Interactions of SR45, an SR-like protein, with spliceosomal proteins and an intronic sequence: insights into regulated splicing

Irene S. Day1,†, Maxim Golovkin1,†, Saiprasad G. Palusa1, Alicia Link1, Gul S. Ali1,‡, Julie Thomas1, Dale N. Richardson2 and Anireddy S. N. Reddy1,*

1Department of Biology, Program in Molecular Plant Biology, Program in Cell and Molecular Biology, Colorado State University, Fort Collins, CO 80523, USA, and 2CIBIO, University of Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal

Received 2 February 2012; revised 27 April 2012; accepted 1 May 2012; published online 28 June 2012.

*For correspondence (e-mail reddy@colostate.edu).
†These authors contributed equally.
‡Present address: Mid-Florida Research & Education Center, Apopka, FL 32703, and Department of Plant Pathology, Institute of Food & Agricultural Sciences, University of Florida, Gainesville, FL 32611, USA.

SUMMARY

SR45 is a serine/arginine-rich (SR)-like protein with two arginine-serine-rich (RS) domains. We have previously shown that SR45 regulates alternative splicing (AS) by differential selection of 5′ and 3′ splice sites. However, it is unknown how SR45 regulates AS. To gain mechanistic insights into the roles of SR45 in splicing, we screened a yeast two-hybrid library with SR45. This screening resulted in the isolation of two spliceosomal proteins, U1-70K and U2AF35b that are known to function in 5′ and 3′ splice site selection, respectively. This screen not only confirmed our prior observation that U1-70K and SR45 interact, but also helped to identify an additional interacting partner (U2AF35). In vitro and in vivo analyses revealed an interaction of SR45 with both paralogs of U2AF35. Furthermore, we show that the RS1 and RS2 domains of SR45, and not the RNA recognition motif (RRM) domain, associate independently with both U2AF35 proteins. Interaction studies among U2AF35 paralogs and between U2AF35 and U1-70K revealed that U2AF35 can form homo- or heterodimers and that U2AF35 proteins can associate with U1-70K. Using RNA probes from SR30 intron 10, whose splicing is altered in the sr45 mutant, we show that SR45 and U2AF35b bind to different parts of the intron, with a binding site for SR45 in the 5′ region and two binding regions, each ending with a known 3′ splice site, for U2AF35b. These results suggest that SR45 recruits U1snRNP and U2AF to 5′ and 3′ splice sites, respectively, by interacting with pre-mRNA, U1-70K and U2AF35 and modulates AS.

Keywords: serine/arginine-rich protein, pre-mRNA splicing, alternative splicing, Arabidopsis, splicing regulator, SR30, SR45.

INTRODUCTION

Precursor mRNA (pre-mRNA) contains protein coding sequences (exons) and intervening non-coding sequences (introns). In order to produce functional mRNA, the introns must be excised precisely from the pre-mRNA. This removal takes place in the spliceosome, a macromolecular complex composed of five (U1, U2, U4, U5 and U6) small nuclear ribonucleoproteins (snRNPs) and many non-snRNP splicing factors (Sharp, 1994; Reddy, 2001; Zhou et al., 2002). In pre-mRNAs with multiple introns, besides constitutive splicing, alternative splicing (AS) can result in multiple transcripts and proteins from a single gene (Reddy, 2007; Barbazuk et al., 2008). Recent genome-wide studies indicate that pre-mRNAs from 95% of intron-containing genes in humans, over 60% in Arabidopsis and about 48% in rice undergo AS (Pan et al., 2008; Filichkin et al., 2010; Lu et al., 2010). Furthermore, AS is regulated by developmental cues and in response to various abiotic stresses in plants (Palusa et al., 2007; Filichkin et al., 2010; Reddy and Ali, 2011). Regulated splicing has been studied extensively in animals using cell-free splicing extracts as well as in vivo assays and it plays an important role in diverse biological processes (Long and Caceres, 2009; Kalsotra and Cooper, 2011). The assembly of the spliceosome and the process of splicing involve many RNA–RNA, RNA–protein and protein–protein interactions.
The recognition of exonic and intronic sequences and selection of 5′ and 3′ splice sites is required for excision of introns and joining of exons (Chasin, 2007; Reddy et al., 2012). In animals it has been shown that recognition of the 5′ splice site involves U1 snRNP, U1-70K and serine–arginine-rich (SR) proteins (Kohtz et al., 1994). The U2 auxiliary factor (U2AF), a heterodimer consisting of U2AF35 and U2AF65 recognizes the 3′ splice site and adjacent polypyrimidine tract, respectively, and this binding is thought to recruit U2snRNP (Wu et al., 1999; Wahl et al., 2009; Mackereth et al., 2011).

Plant introns differ from animal introns in their size, nucleotide composition, branch point sequence, and polypyrimidine tract (Reddy, 2001). Due to these differences, most plant intron-containing transcripts are either not processed or processed incorrectly in mammalian splicing extracts (McCullough et al., 1991; Schuler, 2008). Furthermore, plants and vertebrates differ in prevalence of types of AS events with a high occurrence of intron retention in plants and a high frequency of exon skipping in animals (Reddy et al., 2012). These differences also indicate that some of the initial events involved in splice site recognition are likely to be unique in plants. There is no plant-derived cell-free splicing extract available to study the mechanisms of splicing regulation in plants. In order to dissect the interaction network of proteins involved in splicing in plants, in vivo methods must be used. Many of the factors involved in plant splicing have been identified by their sequence similarity to animal splicing factors. Interactions between plant splicing factors and evidence that they are involved in splicing have been studied using protein–protein interaction studies such as Y2H analysis and pull-down assays. Previously, using U1-70K as bait in Y2H screens, three Arabidopsis SR proteins (RS221, RS222, SCL33) and an SR-like protein (SR45) were identified (Golovkin and Reddy, 1998, 1999).

SR proteins have multiple roles in pre-mRNA splicing (Graveley, 2000; Lorkovic and Barta, 2002; Barta et al., 2008; Reddy and Ali, 2011). They have a C-terminal arginine/serine (RS) rich domain and one or two RNA recognition motifs (RRM) (Barta et al., 2010). The RRM confers RNA-binding specificity and the RS domain mediates protein–protein and protein–RNA interactions (Graveley, 2000; Reddy, 2001; Shen et al., 2004). About 18 SR proteins have been identified in Arabidopsis (Barta et al., 2010). SR45 (AT1g16610) is an SR-like protein with two RS domains, one N-terminal and one C-terminal, whereas SRs have only one C-terminal RS domain (Golovkin and Reddy, 1999). SR45 has orthologs in other flowering plants but was not found in algae (Ali et al., 2007). SR45 was shown to be an essential splicing factor in a complementation assay (Ali et al., 2007) and its interaction with U1-70K along with its localization in nuclear speckles was demonstrated using bimolecular fluorescence complementation (BIFC) (Ali et al., 2008). The SR proteins are involved in constitutive and alternative splicing (AS) and pre-mRNAs of many of them are alternatively spliced themselves (Palusa et al., 2007). SR45 undergoes AS to produce two splice forms, SR45.1 and SR45.2, and its function in AS has been indicated by the altered splicing patterns of several other SR genes in an SR45 mutant (Ali et al., 2007). The sr45-1 mutant plants show delayed flowering, altered leaf morphology, reduced root growth and flowers with abnormal petal and stamen numbers (Ali et al., 2007). A gene complementation study revealed that one splice form of SR45 rescues the flower phenotype while the other splice form rescues the root phenotype (Zhang and Mount, 2009). In addition to the developmental phenotype, the mutant displayed hypersensitivity to glucose and asparagin acid, which was complemented by both long and short isoforms (Carvalho et al., 2010). Recently, defects in DNA methylation were also found in the sr45-1 mutant (Austin et al., 2012).

In this study, we used SR45 as bait in a yeast two-hybrid (Y2H) screen. Isolation of U1-70K in this screen verified the interaction reported previously between these two proteins (Golovkin and Reddy, 1999). Additionally, we identified U2AF35, the small subunit of the U2AF complex involved in recognition of the 3′ splice site in animals (Wu et al., 1999; Zorio and Blumenthal, 1999), as another interacting partner of SR45. As two paralogs of U2AF35 are known in Arabidopsis, U2AF35a and U2AF35b (Wang and Brendel, 2006), we performed Y2H assays, protein pull-down assays and BIFC with both proteins and confirmed the interaction of SR45 with U2AF35b and its paralog U2AF35a. Furthermore, our studies showed interaction of the two paralogs as heterodimers and homodimers. Detailed domain analysis indicates that the two SR45 RS domains (RS1 and RS2) could interact independently with U2AF35a while a C-terminal domain (PSD) of U2AF35b that is present only in photosynthetic eukaryotes and protozoans is not necessary for interaction. Using RNA-binding studies we demonstrate that SR45 binds to the 5′ region of intron 10 in SR30 whereas U2AF35b binds two distinct regions of the same intron, each ending with an experimentally verified 3′ splice site. As SR45 interacts with both U1-70K and U2AF35 and both these proteins bind specific intronic sequences of SR30, it is likely that SR45 functions in 5′ and 3′ site selection and may bridge the 5′ and 3′ components of the spliceosome.

RESULTS

SR45 interacts with U2AF35 in a yeast two-hybrid assay

Using a loss-of-function mutant of SR45, we have shown previously that SR45 regulates AS of the pre-mRNAs of SR genes (Ali et al., 2007). To uncover its mode of action in splicing we searched for the interacting protein partners of SR45 by using it as bait in a Y2H screen. We screened about 50 000 transformants and identified 11 positive clones. U1-
70K, a U1snRNP specific protein, which we reported previously to interact with SR45 (Golovkin and Reddy, 1999) and U2AF35b, another interacting protein partner of SR45 were each represented once in the screen. U2AF35b is one of two U2AF35 paralogs (a and b) in Arabidopsis that form the small subunit of the U2AF complex, which functions in recognition of the 3'-splice site (Wu et al., 1999; Wang and Brendel, 2006). Control β-galactosidase assays for SR45 and the rescued U2AF35b construct alone showed no β-galactosidase activity whereas a double transformant with rescued U2AF35b and SR45 was positive for β-galactosidase activity (Figure 1a). As the isolated U2AF35b in the screen was truncated (aa 125–283), a full-length U2AF35b clone was cloned and its interaction with SR45 was tested using Y2H. Full-length U2AF35b was also shown to interact with SR45 (Figure S1a). Three SR proteins, SCL33, RSZ22 and RSZ21, from two different families of SR proteins, which, like SR45, have been previously confirmed to interact with U1-70K (Golovkin and Reddy, 1998, 1999) were also tested for interaction with U2AF35a and U2AF35b. None of the tested SR proteins interacts with either U2AF35 paralog (Figure S1b), which indicates a specific interaction of U2AF35 proteins with SR45.

A C-terminal domain is present in plant U2AF35 but not in animal U2AF35

Single to multiple copies of U2AF35 genes have been identified in fission yeast and animals, and flowering plants have more than one U2AF35 (Wang and Brendel, 2006). An earlier analysis of 25 sequences of U2AF35 proteins in plants and animals identified a domain in the C-terminal end of the plant proteins that was not present in the animal proteins (Wang and Brendel, 2006). BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) searches using plant U2AF35 sequences also identified this domain in non-photosynthetic protzoans such as Plasmodium and Toxoplasma; however, in most cases the domain was less conserved in the protozoans than in plants. Further analysis of U2AF35 sequences from flowering plants, gymnosperms, moss, algae and animals (84 sequences in total) confirmed the presence of this domain in photosynthetic eukaryotes but not in animals (Figure S2). These 84 U2AF35 homologs form several sub-families in a maximum likelihood gene tree (Figure S3). Although the animal sequences lack the conserved C-terminal PSD, they form a sister group with high bootstrap support values with algal and moss U2AF35 sequences.

Confirmation of SR45 interaction with both U2AF35 paralogs using pull-down assays

To verify the interactions between SR45 and the U2AF35b and to see if SR45 also interacts with U2AF35a we used in vitro pull-down assays with bacterially produced protein. A clone was also isolated for the U2AF35a paralog that includes amino acids 51–296 (see Experimental procedures).

Figure 1. Yeast two-hybrid interactions and confirmation of interaction between SR45 and U2AF35a and U2AF35b. (a) Yeast strain HF7c (cells), cells plus C-terminally truncated SR45/pMC86 (SR45), cells plus U2AF35N-terminally truncated/pACT1 (U2AF35Ntrb) and cells plus both constructs (SR45 + U2AF35Ntrb) were grown on yeast extract-peptone-glucose medium (YPD) and, as controls, on synthetic, minimal medium SD media–W, L, H with 3-amino-1,2,4-triazole [–3(ΑΤ)], SD media–L, and SD media–W. A duplicate plate of yeast on YPD was transferred to a filter medium SD media–W–LYPD. (b) Schematic diagram of U2AF35a and U2AF35b full-length and truncated clones showing the RRM domain (dark green) and C-terminal motif (yellow). (c) Pull-down assays. SR45 protein was expressed as a fusion to S-tag and the U2AF35a, b, Crb, and Ntrb proteins were expressed as fusions to T7-tag. Crude bacterially expressed SR45 protein was incubated with S-protein, beads were washed and then incubated in the U2AF35 bacterially expressed crude proteins. Beads were washed, the proteins were released using loading buffer, electrophoresed on duplicate gels, blotted and probed with T7-antibody or S-protein. Beads incubated in the U2AF35 proteins alone were used as a control.

Expression clones as fusions to T7-tag were made for full-length U2AF35b, the N-terminally truncated U2AF35b isolated in the Y2H screen (U2AF35Ntrb), U2AF35b truncated at amino acid 250 to delete the C-terminal PSD (U2AF35Ctrb), and U2AF35a (amino acids 51–296) (see Figure 1b). The SR45 construct was made previously as a fusion to S-tag (Golovkin and Reddy, 1999). As shown in Figure 1c, all forms of the U2AF35 proteins were pulled down by the SR45 bound to the beads. As a control, S-protein agarose beads were
incubated with crude extracts containing U2AF<sup>35</sup> proteins. None of the U2AF<sup>35</sup> proteins was pulled down by the beads alone. As SR45 could bind to truncated U2AF<sup>35b</sup>, missing the first 125 (Ntrb) or the last 33 (Ctrb) amino acids, the interaction must take place between amino acids 125 and 250 and the PSD is not necessary for this interaction.

**Localization of SR45/U2AF<sup>35</sup> interactions using BiFC**

SR45 has been shown to co-localize with U1-70K in nuclear speckles (Ali et al., 2008). U2AF<sup>35b</sup>a and U2AF<sup>35b</sup>b have also been observed in nuclear speckles (Wang and Brendel, 2006). As an in vivo corroboration of association between SR45 and U2AF<sup>35</sup> and to determine the location of this association, we utilized BiFC (Walter et al., 2004). This approach is based on the reconstitution of fluorescence of the split halves of yellow fluorescent protein (YFP) (N- and C-terminal) when they are fused to two putative interacting proteins. If the proteins are interacting and, therefore, come close enough to reconstitute the YFP, fluorescence can occur. U2AF<sup>35b</sup>a (amino acids 51-296), full-length U2AF<sup>35b</sup>b and U2AF<sup>35b</sup>b Ctrb were cloned into a BiFC vector as fusions to the N-terminal region of YFP (YFP<sup>N</sup>). As we had a full-length U2AF<sup>35b</sup>b clone we did not use the N-terminally truncated clone U2AF<sup>35b</sup>b Ntr but did use U2AF<sup>35b</sup>b Ctrb as we were interested in seeing if the PSD influences the localization of the interacting proteins. SR45 had been cloned earlier into a BiFC vector as a fusion to the C-terminal region of YFP (SR45/ YFP<sup>C</sup>) (Ali et al., 2008). Arabidopsis protoplasts were co-transfected with the SR45/YFP<sup>C</sup> construct and each of the U2AF<sup>35b</sup>/YFP<sup>N</sup> constructs and examined for fluorescence. Figure 2 shows fluorescence observed in representative protoplasts transformed with SR45/YFP<sup>C</sup> and U2AF<sup>35b</sup>a/YFP<sup>N</sup>, U2AF<sup>35b</sup>b/YFP<sup>N</sup> or U2AF<sup>35b</sup>b Ctrb/YFP<sup>N</sup>. The fluorescence is seen mostly in nuclear speckles (concentrated loci of various sizes) with some fluorescence in the nucleoplasm (diffuse fluorescence) as has been shown previously to be the location of SR45 (Ali et al., 2008). The fluorescence resulting from the interaction of SR45/YFP<sup>C</sup> and U2AF<sup>35b</sup>b Ctrb/YFP<sup>N</sup> appeared in most cases to be in smaller speckles and more diffuse (Figure 2, lower panels), although in some occasional nuclei the speckles looked similar to the full-length U2AF<sup>35b</sup>b/YFP<sup>N</sup>:SR45/YFP<sup>C</sup> speckles.

Thus, results from three independent approaches – Y2H interactions and protein pull-down and BiFC assays – support the interaction of SR45 with both U2AF<sup>35a</sup>b and U2AF<sup>35b</sup>b. As the U2AF<sup>35b</sup>b Ctrb constructs lack the PSD, it is apparent that this domain is not necessary for interaction with SR45. However, the C-terminal domain of U2AF<sup>35b</sup>b may have a role in proper localization of the proteins in speckles given that the pattern of fluorescence was more diffuse than with full-length protein.

**RS1 and RS2 of SR45 associate with U2AF<sup>35</sup> proteins independently**

SR45 has a modular structure comprised of an N-terminal RS domain (RS1), a central RRM domain and a C-terminal RS domain (RS2). We BLASTed the Phytozome databases with Arabidopsis SR45 and recovered 24 SR45 homologs (including AtSR45). Figure S4 shows an alignment of the homologs and the location of the domains. These protein sequences formed three major groupings: dicots, monocots and mosses (Figure S5). The domains identified in this alignment were used to determine which part(s) of SR45 interact with the U2AF<sup>35</sup> proteins. A series of deletion mutants of SR45 in a BiFC vector as fusions to YFP<sup>C</sup> was used in BiFC assays with the U2AF<sup>35b</sup>/YFP<sup>N</sup> constructs. The SR45/YFP<sup>C</sup> constructs included SR45 RS1, SR45 RRM, SR45 RS2, SR45 RS1 + RRM and SR45 RRM + RS2 (Figure 3a) (Ali et al., 2008). The SR45/YFP<sup>C</sup> constructs were tested with the YFP<sup>N</sup> constructs of U2AF<sup>35a</sup>b, U2AF<sup>35b</sup>b and U2AF<sup>35b</sup>b Ctrb. The protoplasts transfected with either the SR45 RS1/YFP<sup>C</sup> or SR45 RS2/YFP<sup>C</sup> construct and each U2AF<sup>35b</sup>/YFP<sup>N</sup> construct showed fluorescence, which indicated an in vivo association of the proteins (Figure 3b–d, rows 1 and 3). However, protoplasts transfected with SR45 RRM/YFP<sup>C</sup> and each U2AF<sup>35b</sup>/YFP<sup>N</sup> construct did not show fluorescence (row 2). Furthermore, while the protoplasts transfected with the SR45 RS2 + RRM/YFP<sup>C</sup> and each U2AF<sup>35b</sup>/YFP<sup>N</sup> construct showed fluorescence (row 5), the protoplasts transfected with the SR45 RS1 + RRM/YFP<sup>N</sup> construct showed fluorescence only when co-transfected with U2AF<sup>35b</sup>b Ctrb/YFP<sup>N</sup> (row 4). The fluorescence in protoplasts transfected with the SR45 RS2/YFP<sup>C</sup> and U2AF<sup>35a</sup>b/YFP<sup>N</sup> constructs (Figure 3b, row 2) appeared in smaller diffuse speckles throughout the nucleus as compared with the full-length SR45/YFP<sup>C</sup> construct with the U2AF<sup>35b</sup>b/YFP<sup>C</sup> construct (Fig. 2, upper panels). While in protoplasts transfected with the SR45 RS1/YFP<sup>C</sup> and U2AF<sup>35b</sup>b/YFP<sup>N</sup> constructs, the fluorescence was more diffuse throughout the nucleus and speckles were very fine.
as compared with protoplasts transfected with the full-length SR45/YFPC or SR45RS2/YFPC (Compare Fig. 2 middle panels to Fig. 3c). The fluorescence in the protoplasts transfected with the SR45/YFPC sub-domain constructs and the U2AF35Ctrb appeared in most cases to be in larger speckles than when full-length SR45/YFPC was co-transfected with U2AF35Ctrb (compare Figure 2 lower panels with Figure 2d). These results suggest that although U2AF35 can interact with the SR45RS1 and SR45RS2 domains independently, other domains of the protein modulate the strength and specificity of this interaction. Furthermore, observed fluorescence with the SR45RS1 + RRM and U2AF35Ctrb constructs but not with the U2AF35a and U2AF35b constructs that contained the PSD, indicates that the C-terminal domain is likely to inhibit or interfere with their interaction with the SR45RS1 + RRM. In the case of the SR45RS1 + RRM construct, the RS1 domain is followed by the RRM while in the case of SR45RS2 + RRM construct, the RRM precedes the RS2 domain. It appears that when the C-terminal PSD is present, as in U2AF35a and full-length U2AF35b, U2AF35 cannot bind to the SR45RS1 + RRM. Whereas U2AF35 without this domain (U2AF35Ctrb) can bind, which leads to the possibility that there is steric hindrance between the SR45RRM domain and the U2AF35 C-terminal portion that does not allow association with the SR45RS1 domain but does allow interaction with the SR45RS2 domain. Possibly, U2AF35 binds to the SR45RS2 domain in the full-length protein leaving the SR45RS1 domain to interact with other proteins.

**U2AF35 proteins form hetero- and homodimers**

U2AF is a heterodimer comprised of U2AF35 and U2AF65. However, using fluorescence resonance energy transfer (FRET) microscopy in animal cell lines it was shown that U2AF35 proteins self-interact (Chusainow et al., 2005). Because of this finding with animal U2AF35 proteins, we used Y2H assays to test if plant U2AF35 proteins also form hetero- and homodimers. U2AF35a and U2AF35b were shown to form homo- and heterodimers (Fig. 4, left panels). We then used BiFC to test the association of the U2AF35 proteins with themselves. In addition to the U2AF35 (a, b and Ctrb)/YFPN fusions generated previously for interaction studies with SR45, a YFPc construct was made for each U2AF35 (a, b and Ctrb). Protoplasts transfected with U2AF35a/YFPN together with U2AF35a/YFPc, U2AF35b/YFPc or U2AF35Ctrb/YFPc, and U2AF35b/YFPN with U2AF35b/YFPc or U2AF35Ctrb/YFPc and U2AF35Ctrb/YFPN with U2AF35Ctrb/YFPc were evaluated for the reconstitution of YFP fluorescence. As shown in Figure 4(a–c) (right panels), each set of possible interacting proteins exhibited fluorescence, which indicated that the U2AF35 proteins can form homo- and heterodimers and that the PSD, absent in the U2AF35Ctrb construct, is not necessary for this interaction. However, the localization of the dimer pairs was somewhat different.

**U2AF35 interacts with U1-70K**

As both U1-70K and U2AF35 interact with SR45 and have the same localization pattern in the nucleus, we further investigated the association between U1-70K and U2AF35 using BiFC. The two U2AF35b proteins showed association in the BiFC assays, while U2AF35a did not (Figure 5a). As we expected both paralogs to associate with U1-70K, we used Y2H assays using U1-70K full length, an N-terminal U1-70K (aa 1–172) and a C-terminal region (aa 179–427), each containing about half of the RRM, as fusions to the DNA-binding domain (Golovkin and Reddy, 1998) paired with U2AF35a as
fusions to the activating domain. Earlier we had used these two halves of U1-70K to determine if the interaction between two SR proteins (RSZ21 and RSZ22) and U1-70K was in the region before the RRM or in the region following the RRM (Golovkin and Reddy, 1998). SR45 itself (along with SCL33) had been isolated in a Y2H screen using just the C-terminal part of U1-70K (Golovkin and Reddy, 1999). As shown in Figure 5(b), the U2AF35a paralog could interact with full-length U1-70K and with the N-terminal but not the C-terminal region, a finding that suggested that U2AF35a interacts with U1-70K somewhere in the N-terminal region. The negative BiFC results with U2AF35a could be due to instability of the complex or some post-translational modifications that occur in U2AF35a only in plant cells that might prevent interaction.

Recombinant SR45 and U2AF35b proteins bind to different parts of SR30 intron 10 RNA

Loss of SR45 function in Arabidopsis showed altered splicing patterns of several SR genes, which included changes in AS of SR30 pre-mRNA (Ali et al., 2007). Expression levels of SR30 alternatively spliced products changed in vegetative tissues (root, stem and leaf) of sr45-1 plants, where SR30 isoform 1 (the functional form) and SR30 isoform 3 were reduced to negligible levels whereas SR30 isoform 5 was increased in the mutant (Ali et al., 2007). The generation of these splice variants might involve coordinated selection of alternate 3' splice sites in the tenth intron. One mechanism by which SR45 may regulate AS of SR30 pre-mRNA is by its direct binding to intron 10 and subsequent modulation of splice site choice by interaction with U2AF, which is known to be important in 3' splice site recognition in animals. This would require the binding of SR45 and U2AF35 to the alternatively spliced tenth intron of SR30. To address this situation, we prepared labelled tenth-intron RNA and performed an electrophoretic mobility shift assay (EMSA) to test the binding of purified SR45 and U2AF35b recombinant proteins to intron 10 RNA. Both proteins bound to intron 10 and the binding increased with increasing concentration of protein (Figure 6b, P1). To determine which part of the intron binds to these two proteins, we divided the entire intron into three parts (P2, P3, and P4) (Figure 6a). P2 contains the first 308 nucleotides, P3 contains the middle part of the intron, nucleotides 309–604, with a 3' splice site (AG) at the end, and P4 consists of the remaining 338 nucleotides of SR30 intron 10 with a 3' splice site (AG) at the end (Figure 6a). Five femol of labelled RNA from each part were incubated with increasing concentrations of purified SR45 and U2AF35b proteins and the binding was analyzed by EMSA. As shown in Figure 6(b)
(P2, P3 and P4), the 5′ region of the intron (P2) binds to SR45 and RNA from P3 and P4 did not bind. In contrast, U2AF35b bound to P3 and P4 RNA but no binding was observed with P2. We then addressed the specificity of the P2 RNA binding to SR45 and U2AF35b binding to P3 and P4 RNA by adding an excess of corresponding cold RNA to the binding assay. As shown in Figure 6(c), complex formation was observed between the SR45 protein and P2 RNA (Lane 2), whereas addition of cold competitor RNA completely eliminated the binding (Lane 3). Similarly, when P3 and P4 RNAs were used with U2AF35b, RNA–protein complex formation was observed (Lane 5 and Lane 8), and addition of cold competitor RNA completely abolished the binding (Lanes 6 and 9). These results indicate that the interaction between SR45 and P2 RNA and U2AF35b with P3 and P4 RNAs is specific. Furthermore, our data clearly demonstrate that SR30 intron 10 binds to both SR45 and U2AF35b with SR45 binding to the 5′ region of the intron (P2) and U2AF35b binding to the other two regions (P3 and P4), each containing an experimentally verified 3′ splice site.

**DISCUSSION**

In animals, SR and SR-like proteins play important roles at every step of the coordinated process of splicing, from splice site recognition to the final stages of spliceosome assembly (Long and Caceres, 2009). However, little information is known about how plant SRs and SR-like proteins function in these processes. SR45 is known to bind to U1-70K and, therefore, is thought to be involved in 5′ site identification (Golovkin and Reddy, 1999). Discovery of the U2AF35 interaction with SR45 identifies a connection between the 5′ and 3′ binding site proteins. This interaction was discovered by Y2H studies and the validity of the interaction was further confirmed using protein pull-down assays and in vivo BiFC analysis. Results from these three different approaches strongly support the conclusion that SR45 interacts with U2AF35a and U2AF35b. BiFC results show most of the association in the nuclear speckles. Although speckles are thought to act as storage/assembly/modification sites and contain few if any genes, they have often been observed to be near highly active transcription sites suggesting a functional relationship with gene expression (Spector and Lamond, 2010). Studies with mammalian U2AF65 and U2AF35 found that they interact in nuclear speckles as well as in the nucleoplasm, which led to the idea that U2AF complexes are being preassembled before pre-mRNA binding and recruitment into the spliceosome (Chusainow et al., 2005) as could be the case with SR45 and U2AF35. The phenotypes of mutant SR45 and U2AF35 lines provide evidence that they are involved in similar processes. SR45 knockouts had altered development including delayed flowering in long- and short-day photoperiods and abnormal pistil and stamen numbers (Ali et al., 2007). Plants with low level U2AF35 expression due to a T-DNA insertion (U2AF35a) or RNAi (U2AF35b) were also late flowering under both long- and short-day conditions and their flower morphology was abnormal (Wang and Brendel, 2006). Furthermore, the expression of FLC (Flowering locus C) was much higher in both U2AF35b and SR45 mutants (Wang and Brendel, 2006; Ali et al., 2007).

The U2AF complex in animals is involved in 3′ splice site selection and has been characterized as a heterodimer of U2AF65 and U2AF35 (Mollet et al., 2006). We report here that Arabidopsis U2AF35a and U2AF35b form homo- and heterodimers in vivo. FRET analysis using a human U2AF35 also showed dimerization of the human U2AF35 protein (Chusainow et al., 2005). Our results reinforce the contention put forth by Chusainow et al. (2005) that in vivo, the stoichiometry of the U2AF protein may differ, in part, from the accepted heterodimer paradigm.

A comparison of U2AF35 proteins in plants and animals suggests that there is a C-terminal domain that is highly conserved in photosynthetic eukaryotes and somewhat less conserved in non-photosynthetic protozoans, but is not found in any of the animal U2AF35s. Yeast two-hybrid, protein pull-down and BiFC analyses showed that this domain is not necessary for interaction either with SR45 or for homodimerization. However, there were some differences in the pattern of localization to speckles and nucleoplasm for protein lacking this domain.
SR45 has two RS domains, N- and C-terminal to the RRM domain. BiFC analysis of the SR45 domains/U2AF35 interaction showed that both RS1 and RS2 of SR45 interact independently with both U2AF35 paralogs. There were some differences in the localization of the U2AF35/SR45 interactions with different domains of SR45 and/or different paralogs of U2AF35. The distribution between speckles and nucleoplasm was altered in some of the interaction pairs. Similar results were seen when the same domains of SR45 were used to study SR45/U1-70K interaction (Ali et al., 2008).

However, there was no interaction with any SR45RRM construct with SR45/U1-70K whereas there was fluorescence with U2AF35a, U2AF35b and U2AF35Ctri and RS2 + RRM of SR45 and RS1 + RRM with U2AF35Ctri. This situation suggests that the interactions of SR45 with U1-70K and U2AF35 are regulated somewhat differently. However, it appears in both cases that all three domains are necessary for the specificity found with full-length SR45. Using Y2H assays Tanabe et al. (2009) also examined the interaction of the different domains of another SR-like protein, SR45a, with
proteins that interacted with full-length SR45a. SR45a also has an N-terminal and a C-terminal RS domain, making it, like SR45, an SR-like protein and not an SR protein, which, by definition, have only one RS domain (Tanabe et al., 2007; Barta et al., 2010). Although the nomenclature of this protein implies a close relationship to SR45, they are not related at the sequence level as they share only 26% identity with most of this similarity occurring in the RS domains. As opposed to SR45, its RS1 domain alone did not interact with U2AF35b, which suggests that SR45 and SR45a interact with U2AF35 differently.

Our yeast two-hybrid assays and BiFC using plant U2AF35 and U1-70K proteins suggest that U2AF35a and U2AF35b interact with U1-70K in plant systems. Earlier it had been suggested that as SR proteins in animal systems interact with both U2AF35 and U1-70K, the SR proteins may act as a bridge between these 5’ and 3’ splice site factors (Wu and Maniatis, 1993). No evidence of U1-70K-U2AF35 interaction has been suggested in animals and a FRET analysis of human U1-70K and U2AF35 was negative for interaction (Ellis et al., 2008). However in plants, BiFC results suggest that Arabidopsis U2AF35b associates with Arabidopsis U1-70K and Y2H assays show interaction of U1-70K with U2AF35a. Arabidopsis has three genes that encode the larger functional activities (Kielkopf et al., 2006; Kielkopf et al., 2010). Although the nomenclature of this protein by definition, have only one RS domain (Tanabe et al., 2009) show that SR45a only interacts with U2AF35b. It has been suggested that the five human U2AF35-like proteins (three of which are isomers) and the single U2AF35 may form distinct heterodimers with different functional activities (Kielkopf et al., 2004). This situation may also be the case in plants in which different U2AF35 proteins and SR and SR-like proteins function in different activities.

Although SR and SR-like proteins in plants are thought to bind pre-mRNAs, direct binding of either SR or SR-like proteins to any endogenous pre-mRNA sequences has not been reported so far. Also, the interaction of plant U2AF35 with the 3’ splice site has not been demonstrated in plants. The data presented here provide the first evidence for direct binding of SR45 to the 5’ end of a native intronic sequence of a pre-mRNA, which suggested regulation of AS of this intron by SR45. Furthermore, direct binding of U2AF35b to different fragments, each containing a 3’ splice site, together with SR45 interaction with U2AFs, suggests that a possible regulation of splicing by SR45 recruitment of U2AF35. The previously reported misregulation of SR30 intron 10 AS in the sr45-1 mutant provides further evidence for this fact (Ali et al., 2007).

In Figure 7 we propose a model based on the protein–protein and protein–RNA interaction results presented here. One possibility is that SR45 binds to the 5’ end of SR30 intron 10 and the RS domains of SR45 may interact with splicing factors U1-70K and U2AF35 to recruit them to the 5’ and 3’ splice site, respectively. The RRM-like motifs of U2AF35 and U2AF35b are a class of protein recognition motifs called UHMs (U2AF-homology motifs) that bind RNA weakly and it is known that accessory proteins are necessary to assist proper binding (Kielkopf et al., 2004; Corsini et al., 2007). A series of experiments with U2AF35, SR proteins and enhancer sequences revealed that U2AF35 mediates interactions between U2AF35 and proteins bound to enhancers (Zuo and Maniatis, 1996; Kielkopf et al., 2004). Another possibility is that interaction of both proteins with SR45 may bridge the 5’ and 3’ splice sites. Several human SR proteins were found to interact with U1-70K and U2AF35 and act to bridge the 5’ and 3’ splice sites and this was confirmed for two of the SR proteins (SF2/ASF and SC35) using FRET studies in vivo (Wu and Maniatis, 1993; Ellis et al., 2008). We have shown that SR45 interacts with both U1-70K and U2AF35 and so may also bridge these two sites. Finally, the interaction between U1-70K and U2AF35 proteins may also be involved in bridging the 5’ and 3’ splice sites independent of SR45.
**EXPERIMENTAL PROCEDURES**

**Yeast two-hybrid screen**

SR45 fused to the Gal4 DNA-binding domain was used as a bait to screen a cDNA library of *Arabidopsis thaliana* mature plant and seedling (ARBC, CD4-10) in pACT vector. Details of Y2H screen are provided in the Data S1.

**U2AF<sup>35</sup> and SR 45 sequences acquisition, alignment and gene tree inference**

Arabidopsis U2AF<sup>35</sup>α and U2AF<sup>35</sup>β protein sequences were blasted against all Viridiplantae amino acid sequences of the Phytozome version 7 (Phytozome, http://www.phytozome.net) release, with an e-value cutoff of 2.2e-65 and 1.4e-79, respectively. The resulting hits from each search were cross-referenced and duplicate sequences were eliminated. To find metazoan representatives, human U2AF<sup>35</sup> was used as a blast query at NCBI's nr database. We also included sequences from tomato, barley, *Medicago*, cotton and pine (Wang and Brendel, 2006). Eighty-four U2AF<sup>35</sup> sequences were used in a multiple alignment constructed using *MAFFT* version 6 (Katoh and Toh, 2008), with the E-INS-i iterative refinement method with default parameters. The resulting alignment was then inspected and we removed two N-terminal insertions from Clementine and Orange as they introduce long N-terminal gaps for all other sequences.

Prior to gene tree inference, the multiple alignments were evaluated for the best fitting substitution model and parameter settings using *ProtTest* 3 (Abascal et al., 2005). We used the best scoring model (JTT + I + G + F) for gene tree inference with *RAXML* v7.7.28 (Stamatakis, 2006). The best scoring tree had a final likelihood of -14426.34 and support values from 1000 bootstrap replicate searches were mapped onto this tree.

Similar to the above sequence retrieval process, we blasted AtSR45 against all the organisms in the Phytozome databases. After manual inspection of the results, 24 SR45 protein sequences were aligned using *MAFFT* 6 with the L-INS-i iterative refinement method with default parameters (Katoch and Toh, 2008). The alignment was then fed into *ProtTest* 3 (Abascal et al., 2005) and the JTT + G + F model was chosen (ML: -9168.49) for gene tree inference using *RAxML* v7.7.28 (best scoring tree ML: -9452.05) (Stamatakis, 2006).

**U2AF<sup>35</sup> plasmid constructs and yeast two-hybrid assays**

Gene-specific reverse transcription polymerase chain reaction (RT-PCR) primers (Table S1) were designed for amplification of U2AF<sup>35</sup>α, U2AF<sup>35</sup>β and U2AF<sup>35</sup>Ctrb using *Arabidopsis thaliana* leaf RNA as template. Primers for U2AF<sup>35</sup>α were based on an annotation identifying the second ATG in the actual cDNA sequence as the start codon. Therefore, our constructs for U2AF<sup>35</sup>α are truncated N-terminally by 50 amino acids. PCR was done using *Taq*Ra Ex Taq (Fisher Scientific, www.fishersci.com) according to manufacturer’s specifications (two-step). Amplification products were agarose gel purified (Qiagen, www.qiagen.com) and subcloned into pGEM<sup>TE</sup> for sequence verification. U2AF<sup>35</sup>α/pGEM<sup>TE</sup> was digested with *Bam*HI followed by partial digestion with *Afl*I (compatible to *Nco*I), U2AF<sup>35</sup>β/pGEM<sup>TE</sup> and U2AF<sup>35</sup>Ctrb/pGEM<sup>TE</sup> were digested with *Nco*I/Bg*III* (compatible with *Bam*HI). Vectors pAS1/CYH2 and pACT2 were digested with *Nco*I/Bam*HI*. Vectors and fragments were ligated to give pAS1/CYH2 and pACT2 clones for all three fragments.

**Yeast two-hybrid assays** were performed with pAS1/CYH2 and pACT2 constructs as described previously (Ali et al., 2008). Pull-down assays were prepared from healthy leaves from 4-week-old wild-type (WT) *Arabidopsis thaliana* ecotype Columbia and transfected as described in the Data S1 section. Transfected protoplasts were examined using a Zeiss LSM 510 Meta laser scanner confocal microscope (www.z.eis.de/Confocal-Microscope). Reconstitution of YFP fluorescence was observed using the following YFP filter setup: excitation at 514 nm, 458/514 dichroic, and emission 560–615 BP filter. About 300–400 cells were observed for each set of transfections and representative single protoplasts were photographed using a ×63, N.A. 1.4 oil immersion apochromate objective.

**Preparation of 32P-labelled RNA probes and cold competitor RNAs**

Primers used for preparing the constructs are listed in Table S1. Intron 10 (942 nucleotides) of *SR30* was amplified from genomic DNA and cloned into pCR II-TOPO (Invitrogen). Three subfragments (P2, P3, and P4) were amplified from P1 and cloned into the pGEM4 vector (Promega) at the *Bam*HI and *Hind*III sites for P2, P4 and *Bam*HI and *Sal*I sites for P3. These constructs were linearized with *Nco*I for P1, *Hind*III for P2 and P4, and *Sal*I for P3 and used to generate the P1, P2, P3 and P4 RNA probes.

**RNA probes** were generated using an *in vitro* transcription system and the RNA was internally labelled with 45 μCi of [α-32P] UTP (800 Ci mmol⁻¹, Perkin-Elmer) using SP6 RNA polymerase (Fermentas, www.fermentas.com) in the presence of 500 μM ATP, 400 μM CTP, 400 μM GTP and 400 μM UTP. Full-length and truncated SR45 proteins were amplified with forward and reverse primers that contained *Sal*I and *Bam*HI sites, respectively, and cloned into pSPYNE-35S/pUC-SPYNE and pSPYCE-35S/pUCSPYCE (Ali and Reddy, 2008). Primers listed in Table S1 were used to amplify U2AF<sup>35</sup>β, a and Ctrb using the pAS1 clones as a template. PCR was performed using TaKaRa Ex Taq (Fisher Scientific) according to manufacturer’s specifications (two-step). Fragments were digested with *Sal*I/*Kpn*I and ligated into pSPYNE-35S/pUC-SPYNE and pSPYCE-35S/pUCSPYCE vectors digested with the same enzymes. Full-length *U1-70K* was constructed previously (Ali et al., 2008). Protoplasts were prepared from healthy leaves from 4-week-old wild-type (WT) *Arabidopsis thaliana* from the Columbia and transfected as described in the Data S1 section. Transfected protoplasts were examined using a Zeiss LSM 510 Meta laser scanning confocal microscope (www.z.eis.de/Confocal-Microscope). Reconstitution of YFP fluorescence was observed using the following YFP filter setup: excitation at 514 nm, 458/514 dichroic, and emission 560–615 BP filter. About 300–400 cells were observed for each set of transfections and representative single protoplasts were photographed using a ×63, N.A. 1.4 oil immersion apochromate objective.
500 μM CTP, 50 μM GTP, 50 μM UTP, and 7mGpppG. Unlabelled competitor RNAs were prepared as above with slight modifications. The 7mGpppG and radiolabelled nucleotide were excluded, and 500 μM UTP and 500 μM GTP were added. Two to five fmol of the radiolabelled RNA probe (P1, P2, P3 or P4) was incubated with increasing amounts of purified recombinant SR45 and U2AF35b proteins in the presence of 20 units of RNase inhibitor (Fermentas), 0.15 mM spermidine and gel-shift buffer (15 mM HEPES (pH 7.9), 8% glycerol, 100 mM KCl and 2 mM MgCl2) for 5 min at 30°C in a 14 μl reaction volume. Following incubation, 4 μg/μl of heparin sulfate (Sigma, www.sigmaaldrich.com) was added to the reaction mixture. Samples were chilled for 5 min on ice and 1.5 μl of 6x loading dye (30% glycerol, 0.5% bromophenol blue and 0.5% xylene cyanol) was added. RNA–protein complexes were separated from free probe on a 5% native polyacrylamide gel at 4°C in 1x TBE buffer at 200 volts for 3-6 h. Gels were dried, exposed to a phosphor screen, and visualized by Phosphor Imaging using Storm 840 (Molecular Dynamics, www.gehealthcare.com).

Expression and purification of recombinant SR45 and U2AF35b proteins

Recombinant SR45 and U2AF35b proteins were induced as described above under the ‘pull-down assay’ section. The cell pellet was resuspended in binding buffer [50 mM Tris–HCl (pH 8.0), 2 mM EDTA, 100 μg/ml lysozyme and 0.1% Triton X-100]. Complete protease inhibitor cocktail from Roche (www.roche.com) equivalent to one-tenth of culture volume and incubated at 4°C for 15 min. The samples were then sonicated, centrifuged, and the supernatant was collected. S-protein agarose beads for SR45 and T7 beads for U2AF35b were added to the supernatant and the mixture was incubated for 1 h at 4°C. After washing the beads four times with binding buffer, bound protein was eluted with 0.2 M citrate buffer (pH 2) neutralized by adding one-twentieth the volume of 2 M Tris base pH 10.4. The eluted proteins were dialyzed using the phosphate buffer (10 mM Na2HPO4, 2 mM KH2PO4, 2.7 mM KCl, 137 mM NaCl pH 7.4). S-protein and T7 antibodies were used for immunodetection of the purified SR45 and U2AF35b proteins (Golovkin and Reddy, 1999).

ACKNOWLEDGEMENTS

We thank Alyssa Cohen and Karen Markley for their help on this work. This work was supported by a grant from the National Science Foundation (DBI 0743097).

AUTHOR CONTRIBUTIONS

This study was conceived by A.S.N.R., I.D and M.G. All authors were involved in experimental design. Yeast two-hybrid analyses and in vitro interaction studies were performed by I.D. and M.G. A.L. and G.S.A performed BiFC experiments. S.G.P and JT have done RNA-binding studies with SR45 and U2AF35b. DNR performed phylogenetic analyses. I.D and A.S.N.R. wrote the manuscript with contributions from M.G, S.G.P, A.L, G.S.A, D.N.R. and JT.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. V2H interaction assays.

Figure S2. Multiple alignment of U2AF35b protein sequences from flowering plants, mosses, algae and animals.

Figure S3. Cladogram of U2AF35b proteins.

Figure S4. Multiple alignment of SR45 sequences from flowering plants and mosses.

Figure S5. Cladogram of SR45 proteins.

Table S1. Primers used for preparing clones for protein expression and for generating SR30 intron RNA probes.

Data S1. Methods.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES


Interaction of SR45 with spliceosomal components


