacement of the MBP-CPSF30 with purified MBP eliminated this copurification (middle). This result indicates that AtCPSF30 interacts with itself.

To define the domain of AtCPSF30 that mediates the self-association, similar experiments were done with labeled portions of AtCPSF30. As shown in the top

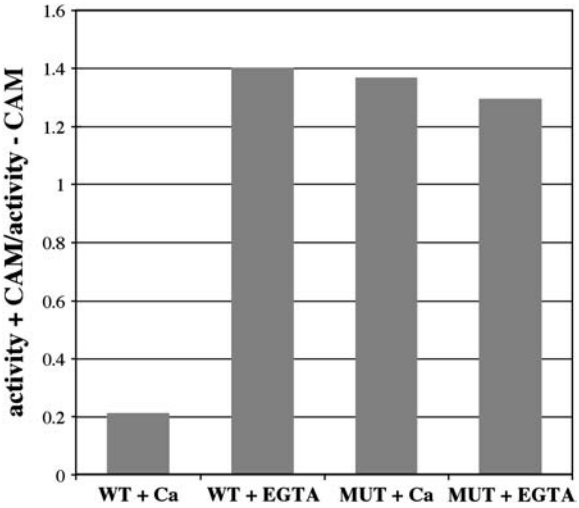


**Figure 7.** Effects of calmodulin on RNA binding by AtCPSF30 and the 30M mutant. A total of 1.25 pmol of uniformly labeled RNA containing the polyadenylation signal of the pea *rbcS-E9* gene was incubated with 40 pmol of MBP-AtCPSF30 (WT) or the purified 30M mutant protein (MUT), and with various combinations of calmodulin (60 pmol; CAM), calcium chloride (10 $\mu$M), and EGTA (1 mM), and RNA binding assessed by electrophoresis on native gels and autoradiography. The ratios of binding activity observed in the presence or absence of calmodulin (activity + CAM/activity − CAM) were plotted for each of four sets of conditions: wild-type protein + calcium chloride (WT + Ca), mutant protein + calcium chloride (MUT + Ca), wild-type protein + EGTA (WT + EGTA), and the mutant protein + EGTA (MUT + EGTA). Purified calmodulin does not bind RNA in this assay and has no effect on the added RNA (data not shown).

section of Figure 8B, no apparent copurification of the plant-specific N and C termini of AtCPSF30 (m6, m7, and m8, respectively) with MBP-AtCPSF30 could be seen in the copurification assay. In contrast, significant quantities of the evolutionarily conserved central domain (m2) did copurify with MBP-AtCPSF30. Replacement of MBP-AtCPSF30 with MBP eliminated this copurification (middle section of Fig. 8B). These results indicate that the central domain of AtCPSF30 is the part of the protein responsible for the self-association of this protein.

The cooperativity that is suggested by the binding curve (Fig. 5) and direct demonstration of protein-protein interaction (Fig. 8B) suggests a possible mechanism for the inhibition by calmodulin of RNA binding by AtCPSF30. Specifically, it is possible that calmodulin inhibits the self-interaction of AtCPSF30, thus eliminating the possibility of cooperative binding of the protein to RNA. To test this, the effects of a large molar excess of calmodulin on the self-association were examined. As shown in Figure 8C, calmodulin had no discernible effect on the copurification of labeled AtCPSF30 with MBP-AtCPSF30. This result suggests that calmodulin does not affect the self-association and, thus, that the inhibition of RNA binding by calmodulin is not mediated by an inhibition of self-association of AtCPSF30.

## DISCUSSION

### A Functional Map of a Putative Plant Polyadenylation Factor Subunit

The results presented in this study permit the construction of a functional map (Fig. 9) of AtCPSF30, the plant ortholog of the 30-kD subunit of the CPSF. Specifically, RNA binding by this protein requires much of the central zinc-finger domain that is conserved in all eukaryotic CPSF30 orthologs (Fig. 1A), as well as some 11 amino acids immediately next to the N terminus of the three zinc-finger motifs. This latter domain is conserved in plant CPSF30-like proteins but not in its eukaryotic counterparts (Figs. 1A and 4). Calmodulin binding involves the same plant-specific domain that is adjacent to the first zinc-finger motif. However, the observation that the m7 mutant is impaired in calmodulin binding (Fig. 6), even though it contains this motif, indicates that some sequences within the central zinc-finger domain are also needed for the interaction with calmodulin. While there is some overlap in the calmodulin- and RNA-binding domains, these functions can be separated; thus, the four Ala substitutions in the 30M mutant affect calmodulin binding, but not RNA binding itself. This suggests that these two functions involve different amino acids. It also provides a model for the inhibition of RNA binding by calmodulin; specifically, it suggests that calmodulin and RNA occupy the same surface of AtCPSF30, such that occupancy by calmodulin prevents access to the protein by RNA. A somewhat similar mechanism has been suggested for the inhibition by calmodulin of DNA binding by the E-protein family of mammalian basic helix-loop-helix proteins (Onions et al., 1997; Saarikettu et al., 2004).

The third function that has been mapped in this study is the self-association of AtCPSF30. This activity can also be attributed to the central zinc-finger domain of the protein. However, the apparently cooperative nature of binding of RNA by AtCPSF30 suggests that RNA binding and self-association are not mutually exclusive; indeed, they may be reinforcing. Additionally, while calmodulin binding involves at least part of the zinc-finger domain, it does not seem to have an impact on the self-association of AtCPSF30. Thus, while calmodulin- and RNA-binding are mutually exclusive activities of AtCPSF30, calmodulin binding and the self-association are not. These considerations permit a tentative demarcation of the central zinc-finger domain, with the N-terminal portion being involved in RNA and calmodulin binding, and the C terminal part probably important for self-association. That self-association involves one or more zinc-finger motifs is consistent with the frequent association of zinc-finger domains with protein-protein interactions (Matthews and Sunde, 2002); that the zinc-finger array in AtCPSF30 seems to consist of motifs that act either in RNA binding or protein-protein interaction is more interesting and may suggest that similar arrays in diverse

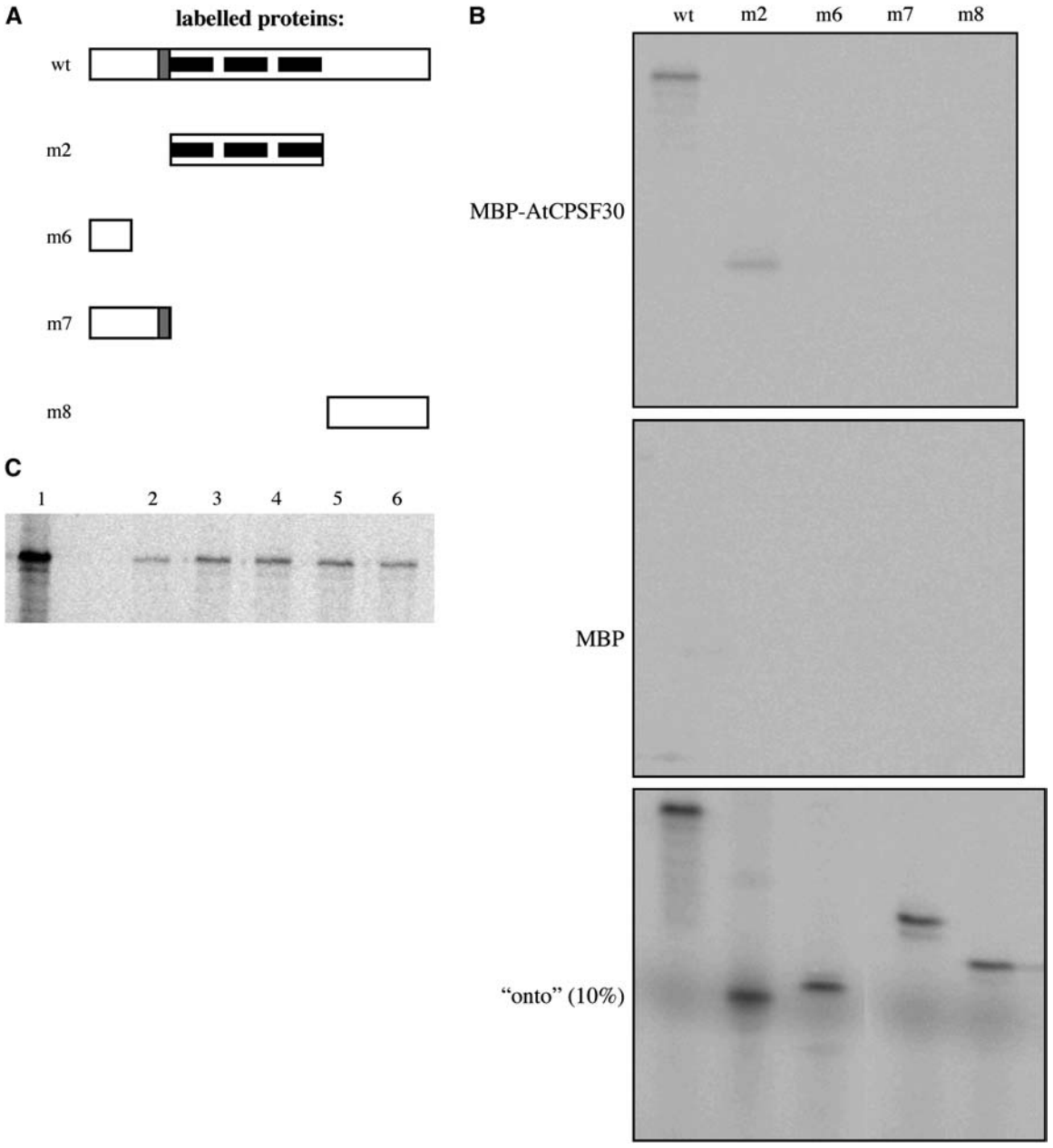


**Figure 8.** AtCPSF30 interacts with itself. A, Depiction of the AtCPSF30 variants that were produced as labeled proteins. The symbols used are described in the legend for Figure 5. B, Binding of labeled AtCPSF30 and deletion derivatives to MBP-AtCPSF30 and MBP. Experiments were performed as described in ''Materials and Methods.'' The identity of the labeled protein is indicated at the top. Gels showing results with MBP-AtCPSF30 (top) or MBP (middle) as baits are shown. Ten percent of the quantity of each translation mixture used in the copurification assays is shown in the bottom section [''onto'' (10%)]. C, Calmodulin has no effect on the self-interaction. Copurification assays with labeled AtCPSF30 and various combinations of calmodulin, $CaCl_2$, and EGTA were performed. Lane 1, 10% of the labeled input protein. Lanes 2 to 6, All had the same labeled input (AtCPSF30). Lane 2, No addition. Lanes 3 and 5, 2.5 mM Ca $Cl_2$. Lanes 4 and 6, 2.5 mM EGTA. Lanes 5 and 6, + calmodulin.

zinc-finger protein families may likewise be involved in multiple activities.

### Implications of the Properties of AtCPSF30 for RNA Processing

AtCPSF30 is the only Arabidopsis protein with a degree of sequence similarity to other eukaryotic CPSF30 proteins that extends beyond the typical spacing of Cys and His residues in the CCCH zinc-finger motif, and a number of lines of evidence support the conclusion that AtCPSF30 is an authentic polyadenylation factor subunit. As is the case with its yeast counterpart (Yth1p; Barabino et al., 1997), AtCPSF30 interacts with another Arabidopsis polyadenylation subunit homolog, AtFip1(V) (Forbes et al., 2006); AtFip1(V) also


**Figure 9.** A functional map of AtCPSF30. The complete AtCPSF30 coding region is represented as the long, lightly shaded rectangle. The three zinc-finger motifs are indicated with thick black lines lying within the AtCPSF30 coding region. Above this box are shown the relative locations of the domains responsible for self-association (CPSF30) and RNA binding (RNA); shading near the ends of these boxes indicates uncertainties in the possible extents of these two domains. The location of the calmodulin-binding site is shown beneath the light gray box (C).

interacts with poly(A) polymerase and thereby provides a conceptual link between AtCPSF30 and poly(A) polymerase. The results presented in this study show that AtCPSF30 is present in the nucleus (Fig. 2B), as would be expected of a polyadenylation factor subunit. The coimmunoprecipitation of AtCPSF30 by antibodies raised against AtCPSF100 (Fig. 2C) indicates that AtCPSF30 resides, at least in part, in a complex with another Arabidopsis polyadenylation factor subunit. AtCPSF30 is an RNA-binding protein, as are its eukaryotic counterparts. AtCPSF30 shares with other eukaryotic CPSF30 proteins a central domain that consists of three CCCH-type zinc-finger motifs (Fig. 1A). The central domain (that includes the second, third, and fourth of the five zinc-finger motifs in the yeast and animal proteins) is involved in RNA binding by the yeast homolog Yth1p and in interactions of Yth1p with other polyadenylation factor subunits (Tacahashi et al., 2003). These commonalities suggest that AtCPSF30 and Yth1p interact with RNA in similar manners.

In yeast, the gene that encodes its CPSF30 ortholog, *YTH1*, is essential. In contrast, the At1g30460 gene in Arabidopsis is not absolutely essential for the growth of the plant, as adult, fertile plants that bear a T-DNA insertion within the first exon can be obtained. While it is not apparent from similarity searches of Arabidopsis databases, it is nonetheless possible that other distantly related proteins (perhaps one or more of the family of CCCH motif-containing proteins listed in Fig. 3) may serve analogous roles in mRNA 3′-end formation in Arabidopsis. Alternatively, AtCPSF30 may function in a somewhat different fashion in mRNA 3′-end formation than its yeast counterpart, serving a nonessential role.

The central zinc-finger domain of AtCPSF30 is also involved in the self-association of this protein (Fig. 8B). There have been no reports of self-association of CPSF30 in other systems. Whether this reflects a difference between plants and other eukaryotes is an open issue. Nonetheless, the self-association described in this report raises interesting possibilities with respect to mRNA 3′-end formation in plants and to the role(s) that AtCPSF30 may play in RNA processing and metabolism. For example, RNA recognition by a multimer (such as a dimer) would permit AtCPSF30 to recognize more than one sequence element in a plant polyadenylation signal.

The At1g30460 gene encodes two polypeptides, AtCPSF30 and a larger polypeptide that contains all but the last 13 amino acids of AtCPSF30 fused to the N terminus of another domain (Fig. 1A). The latter includes a motif (the so-called YT521-B motif) that was first described in proteins that are associated with pre-mRNA splicing in mammals (Stoilov et al., 2002). In mammals, YT521-B proteins interact with splicing factors (Imai et al., 1998), a nuclear membrane protein (emerin; Wilkinson et al., 2003), and the tissue-specific splicing factor rSLM-1 (Stoss et al., 2004), and alterations in the expression of the prototypical member of this family of proteins (YT521) alter the alternative

**Table I.** *DNA oligonucleotides used in this study*

| Primer Name | Sequence (5′ → 3′) | Used for: |
|---|---|---|
| CPSF4-Nter | TTTAGATCTACCATGGAGGATGCTGATGGACTTAGC | Cloning into pMAL-C2 and pCITE[a]: wild type, m1, m4, m6, m7 |
| CPSF4-B | TTTAGATCTTACTGCTGATGGAGAAACCCACAGGCGTC | Cloning into pMAL-C2 and pCITE: m1 |
| CPSF4-C | TTTAGATCTGTTTGTAGACACTGGCTTCGA | Cloning into pMAL-C2 and pCITE: m2, m5 |
| CPSF4-D | TTTAGATCTTAAGGTCCAGGAAGCTTTGCATGCCTGTACC | Cloning into pMAL-C2 and pCITE: m2, m4 |
| CPSF4-E | TTTAGATCTATCAAAGAATGCAATATGTACAAGCTG | Cloning into pMAL-C2 and pCITE: m3 |
| CPSF4-F | TTTAGATCTTACAGAACCCAATTAAAAACCTTAG | Cloning into pMAL-C2 and pCITE: wild type, m3, m5, m8, m9 |
| CPSF4-G | CCCAGATCTGGTGTTGAGGAAGTTCTTCAGAAGATACAA-CAATT | Cloning into pMAL-C2 and pCITE: m8 |
| CPSF4-H | TTTAGATCTGTTTTCAGGCGGAGCAACCGGGACGGAAG | Cloning into pMAL-C2 and pCITE: m6 |
| CPSF4-I | TTTAGATCTCTACAAACAGTTTGACGGAAACTTCTACCC | Cloning into pMAL-C2 and pCITE: m7 |
| CPSF4-K | TTTAGATCTCGAGCTGGGAGGGGTAGAAGTTTCCGTCAA | Cloning into pMAL-C2 and pCITE: m9, m10 |
| C30CBD-F | CGGTGGCTGGAGCTGGGGCGGGTGCAAGTGCACGTG-CAACTGTTTGTAGACACTGG | Site-directed mutagenesis for 30M[b] |
| C30CBD-R | CCAGTGTCTACAAACAGTTGCACGTGCACTTGCACCC-GCCCCAGCTCCAGCCACCG | Site-directed mutagenesis for 30M |
| T7-E9 | TAATACGACTCACTATAGGGAGTATTATGGCATTGGGAA | PCR production of *rbcS-E9* in vitro transcription template |
| E9 61-80 | AAATGTTTGCATATCTCTTA | PCR production of *rbcS-E9* in vitro transcription template |

[a]See Figures 5, 7, and 9 for descriptions of the CPSF30 mutant designations (m2, m3, etc.). [b]See Figure 6 for a description of the 30M mutant.

splicing of a number of mRNAs (Hartmann et al., 1999). That one of the two products of the At1g30460 gene contains the RNA-binding domain of AtCPSF30 fused to a potential splicing factor raises the interesting possibility that similar or identical RNA-binding domains may facilitate splicing and polyadenylation in plants.

The potential commonalities or links between splicing and polyadenylation extend beyond the presence in plants of a protein with polyadenylation- and splicing-related functionalities. Given that the central domain of AtCPSF30 is capable of interacting with itself (Fig. 8), it follows that the larger At1g30460-encoded protein is expected to be capable of interactions with itself and with AtCPSF30. Thus, three different classes of multimeric complexes containing the AtCPSF30 polypeptide can be formed in the cell. These different variants have the possibility of acting in somewhat different ways; complexes consisting of just AtCPSF30 may act in polyadenylation, while complexes containing the AtCPSF30-YT521-B polypeptide may work in splicing. The possibility also exists that heteromeric AtCPSF30-containing complexes may mediate communication between splicing and polyadenylation. These considerations provide a number of potential conceptual links between pre-mRNA splicing and polyadenylation in plants. Further biochemical characterization as well as studies of mutants bearing defects in the At1g30460 gene promise to shed interesting new insight into the interplay between different RNA-processing events in plants.

Finally, the effects of calmodulin on RNA binding by AtCPSF30 are of interest for several reasons. In animals, CPSF30 is a target of the influenza virus-encoded protein NS1; binding of this protein to CPSF30 inhibits polyadenylation, thereby contributing to the shutdown of host gene expression in virus-infected cells (Nemeroff et al., 1998). This suggests that, in mammals, CPSF30 has functions in polyadenylation that can be accessed and altered by regulatory factors. While it is established that alternative RNA processing is widespread in plants, the mechanisms by which alternative processing may be regulated are less well understood. The likelihood that calmodulin can inhibit the RNA-binding activities of all of the hypothetical forms of AtCPSF30 that may exist in the nucleus provides a direct conceptual link between processes that increase intracellular calcium levels and a number of RNA-processing events, and thus a different means by which RNA processing may be regulated by various stimuli. Further characterization of AtCPSF30 and of the larger At1g30460-encoded protein promises to add interesting new insight into the links between environment and RNA processing.

## CONCLUSION

In conclusion, these studies demonstrate that an Arabidopsis protein (AtCPSF30) that is related to the eukaryotic polyadenylation factor subunit CPSF30 is an RNA-binding protein. This protein also interacts with calmodulin such that, in the presence of calmodulin and calcium, its RNA-binding activity is inhibited. AtCPSF30 also interacts with itself, raising the possibility that it acts in RNA processing as a multimer. These results indicate that RNA processing in plants is probably regulated by stimuli that signal through calmodulin.

## MATERIALS AND METHODS

### Sequence Analysis

Sequence data were analyzed using Vector NTI software (Informax). Multiple amino acid sequence alignments were performed using ClustalX 1.83 and unrooted trees generated using the same software package. Trees were drawn using TreeView version 1.6.2. Calmodulin-binding predictions were performed using the service at http://calcium.uhnres.utoronto.ca/ctdb/ctdb/sequence.html (Yap et al., 2000).

### Isolation of the *oxt6* Mutant

The *oxt6* mutant, bearing a T-DNA insertion in the first exon of At1g30460, was identified in a collection of Arabidopsis (*Arabidopsis thaliana*) Columbia that had been mutagenized by T-DNA (pROK2) insertion; this collection was obtained from the Arabidopsis Biological Resource Center (Ohio State University). The position of the insertion and genotype of the plants were confirmed by PCR and RNA-blot analyses. This mutant is viable but has a somewhat dwarfed growth stature. Details of the isolation of the mutant and its characteristics will be published elsewhere.

### RNA Expression Analysis

RNA-blot analysis was conducted as described previously (Xu et al., 2004), using a probe derived from the first At1g30460 exon. Hybridizing signals were detected after exposure to a phosphor screen with a PhosphorImager (model 445SI; Molecular Dynamics). RT-PCR analysis of the smaller At1g30460-derived RNA was performed using the CPSF4-Nter and CPSF4-F primers (Table I). The CPSF4-F primer is situated downstream of the alternative splice site in this gene and will amplify only the smaller of the two At1g30460-encoded RNAs. This combination of primers spans the first intron in the gene and thus permits differentiation of transcripts from contaminating genomic DNA.

### Arabidopsis Nuclei Isolation and Nuclear Protein Extraction

Leaves of 3- to 4-week-old plants were used for nuclei isolation. Leaves were frozen and ground in liquid nitrogen to fine powder with a mortar and pestle. All subsequent steps were carried out on ice or at 4°C. About 20 g of tissue powder were mixed with 100 mL of extraction buffer (250 mm Suc, 10 mm PIPES-KOH, pH 7.0, 10 mm KCl, 10 mm MgCl2, 0.3% Triton X-100, 1 mm dithiothreitol, and 0.2 mm phenylmethylsulfonyl fluoride) on ice for 30 min, then filtered through two layers of Miracloth and subjected to centrifugation at 2,000*g* for 10 min. The pellet was resuspended in 15 mL of nuclei washing buffer (NWB; 0.4 M hexylene glycol, 10 mM PIPES-KOH, pH 7.0, 10 mM $MgCl_2$, 0.3% Triton X-100), filtered through a 100-$\mu$m nylon mesh, then layered on 15 mL of 30% Percoll in NWB and centrifuged again at 1,000*g* for 30 min. The pellet was resuspended in 20 mL of NWB and filtered through a 100-$\mu$m nylon mesh. Nuclei were further purified in a discontinuous Percoll gradient made by layering 10 mL of 30% Percoll in NWB on 10 mL of floating buffer (0.44 M Suc, 80% Percoll, 10 mM PIPES-KOH, pH 7.0, 10 mM $MgCl_2$). The gradient was centrifuged at 200*g* for 5 min. Most of the nuclei formed a layer just above floating buffer. They were removed, washed once with NWB (without addition of Triton X-100), and finally collected by centrifugation. This nuclei-enriched fraction was resuspended in protein sample buffer (0.175 M

Tris-HCl, pH 8.8, 5% SDS, 15% glycerol, 30 mM dithiothreitol) subjecting to SDS-PAGE gels for western-blot analysis. Nuclear protein extract used for coimmunoprecipitation was prepared using buffer X (20 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 150 mM NaCl, 0.2 mM EDTA, 0.5% Triton X-100, 25% glycerol, protease inhibitor cocktail [Sigma]) and treated twice with sonication (5 s each), followed by an incubation on ice for 4 h and centrifugation. Protein concentration was determined using the Bradford reagent (Sigma).

## Antibodies, Coimmunoprecipitation, and Immunoblotting

Antibodies were prepared against synthetic peptides (RFFRLYGECREQDC for AtCPSF30 and YNHRKERHLNGTVLC for AtCPSF100) from the predicted protein sequences of AtCPSF30 and AtCPSF100 (with an additional C each for conjugation purpose). After conjugation with KLH carrier protein using the Imject maleimide activated mcKLH kit (Pierce), the conjugant was injected into two rabbits (200 $\mu$g/injection). Each rabbit was given four booster injections about 2 weeks apart. Peptide-specific antibodies were purified using an affinity-purification procedure as described (Li et al., 1998). Briefly, the protein gels were transferred to Immobilon-P membranes (Millipore) and probed with affinity-purified antibodies. Detection was based on a color reaction of alkaline phosphatase-conjugated goat anti-rabbit antibodies and nitro blue tetrazolium chloride/5-bromo-4-chloro-3-indolyl phosphate.

For coimmunoprecipitation experiments, 20 $\mu$L of affinity-purified antibody was added to 200 $\mu$g of nuclear protein in 300 $\mu$L of PBST buffer (0.14 M NaCl, 0.008 M sodium phosphate, 0.002 M potassium phosphate, 0.01 M KCl, pH 7.4, 0.5% Triton X-100) and incubated overnight with gentle shaking at 4°C. Forty microliters of Affi-Gel protein A (Bio-Rad) was then added and incubated for 2 to 4 h at 4°C. The beads were collected by centrifugation and washed six times with 400 $\mu$L of PBST buffer. The affinity-bound proteins were eluted from the beads by boiling for 5 min in a SDS sample buffer (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS, 0.7 M $\beta$-mercaptoethanol, 0.1% bromphenol blue) before loading onto 10% SDS-PAGE for electrophoresis.

## Production of AtCPSF30 in *Escherichia coli*

DNAs encoding the AtCPSF30 coding region were isolated from total RNA by RT-PCR as described elsewhere (Addepalli et al., 2004). The coding region for the smaller At1g30460-encoded protein (Fig. 1A) was amplified from cDNA using primers Nter and F (Table I); primer F is unique to the smaller of the two At1g30460 transcripts and thus does not amplify the coding region for the N terminus of the AtCPSF30-YT521-B polypeptide. PCR products were cloned into pGEM and the inserts sequenced as described (Addepalli et al., 2004). The AtCPSF30 fragment was excised with *Bgl*II and cloned into *Bam*HI-digested pMALC2 (New England Biolabs). The resulting recombinant plasmid (pMALC2-AtCPSF30) was introduced into Rosetta (D3) cells (Novagen) for the production of protein.

Deletion derivatives of AtCPSF30 were amplified by PCR using the full-length AtCPSF30 cDNA clone as a template and the primers listed in Table I. The sequence of cloning into pGEM, sequencing, and cloning into pMALC2 described in the preceding paragraph was followed for each clone. The calmodulin-insensitive AtCPSF30 mutant was generated from the pMALC2-AtCPSF30 clone using the QuickChange XL site-directed mutagenesis kit (Stratagene) and the oligonucleotides indicated in Table I; several independent clones were isolated, the mutation confirmed by DNA sequencing, and further analysis performed as indicated in the following.

To produce MBP fusion proteins, 200 mL of LB media were inoculated with 10 mL of an overnight culture of transformed Rosetta (D3) cells and the 200-mL cultures grown at 37°C for 3 to 4 h. Expression of the fusion protein gene was induced by addition of 200 $\mu$L of 1 M isopropylthio-$\beta$-galactoside. After an additional 2 h of growth at 37°C, cells were harvested and resuspended in 5 mL of lysis buffer (GLB; 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, and 1 mm phenylmethylsulfonyl fluoride). Cells were disrupted by sonication (three bursts, 30 s each) and debris was removed by centrifugation. Extracts were then passed over columns of amylose resin (New England Biolabs; approximately 0.5–1 mL bed volume) that had been equilibrated in GLB. The columns were washed with 10 mL of GLB + 2 M NaCl, then 5 mL of GLB. MBP fusion proteins were eluted in 200 to 1,000 $\mu$L of GLB + 20 mM maltose. Purified proteins were dialyzed at 4°C for 12 h against 2,000 to 6,000 volumes of GLB or, for the experiment shown in Figure 7, 50 mM Tris-HCl, pH 8.0 + 150 mM NaCl. Protein preparations were analyzed by SDS-PAGE and staining with Coomassie Brilliant Blue, and quantities were estimated by comparison with known quantities of bovine serum albumin. MBP was purified from Rosetta (D3) cells transformed with pMAL-C2x using the same procedure.

## RNA-Binding Studies

For RNA binding, the electrophoretic mobility shift assays described elsewhere (Das Gupta et al., 1998) were adapted for this study. The labeled RNA was derived from the pea (*Pisum sativum*) *rbcS-E9* polyadenylation signal, and contained nucleotides extending from 145 nts upstream to 80 nts downstream from the site noted as "+1" by Mogen et al. (1992). To prepare the labeled RNA, templates were produced by PCR using plasmids carrying the full *rbcS-E9* polyadenylation signal (Mogen et al., 1992) as a template and the primers noted in Table I. Labeled RNA was prepared using the Ampliscribe kit (Epicentre). RNA-binding reactions contained varying quantities (ranging from 0.3–150 pmol) of purified protein and 4 to 5 pmol of labeled RNA in a volume of 20 $\mu$L of GLB. Typically, 5 $\mu$L of a 20-$\mu$L reaction was loaded on a nondenaturing polyacrylamide gel (4% acrylamide, 0.08% bisacrylamide, cast in 1× TBE) and separated in the cold. Gels were dried and exposed to a phosphor imager screen. Autoradiographs were analyzed using ImageJ software and the percentage of RNA bound in low-mobility complexes calculated.

For the experiment shown in Figure 7, RNA-binding reactions were supplemented with $CaCl_2$ (10 $\mu$M) or EGTA (1 mM), or purified Arabidopsis calmodulin (Cam6, the product of the Arabidopsis gene At5g21274, provided by Dr. Ray Zielinski, University of Illinois).

## Detection of Calmodulin-Binding Proteins

The binding of recombinant proteins to calmodulin was assessed using the Stratagene Affinity CBP detection system (catalog no. 200370). In some cases, biotinylated calmodulin from STI Signal Transduction Products was used in place of the same reagent from Stratagene.

## In Vitro Translations and Protein-Protein Interaction Assays

DNA fragments encoding various parts of AtCPSF30 were cloned into pCITE (Novagen) as the same *Bgl*II fragments used to clone them into pMALC2. Recombinant pCITE plasmids were used to program in vitro translation reactions using the STP3 (Novagen) kit, exactly as recommended by the manufacturer. (Note that this protocol includes a postreaction treatment with RNAse A.) Five microliters of the translation reactions were added to 1 $\mu$L of MBP or MBP-AtCPSF30 (containing approximately 1 $\mu$g of purified protein, in GLB). After 30 min at 30°C, the reactions were added to a mixture containing 25 $\mu$L (packed resin volume) of amylose resin (New England Biolabs) in 100 $\mu$L of GLB + 0.1% Nonidet P-40. After 5 min of gentle rocking at room temperature, the resin was collected by centrifugation (5 s in a microcentrifuge), washed three times with GLB + 0.1% Nonidet P-40, and suspended in 25 $\mu$L of SDS-PAGE sample buffer. After boiling for 5 min, 10-$\mu$L aliquots were separated by SDS-PAGE, and the gels were dried and analyzed by autoradiography using a phosphorimager. For the experiment shown in Figure 7C, calmodulin (1 $\mu$g), $CaCl_2$ (final concentration of 3 mM), and EGTA (final concentration of 3 mM) were added in additional volumes of 3 $\mu$L total. The volumes of the controls in this experiment were adjusted accordingly.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession number AY140901.

## ACKNOWLEDGMENTS

We thank Carol Von Lanken for technical assistance, and Dr. Balasubrahmanyam Addepalli and Dr. Ray Zielinski for gifts of MBP and calmodulin, respectively.

Received September 29, 2005; revised December 20, 2005; accepted January 27, 2006; published March 10, 2006.

## LITERATURE CITED


**Addepalli B, Meeks LR, Forbes KP, Hunt AG** (2004) Novel alternative splicing of mRNAs encoding poly(A) polymerases in *Arabidopsis*. Biochim Biophys Acta **1679:** 117–128

**Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ** (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389–3402

**Amasino RM** (2003) Flowering time: a pathway that begins at the 3′ end. Curr Biol **13:** 670–672

**Bai C, Tolias PP** (1998) Drosophila clipper/CPSF 30K is a post-transcriptionally regulated nuclear protein that binds RNA containing GC clusters. Nucleic Acids Res **26:** 1597–1604

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

**Barabino SM, Ohnacker M, Keller W** (2000) Distinct roles of two Yth1p domains in 3′-end cleavage and polyadenylation of yeast pre-mRNAs. EMBO J **19:** 3778–3787

**Bond GL, Prives C, Manley JL** (2000) Poly(A) polymerase phosphorylation is dependent on novel interactions with cyclins. Mol Cell Biol **20:** 5310–5320

**Bourgeois CF, Lejeune F, Stevenin J** (2004) Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. Prog Nucleic Acid Res Mol Biol **78:** 37–88

**Colgan DF, Murthy KG, Prives C, Manley JL** (1996) Cell-cycle related regulation of poly(A) polymerase by phosphorylation. Nature **384:** 282–285

**Colgan DF, Murthy KG, Zhao W, Prives C, Manley JL** (1998) Inhibition of poly(A) polymerase requires p34cdc2/cyclin B phosphorylation of multiple consensus and non-consensus sites. EMBO J **17:** 1053–1062

**Das Gupta J, Li Q, Hunt AG** (1998) Characterization of two RNA binding factors from pea nuclei. J Plant Biochem Biotechnol **7:** 1–5

**Dass B, Attaya EN, Michelle Wallace A, MacDonald CC** (2001a) Over-expression of the CstF-64 and CPSF-160 polyadenylation protein messenger RNAs in mouse male germ cells. Biol Reprod **64:** 1722–1729

**Dass B, McMahon KW, Jenkins NA, Gilbert DJ, Copeland NG, MacDonald CC** (2001b) The gene for a variant form of the polyadenylation protein CstF-64 is on chromosome 19 and is expressed in pachytene spermatocytes in mice. J Biol Chem **276:** 8044–8050

**Dean C, Tamaki S, Dunsmuir P, Favreau M, Katayama C, Dooner H, Bedbrook J** (1986) mRNA transcripts of several plant genes are polyadenylated at multiple sites in vivo. Nucleic Acids Res **14:** 2229–2240

**Elliott BJ, Dattaroy T, Meeks-Midkiff LR, Forbes KP, Hunt AG** (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

**Forbes KP, Addepalli B, Hunt AG** (2006) An Arabidopsis Fip1 homologue interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. J Biol Chem **281:** 176–186

**Hartmann AM, Nayler O, Schwaiger FW, Obermeier A, Stamm S** (1999) The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). Mol Biol Cell **10:** 3909–3926

**Hayashi M, Nishimura M** (1999) Light regulates alternative splicing of hydroxypyruvate reductase in pumpkin. Plant J **17:** 309–320

**Hunt A** (1994) Messenger RNA 3′ end formation in plants. Annu Rev Plant Physiol Plant Mol Biol **45:** 47–60

**Iida K, Seki M, Sakurai T, Satou M, Akiyama K, Toyoda T, Konagaya A, Shinozaki K** (2004) Genome-wide analysis of alternative pre-mRNA splicing in *Arabidopsis thaliana* based on full-length cDNA sequences. Nucleic Acids Res **32:** 5096–5103

**Imai Y, Matsuo N, Ogawa S, Tohyama M, Takagi T** (1998) Cloning of a gene, YT521, for a novel RNA splicing-related protein induced by hypoxia/reoxygenation. Brain Res Mol Brain Res **53:** 33–40

**Isshiki M, Nakajima M, Satoh H, Shimamoto K** (2000) dull: rice mutants with tissue-specific effects on the splicing of the waxy pre-mRNA. Plant J **23:** 451–460

**Jordan T, Schornack S, Lahaye T** (2002) Alternative splicing of transcripts encoding Toll-like plant resistance proteins: What's the functional relevance to innate immunity? Trends Plant Sci **7:** 392–398

**Kazan K** (2003) Alternative splicing and proteome diversity in plants: the tip of the iceberg has just emerged. Trends Plant Sci **8:** 468–471

**Klahre U, Hemmings-Mieszczak M, Filipowicz W** (1995) Extreme heterogeneity of polyadenylation sites in mRNAs encoding chloroplast RNA-binding proteins in Nicotiana plumbaginifolia. Plant Mol Biol **28:** 569–574

**Kong J, Gong JM, Zhang ZG, Zhang JS, Chen SY** (2003) A new AOX homologous gene OsIM1 from rice (Oryza sativa L.) with an alternative splicing mechanism under salt stress. Theor Appl Genet **107:** 326–331

**Li QQ, Das Gupta J, Hunt AG** (1998) Polynucleotide phosphorylase is a component of a novel plant poly(A) polymerase. J Biol Chem **273:** 17539–17543

**Manen JF, Simon P** (1993) A possible explanation for the multiple polyadenylation sites in transcripts coding for a winged-bean leghemoglobin. Planta **191:** 289–292

**Mano S, Hayashi M, Nishimura M** (1999) Light regulates alternative splicing of hydroxypyruvate reductase in pumpkin. Plant J **17:** 309–320

**Mano S, Hayashi M, Nishimura M** (2000) A leaf-peroxisomal protein, hydroxypyruvate reductase, is produced by light-regulated alternative splicing. Cell Biochem Biophys **32:** 147–154

**Marrs KA, Walbot V** (1997) Expression and RNA splicing of the maize glutathione S-transferase Bronze2 gene is regulated by cadmium and other stresses. Plant Physiol **113:** 93–102

**Matthews JM, Sunde M** (2002) Zinc fingers—folds for many occasions. IUBMB Life **54:** 351–355

**Meyers BC, Vu TH, Tej SS, Ghazal H, Matvienko M, Agrawal V, Ning J, Haudenschild CD** (2004) Analysis of the transcriptional complexity of *Arabidopsis thaliana* by massively parallel signature sequencing. Nat Biotechnol **22:** 1006–1011

**Mogen BD, MacDonald MH, Leggewie G, Hunt AG** (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

**Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM** (1998) Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3′ end formation of cellular pre-mRNAs. Mol Cell **1:** 991–1000

**Onions J, Hermann S, Grundstrom T** (1997) Basic helix-loop-helix protein sequences determining differential inhibition by calmodulin and S-100 proteins. J Biol Chem **272:** 23930–23937

**Quesada V, Macknight R, Dean C, Simpson GG** (2003) Autoregulation of FCA pre-mRNA processing controls *Arabidopsis* flowering time. EMBO J **22:** 3142–3152

**Reddy AS** (2004) Plant serine/arginine-rich proteins and their role in pre-mRNA splicing. Trends Plant Sci **9:** 541–547

**Rothnie HM** (1996) Plant mRNA 3′-end formation. Plant Mol Biol **32:** 43–61

**Saarikettu J, Sveshnikova N, Grundstrom T** (2004) Calcium/calmodulin inhibition of transcriptional activity of E-proteins by prevention of their binding to DNA. J Biol Chem **279:** 41004–41011

**Sanford JR, Longman D, Caceres JF** (2003) Multiple roles of the SR protein family in splicing regulation. Prog Mol Subcell Biol **31:** 33–58

**Simpson GG, Dijkwel PP, Quesada V, Henderson I, 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

**Staiger D, Zecca L, Wieczorek Kirk DA, Apel K, Eckstein L** (2003) The circadian clock regulated RNA-binding protein AtGRP7 autoregulates its expression by influencing alternative splicing of its own pre-mRNA. Plant J **33:** 361–371

**Stoilov P, Rafalska I, Stamm S** (2002) YTH: a new domain in nuclear proteins. Trends Biochem Sci **27:** 495–497

**Stoss O, Novoyatleva T, Gencheva M, Olbrich M, Benderska N, Stamm S** (2004) p59(fyn)-mediated phosphorylation regulates the activity of the tissue-specific splicing factor rSLM-1. Mol Cell Neurosci **27:** 8–21

**Tacahashi Y, Helmling S, Moore CL** (2003) Functional dissection of the zinc finger and flanking domains of the Yth1 cleavage/polyadenylation factor. Nucleic Acids Res **31:** 1744–1752

**Takagaki Y, Manley JL** (1998) Levels of polyadenylation factor CstF-64 control IgM heavy chain mRNA accumulation and other events associated with B cell differentiation. Mol Cell **2:** 761–771

**Veraldi KL, Arhin GK, Martincic K, Chung-Ganster LH, Wilusz J, Milcarek C** (2001) hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. Mol Cell Biol **21:** 1228–1238

**Wallace AM, Dass B, Ravnik SE, Tonk V, Jenkins NA, Gilbert DJ, Copeland NG, MacDonald CC** (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 USA **96:** 6763–6768

**Wallace AM, Denison TL, Attaya EN, MacDonald CC** (2004) Developmental distribution of the polyadenylation protein CstF-64 and the variant tauCstF-64 in mouse and rat testis. Biol Reprod **70:** 1080–1087

**Wilkinson FL, Holaska JM, Zhang Z, Sharma A, Manilal S, Holt I, Stamm S, Wilson KL, Morris GE** (2003) Emerin interacts *in vitro* with the splicing-associated factor, YT521-B. Eur J Biochem **270:** 2459–2466

**Xu R, Ye X, Li QQ** (2004) AtCPSF73-II gene encoding an Arabidopsis homolog of CPSF 73 kDa subunit is critical for early embryo development. Gene **324:** 35–45

**Yap KL, Kim J, Truong K, Sherman M, Yuan T, Ikura M** (2000) Calmodulin target database. J Struct Funct Genomics **1:** 8–14

**Zhou Y, Zhou C, Ye L, Dong J, Xu H, Cai L, Zhang L, Wei L** (2003) Database and analyses of known alternatively spliced genes in plants. Genomics **82:** 584–595

**Zhu W, Schlueter SD, Brendel V** (2003) Refined annotation of the Arabidopsis genome by complete expressed sequence tag mapping. Plant Physiol **132:** 469–484