Cell, Vol. 89, 403–412, May 2, 1997, Copyright ©1997 by Cell Press

Cross-Intron Bridging Interactions in the Yeast Commitment Complex Are Conserved in Mammals

Nadja Abovich and Michael Rosbash
Howard Hughes Medical Institute
Department of Biology
Brandeis University
Waltham, Massachusetts 02254

Summary

The commitment complex is the first defined step in the yeast (S. cerevisiae) splicing pathway. It contains U1 snRNP as well as Mud2p, which resembles human U2AF65. In a genetic screen, we identified the yeast gene *MSL-5*, which is a novel commitment complex component. Genetic and biochemical criteria indicate a direct interaction between Msl5p and both Mud2p and the U1 snRNP protein Prp40p. This defines a bridge between the two ends of the intron. Msl5p (renamed BBP for branchpoint bridging protein) has a mammalian ortholog, the splicing factor SF1. Our results show that SF1 interacts strongly with human U2AF65, and that SF1 is a bona fide E complex component. This implies that aspects of these novel cross-intron protein–protein interactions are conserved between yeast and mammals.

Introduction

Pre-mRNA splicing can be divided into three parts. The late stage is largely preoccupied with the chemistry, the efficiency, and the specificity of the two cleavage and ligation steps. The middle stage consists of the spliceosome assembly and maturation steps, including conformational changes that generate the active spliceosome. The early stage encompasses the initial intron recognition events (Moore et al., 1993).

In vitro spliceosome assembly is initiated by formation of the U1 snRNP–pre-mRNA commitment complex in yeast (Rosbash and Séraphin, 1991) and in mammals (Michaud and Reed, 1991; Jamison and Garcia-Blanco, 1992), in which the E complex is a likely mammalian counterpart (Michaud and Reed, 1993). Both complexes involve base pairing of the 5′ end of U1 RNA to the 5′ splice site region (Zhuang and Weiner, 1986; Séraphin et al., 1988; Siliciano and Guthrie, 1988).

Indirect experiments originally suggested that the yeast U1 snRNP–containing commitment complex involves the highly conserved branchpoint sequence on the 3′ side of the intron region as well as the 5′ splice site region (Legrain et al., 1988; Ruby and Abelson, 1988). (There is no evidence in the yeast system for any substantial contribution of substrate RNA 3′ to the branchpoint region, i.e., the polypyrimidine-rich region [Rymond and Rosbash, 1985].) A direct assay for U1 snRNP binding verified this prediction by defining two complexes: a basal complex only dependent on a 5′ splice site (CC1) and a second complex of lower mobility dependent on a branchpoint as well as a 5′ splice site region (CC2) (Séraphin and Rosbash, 1989). A synthetic-lethal screen then identified factors interacting with U1 RNA (Liao et al., 1992, 1993). These included the *MUD2* gene product (Mud2p), which is a splicing factor and a component of the CC2 complex (Liao, 1992). As the association of Mud2p with commitment complexes was also dependent on a proper substrate branchpoint sequence (Abovich et al., 1994), this protein had characteristics appropriate for interacting directly with the 3′ side of the intron substrate and indirectly with the 5′ side of the intron via a U1 snRNP contact.

Contemporary work in mammalian systems reinforced the generality of this picture. Characterization of the E complex (the early mammalian U1 snRNP complex) by Reed and her colleagues identified the essential splicing factor U2AF65 as an E complex component (Michaud and Reed, 1991; Bennett et al., 1992; Michaud and Reed, 1993). U2AF65 interacts directly with the mammalian intron polypyrimidine region, which is important for early complex formation and usually lies adjacent to the less well conserved mammalian branchpoint sequence (Zamore and Green, 1989; Singh et al., 1995). As the C-terminal region of yeast Mud2p has substantial sequence similarity to the third RBD (RNA-binding domain) of mammalian U2AF65 (Abovich et al., 1994), both systems indicate similar players associated with the two ends of the intron: U1 snRNP at the 5′ side for yeast as well as humans, and yeast Mud2p or mammalian U2AF65 at the 3′ side.

In the yeast system, the cross-intron contacts between Mud2p and U1 snRNP are completely unknown (Abovich et al., 1994). Based on the assumption that a commitment complex bridging factor might collaborate with Mud2p in binding to the 3′ side of the intron, we employed a synthetic-lethal strategy to identify an essential splicing factor with the predicted properties. Biochemical and genetic criteria indicate that Msl5p (MUD synthetic-lethal 5p) or BBP (branchpoint bridging protein) does indeed interact with Mud2p as well as with the recently identified U1 snRNP protein Prp40p (Kao and Siliciano, 1996), thereby defining a bridging interaction between the two ends of the intron. The newly identified yeast splicing factor has a mammalian ortholog, the recently defined splicing factor SF1 (Arning et al., 1996). Remarkably, our results show that the biochemical interactions of SF1 with U2AF65 parallel those of BBP with Mud2p, suggesting that aspects of these novel protein–protein interactions in the yeast splicing system are also conserved in mammals.

Results

To search for commitment complex components that make a specific contribution to interactions with the 3′ end of the intron and or that bridge these interactions with U1 snRNP, we undertook a genetic synthetic-lethal approach (Frank et al., 1992; Liao et al., 1993). The strategy exploited two features of *MUD2*: that it is inessential

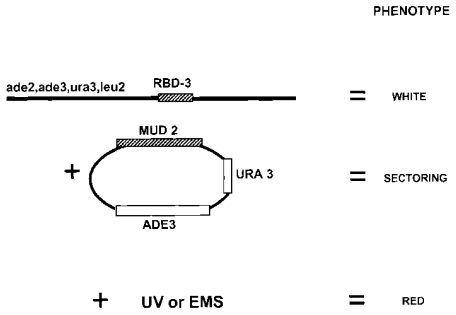

Figure 1. Sectoring Assay

The replacement of *MUD2* by *RBD-3* at its locus in strain CH1305 was done in two steps, as described in Experimental Procedures. The construction of the plasmid carrying the *MUD2* as well as the *URA3* and *ADE3* genes (pCH1122-MUD2) is also described. An *ade2,ADE3* strain forms red colonies on YEPD plates. In nonselective conditions, the plasmid is lost at a low frequency during cell division so that colonies have white sectors. This is the phenotype of the parental strain CH1305-RBD-3 (pCH1122-MUD2). After mutagenesis by UV irradiation or EMS treatment, solid red (nonsectoring) candidates, presumably dependent on a functional *MUD2* gene for survival, were picked.

for viability, and that its conserved C-terminal RBD (*RBD-3*) contributes to in vivo splicing in the absence of the more divergent N-terminal two-thirds of the protein. Overexpression of *RBD-3* rescues the synthetic lethality of a Δ*MUD1* and a Δ*MUD2* combination, indicating that it provides some biological activity without the N-terminal two-thirds of the protein (our unpublished data). Moreover, point mutants within the RNP1 motif of *RBD-3* eliminate gene activity as well as cross-linking to the pre-mRNA substrate (Abovich et al., 1994; data not shown), suggesting that this RBD binds directly to pre-mRNA. If the missing N-terminal two-thirds of the protein contributes principally to protein–protein interactions, the screen might target splicing factors that interact with Mud2p at this early step of splicing. Based on this rationale, we searched for mutants that would survive in the presence of the complete gene but would die with only *RBD-3* (Figure 1 and Experimental Procedures).

Mutagenesis with UV and EMS generated five and two candidate mutants, respectively (Table 1; see Experimental Procedures). To identify which of these mutants might affect commitment complex formation, the seven heterozygous diploids (mutant/+) were sporulated. Seven pairs of sister spores (a mutant and a wild-type segregant) from each diploid were grown under selective conditions to maintain *MUD2* in both mutant and wild-type cells. Extracts were then made from all 14 strains and assayed for effects on commitment complex formation (Figure 2). Only *msl-5* had a prominent effect on complex formation (Figure 2, lane 4); CC2, the branchpoint-sensitive U1 snRNP–pre-mRNA complex, was absent. Extract from a wild-type sister spore was indistinguishable from the wild-type control strain (Figure 2, lanes 3 and 1, respectively). As expected from the genetics, the biochemical effect of the mutant was recessive; CC2 formation was restored to the parental pattern (Figure 2, lane 2) in an extract from a rescued strain (*msl-5* plus *MSL-5* on a *LEU* plasmid; Figure 2, lane 5). In contrast, the mutant had little obvious impact on CC1, the basal U1 snRNP–pre-mRNA complex (Figure 2).

Table 1. Summary of the *MUD2* Synthetic-Lethal Mutations

UV[1]	EMS[2]	Gene
msl-1		*prp40*
msl-2		
msl-3		
msl-4	*msl-6*	*mud1(U1A)*
msl-5		
	msl-7	

Five (*msl-1–5*) and two (*msl-6 and 7*) candidate nonsectoring strains were obtained by UV and EMS mutagenesis, respectively. These seven strains also failed to grow on 5-fluororotic acid (FOA)-containing plates. They failed to sector when transformed with the *LEU2*-carrying vector pRS315 but regained sectoring and the ability to grow on FOA when pRS315 also carried the *MUD2* gene. The seven mutants were crossed to strain CH1462-MUD2KOLEU, and all diploids displayed the sectoring phenotype. They were sporulated, and tetrads were dissected. In all cases where complete tetrads were recovered, the sectoring phenotype and the FOA lethality segregated 2:2. Nonsectoring haploids of the appropriate mating types were crossed to each other and tested for complementation on FOA. The only two mutants that failed to complement were *msl-4* and *msl-6*. Nonsectoring haploids from each of the above sporulated diploids were used to clone the genes from a genomic library on a *LEU2* centromeric vector, as described previously (Liao et al., 1993). *msl-3* was impossible to clone by rescue due to a high frequency of reversion of the FOA phenotype. A single ORF has not yet been unambiguously defined for *msl-7*. *msl-4* and *msl-6* were both rescued with MUD1, the U1A protein gene (Liao et al., 1993). *msl-2* was rescued by an ORF of unknown function, and no further studies on this ORF are presented here. *msl-1* was rescued by PRP40 (Kao and Siliciano, 1996), an essential gene coding for a U1 snRNP protein. The mutation in *msl-1* was determined to be caused by a stop codon at amino acid 521. *msl-5* was rescued by ORF YL116w, and the *msl-5* mutation was determined to be Gly230-Ser.

[1] Total colonies screened = 155,000; survival = 66%.

[2] Total colonies screened = 71,500; survival = 59%.

Immunoprecipitation experiments were then carried out with an extract containing an epitope-tagged Msl5p. An *MSL-5* gene deletion established that the *MSL-5* gene is essential. Growth of a rescued strain containing Msl5p-HA was indistinguishable from the strain with wild-type Msl5p.

Incubation of the Msl5p-HA extract with anti-HA antibodies immunoprecipitated only small amounts of U1 RNA, as compared to the negative control untagged extract (Figure 3A, lanes 2 and 1, respectively). Similar amounts of U1 were recovered with the identified CC2 component Mud2p-HA (lane 3). As U1 levels were much lower than with a bona fide U1 snRNP protein, U1 70K-HA (lane 4), it suggests that Msl5p, like Mud2p, is not a tightly associated U1 snRNP protein. However, splicing complex formation (caused by the addition of cold pre-mRNA and incubation; lanes 5–12) resulted in the recovery of larger amounts of U1 RNA, both for Msl5p-HA and for Mud2p-HA extracts; as expected, RNA recovery from the U170K-HA extract was less affected by complex formation (lanes 4, 8, and 12). The results indicate that Msl5p is a splicing complex component. As the effect was similar without U2 snRNP activity (lanes 9–12), it suggests that Msl5p, like Mud2p, is a commitment complex component.

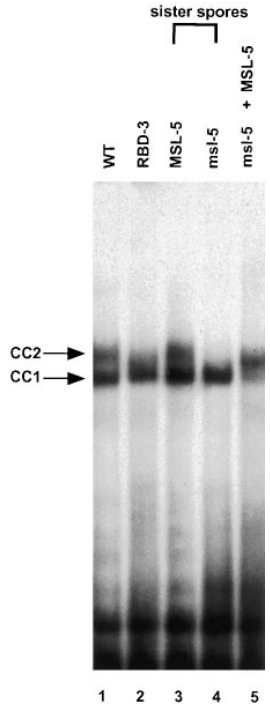

Figure 2. Native Gel Electrophoresis of Commitment Complexes

Splicing extracts (4 μl) prepared from the indicated strains were incubated with radiolabeled pre-mRNA substrate Δ2-3 under the conditions described to block prespliceosome formation (Liao et al., 1992). Lane 1 contains the parental strain for mutagenesis CH1305-RBD-3 (pCH1122-MUD2). Lane 2 contains CH1305-RBD-3. Lanes 3 and 4 contain two sister spores representing a wild-type (*MSL-5*) and a mutant (*msl-5*) product, respectively, of the original *msl-5* isolate crossed to CH1462-MUD2KO LEU. Lane 5 is the same as lane 4, except after transformation with pMSL-5 LEU and FOA selection. The strains in lanes 1 and 3 were grown in uracil-selective medium to maintain the *MUD2* gene.

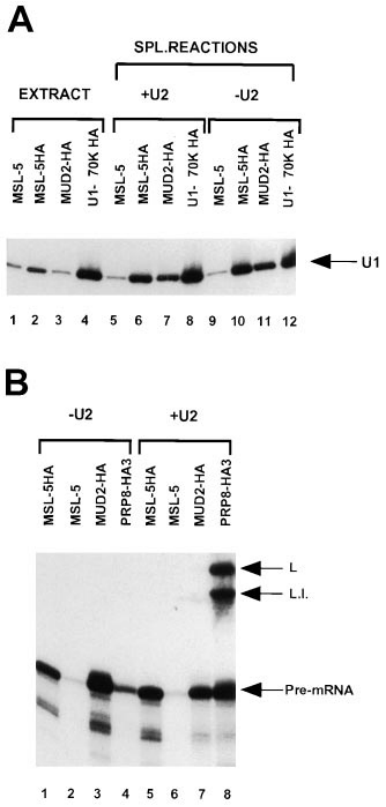

Figure 3. Immunoprecipitation of U1 snRNP and Pre-mRNA

(A) Immunoprecipitation of U1 snRNP. Splicing extracts (10 μl) or splicing reactions (containing 8 μl of extract) as indicated over each lane were immunoprecipitated with the anti-HA tag 12CA5 antibody (Boehringer) and GammaBind Plus (Pharmacia), as previously described (Abovich et al., 1994). After the beads were washed with NET-200 (50 mM Tris [pH 7.5], 200 mM NaCl, 0.1% NP-40) buffer, the RNA was extracted and used for cDNA synthesis with Reverse Transcriptase and a radiolabeled oligonucleotide complementary to U1 snRNA. Splicing reactions contained 100 ng of low specific activity Δ2–3 pre-mRNA substrate. Lanes 9–12 contained an oligonucleotide complementary to U2 snRNA to block prespliceosome assembly. Lanes 1, 5, and 9 contained a control extract without an HA-tagged protein.

(B). Immunoprecipitation of pre-mRNA. Splicing reactions (containing 4 μl of extract and 1 ng of high specific activity pre-mRNA substrate) were immunoprecipitated as above, except that the beads were washed with NET buffer containing 150 mM NaCl. The RNA was extracted and the labeled species (lariat intermediate, lariat intron, and pre-mRNA) were separated on a 15% acrylamide denaturing gel.

The immunoprecipitation results with radioactive pre-mRNA reinforce this conclusion (Figure 3B). RNA is recovered to similar extent from Msl5p-HA and from Mud2p-HA extracts, and these interactions do not require U2 snRNP activity (lanes 1,3, 5, and 7). As perhaps expected of commitment complex interactions, splicing intermediates and products were not recovered. Although we cannot exclude epitope masking, the results suggest that these protein–pre-mRNA interactions do not persist through the late stages of splicing: a Prp8p-HA extract served as a positive control for this aspect of the immunoprecipitation experiment (lane 8). Taken together, the three sets of extract results (Figures 2, 3A, and 3B) complement the genetic synthetic lethality and suggest that, like Mud2p, Msl5p is a commitment complex component.

Msl5p is a 476 aa polypeptide (ORF YLR116W) with intriguing features for a splicing factor candidate. There is a KH domain, first characterized in a subset of HnRNP proteins (HnRNPK; Siomi et al., 1993). This domain is also present in the fragile X syndrome gene products (Zhang et al., 1995), as well as in a diverse group of proteins that function in close contact with RNA (Musco et al., 1996). Importantly, this group also includes the yeast splicing factor Mer1p (Nandabalan et al., 1993) and the Drosophila splicing factor PSI (Siebel et al., 1996). Splicing regulation by both of these proteins involves an interaction with U1 snRNP or effects on early complex formation. The relevance of the Msl5p KH domain to function is suggested by the sequence of the *msl5-1* allele: there is a single nucleotide alteration that changes a glycine to a serine within a helical region of the KH domain (Figure 4, legend). There are also two zinc fingers of the retroviral Zn knuckle C2C4H4C family (Darlix et al., 1995); this motif is also present in SLU7, a late splicing factor (Frank and Guthrie, 1992).

There are two additional regions that are even more relevant for this study. Msl5p contains a proline-rich C

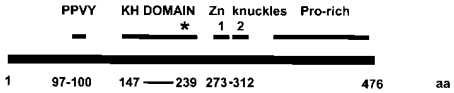

Figure 4. Structure of the *MSL-5* Protein

The *MSL-5* protein sequence contains several key features. The peptide PPVY (residues 97–100) is indicated, as well as the KH domain. The asterisk indicates the position of the *msl-5* mutation Gly230/Ser within this domain. Two Zn knuckles of the retroviral type (C2C4H4C) are indicated, as well as the proline-rich carboxy terminus, which contains several copies of the peptide PPXL.

terminus as well as the sequence PPVY at amino acids 97–100 (Figure 4). The reason for pointing out these potential domains is the identification of the U1 snRNP component Prp40p as another synthetic-lethal mutant in the initial screen (Table 1). Prp40p contains two tandem copies of a protein interaction motif (WW), which is characterized by two tryptophans and specific intertryptophan spacing (Bork and Sudol, 1994). WW domains are known to interact with proline-rich sequences (Chen and Sudol, 1995), and an interaction of this type has been analyzed at high resolution in the case of the YAP65 (yes kinase–associated protein) WW domain and its ligand core motif PPxY (Macias et al., 1996). Although the putative ligands of the Prp40p WW domains are unknown, a physical interaction between the proline-rich region of Msl5p and a WW domain of Prp40p might take place. This would then be relevant to the synthetic lethality between *prp40-1* and the starting *RBD-3* allele of *MUD2*, as well as between *msl5-1* and the *RBD-3* allele.

Indeed, two-hybrid assays with these three proteins verify that a set of physical interactions underlies some of the genetic interactions (Figure 5). Msl5p interacts with Prp40p, consistent with the suggested PPxY–WW interaction. As Prp40p contains two WW domains, we searched for additional PPxY-containing splicing factors that might also interact with Prp40p. Yeast Prp8p (a U5 snRNP protein) (Jackson et al., 1988) contains a PPxY motif (amino acids 26–29) embedded in a proline-rich N terminus, and Prp40p also interacts with a Prp8p N-terminal fusion (Experimental Procedures; Figure 5). As a U1 snRNP protein, Prp40p may therefore be a contact point between U1 snRNP and U5 snRNP (Figure 5A). The multiple interactions of Prp40p are not due to an obvious promiscuity, as the protein does not significantly interact with several other splicing factors in this assay (Figure 5B and data not shown). These include Mud2p, suggesting that the synthetic lethality between mutant alleles of *MUD2* and *PRP40* probably reflects an indirect interaction. Consistent with this notion is the prominent physical interaction between Msl5p and Mud2p (Figure 5B). This explains the synthetic lethality not only between *msl5-1* and *rbd-3* but also between *prp40-1* and *rbd-3*. As both Prp40p and Mud2p interact with Msl5p, it is an excellent candidate for a physical bridge between these two indirect interactors (Figure 5A). As Prp40p is a U1 snRNP protein linked to the 5′ end of the intron and Mud2p is linked to the 3′ end of the intron, Msl5p also serves as a bridge between the two ends of the intron (Figure 8).

To verify these interactions by an independent method,

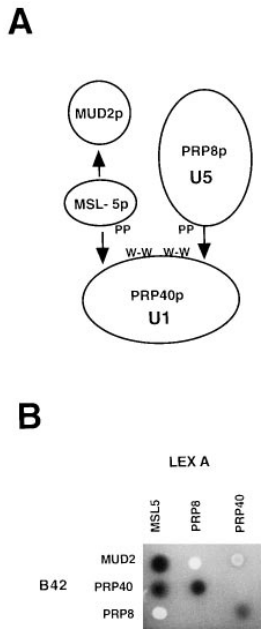

Figure 5. Protein Interactions in the Yeast Two-Hybrid System

(A) The arrows indicate positive interactions detected among the indicated proteins fused to LexA and the B42 activator, as described in Experimental Procedures and shown in (B). PRP40 and PRP8 are integral proteins of the U1 and U5 snRNPs, respectively, as indicated. The two WW domains at the N terminus of PRP40 are shown facing the proline-rich surfaces of *MSL-5* and PRP8.

(B) Yeast cells EGY48 carrying the indicated LexA and B42 fusion proteins were grown in liquid selective medium (Ura, His, Trp) containing 2% glucose. The cells were spotted on selective indicator plates containing X-Gal, 3% galactose, and 1% raffinose for induction of the B42 fusions. This plate was incubated at 30°C for 36 hrs. All spots were white on the control indicator plates containing glucose, even after prolonged incubation (not shown).

various GST fusion proteins were expressed in E. coli and purified on glutathione-agarose beads. The beads and fusion proteins were added to a number of yeast splicing extracts, incubated, and washed extensively. The extracts contained relevant HA-tagged proteins: Mud2p-HA, Prp40p-HA(3), Prp8p-HA(3), or Msl5p-HA(3). Interactions between these proteins and the GST fusion proteins were revealed by Western blotting with an anti-HA antibody (Figure 6A and 6B).

As predicted from the two-hybrid results, the GST-MSL-5-N fusion from aa 1–369 (see Figure 4) interacted well with Mud2p and Prp40p; the two direct interactions are consistent with the strength of the signals in lanes 13 and 15, as compared to those in lanes 1 and 3 (Figure 6A). A much weaker interaction takes place with Prp8p and Msl5p, evidenced by the signals in lanes 14 and 16, compared to those in lanes 2 and 4. Since even these weak signals are absent with the GST control protein in lanes 5–8, they are probably not due to background noise but rather to weak or indirect interactions. The GST-MSL-5-C fusion consisting of amino acids 353–476 and containing the proline-rich region also did not interact detectably with any of the proteins in this assay (lanes 9–12). A GST fusion protein containing both WW domains of Prp40p interacted with three of the four

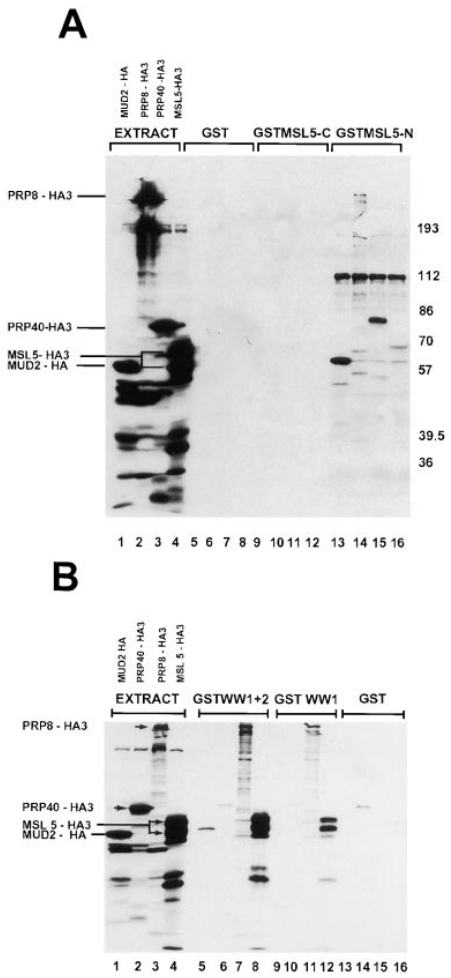

Figure 6. Protein Interactions with Recombinant GST Fusions

(A) A GST-MSL-5 fusion protein interacts with MUD2 and PRP40 from splicing extracts. Yeast splicing extracts were prepared from strains carrying the indicated HA-tagged proteins, including MUD2-HA (lanes 1, 5, 9, and 13), PRP8-HA(3) (lanes 2, 6, 10, and 14), PRP40-HA(3) (lanes 3, 7, 11, and 15), and MSL-5-HA(3) (lanes 4, 8, 12, and 16). Each extract (10 μl) was precleared as described in Experimental Procedures and incubated with glutathione-Sepharose beads containing the same amount of either GST alone (lanes 5–8), a C-terminal *MSL-5* GST fusion (aa 353–476; lanes 9–12), or an N-terminal *MSL-5* GST fusion (aa 1–369; lanes 13–16). The washed beads were resuspended in SDS sample buffer, and the proteins were separated on an 8% acrylamide gel. The proteins were transferred to a nitrocellulose filter, and the HA-tagged proteins were visualized with the 12CA5 antibody. In lanes 1–4, 10 μl of each extract was loaded directly onto the gel. The arrows indicate each tagged protein. The molecular weights (in kDa) of prestained molecular weight markers (Sigma) are indicated on the left (the identity of the cross-reacting band [100 kDa] is unknown).

(B) GST fusions of the isolated WW motifs of PRP40 interact with MSL-5 and PRP8 in yeast extracts. GST fusion proteins contained either the two WW domains of PRP40 (GST WW 1+2) or only the N-terminal one (GST WW 1). Lanes 1, 5, 9, and 13 contained MUD2-HA extract; lanes 2, 6, 10, and 14, PRP40-HA(3) extract; lanes 3, 7, 11, and 15, PRP8-HA(3) extract; and lanes 4, 8, 12, and 16, MSL-5-HA(3) extract.

tagged extract proteins: Msl5p-HA(3), Prp8p-HA(3), and Mud2p-HA (Figure 6B; lanes 5, 7, and 8). There was also a weak signal with the fourth extract protein, Prp40p-HA(3), but a similar signal was obtained with the GST nonfusion control beads (lane 14). This indicates that the putative homotypic interaction (Prp40p-HA:GST-Prp40p) reflects nonspecific background. The three positive proteins interacted more weakly with the single WW–containing fusion protein (lanes 9–12), presumably due to the absence of a second domain. The strong signal intensities are consistent with direct interactions between Prp40p and Msl5p as well as between Prp40p and Prp8p; the weaker interaction between Prp40p and Mud2p is consistent with the notion that it is indirect and bridged by the direct Prp40p–Msl5p and Msl5p–Mud2p interactions. Based on the fact that Msl5p appears to contact both the branchpoint region (directly or via Mud2p) and U1 snRNP (via Prp40p), we have renamed this protein BBP (Figure 8).

Because BBP appears to be a novel yeast splicing factor, we searched the data base for potential mammalian ortholog proteins. One protein, named ZFM1 (Toda et al., 1994) and with no known biochemical function, has substantial sequence conservation with BBP over its entire length. While this work was in progress, we learned that the mammalian splicing factor SF1 (a protein purified from HeLa nuclear extract and required for the assembly of a prespliceosome complex; Kramer, 1992) corresponds to ZFM1 (Arning et al., 1996; see this paper for an alignment of SF1 and BBP).

To ask whether SF1 might play a role in mammalian splicing similar to the one BBP plays in yeast, we searched for an in vitro interaction between SF1 and U2AF65 that would parallel the demonstrated interaction between BBP and Mud2p. A GST-SF1 fusion protein was purified from E. coli and added to mammalian proteins: a HeLa splicing nuclear extract, an S-100 fraction, or a poly(U)–Sepharose–purified fraction containing U2AF65 (Figure 7A). The results indicate that GST-SF1 interacts well with U2AF65 in nuclear extract as well as in S-100 (lanes 7 and 8). The purified U2AF65 interacts very strongly with the GST-SF1 protein (fraction of input protein bound to the beads; lane 9). There was no detectable binding to GST alone (lanes 4–6). To control for specificity, these pull-down experiments were repeated, and the filters were probed with antibodies to U2AF35 and U170K (Figure 7B). As expected, a substantial amount of U2AF35 was recovered (lane 3). The absence of a comparable U170K protein signal (lane 6) provides a control for specificity; there was also no recovery of mPRP8 (data not shown). The two-hybrid assay verified the interaction between SF1 and U2AF65 (Figure 7C).

To test whether SF1 actually participates in mammalian commitment complex (E complex) formation, purified fractions of E as well as H (H = nonspecific) complexes (generously provided by Robin Reed) were prepared. These fractions were probed sequentially with antibodies against U2AF65, SF1, and U2AF35 (Figure 7D). The results show that SF1 is indeed a specific E complex component, similar to the well characterized U2AF65 and U2AF35 proteins (Bennett et al., 1992; Michaud and Reed, 1993).

Because there is no evidence for a splicing-relevant Prp40p-like protein in mammalian systems, we sought to provide an indication of this third protein and its direct interaction with SF1. To this end, we incubated the GST-Prp40p fusion protein containing both yeast WW domains in a mammalian nuclear extract. We also used

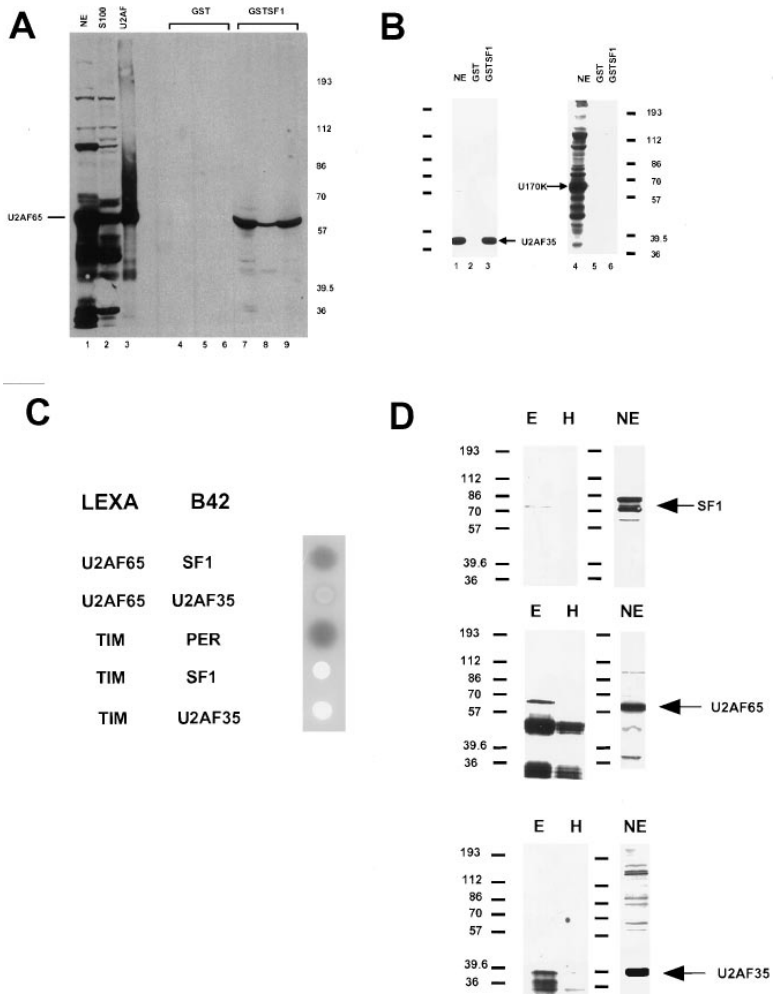

Figure 7. The Mammalian Splicing Factor SF1 Interacts with U2AF65 and Is a Component of the E Complex

(A) The GST-SF1 fusion protein interacts with U2AF65. Either 10 μl of HeLa nuclear extract or S100 (lanes 4 and 7 or 5 and 8, respectively) or 1 μl of a poly(U)–Sepharose–purified U2AF fraction (lanes 6 and 9) was incubated after preclearing as described in Experimental Procedures with glutathione-Sepharose beads carrying either GST alone or GST-SF1 fusion proteins as indicated over the lanes. After washing, the beads were resuspended in SDS sample buffer and loaded onto an 8% acrylamide gel. Lanes 1–3 contain 10 μl of nuclear extract (NE), 10 μl of S100, and 1 μl of the U2AF fraction described above, loaded directly onto the gel. The separated proteins were transferred to a nitrocellulose filter, and U2AF65 was visualized with anti Pep-A antibody.

(B) NE (10 μl) was incubated as described above with GST (lanes 2 and 5) or GST-SF1 (lanes 3 and 6). In lanes 1 and 4, 7.5 μl of NE was loaded. The filter on the left was probed with αU2AF35 antibody and the one on the right with anti U170K antibody.

(C) Yeast two-hybrid interaction. Yeast strain EGY48-pSH18-34 was cotransformed with LexAU2AF65 and the B42 fusion proteins of SF1 or U2AF35. LexA*tim* was cotransformed with B42*per* (positive control) or with the B42 fusion proteins of SF1 or U2AF35 (negative control). Cells were grown in glucose containing selective media and were spotted on the indicator plates containing X-Gal, 3% galactose, and 1% raffinose and incubated at 30°C for 40 hr.

(D) Total protein isolated from affinity-purified E and H complexes assembled on 250 ng of (AdML) pre-mRNA was fractionated on an 8% acrylamide SDS gel and transferred to nitrocellulose. The filter was probed sequentially with antibodies against SF1, U2AF65, and U2AF35, as indicated in the three panels. The lane labeled (NE) contains 2.5 μl of nuclear extract.

two additional GST fusion proteins, containing both WW domains of FBP11 and FBP21. These two mammalian proteins were isolated from cDNA expression libraries by their ability to bind a proline-rich peptide derived from formin (Chan et al., 1996). As pointed out by the authors, the spacing between the two WW domains in both proteins is identical to that between the two domains in Prp40p. SF1, U2AF65, and mPrp8p all interacted specifically with all three WW-containing fusion proteins (data not shown). The interaction with SF1 was virtually quantitative, consistent with a strong and direct interaction between this extract protein and the recombinant WW domains. This is also consistent with the isolation of SF1 in a screen for ligands of FBP11 (Bedford et al., 1997). Taken together, the results support a parallel set of bridging interactions in yeast and mammalian systems (Figure 8).

Discussion

To identify novel components that might play a role in yeast commitment complex formation, we undertook a synthetic-lethal, genetic approach focused on protein–protein contacts within the complex. The starting strain contained a *mud2* allele encoding only the C-terminal third of Mud2p (*RBD-3*). The mutant allele of *BBP* is synthetic lethal with *RBD-3*, defining a genetic interaction between the two genes. This might not reflect a physical interaction between the wild-type proteins, but a number of results show that BBP, like Mud2p, is a commitment complex component, indicating that they are close together in space during this early step of splicing: (1) There is no detectable CC2 formation in the BBP mutant extract; (2) Pre-mRNA binding to BBP does not require U2 snRNP activity; and (3) U1 snRNP binding to BBP is potentiated or stabilized by the addition of pre-mRNA. All of these criteria are similar to those previously fulfilled for *mud2* mutant and Mud2p-HA extracts (Abovich et al., 1994). In fact, there are two arguments that suggest that BBP is even more important to commitment complex formation than Mud2p: (1) *MSL-5* is an essential gene, and *MUD2* is inessential; and (2) There is CC2 formation in a Δ*MUD2* extract (due to BBP?; Abovich et al., 1994) but no detectable CC2 in the BBP mutant

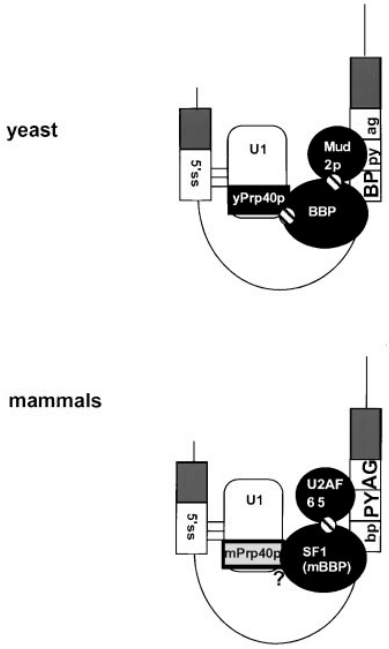

Figure 8. Models of Cross-Intron Interactions

Proteins are indicated by closed ovals and are defined in the text. The exception is mPrp40p, which is shaded to indicate its speculative nature. Indicated on the RNA substrate are the 5′ and 3′ exons (shaded rectangles), 5′ ss region (open rectangle, horizontal lines indicating base pairing to U1), the branchpoint region (bp), polypyrimidine region (py), and the 3′ ss region (ag). In the yeast diagram (top), the (bp) region is in bold, to indicate its relative importance for yeast. In the mammals diagram, the (py) and (ag) regions are in bold, to indicate their relative importance to mammalian splicing. The small cross-hatched circles indicate protein–protein contacts addressed in this work. As described in the text, protein contacts to the 3′ splice site regions are speculative, except for U2AF65 (Valcárcel et al., 1993).

extract (Figure 2). Because BBP interacts well with Mud2p in the two-hybrid assay and in the in vitro interaction assay, we assume that both proteins not only contribute to the same CC2 splicing complex but also physically interact in the process.

In contrast to the BBP mutation, the isolation of a *prp40* allele in the screen probably does not reflect a direct physical interaction between Prp40p and Mud2p because there is no significant two-hybrid interaction between them. As both proteins interact well with BBP, the genetic interaction likely reflects a physical interaction between Prp40p and Mud2p, but an indirect one that is bridged by BBP. This interpretation is also consistent with the in vitro associations between the GST-Prp40p fusion proteins and the yeast extract proteins: the fusion proteins interact with both BBP-HA and Mud2p-HA; the former is presumably direct and the latter indirect. This leads to a coherent picture that includes interactions with both ends of the intron: the 3′ side is contacted by Mud2p and the 5′ side by base pairing to the 5′ end of U1 RNA.

A multicontact bridging model implies that BBP is simultaneously linked to Prp40p and to Mud2p. The GST-Prp40p yeast extract results also suggest that these two interactions of BBP need not be mutually exclusive: BBP must bridge the in vitro interaction between Prp40p and Mud2p. Although genetic as well as biochemical data support this coherent picture, proving that these contacts are direct and relevant to splicing will require purifying all interacting proteins from E. coli as well as performing protein–protein cross-linking experiments in splicing extracts.

Although the interacting regions of BBP and Mud2p are completely unknown, a proline-rich region of BBP is probably involved in binding to a WW domain of Prp40p. Similarly, a proline-rich region of Prp8p is a good candidate for interacting with Prp40p. The effects of point mutants in these regions, in vitro and in yeast, will be required to show that they interact in a functionally significant manner. Although we do not know whether the two WW domains of PRP40 have intrinsic binding preferences, a single domain is able to associate in vitro with both BBP and Prp8p (Figure 6B). A significant restriction in binding partner specificity may therefore require the intact proteins if not the biological context of genuine splicing complexes. The association of Prp40p and Prp8p defines a potential U1snRNP–U5snRNP interaction, for which there is precedent in both yeast and mammalian systems (Newman and Norman, 1992; Wyatt et al., 1992; Ast and Weiner, 1996). This may serve to recruit U5 snRNP to the early spliceosome, or the U5 snRNP contact may aid in destabilizing U1 snRNP during the transition to more mature spliceosome forms. One can imagine that the U1–U5 interaction occurs at the same time as the cross–intron bridging interactions, or these interactions may be mutually exclusive and therefore sequential.

The results also suggest that some of these yeast protein–protein interactions are conserved in mammalian systems. The mammalian protein SF1 functions at an early stage of in vitro splicing (Kramer, 1992; Arning et al., 1996), and we have shown that it is an E complex component (Figure 7D). Our biochemical results also show that a substantial fraction of U2AF65 in a mammalian extract can associate with SF1 (Figure 7A). We therefore suggest that SF1 is the mammalian ortholog of yeast BBP (mBBP; Figure 8). This implies that BBP and mBBP participate in parallel interactions and have comparable functions. It also implies that mBBP and U2AF65 might interact with each other at the same time that they interact with the mammalian pre-mRNA 3′ splice site region, similar to the proposed three-way interaction between BBP, Mud2p, and the yeast 3′ splice site region (Figure 8).

The contributions of these multiple protein–protein and protein–RNA interactions might be germane to the different requirements for the branchpoint and polypyrimidine track regions in the two systems. In yeast, the branchpoint sequence is highly conserved and the polypyrimidine track of minor importance, whereas the situation is reversed in mammals (Moore et al., 1993; see Figure 8). A parallel situation may be emerging at the protein level: Mud2p is inessential and has only one bona fide RBD (Abovich et al., 1994), whereas U2AF65 is very important, has three RBDs, and interacts strongly

with a polypyrimidine track (Zamore et al., 1992). BBP, in contrast to Mud2p, is essential. If BBP undergoes a direct or indirect interaction with the more highly conserved yeast branchpoint sequence, it may contribute to the branchpoint-dependent activities that persist in the absence of Mud2p (Abovich et al., 1994; these data). SF1 may rely more on its interaction with U2AF65, which would then allow pre-mRNA binding without strong branchpoint sequence conservation. Consistent with this picture, BBP and SF1 contain RNA-binding motifs, and both purified proteins interact with pre-mRNA (Arning et al., 1996; A. Berglund, unpublished data). Based on these considerations, we have drawn BBP and SF1 as interacting directly with pre-mRNA in the vicinity of the branchpoint region (Figure 8).

The interactions of recombinant yPrp40p and FBP fusion proteins with the three mammalian proteins SF1, U2AF65, and Prp8p (data not shown) suggest that one or more mammalian Prp40-like splicing factors also exist. These may include FBP11 or FBP21 (Chan et al., 1996; Bedford et al., 1997). More generally, there are only three mammalian U1 snRNP–specific proteins (U1 A, U1 C, and U1 70K) and many more yeast U1 snRNP–specific proteins (Fabrizio et al., 1994). Their mammalian counterparts may be simply more loosely associated with U1 snRNP.

There is a previously proposed mammalian bridging model with a very different set of cross-intron contacts (Wu and Maniatis, 1993): the polypyrimidine stretch is linked to U2AF65; U2AF65 contacts U2AF35; the SR domain of U2AF35 contacts an SR protein (i.e., SC35); this protein forms a simultaneous contact with the SR tail of the U1 snRNP 70K protein; and U1 snRNP is base paired to the 5′ splice site. These contacts are not incompatible with those described here, indicating that both could contribute simultaneously to bridge the two ends of a mammalian intron. There are, however, two less attractive features of this previous cross-intron model. First, it is unlikely to be universal and relevant to yeast splicing. This is because there are no bona fide SR proteins in yeast, and there is no obvious yeast ortholog of U2AF35; even the SR tail of mammalian U170K protein is missing or much shorter in yeast U1 70K (Smith and Barrell, 1991). Second, the U2AF65–U2AF35–SR-U170K interactions were also proposed to function as cross-exon contacts in constitutive and regulated mammalian splicing (Reed, 1996; Zuo and Maniatis, 1996). The use of the SR domain of U1 70K for both sets of contacts may render the cross-intron and cross-exon functions incompatible; one might preclude the other. In contrast, our novel cross-intron bridging model avoids both difficulties: the proposed bridging proteins are ubiquitous, and the mammalian cross-intron contacts can exist simultaneously with the cross-exon contacts (Zuo and Maniatis, 1996). As exonic enhancers and more general features of exon definition are unknown in S. cerevisiae, the metazoan-specific proteins would be dedicated to these functions. Intron definition would require a universal set of proteins (Figure 8). Alternatively, the previous model may apply to the large mammalian introns and this new set of interactions to smaller, yeast-like introns (Reed, 1996).

If there is a mammalian Prp40p–BBP–U2AF65 bridge, then yeast genetics should contribute to defining which regions of these proteins make physical contact. This approach should also aid in identifying the subset of contacts important for the splicing process itself, as well as in identifying additional components that contribute to these bridging interactions. Finally, the initial branchpoint recognition events remain undefined in both yeast and mammalian systems. Biochemistry as well as genetics will undoubtedly be required to elucidate the precise nature of this missing contact.

Experimental Procedures

Isolation of Mutants and Yeast Strains

The original strains for the colony-sectoring assay, CH1305 (*Mata, ade2, ade3, leu2, ura3, lys2,* and *Can1*) and CH1462 (*Matα, ade2, ade3, leu2, ura3, his3*), were obtained from C. Holmes (Kranz and Holm, 1990). The replacement of *MUD2* by *MUD2*-RBD3 was performed according to standard procedures. The *MUD2* gene in the strain of the opposite mating type, CH1462, was deleted and replaced with the *LEU2* gene as previously described (Abovich et al., 1994). Nonsectoring strains were mated to CH1462-M2KO *LEU2*. After sporulation and tetrad dissection, haploids of the appropriate genotype were used for complementation tests, for biochemical analyses, and for cloning the genes by rescue of the synthetic-lethal phenotype (see legend to Table 1).

MSL-5KO *LEU*

A diploid strain obtained by crossing strains MGD 353-13D (*Mata, trp1-289, ura3-52, arg4, leu2, ade2*) and MGD 353-46D (*Matα, trp1-289, ura3-52, leu2, his3, CyhR*) was transformed with a linear fragment into which the *MSL-5* coding sequences were replaced by the *LEU2* gene (Rothstein, 1983). After sporulation and tetrad dissections, only 2 LEU^- spores per tetrad were viable, indicating that the *MSL-5* gene is essential. This diploid was transformed with *URA3* plasmids (see below) carrying either the wild-type *MSL-5* gene or HA-tagged versions [HA or HA(3)]. Tetrads with four viable spores were obtained in all cases. The LEU^+ spores were also FOA^-, confirming that the gene is essential. The growth rates of spores carrying tagged versions were indistinguishable from those carrying the wild-type gene. The haploid strains MSL-5, MSL-5-HA, and MSL-5-HA(3) were thus obtained. The PRP40-HA(3) strain was obtained by performing the same manipulations. The strain YJU77 carrying the PRP8-HA(3) protein was a gift of J. Umen and C. Guthrie (Umen and Guthrie, 1995).

The strain EGY48 transformed with the reporter pSH 18-34 for analysis of protein interactions in the yeast 2-hybrid system as well as the vectors pEG202 and pJG4-5 were a gift of Roger Brent (Harvard Medical School, Boston, MA).

Plasmids

The screening plasmid pCH1122-MUD2 was constructed in a four-way ligation. The vector pCH1122 (a gift from C. Holmes) was cut with NruI and SmaI. The three other fragments were: an EcoRI-filled BamHI fragment carrying the *URA3* gene and the Gal *UAS* sequences from pLGSD5 (Guarente et al., 1982), a BamHI–SalI fragment carrying the *MUD2* HA sequences, and a SalI SmaI fragment of *MUD2* downstream sequences.

The testing plasmid *MUD2 LEU2* contains a 3 kb insert carrying the wild-type *MUD2* gene in pRS315. Plasmid *MSL-5* LEU contains a 2.3 kb SpeI–EcoRV fragment in pRS315, carrying ORF YL116W, which rescues the synthetic lethality of the *msl-5* mutant strain. For MSL5-HA, one copy of the 9 aa HA tag was introduced into *MSL-5* after the last amino acid by standard procedures, and the coding sequence was replaced into Gal MUD2-HA (Abovich et al., 1994). Gal PRP40 was obtained by replacing the wild-type PRP40 sequence into the same vector.

MSL-5-HA(3) and PRP40-HA(3) were generated by placing the NotI cassette containing 3 copies of the HA epitope [from PRP8-HA(3), as described (Umen and Guthrie, 1995)] after the last amino acid of MSL-5 or PRP40, respectively, in the above vectors.

The LexA and B42 fusion proteins for the analysis of protein–protein interactions in yeast were constructed in the pEG-202 and pJG 4–5 vectors, respectively. For LexA constructs, BamHI and SalI sites were introduced by PCR at the N and C termini, respectively, except for PRP8, where an EcoRI site was placed at the N terminus of the fusion. For the B42 constructs, EcoRI and SalI sites were employed and cloned into the EcoRI and XhoI sites in the vector. LexAU2AF65, LexAMSL-5, and LexAPRP40 contained the complete proteins; LexAPRP8 contained amino acids 1–349. B42-U2AF35 was a gift from T. Maniatis (Wu and Maniatis, 1993). B42-MUD2 and B42-PRP40 contained the complete proteins; B42-PRP8 was as described above; and B42-SF1 contained amino acids 1–361. The LexA*tim* and B42*per* were previously described (Rutila et al., 1996).

For GST-SF1, a sequence designated ZFM1 (accession number D26120) (Toda et al., 1994) was identified in our search of the database for protein homologies with *MSL-5*. These two sequences (*MSL-5* and *ZFM1*) are 37% identical and 54% homologous. We later learned that ZFM1 was the human splicing factor SF1 (Arning et al., 1996). An oligonucleotide 5′ AAAAGGAAGCCTTCCAGGAAG GAGAGGA 3′ was designed to prime cDNA synthesis from HeLa cell poly(A)$^+$ RNA with Superscript reverse transcriptase (GIBCO BRL) as recommended by the manufacturer. This cDNA was used as the template for PCR cloning amino acids 1–361 of SF1, where the homology to *MSL-5* is greatest. The oligonucleotides contained synthetic SalI sites for the expression of SF1 amino acids 1–361 as a GST fusion in pGEX4T-1. The fusion protein was purified on glutathione-Sepharose beads (Pharmacia) and eluted with glutathione. This protein was used as an immunogen in mice to generate anti-SFI antibodies.

A GST-MSL-5-N fusion protein including amino acids 1–369 and comparable to the GST-SF1 fusion described above was made in GEX2TK (Pharmacia). The same strategy was employed to generate the GST-MSL-5(C) fusion protein, comprising amino acids 353–476. The PRP40 GST WW 1 as well as the WW 1 and 2 fusion proteins encompassed amino acids 1–41 and 1–76 of PRP40 and contained either one or both copies of the WW domains, respectively.

The FBP11 and FBP21 GST-2TK fusions contained both copies of the mouse WW domains (Chan et al., 1996) and were a gift from Mark Bedford and Phil Leder.

GST Fusion Precipitation

Expression of GST fusion proteins in E. coli strain BL21(DE3) (Novagen) and preparation of cell lysates were performed as described (Smith and Johnson, 1988). Lysates were incubated with glutathione-Sepharose at room temperature for 1 hr in PBS and the beads washed in ELB as described (Belanger et al., 1994). Yeast splicing extracts were prepared as described by J. Umen (Umen and Guthrie, 1995). HeLa nuclear extract (NE) and the S100 fraction were a gift from Claudio Pikielny. The poly(U)–Sepharose–purified U2AF fraction was a gift from P. Zamore and M. Green (Zamore and Green, 1989). Either yeast extract NE or S100 (10 μl) or purified U2AF fraction (1 μl) in 100 μg of bovine serum albumin (BSA) were precleared for 1 hr at 4°C by incubation with GST in 400 μl of ELB containing 1% milk. Precleared extracts were incubated with GST or GST fusion Sepharose for 1 hr at 4°C, after which the beads were washed five times with ELB. The beads were placed in a new tube prior to the final wash. Bound proteins were eluted in SDS sample buffer and subjected to electrophoresis. The separated proteins were electrophoretically transferred to nitrocellulose and detected, after incubation with the appropriate antibodies, using the ECL detection system (Amersham Corp., Arlington Heights, IL).

Mammalian E and H Complexes

Mammalian E (early commitment) and H (nonspecific hnRNP) complexes were the generous gift of Robin Reed. They were assembled on biotynilated AdML pre-mRNA, isolated by gel filtration, and affinity purified as previously described (Bennett et al., 1992; Michaud and Reed, 1993).

Antibodies

12CA5 anti-HA antibodies were purchased from Boehringer Mannheim. The SF1 antibodies were prepared as described above. The anti-U2AF65 antibody anti Pep-A (Zamore and Green, 1991) was a gift of P. Zamore and M. Green, the anti-U170K and U2AF35 were a gift from Robin Reed, and the mPRP8 antibody was a gift from Melissa Moore. In vitro splicing reactions, immunoprecipitation with the 12CA5 antibody, and native gel electrophoresis were previously described (Abovich et al., 1994).

Acknowledgments

Correspondence regarding this paper should be addressed to M. R. We gratefully acknowledge Robin Reed's generous gift of purified mammalian E and H complexes, antibodies to U2AF35 and U170K, as well as important suggestions on the manuscript. We thank P. Zamore and M. Green for their gifts of purified HeLa U2AF65 and anti-PepA antibody, M. Bedford and P. Leder for providing the plasmids GST-FBP11 and GST-FBP21, T. Maniatis for the pJG-U2AF35 DNA, J. Umen and C. Guthrie for the PRP8-HA(3) strain YJU77, and C. Holmes for the strains and the plasmid used in the sectoring assay. We thank M. Moore for mammalian anti-PRP8 antibody and critical reading of the manuscript, and D. Black for a helpful suggestion.

We are grateful to members of the Rosbash laboratory and especially to H. Colot for helpful suggestions and critical comments on the manuscript. N. A. thanks L. Liu for cloning the GST-SF1 fusion protein and raising anti-SF1 antibodies, and L. A. Monaghan for excellent secretarial assistance. This work was supported by a grant from the NIH (GM-23549) to M. R.

Received December 6, 1996; revised March 21, 1997.

References

Abovich, N., Liao, X.C., and Rosbash, M. (1994). The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. Genes Dev. *8*, 843–854.

Arning, S., Gruter, P., Bilbe, G., and Kramer, A. (1996). Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. RNA *2*, 794–810.

Ast, G., and Weiner, A.M. (1996). A U1/U4/U5 snRNP complex induced by a 2′*-0*-methyl-oligoribonucleotide complementary to U5 snRNA. Science *272*, 881–884.

Bedford, M.T., Chan, D.C., and Leder, P. (1997). FBP WW domains and the Ab1 SH3 domain bind to a specific class of proline-rich ligands. EMBO J., in press.

Belanger, K.D., Kenna, M.A., Wei, S., and Davis, L.I. (1994). Genetic and physical interactions between Srp1p and nuclear pore complex proteins Nup1p and Nup2p. J. Cell Biol. *126*, 619–630.

Bennett, M., Michaud, S., Kingston, J., and Reed, R. (1992). Protein components specifically associated with prespliceosome and spliceosome complexes. Genes Dev. *6*, 1986–2000.

Bork, P., and Sudol, M. (1994). The WW domain: a signalling site in dystrophin? Trends Biochem. Sci. *19*, 531–533.

Chan, D.C., Bedford, M.T., and Leder, P. (1996). Formin binding proteins bear WWP/WW domains that bind proline-rich peptides and functionally resemble SH3 domains. EMBO J. *15*, 1045–1054.

Chen, H.I., and Sudol, M. (1995). The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3–binding modules. Proc. Natl. Acad. Sci. USA *92*, 7819–7823.

Darlix, J.-L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B.P. (1995). First glimpses at structure–function relationships of the nucleocapsid protein of retroviruses. J. Mol. Biol. *254*, 523–537.

Fabrizio, P., Esser, S., Kastner, B., and Luhrmann, R. (1994). Isolation of S. cerevisiae snRNPs: comparison of U1 and U4/U6/U5 to their human counterparts. Science *2645*, 261–265.

Frank, D., Patterson, B., and Guthrie, C. (1992). Synthetic lethal mutations suggest interactions between U5 small nuclear RNA and four proteins required for the second step of splicing. Mol. Cell. Biol. *12*, 5197–5205.

Frank, D., and Guthrie, C. (1992). An essential splicing factor, *SLU7*, mediates 3′ splice-site choice in yeast. Genes Dev. *6*, 2112–2124.

Guarente, L., Yocum, R.R., and Gifford, P. (1982). A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. Proc. Natl. Acad. Sci. USA *79*, 7410–7414.

Jackson, S.P., Lossky, M., and Beggs, J.D. (1988). Cloning of the *RNA8* gene of *Saccharomyces cerevisiae*, detection of *RNA8* protein, and demonstration that it is essential for nuclear pre-mRNA splicing. Mol. Cell. Biol. *8*, 1067–1075.

Jamison, S.F., and Garcia-Blanco, M.A. (1992). An ATP independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. Proc. Natl. Acad. Sci. USA *89*, 5482–5486.

Kao, H.-Y., and Siliciano, P.G. (1996). Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. Mol. Cell. Biol. *16*, 960–967.

Kramer, A. (1992). Purification of splicing factor SF1, a heat-stable protein that functions in the assembly of a presplicing complex. Mol. Cell. Biol. *12*, 4545–4552.

Kranz, J.E., and Holm, C. (1990). Cloning by function: an alternative approach for identifying yeast homologs of genes from other organisms. Proc. Natl. Acad. Sci. USA *87*, 6629–6633.

Legrain, P., Séraphin, B., and Rosbash, M. (1988). Early commitment of yeast pre-mRNA to the spliceosome pathway. Mol. Cell. Biol. *8*, 3755–3760.

Liao, X.C. (1992). Pre-mRNA splicing in yeast: roles of U1 snRNA and its associated protein. PhD Thesis, Brandeis University, Waltham, Massachusetts.

Liao, X.C., Colot, H.V., Wang, Y., and Rosbash, M. (1992). Requirements for U2 snRNP addition to yeast pre-mRNA. Nucl. Acids Res. *20*, 4237–4245.

Liao, X.C., Tang, J., and Rosbash, M. (1993). An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. Genes Dev. *7*, 419–428.

Macias, M.J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. (1996). Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. Nature *382*, 646–649.

Michaud, S., and Reed, R. (1991). An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. Genes Dev. *5*, 2534–2546.

Michaud, S., and Reed, R. (1993). A functional association between the 5′ and 3′ splice sites is established in the earliest prespliceosome complex (E) in mammals. Genes Dev. *7*, 1008–1020.

Moore, M.J., Query, C.C., and Sharp, P.A. (1993). Splicing of precursors to mRNAs by the spliceosome. In The RNA World, R.F. Gesteland and J.F. Atkins, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 303–357.

Musco, G., Stier, G., Joseph, C., Antonietta, M., Morelli, C., Nilges, M., Gibson, T.J., and Pastore, A. (1996). Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. Cell *85*, 237–245.

Nandabalan, K., Price, L., and Roeder, G.S. (1993). Mutations in U1 snRNA bypass the requirement for a cell type–specific RNA splicing factor. Cell *73*, 407–415.

Newman, A.J., and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. Cell *68*, 743–754.

Reed, R. (1996). Initial splice-site recognition and pairing during pre-mRNA splicing. Curr. Opin. Genet. Dev. *6*, 215–220.

Rosbash, M., and Séraphin, B. (1991). Who's on first? The U1 snRNP-5′ splice site interaction and splicing. Trends Biol. Sci. *16*, 187–190.

Rothstein, R.J. (1983). One step gene disruption. Meth. Enzymol. *101*, 202–211.

Ruby, S.W., and Abelson, J. (1988). An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. Science *242*, 1028–1035.

Rutila, J.E., Zeng, H., Le, M., Curtin, K.D., Hall, J.C., and Rosbash, M. (1996). The tim^{SL} mutant of the *Drosophila* rhythm gene *timeless* manifests allele-specific interactions with *period* gene mutants. Neuron *17*, 407–415.

Rymond, B.C., and Rosbash, M. (1985). Cleavage of 5′ splice site and lariat formation are independent of 3′ splice site in yeast mRNA splicing. Nature *317*, 735–737.

Séraphin, B., Kretzner, L., and Rosbash, M. (1988). A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5′ cleavage site. EMBO J. *7*, 2533–2538.

Séraphin, B., and Rosbash, M. (1989). Identification of functional U1 snRNA–pre-mRNA complexes committed to spliceosome assembly and splicing. Cell *59*, 349–358.

Siebel, C.W., Admon, A., and Rio, D.C. (1996). Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. Genes Dev. *9*, 269–283.

Siliciano, P.G., and Guthrie, C. (1988). 5′ splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev. *2*, 1258–1267.

Singh, R., Valcarcel, J., and Green, M.R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine-tract binding proteins. Science *268*, 1173–1176.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucl. Acids Res. *21*, 1193–1198.

Smith, D.B., and Johnson, K.S. (1988). Single step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene *67*, 31–40.

Smith, V., and Barrell, B.G. (1991). Cloning of a yeast U1 snRNP 70K protein homologue: functional conservation of an RNA-binding domain between humans and yeast. EMBO J. *10*, 2627–2634.

Toda, T., Lida, A., Miwa, T., Nakamura, Y., and Imai, T. (1994). Isolation and characterization of a novel gene encoding nuclear protein at a locus (D11S636) tightly linked to multiple endocrine neoplasia type 1 (MEN1). Hum. Mol. Genet. *3*, 465–470.

Umen, J.G., and Guthrie, C. (1995). A novel role for a U5 snRNP protein in 3′ splice site selection. Genes Dev. *9*, 855–868.

Valcárcel, J., Singh, R., Zamore, P.D., and Green, M.R. (1993). The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. Nature *362*, 171–175.

Wu, J., and Maniatis, T. (1993). Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. Cell *75*, 1061–1070.

Wyatt, J.R., Sontheimer, E.J., and Steitz, J.A. (1992). Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. Genes Dev. *6*, 2542–2553.

Zamore, P.D., Patton, J.G., and Green, M.R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. Nature *355*, 609–614.

Zamore, P.D., and Green, M.R. (1989). Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc. Natl. Acad. Sci. USA *86*, 9243–9247.

Zamore, P.D., and Green, M.R. (1991). Biochemical characterization of U2 snRNP auxiliary factor: an essential pre-mRNA splicing factor with a novel intranuclear distribution. EMBO J. *10*, 207–214.

Zhang, Y., O'Connor, J.P., Siomi, M.C., Srinivasan, S., Dutra, A., Nussbaum, R., and Dreyfuss, G. (1995). The fragile X mental retardation syndrome protein interacts with novel homologs FXR1 and FXR2. EMBO J. *14*, 5358–5366.

Zhuang, Y., and Weiner, A.M. (1986). A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell *46*, 827–835.

Zuo, P., and Maniatis, T. (1996). The splicing factor $U2AF^{35}$ mediates critical protein–protein interactions in constitutive and enhancer-dependent splicing. Genes Dev. *10*, 1356–1368.