3260–3267 Nucleic Acids Research, 1995, Vol. 23, No. 16

U1 snRNP–ASF/SF2 interaction and 5′ splice site recognition: characterization of required elements

Sharon F. Jamison[1], Zvi Pasman[1,2], Jin Wang[4], Cindy Will[5], Reinhard Lührmann[5], James L. Manley[4] and Mariano A. Garcia-Blanco[1,2,3,*]

Departments of [1]Molecular Cancer Biology, [2]Microbiology and [3]Medicine, Levine Science Research Center, Duke University Medical Center, Durham, NC 27710, USA, [4]Department of Biological Sciences, Columbia University, New York, NY 10027, USA and [5]Institut fur Molekularbiologie und Tumorforschung, Emil-Mannkopff-Strasse 2, D-35037 Marburg, Germany

Received April 21, 1995; Revised and Accepted July 4, 1995

ABSTRACT

Members of the SR family of proteins, can collaborate with U1 snRNP in the recognition of 5′ splice sites in pre-messenger RNAs. We have previously shown that purified U1 snRNP and ASF/SF2 form a ternary complex with pre-mRNA, which is dependent on a functional 5′ splice site. In this manuscript we dissect the requirements for the formation of this complex. Sequences in the pre-mRNA, domains in ASF/SF2 and components of the U1 snRNP particle are shown to be required for complex formation. We had shown that sequences at the 5′ splice site of PIP7.A are necessary and now we show these are sufficient for complex formation. Furthermore, we show that one functional RNA binding domain and the RS domain are both required for ASF/SF2 to participate in complex formation. The RNA binding domains were redundant in this assay, suggesting that either domain can interact with the pre-messenger RNA. Finally, our experiments show no function for the U1-specific A protein in complex formation, whereas a function for U1-specific C protein was strongly suggested. The study of the earliest interactions between pre-mRNA and splicing factors suggests a model for 5′ splice site recognition.

INTRODUCTION

Genes in vertebrates are complex units, usually composed of multiple and sometimes alternatively utilized exons and introns (1). Thus, primary transcripts or pre-messenger RNAs (pre-mRNAs), contain many 5′ splice site and 3′ splice site pairs, which must be properly recognized and processed. Each splice site pair encompasses an intron that will be removed by spliceosomes, macromolecular enzymes that undergo step-wise assembly with pre-mRNAs (2–6). Early recognition of splice sites is conserved from yeast to vertebrates: formation of the commitment complex, the first stable complex formed between pre-mRNA and splicing factors, usually requires both 5′ splice site and 3′ splice site sequences (6,7). In many cases pre-mRNAs with mutations in either the 5′ or the 3′ splice site show defects in formation of the commitment complex or the pre-spliceosome (6,8,9). Several required splicing factors are known to interact with pre-mRNA in early complexes: members of the SR protein family (10), U2AF (11) and U1 and U2 snRNPs (12,13). These factors provide a platform for assembly of the full spliceosome. In cases of alternative splicing, regulatory factors are found interacting with the constitutive splicing machinery in these early complexes (14,15). Alternative splicing decisions are established early in the case of the *Drosophila* developmental regulators Sxl (16), tra and tra-2 (17) and in the case of adenovirus early to late switch (18).

Protein factors of the SR family have been shown to be required for early steps of spliceosome assembly. These proteins share N-terminal RNA recognition motifs (RRM) and homologous C-terminal domains, rich in serines and arginines, RS domains (19). SR proteins of 20, 30a/b, 40, 55 and 75 kDa can be purified to apparent homogeneity by two sequential salt precipitations. SR_p30a, ASF/SF2, was shown to be required for constitutive splicing and to be involved in alternative selection of proximal 5′ splice sites (20–22). A second member of the SR family, SR_p30b or SC35, was identified first as a spliceosome component and later was shown to be a required splicing factor (23). S100 extracts are deficient in SR proteins and because of this cannot splice pre-mRNAs. Addition of a mixture of purified SR proteins or addition of any one of six SR proteins SRp20, SRp30a, SRp30b, SRp40, SRp55 or SRp75 can reconstitute pre-mRNA splicing in S100 extracts, suggesting that SR proteins have overlapping activities (24,25). There is, however, evidence that supports the idea of preferred functions for each SR protein (25–27).

Several lines of evidence argue that the SR proteins may be the first components to bind the pre-mRNA. First, pre-mRNAs that are pre-incubated with SC35, ASF/SF2 or SRp55 can be spliced in the presence of competitor RNA (26). Secondly, ASF/SF2, U1 snRNP and pre-mRNA form stable and specific complexes, but only when ASF/SF2 is added first to the pre-mRNA (28). Thirdly, SR proteins potentiate the formation of E complex, which may be

*To whom correspondence should be addressed at: Department of Molecular Cancer Biology, Box 3686, DUMC, Durham, NC 27710, USA

equivalent to the commitment complex (15). Moreover, SR proteins collaborate with U1 snRNP in 5′ splice site recognition: purified ASF/SF2 and U1 snRNP form complexes on PIP7.A pre-mRNA, but not on a 5′ splice site mutant pre-mRNA (28). This protein has been postulated to enhance the affinity of U1 snRNP for 5′ splice sites (29). This effect is likely to be mediated directly by an interaction between ASF/SF2 and 70 kDa protein of U1 snRNP, which requires the arginine-serine rich domains in both ASF/SF2 and U1–70 kDa protein (28,30). In keeping with its very early role in 5′ splice site recognition, ASF/SF2 has alternative splicing activity both *in vitro* (20,22) and *in vivo* (31).

U1 snRNP is the earliest snRNP to assemble on pre-mRNAs (5,32) defining the 5′ splice site via RNA–RNA and RNA–protein interactions (33,34). The U1 snRNP is also the only spliceosomal snRNP required for commitment complex formation (6,7). The subsequent arrival of U2 snRNP and the interactions between U2 and U1 snRNPs may be mediated by U2AF and SC35 (11,35). Direct and indirect interactions among these factors has been suggested by two-hybrid experiments (30). For several pre-mRNAs, the requirement for U1 snRNP can be relieved by very high levels of SR proteins (36,37). This implies that the interactions of U1 snRNP with U2 snRNP and other factors arriving later may be mediated by SR proteins. Cleavage at a 5′ splice site during the first transesterification reaction can be determined by subsequent interactions between this site and U5 or U6 snRNPs (38–42). The U1 independent reactions have lost fidelity for the authentic 5′ splice site, however, and show a preponderance of cryptic site utilization (36,37). Therefore, it is possible to dissect the early roles of the SR proteins and U1 snRNP, the former nucleate the assembly of the splicing machinery and the latter provides specific selection of authentic 5′ splice site sequences.

Previously we have shown that highly purified U1 snRNP and ASF/SF2 cooperated to form complexes with pre-mRNAs (28). These complexes could distinguish between a functional 5′ splice site and an inactive site where the GU dinucleotide was mutated to AU. Moreover, we showed that this interaction required the arginine-serine domains of ASF/SF2 and probably also an arginine-serine rich domain of the 70 kDa protein of U1 snRNP. Here we characterize these complexes further: we show that 5′ splice site sequences are necessary and sufficient for complex formation. We also show that deletions in the RNA binding domains of ASF/SF2 abolish complex formation. In addition our data demonstrate that the C protein of U1 snRNP is required for this interaction. The data presented here and previous findings suggest a model for 5′ splice site recognition and spliceosome assembly mediated by cooperation between SR proteins and U1 snRNP.

MATERIALS AND METHODS

Plasmids and RNAs

pPIP7.A was described previously (43). pPIP7.A Δ5′E was constructed by removing the *Eco*RI–*Kpn*I fragment of pPIP7.A, making ends blunt with T4 DNA polymerase in the presence of deoxynucleotide triphosphates and ligating, pPIP7.A Del I was constructed by removing the *Sac*I–*Xho*I fragment of pPIP7.A. In pPIP7.A ΔPy, a mutant *Xho*I–*Pst*I fragment derived from oligonucleotides, replaced the PIP7.A *Xho*I–*Pst*I fragment and changed the polypyrimidine tract from TTTCCCTTTTTTTT to ATACACATATATAT. pPIP7.A$(AG)_n$ and pPIP7.A$(\Delta AG)_n$ have the sequence AGGACAGAGC or the sequence AGUACUAUCU respectively replacing the pPIP7.A sequence: AGGACAAACU, which is located in the 3′ exon bases 197–207 in the RNA. These changes were introduced by replacing the *Xho*I–*Pst*I fragment with mutated oligodeoxyribonucleotides. The sequences of all plasmid inserts and deletions were verified by dideoxy-sequencing using Sequenase 2.0 (US Biochemical).

All RNAs were synthesized using T7 RNA polymerase (Stratagene) as described previously (44,45) except that RNAs were labeled with [α-^{32}P]UTP at a final specific activity of 300 Ci/mmol in the labeling reaction. PIP7.A *Bam*HI, PIP7.A *Sal*I, PIP7.A *Pst*I and PIP7.A Δ5′E-*Bam*HI RNAs were synthesized by linearizing pPIP7.A or pPIP7.A Δ5′E with the indicated restriction endonucleases. Otherwise all RNAs were synthesized off *Hin*dIII linearized templates. Competitor RNAs were labeled in a reaction mixture containing ~3.75 mCi of [α-^{32}P]UTP per mmol.

ASF/SF2 and U1 snRNP

Recombinant ASF/SF2 and mutant derivatives were described previously (28,46). To construct the double mutant with deletions in both RBDs, pDS-ASF-NR-g = 35–63, pDS-ASF-NR-d = 172–198 (46) were digested with *Bst*BI and *Apa*I and religated following treatment with Klenow. Proteins were expressed and purified as described (46).

Purification of U1 snRNPs was accomplished by using anti-m_3G affinity chromatography as described previously (47). In order to produce U1 ΔC, U1 ΔA or core U1 snRNP particles partially purified U1 snRNPs were fractionated on Mono-Q columns at various temperatures as described by Bach *et al.* (47).

Complex reactions and native gel electrophoresis

ASF/SF2 and U1 snRNP complex formation assays were described previously (28). Briefly, ASF/SF2 protein (240 ng; 7.3 pmol), U1 snRNP (800 ng; 2.7 pmol), pre-mRNA (16 pg; 2.4 fmol), 62 mM KCl and 1.6 mM $MgCl_2$ were incubated for 5 min at 30°C. Heparin was added to a final concentration of 0.5 $mgml^{-1}$ before loading onto native gels. Samples were loaded on non-denaturing 4% acrylamide:bisacrylamide (80:1) gels containing 50 mM Tris–50 mM glycine (3). Electrophoresis was at 14 Vcm^{-1} for 2.5 h. Complexes were visualized by autoradiography with hyperfilm (Amersham).

Antisera

The normal human serum (J_m) and the systemic lupus erythematosus patient serum (AW), a kind gift of Dr J. Keene, were both described in Jamison *et al.* (6). Reactions containing ASF/SF2, U1 snRNP and radiolabeled PIP7.A and supplemented with Heparin, were incubated with 1:160 (v/v) dilution of antisera in Buffer D (48) in the presence of 600-fold molar excess of competitor PIP7.A RNA, 5 mg/ml BSA and 1% (w/v) NP40 for 5 min at 30°C. These were then loaded onto native gels as described above.

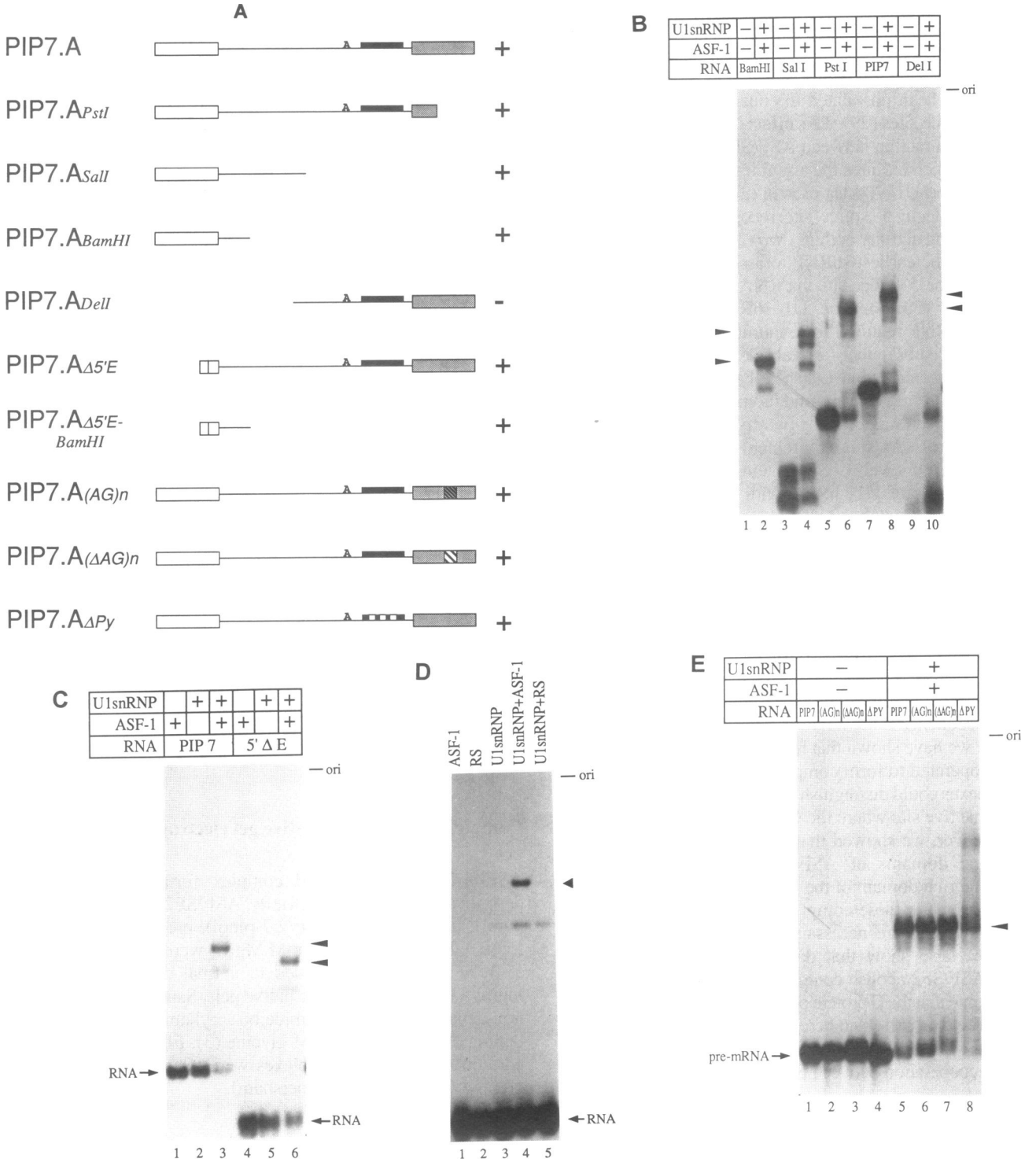

Figure 1. Sequences at the 5′ splice site are necessary and sufficient for U1 snRNP and ASF/SF2 binding. (**A**) A schematic showing the pre-mRNAs used to analyze the *cis*-element requirement for the ASF/SF2 and U1 snRNP dependent gel shift. PIP7.A RNA has been described previously by Gil *et al.* (43). The deletion mutants are indicated clearly in the diagram. PIP7.A$(AG)_n$ substituted the sequence 5′-AGGACAAACU-3′ in the second exon of PIP7.A for 5′-AGGACAGAGC-3′. This sequence was among those selected by ASF/SF2 in *in vitro* selection experiments. In PIP7.A$(\Delta AG)_n$ this sequence was mutated in order to make it pyrimidine-rich: 5′-AGUACUAUCU-3′. The polypyrimidine tract in the intron of PIP7.A was substituted by a tract containing an adenine at every other position in PIP7Δ*Py*. (**B**) Deletion mutants of PIP7.A RNA, which are described in (A), were assayed for complex formation in the presence of ASF/SF2 and purified U1 snRNP (even number lanes). The RNAs were loaded in odd number lanes in this non-denaturing gel. ASF-1 is wild type recombinant ASF/SF2. (**C**) A deletion in the 5′ exon (PIP7.A Δ*5′E*) of PIP7.A was tested for complex formation in the presence of ASF/SF2 and/or U1 snRNP (lanes 4–6). PIP7.A is shown as a control (lanes 1–3). (**D**) A short RNA spanning the 5′ splice site consensus sequence (PIP7.A Δ*5′E-Bam*HI) of PIP7.A was tested for complex formation in the presence of ASF/SF2 (ASF-1) (lane 1) or RS, an ASF deleted of the RS domain (lane 2), U1 snRNP (lane 3). U1 snRNP and ASF-1 (lane 4) or U1 snRNP and RS (lane 5). (**E**) Mutant PIP7.A RNAs with changes in either the second exon, PIP7$(AG)_n$ and PIP7$(\Delta AG)_n$ or in the intronic polypyrimidine tract, PIP7Δ*Py*, were assayed as in B and C (lanes 6–8); RNAs are shown in lanes 1–4. Ori denotes the origin of electrophoresis and complexes are indicated by arrowheads.

RESULTS

The 5′ splice site is necessary and sufficient for complex formation

We first tested the effect of deletions in PIP7.A pre-mRNA in order to identify the minimum sequence required for formation of the ASF/SF2 and U1 dependent complexes (28) (Fig. 1A). Efficient formation of the complexes was seen using the full length PIP7.A pre-mRNA (Fig. 1B, lane 8) and with all deletions that retained the 5′ splice site, PIP7.A*Bam*HI, PIP7.A*Sal*I, PIP7.A*Pst*I (Fig. 1B, lanes 2, 4 and 6). A hemi-intron RNA missing the 5′ splice site, PIP7.A*Del*I, however, could no longer form complexes (Fig. 1B, lane 10). A deletion that removed most of the 5′ exon sequence but left the 5′ splice site consensus sequence intact, PIP7.AΔ*5′E*, also formed complexes efficiently (Fig. 1C, lane 6). Furthermore, a 26 nucleotide-long RNA spanning the 5′ splice site consensus sequence, PIP7.AΔ*5′E-Bam*HI, was competent to form complexes (Fig. 1D). Our previously published report indicates that the 5′ splice site is necessary for ternary complex formation. The data above showed that sequences at or around the 5′ splice site are necessary and sufficient for complex formation.

We also tested the possibility that inserting a purine-rich sequence in the downstream exon might improve complex formation. Purine-rich elements have been shown to bind SR proteins and the sequence 5′-AGGACAGAGC-3′ which we introduced in the second exon of PIP7.A*(AG)n*, corresponds to one of two high affinity consensus sequences derived from a combinatorial selection experiment with ASF/SF2 (R. Tacke and J. Manley, submitted). As a control for the purine-rich sequence, we constructed a mutant sequence: 5′-AGUACUAUCU-3′ in the second exon of RNA PIP7.A*(ΔAG)n*. When tested in our assay, both PIP7.A*(AG)n* and PIP7.A*(ΔAG)n* formed complexes with the same efficiency as PIP7.A (Fig. 1E, lanes 5–7). It is also noteworthy that these two RNAs were spliced equally well (data not shown), which is consistent with the idea that purine-rich elements display their function primarily on weak introns. A mutation in PIP7.A that changed the polypyrimidine tract by inserting adenines at every other position, PIP7.AΔ*Py*, was also able to form complexes (Fig. 1E, lane 8). The very slight advantage seen for PIP7.AΔ*Py* over PIP7.A seen in this experiment was not reproducible. These data indicated that in the presence of a strong 5′ splice site, other elements are not necessary for recognition by ASF/SF2 and U1 snRNP.

A functional RNA binding domain and an arginine-serine domain are both required for recognition of the 5′ splice site

ASF/SF2 has three well characterized primary sequence motifs. Two RNA recognition motifs are found in the N-terminal 200 amino acids (Fig. 2A). These motifs have been empirically shown to bind RNA in isolation and have been thus named RNA binding domains 1 and 2 (RBD1 and RBD2) (46,49). The C-terminal 50 amino acids constitute the arginine-serine (RS) domain, which we had previously shown is required for 5′ splice site recognition (28). We decided to test mutants in the RNA binding domains of ASF/SF2 to determine their requirements in 5′ splice site recognition in collaboration with U1 snRNP. Two mutants Δ35–63 and Δ1–35, which disrupt the RNP-1 and RNP-2 of RBD1 respectively, did not diminish the capacity of the ASF/SF2

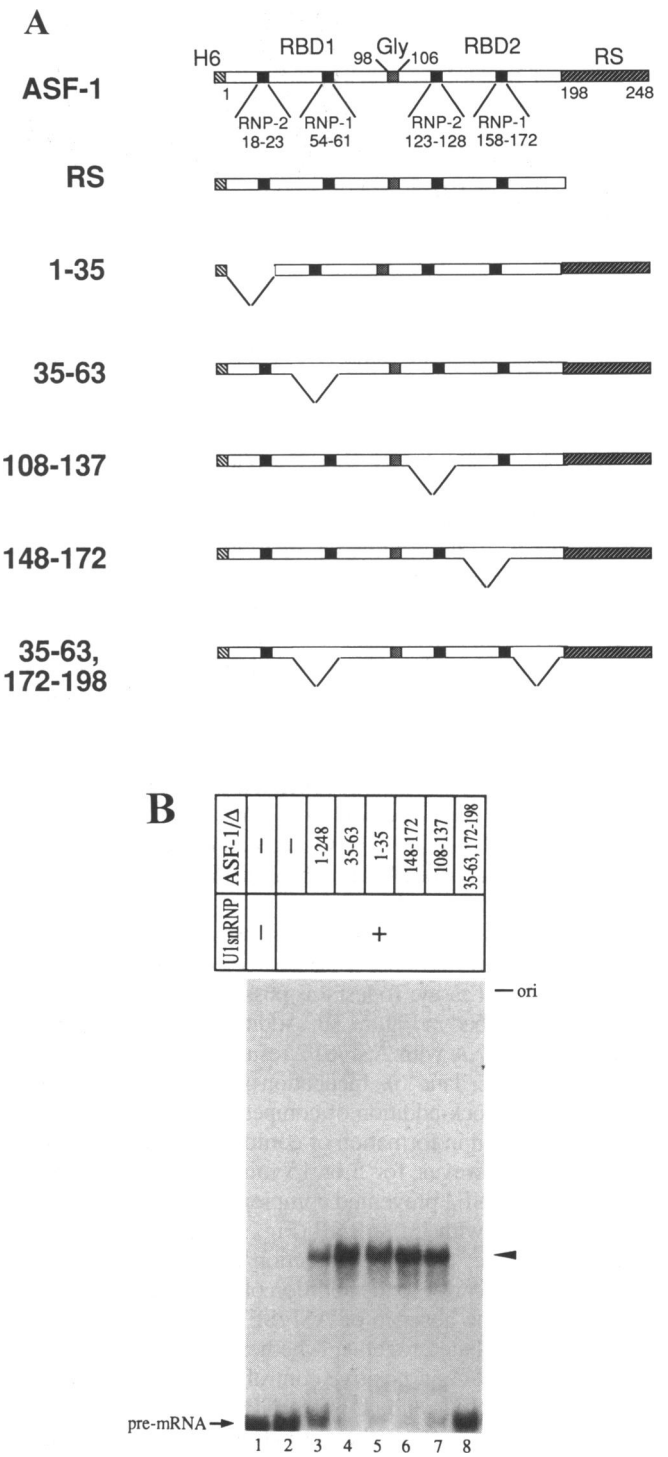

Figure 2. One functional RBD in ASF/SF2 is sufficient for 5′ splice site recognition. (**A**) Schematic showing ASF/SF2 deletion mutants that were used, similar levels of protein were confirmed by limiting dilution series and silver staining of purified recombinant proteins. (**B**) Full length recombinant and His tagged ASF/SF2, ASF-1, (1–248) or mutants of ASF/SF2 were assayed for complex formation with U1 snRNP and PIP7.A pre-mRNA. Deletions in RBD1 were 35–63 and 1–35, deletions in RBD2 were 148–172 and 108–137 and a deletion in both RBDs was 35–63, 172–198. Complexes were detected by native gel electrophoresis as described in the Materials and Methods. Ori denotes the origin of electrophoresis and complexes are indicated by arrowheads.

to interact with U1 snRNP and the 5′ splice site (Fig. 2B, lanes 4 and 5). The same was observed when either the RNP-1 or RNP-2 sequences of RBD-2 were disrupted in mutants Δ148–172 and Δ108–137 respectively (Fig. 2B, lanes 6 and 7). When both RBD-1 and RBD-2 were disrupted, however, no complex formation was observed (Fig. 2B, lane 8). Thus, we conclude that only one of the two RBDs in ASF/SF2 is required for 5′ splice site recognition.

The U1 snRNP stabilizes the interaction of ASF/SF2 with pre-mRNA

We have shown previously that 5′ splice site recognition by purified ASF/SF2 and U1 snRNP was absolutely dependent on the order of addition (28). These experiments indicated that ASF/SF2 must interact with the pre-mRNA first, although this interaction could not be detected by gel shift assays. Presumably ASF/SF2 promotes subsequent U1 snRNP binding. We show here that the complexes seen on native gels were remarkably stable (Fig. 3A). Incubations with either U1 snRNP or ASF/SF2 alone did not result in efficient complex formation, which is consistent with our previous data (Fig. 3A, lanes 1–4). Preincubation with 1500-fold molar excess of unlabeled competitor PIP7.A inhibited formation of the ASF/SF2 and U1 snRNP dependent complexes (lane 5). If these complexes were allowed to form, however, they were recalcitrant to challenge with the same amount of competitor for up to 1 h of incubation (lanes 6–11). Therefore the ternary complexes are remarkably stable.

It was possible that stable complexes formed between ASF/SF2 and pre-mRNA that were not resistant to electrophoresis and thus invisible in our gel assay. To test this possibility we performed the experiment described in Figure 3B. Addition of competitor before incubation of PIP7.A with ASF/SF2 resulted in no complexes, as expected (Fig. 3C, lane 6). Incubation with ASF/SF2 followed sequentially by mock-addition of competitor and incubation with U1 snRNP resulted in formation of complexes (lane 7). Addition of competitor, however, for 5 or 15 min, after incubation with PIP7.A and ASF/SF2 prevented complex formation upon subsequent incubation with U1 snRNP (Fig. 3C, lanes 8 and 9). It is unlikely, therefore, that the interaction between ASF/SF2 and PIP7.A was stable prior to the addition of U1 snRNP. Complexes did not form in the absence of ASF/SF2 when PIP7.A and U1 snRNP were incubated together whether or not competitor was added (lanes 2–5). As a positive control, PIP7.A was incubated with both ASF/SF2 and U1 snRNP prior to addition of competitor, demonstrating the formation of stable complexes (lanes 11–13). Together, these results suggest an early but unstable interaction between ASF/SF2 and PIP7.A pre-mRNA that is essential for subsequent complex formation.

We characterized the binding of U1 snRNP and ASF/SF2 further by nuclease footprinting. RNase T1 and RNase A gave reproducible cleavage patterns on 5′ end labeled PIP7.A RNA (data not shown). Addition of increasing concentrations of ASF/SF2 resulted in equivalent protection from RNase T1 cleavage at all guanines (G). No preference for protection of G_{55} and G_{56} at the 5′ splice site was observed, whereas U1 snRNP specifically protected the RNase T1 cleavage sites at the 5′ splice site, G_{55} and G_{56} (data not shown).

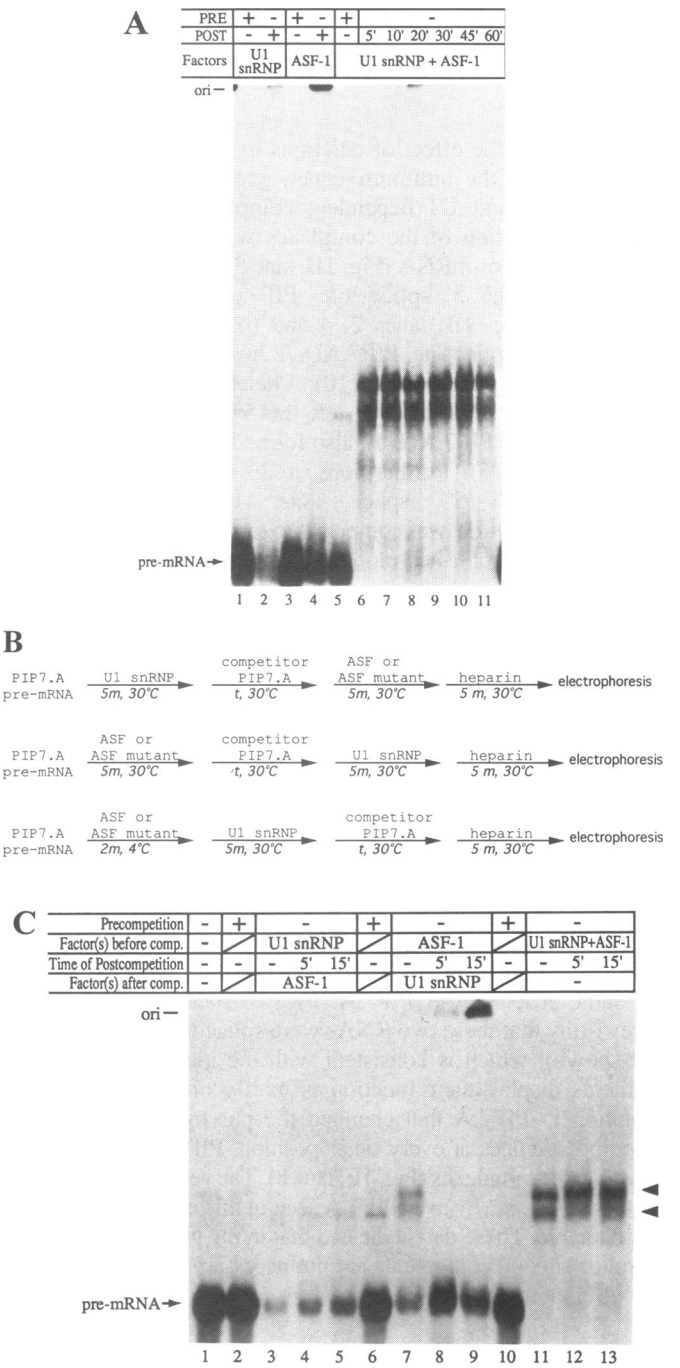

Figure 3. The ASF/SF2 and U1 snRNP dependent complexes are stable. (**A**) Complex formation was assayed by native gel electrophoresis as described above. In this experiment, a 1500-fold molar excess of unlabeled PIP7.A pre-mRNA was added either before or after the complexes were formed 'PRE' and 'POST' respectively. PIP7.A RNA was incubated with either ASF-1 or U1 snRNP (lanes 1–4) or both together (lanes 5–11). Unlabeled PIP7.A was added before complex formation (lane 5) or after complex formation (lanes 6–11). (**B**) A schematic that explains the experiment shown in panel C. The time of addition of excess unlabeled PIP7.A RNA (competitor PIP7.A relative to the addition of ASF/SF2 and U1 snRNP is indicated. '*t*' indicates where times of incubation were varied in the experiment (see C). (**C**) Complex formation was assayed as described above, competitor RNA was added before ASF/SF2 or U1 snRNP in lanes 2, 6 and 10. ASF/SF2 or U1 snRNP or both were added to labeled PIP7.A RNA before the competitor in lanes 3–5, 7–9 or 11–13 respectively. A minus sign for time of post-competition indicates that no competitor RNA was added to those reactions. Ori denotes the origin of electrophoresis and complexes are indicated by arrowheads.

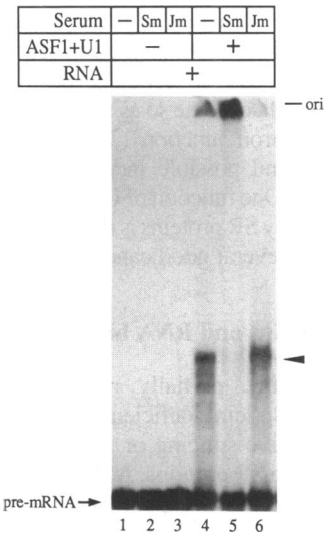

Figure 4. U1 snRNP is found in the complexes after electrophoresis. ASF/SF2 and U1 snRNP were incubated with labeled PIP7.A RNA as described above. After complex formation, the reactions were incubated with Buffer D (lane 4) or S_m (lane 5) or a J_m (lane 6) (see Materials and Methods). The free RNAs were also incubated with the antisera as a negative control (lanes 1–3). These reactions were then subjected to non-denaturing gel electrophoresis as described above. Ori denotes the origin of electrophoresis and complexes are indicated by arrowheads.

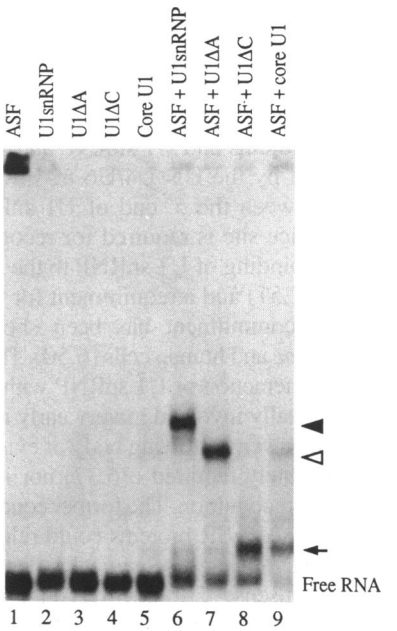

Figure 5. U1 snRNP specific C protein is necessary for complex formation. Complex formation was assayed as above except that incubations contained incomplete U1 snRNP particles: U1 snRNP missing the U1-specific A protein (U1 ΔA), U1-specific C protein (U1ΔC) or all U1-specific proteins (core). Incubations with either ASF/SF2 (lane 1) or U1 snRNP particles (lanes 2–5) as well as incubations with both (lanes 6–9) are shown. Ori denotes the origin of electrophoresis and complexes are indicated by arrowheads.

Role of U1 snRNP proteins in recognition of the 5′ splice site

We had previously shown that in order to form complexes both ASF/SF2 and U1 snRNP were required (28). It was not clear, however, if U1 snRNP was in the complexes in the native gels. To date it has been difficult to identify a U1 snRNP pre-mRNA complex in native gels when using mammalian nuclear extracts. In contrast yeast CC1 and CC2 complexes are U1 snRNP-containing complexes observed in native gels (50). To determine whether or not U1 snRNP was in the U1-ASF/SF2 dependent complexes, we tested the effect of a highly specific anti-U snRNP antiserum on their mobility. A mobility shift of the complexes was observed with this human antiserum (S_m) directed at U snRNP proteins (Fig. 4, lane 5), whereas a normal human serum (J_m) did not shift these complexes (Fig. 4, lane 6). Further evidence that U1 snRNP was in the complex was obtained by observing complex mobility changes with variant U1 snRNP particles (see below and Fig. 5).

A U1 snRNP particle missing the U1 specific A protein (U1 ΔA) was capable of collaborating with ASF/SF2 to bind PIP7.A (Fig. 5, lane 7). The resulting complex had a faster mobility than the one formed with U1 snRNP in the native gels. The complex formed by U1 ΔA required a functional 5′ splice site because it was not observed with PIP7.A 5′ AU where the first base of the intron has been changed to an A (data not shown). Moreover, this complex did not form with ASF ΔRS, a truncated ASF/SF2 missing the RS domain (data not shown).

A U1 snRNP particle missing the U1 specific C protein U1 ΔC did not form complexes with mobilities similar to the U1 ΔA and U1 snRNP complexes (Fig. 5, lane 8). A small shift, however, was reproducibly observed when PIP7.A was incubated with ASF/SF2 and U1 ΔC. This band was not observed with a pre-mRNA mutated in the 5′ splice site or when the ASF ΔRS was used, suggesting that it required an interaction between ASF/SF2 and U1 snRNP (data not shown). The shifted band, however, has a mobility that is not consistent with the presence of U1 snRNP (data not shown). A U1 snRNP preparation known as core U1, which is missing the A, C and 70 kDa proteins, resulted in low levels of complexes, mostly with the same mobility as those seen with U1 ΔC (Fig. 5, lane 9). Results with the core U1 preparation were more variable leading us to believe it was a mixed population of U1 snRNP particles (51). These results and our previously published data are consistent with the idea that the U1-C and 70 kDa proteins, but not the U1-A protein, are required for complex formation.

DISCUSSION

We have shown previously that purified SR protein ASF/SF2 collaborates with purified U1 snRNP to recognize a functional 5′ splice site (28). Here we show that a complex formed by ASF/SF2, U1 snRNP and pre-mRNA requires only sequences at and around the 5′ splice site. Only one functional RNA binding domain and the RS domain in ASF/SF2 are required for 5′ splice site recognition in cooperation with U1 snRNP. Furthermore, we show that U1 snRNP C protein is required for proper 5′ splice site recognition, whereas U1 A protein is dispensable.

Recognition and selection of 5′ splice sites

Recognition and selection of 5′ splice site sequences and eventual selection of authentic sites is a two phase process. The early phase is determined by SR proteins and U1 snRNP and the late phase is probably determined by the U5 · U4/U6 tri-snRNP. Watson–Crick base pairing between the 5′ end of U1 snRNA and the sequences at the 5′ splice site is required for recognition of the splice site (7,33). The binding of U1 snRNP to the 5′ splice sites is well documented (32,51) and a requirement for U1 snRNP at the level of splicing commitment has been shown for both *Saccharomyces cerevisiae* and human cells (6,50). Therefore there is little doubt that the interaction of U1 snRNP with sequences at the 5′ splice site is normally involved in very early recognition of these sites. There is, however, a growing body of evidence that U1 snRNP is neither absolutely required (36,37) nor is it sufficient (52,53) for 5′ splice site recognition. The former conclusion comes from studies in which excess SR proteins could relieve blocks in extracts where U1 snRNP was biochemically depleted (36) or debilitated with an antisense 2′-*O*-methyl oligoribonucleotide directed to the 5′ end of the U1 snRNA (37). The SR proteins had already been shown to be required splicing factors that acted at the earliest measurable steps in the splicing pathway (26) and could collaborate with U1 snRNP to specifically recognize a functional 5′ splice site (28). None of these reports, however, predicted that SR proteins could replace the splicing requirement for U1 snRNP and therefore all previous models of SR function must be modified to explain the data of Crispino *et al.* (36) and Tarn and Steitz (37). Our data both in this report and in Kohtz *et al.* (28) help establish such a model of SR function. Tarn and Steitz have offered an alternative (37) but similar explanation based on the ability of SR proteins to directly recognize 5′ splice sites (54). Our model of SR protein action does not deny a possible preference for 5′ splice site sequences but it does not require it.

What is the mechanism of action of SR proteins? We envision that SR proteins bind the pre-mRNA first and promote a stabilization of U1 snRNP binding at 5′ splice sites. It is clear that purified U1 snRNP can bind pre-mRNA in the absence of SR proteins as documented by RNase T1 footprinting (data not shown). This U1–5′ splice site interaction does not withstand native gel electrophoresis, however, whereas a U1–5′ splice site interaction in the presence of ASF/SF2 results in the formation of stable complexes. In agreement with this, addition of ASF/SF2 to nuclear extracts increases the apparent affinity of U1 snRNP to all 5′ splice site sequences in a pre-mRNA (29). Our data also suggest that U1 snRNP changes the interaction between ASF/SF2 and pre-mRNA. Our data argue that U1 snRNP and SR proteins mutually enhance the affinity of the other for 5′ splice sites. We can define then a role for U1 snRNP as a 5′ splice site dependent stabilizer of SR protein–pre-mRNA interactions. This is a role that could be replaced by addition of a large excess of SR proteins. This role must be one acted out early by U1 snRNP given that an excess of SR proteins can replace the requirement for U1 snRNP in splicing commitment (26).

Following initial pre-mRNA binding and 5′ splice site definition in collaboration with U1 snRNP, SR proteins promote an interaction across the intron (14) with the splicing factor U2AF (11). This heterodimeric factor has two subunits, one of which, a 65 kDa polypyrimidine tract binding subunit with an N-terminal arginine-serine domain, is absolutely required for splicing (55). U2AF, probably in combination with SR proteins (23), promotes the binding of U2 snRNP to branch point sequences (11).

The U6 and U5 snRNAs can recognize the 5′ splice site a second time and situate the spliceosome so as to cleave the phosphodiester bond at the exon–intron junction (38–42,56–59). This would represent a second and possible independent recognition of 5′ splice site sequences. One function of U1 snRNP, which apparently cannot be duplicated by SR proteins is the selection of the authentic 5′ splice site among several good candidate sites (37).

RNA binding domains and RNA binding by SR proteins

The SR proteins have partially redundant activity in the complementation of splicing deficient S100 extracts (24,25). In regulation of alternative splicing or preference for proximal 5′ splice sites at least two SR proteins, ASF/SF2 and SC 35, can also have similar activity (35). This apparent conservation of function is paralleled by conservation of structure. Zahler *et al.* (25) divide the SR family into two subgroups, those containing one RBD, SRp20 and SRp30b (SC35) and those with one classic RBD and a distant homologue of an RBD: SR30a(ASF/SF2), SRp40, SRp55 and SRp75 (24). Both RBD1 and RBD2 in isolation can bind a 95 nt SV40 pre-mRNA fragment (46) and can be cross-linked to a human β-globin pre-mRNA (49), albeit with low affinity relative to proteins containing both domains. Mutations in either RBD can diminish the constitutive splicing activity of ASF/SF2, however, the mutational analysis has not clearly established separable functions for the domains. Our data demonstrate that disruption of one or another of the RBDs does not diminish the ability of ASF/SF2 to recruit U1 snRNP to the 5′ splice site. On the other hand disruption of both completely prevents this activity. The RBD redundancy may be consistent with evolution of the SR family, in which some members only have one RBD while others have two (24). These results, coupled with previous experiments (43; Tacke and Manley, submitted), suggest that low affinity, non-specific interactions with the RNA are sufficient for complex formation with U1 snRNP. It is important to note, however, that all of the RBD mutants tested here are completely inactive as essential splicing factors (46), indicating that the ability to form ternary complexes is not sufficient for splicing activity. Given that SR proteins interact with several different components of the spliceosome it is not surprising that the splicing assay is more sensitive than our assay to single RBD mutations. Our previous work (28) and this report establish that the RS domain is absolutely required for U1 snRNP recruitment to 5′ splice sites. This central activity of RS domain is conserved throughout the family and may reflect an ability to interact with the U1 70 kDa protein by most SR members (28,30).

ASF/SF2 has been shown to bind SV40 and adenovirus pre-mRNAs in a 5′ splice site dependent fashion (54), suggesting an ability for the protein to recognize the site. Recently, *in vitro* selection experiments have revealed binding preferences for ASF/SF2 and SC35 to distinct purine-rich sequences (Tacke and Manley, submitted). The sequences recognized by ASF/SF2 are very similar to those described in certain exon splicing enhancer elements (14,60,61). In footprinting experiments, we found no evidence of preferential affinity for sequences at the 5′ splice site of PIP7.A RNA. It is possible that purine-rich sites or 5′ splice sites nucleate the initial binding, which can then be observed over most of the pre-mRNA. The 5′ splice site consensus in mammals AG:GURAGU is purine-rich and thus may favor the binding of

SR proteins, as originally suggested by Zuo and Manley (54). Our view of SR protein function, however, does not require independent binding to 5′ splice sites or to purine-rich sequences. As long as these proteins bind the pre-mRNA and recruit U6/4 and U5 snRNPs to the vicinity of an intron, these latter factors could 'search' the pre-mRNA to select an active 5′ splice site. In the presence of U1 snRNP, this process is more selective; the functional concentration of SR proteins at the authentic 5′ splice site is high and subsequent reactions are thus driven to the correct site. An attractive model, consistent with all the data, is that for introns with weak splicing signals, high affinity SR-protein binding sites, and hence a high concentration of SR proteins, are required for recruitment of snRNPs and activation of splicing. In contrast, when splicing signals are strong, U1 may require less 'assistance' from SR proteins and the non-specific binding, such as we observed with the PIP7.A pre-mRNA, may be sufficient to allow stable complex formation. Indeed, in this case the snRNP, in addition to helping define the 5′ splice site, may function to stabilize binding of SR proteins, which are then required for subsequent steps in spliceosome assembly.

ACKNOWLEDGEMENTS

We thank J. Keene (Duke University) for much encouragement and for providing us with S_m antisera. We are grateful to Sabina Sager for assistance in the preparation of this manuscript. This work was funded by research grants from NIH (M.A.G.-B. and J.L.M). M.A.G.-B. is funded by a JFRA from the ACS and would also like to acknowledge the support of the Keck Foundation to the Levine Science Research Center.

REFERENCES

1 Sharp, P.A. (1994) *Cell*, **77**, 805–815.
2 Moore, J.M., Query, C.C. and Sharp, P.A. (1993) In R.F. Gesteland and J.F. Atkins (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.
3 Konarska, M.M. and Sharp, P.A. (1986) *Cell*, **46**, 845–855.
4 Pikielny, C.W., Rymond, B.C. and Rosbash, M. (1986) *Nature*, **324**, 341–345.
5 Ruby, S.W. and Abelson, J. (1988) *Science*, **242**, 1028–1035.
6 Jamison, S.F., Crow, A. and Garcia-Blanco, M.A. (1992) *Mol. Cell. Biol.*, **12**, 4279–4287.
7 Seraphin, B. and Rosbash, M. (1989) *Gene*, **82**, 145–151.
8 Lamond, A.I., Konarska, M.M. and Sharp, P.A. (1987) *Genes Dev.*, **1**, 532–543.
9 Roscigno, R.F., Weiner, M. and Garcia-Blanco, M.A. (1993) *J. Biol. Chem.*, **268**, 11 222–11 229.
10 Zahler, A.M., Lane, W.S., Stolk, J.A. and Roth, M.B. (1992) *Genes Dev.*, **6**, 837–847.
11 Zamore, P.D. and Green, M.R. (1989) *Proc Natl Acad Sci USA*, **86**, 9243–9247.
12 Black, D.L., Chabot, B. and Steitz, J.A. (1985) *Cell*, **42**, 737–750.
13 Krainer, A.R. and Maniatis, T. (1985) *Cell*, **42**, 725–736.
14 Tian, M. and Maniatis, T. (1993) *Cell*, **74**, 105–114.
15 Staknis, D. and Reed, R. (1994) *Mol. Cell. Biol.*, **14**, 2994–3005.
16 Valcarcel, J., Singh, R., Zamore, P.D. and Green, M.R. (1993) *Nature*, **362**, 171–175.
17 Tian, M. and Maniatis, T. (1992) *Science*, **256**, 237–240.
18 Kreivi, J.P. and Akusjarvi, G. (1994) *Nucleic Acids Res.*, **22**, 332–337.
19 Kenan, D.J., Query, C.C. and Keene, J.D. (1991) *Trends Biochem Sci.*, **16**, 214–220.
20 Ge, H. and Manley, J.L. (1990) *Cell*, **62**, 25–34.
21 Krainer, A.R., Conway, G.C. and Kozak, D. (1990) *Genes Dev.*, **4**, 1158–1171.
22 Krainer, A.R., Conway, G.C. and Kozak, D. (1990) *Cell*, **62**, 35–42.
23 Fu, X.D. and Maniatis, T. (1992) *Science*, **256**, 535–538.
24 Zahler, A.M., Neugebauer, K.M., Stolk, J.A. and Roth, M.B. (1993) *Mol. Cell. Biol.*, **13**, 4023–4028.
25 Zahler, A.M., Neugebauer, K.M., Lane, W.S. and Roth, M.B. (1993b) *Science*, **260**, 219–222.
26 Fu, X.D. (1993) *Nature*, **365**, 82–85.
27 Kim, Y.J., Zuo, P., Manley, J.L. and Baker, B.S. (1992) *Genes Dev.*, **6**, 2569–2579.
28 Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Luhrmann, R., Garcia-Blanco, M.A. and Manley, J.L. (1994) *Nature*, **368**, 119–124.
29 Eperon, I.C., Ireland, D.C., Smith, R.A., Mayeda, A. and Krainer, A.R. (1993) *EMBO J.*, **12**, 3607–3617.
30 Wu, J.Y. and Maniatis, T. (1993) *Cell*, **75**, 1061–1070.
31 Caceres, J.F., Stamm, S., Helfman, D.M. and Krainer, A.R. (1994) *Science*, **265**, 1706–1709.
32 Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A. and Steitz, J.A. (1983) *Cell*, **33**, 509–518.
33 Zhuang, Y. and Weiner, A.M. (1986) *Cell*, **46**, 827–835.
34 Zhuang, Y., Leung, H. and Weiner, A.M. (1987) *Mol. Cell. Biol.*, **7**, 3018–3020.
35 Fu, X.D., Mayeda, A., Maniatis, T. and Krainer, A.R. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 11 224–11 228.
36 Crispino, J.D., Blencowe, B.J. and Sharp, P.A. (1994) *Science*, **265**, 1866–1869.
37 Tarn, W.Y. and Steitz, J.A. (1994) *Genes De.v*, **8**, 2704–2717.
38 Newman, A.J. and Norman, C. (1991) *Cell*, **65**, 115–124.
39 Newman, A.J. and Norman, C. (1992) *Cell*, **68**, 743–754.
40 Lesser, C.F. and Guthrie, C. (1993) *Science*, **262**, 1982–1988.
41 Kandels-Lewis, S. and Seraphin, B. (1993) *Science*, **262**, 2035–2039.
42 Konforti, B.B. and Konarska, M.M. (1994) *Genes Dev.*, **8**, 1962–1973.
43 Gil, A., Sharp, P.A., Jamison, S.F. and Garcia-Blanco, M.A. (1991) *Genes Dev.*, **5**, 1224–1236.
44 Garcia-Blanco, M.A., Jamison, S.F. and Sharp, P.A. (1989) *Genes Dev.*, **3**, 1874–1886.
45 Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
46 Zuo, P. and Manley, J.L. (1993) *EMBO J*, **12**, 4727–4737.
47 Bach, M., Bringmann, P. and Luhrmann, R. (1990) *Methods Enzymol.*, **181**, 232–257.
48 Dignam, J.D., Lebovitz, R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
49 Caceres, J.F. and Krainer, A.R. (1993) *EMBO J.*, **12**, 4715–4726.
50 Seraphin, B. and Rosbash, M. (1989b) *Cell*, **59**, 349–358.
51 Heinrichs, V., Bach, M., Winkelmann, G. and Luhrmann, R. (1990) *Science*, **247**, 69–72.
52 Siliciano, P.G. and Guthrie, C. (1988) *Genes Dev.*, **2**, 1258–1267.
53 Seraphin, B. and Rosbash, M. (1990) *Cell*, **63**, 619–629.
54 Zuo, P. and Manley, J.L. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 3363–3367.
55 Zamore, P.D., Patton, J.G. and Green, M.R. (1992) *Nature*, **355**, 609–614.
56 Cortes, J.J., Sontheimer, E.J., Seiwert, S.D. and Steitz, J.A. (1993) *EMBO J.*, **12**, 5181–5189.
57 Wyatt, J.R., Sontheimer, E.J. and Steitz, J.A. (1992) *Genes Dev.*, **6**, 2542–2553.
58 Sawa, H. and Shimura, Y. (1992) *Genes Dev.*, **6**, 244–254.
59 Sontheimer, E.J. and Steitz, J.A. (1993) *Science*, **262**, 1989–1996.
60 Watakabe, A., Tanaka, K. and Shimura, Y. (1993) *Genes Dev.*, **7**, 407–418.
61 Inoue, K., Hoshijima, K., Higuchi, I., Sakamoto, H. and Shimura, Y. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 8092–8096.