

TABLE 1. Quantitation of AG selection in Fig. 1 and 2

Pre-mRNA (nt) ^a	Ratio of distal AG products to proximal AG products
15/18	4.2
15/19	1.7
15/20	2.1
15/21	0.1
15/22	0.2
15/23	0 ^b
16/19	3.0
16/20	2.4
16/21	1.3
16/22	0 ^b
17/20	2.0
17/21	0.6
17/22	0.5
17/23	0 ^b
18/21	1.6
18/22	0.3
18/23	0.5
18/24	0 ^b
19/22	1.2
19/23	0.2
23/26	0.4
23/27	0 ^b

^a Distance of proximal AG downstream from the BPS/distance of distal AG from the BPS.

^b The levels of the distal lariat were not greater than background.

UUUUCUUUUUUUUUCAG-3'; and 40AG, 5'-aCUUUUUUUUUUUUUUUCUUUUUUUUUUUUUUUUUUUUUCAG-3'.

Splicing reactions. ³²P-labeled pre-mRNAs (10 ng) were incubated in 25- μ l splicing reaction mixtures for 1 h unless otherwise indicated. Total RNA was prepared and fractionated on 6.5, 7.5, or 15% denaturing polyacrylamide gels. RNAs were quantitated by phosphorimager (Quantityone; Bio-Rad). The relative use of each AG is indicated in each figure and is also presented in Table 1. The splicing intermediates and products were identified by comigration with known markers on the gels with different percentages of denaturing polyacrylamide.

RESULTS

AG selection depends on the BPS-to-AG and AG-to-AG distances. Previous analysis of a pre-mRNA containing duplicated AGs demonstrated that an AG located 11 nt from the BPS is outcompeted by an AG located 15 nt from the BPS (5). To determine whether an AG at 15 nt is optimally located, we asked whether this AG can be outcompeted by distal AGs at various positions downstream (see schematic in Fig. 1A). As shown in Fig. 1B, the proximal AG at 15 nt is efficiently outcompeted by a distal AG located 3 nt downstream (Fig. 1B, lane 1 [compare levels of distal and proximal lariat]; see Fig. 1C, lane 1, for mRNAs). However, as the distal AG is moved further downstream (Fig. 1A), the AG at 15 nt becomes more competitive until it is used almost exclusively (Fig. 1B and C, lanes 2 to 6). The same pattern of AG selection is observed at late time points (90 min) in the splicing reaction (data not shown). We conclude that distal AGs located less than 6 nt downstream from the proximal AG can efficiently compete with the proximal AG.

To test whether it is a general rule that distal AGs can compete with proximal AGs, we examined pre-mRNAs containing a proximal AG at increasing BPS-to-AG distances and distal AGs at increasing AG-to-AG distances (Fig. 2). This analysis revealed that distal AGs can indeed compete with

proximal AGs located 16, 17, or 18 nt from the BPS (Fig. 2A, B, and C, respectively). Again, the AG-to-AG distance must be less than 6 nt in order for the distal AGs to compete. Within this AG-to-AG distance, distal AGs compete better the shorter the BPS-to-AG distance. For example, a distal AG at an AG-to-AG distance of 5 nt competes optimally when the proximal AG is at 15 nt and does not compete well when the proximal AG is at 17 or 18 nt (Fig. 1 and 2, lanes 3).

Our assay using duplicated AGs shows that AG selection depends on both the BPS-to-AG distance and the distance between the two adjacent AGs. Moreover, the data indicate that there is probably a maximal BPS-to-AG distance beyond which the proximal AG outcompetes any distal AG. To find this distance, we increased the BPS-to-AG distance to 19 or 23 nt. At a BPS-to-AG distance of 19 nt, the distal AG 3 nt away was still preferentially used (Fig. 2D, lane 1). Efficient use of the proximal AG was observed only when the AG-to-AG distance was further increased to 4 nt (Fig. 2D, lane 2). Significantly, however, when the BPS-to-AG distance was 23 nt, a switch occurred. For the first time, the proximal AG was pre-

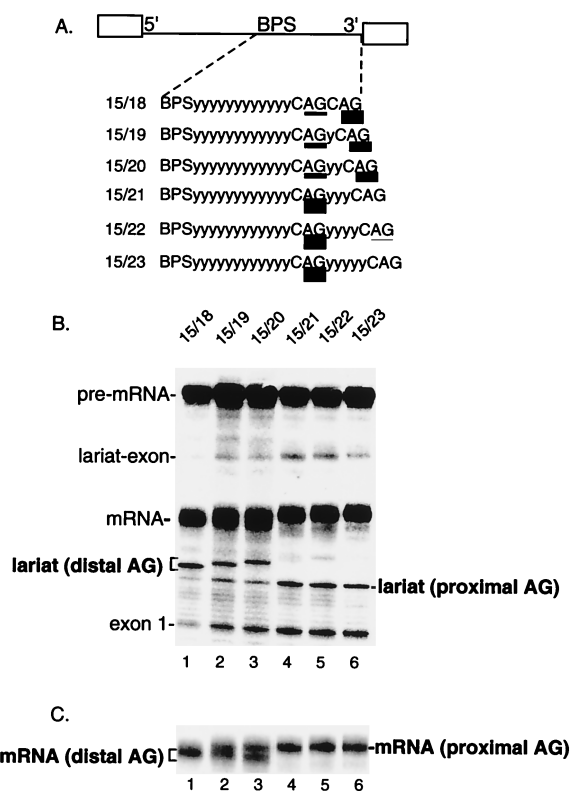


FIG. 1. Distal AGs can compete with a proximal AG located 15 nt from the branch site. (A) Schematic of AdML pre-mRNA derivatives. The proximal AG is 15 nt downstream from the BPS, and the distal AG is located between 18 and 23 nt from the BPS. The relative use of the competing AGs is indicated by the thickness of the underlining and is presented in Table 1. y, pyrimidine. (B) Total RNA from splicing reactions was isolated and fractionated on a 7.5% denaturing polyacrylamide gel. The mRNA and lariat introns corresponding to use of the proximal and distal AGs were identified by comigration with known markers on gels with different percentages of polyacrylamide (data not shown). (C) A lighter exposure of the gel in panel B shows the mRNAs generated using proximal versus distal AGs.

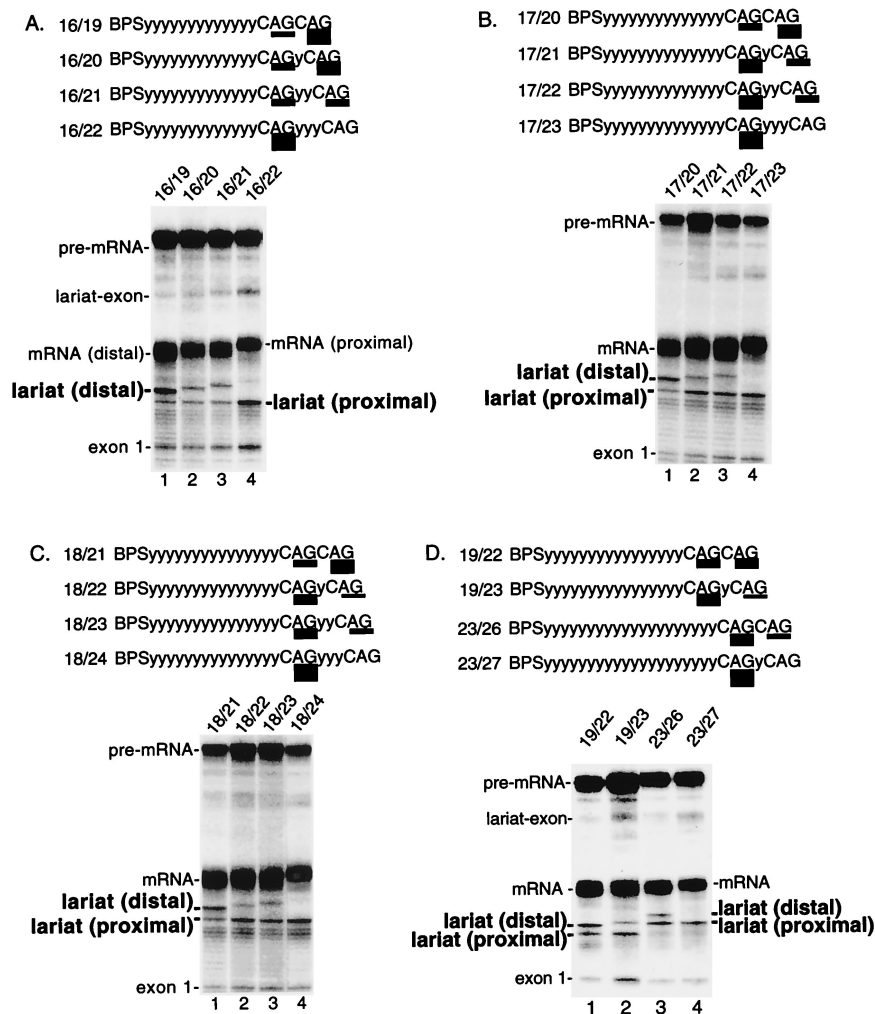


FIG. 2. Competition between duplicated AGs depends on the BPS-to-AG distance and the AG-to-AG distance. Pre-mRNAs contain a proximal AG at 16 nt (A), 17 nt (B), 18 nt (C), and 19 or 23 nt (D) from the BPS. In each panel, a distal AG was located from 3 to 6 nt downstream of the proximal AG. Pre-mRNAs were spliced and analyzed on a 7.5% (A to C) or 6.5% (D) denaturing polyacrylamide gel. Schematics of the pre-mRNAs are shown in each panel. The relative use of each AG is indicated by the thickness of the underlining and is presented in Table 1.

ferred even when the AG-to-AG distance was only 3 nt (Fig. 2D, lane 3). Moreover, the proximal AG was used exclusively when the AG-to-AG distance was 4 nt (Fig. 2D, lane 4). We conclude that distal AGs can no longer outcompete the proximal AG at a BPS-to-AG distance of 23 nt. Together, these data define the maximal BPS-to-AG distance beyond which a proximal AG can no longer be outcompeted by a distal AG. This distance is between 19 and 23 nt.

An optimal BPS-to-AG distance for step II. Our assay using duplicated AGs indicated that distal AGs can compete when a proximal AG is located 19 to 23 nt or less from the BPS. The observation that distal AGs can be used at all suggests that proximal AGs are sterically hindered by being too close to the BPS. If this is the case, then we would predict that the efficiency of step II should be impeded when a single AG is located 19 to 23 nt or less from the BPS. To test this possibility, we examined pre-mRNAs containing single AGs with BPS-to-AG distances ranging from 15 to 40 nt (Fig. 3A). As shown in Fig. 3B and C, step II efficiency is indeed impeded when a single AG is located 15 or 18 nt from the BPS (the level of step

II was quantitated on a darker exposure of the gel [data not shown]). Significantly, however, when the AG is moved further than 21 nt from the BPS, the efficiency of step II also progressively decreases (Fig. 3B and C). Thus, these data indicate that there is an optimal BPS-to-AG distance of between 19 and 23 nt for step II. Our data with the duplicated AGs indicate that upstream of this distance AGs are sterically hindered by being too close to the BPS. The data with the single AGs indicate that beyond this distance AGs are not used efficiently because they are too far from the BPS. We conclude that step II is optimally efficient over a very narrow BPS-to-AG distance (19 to 23 nt). We note that this distance may vary with other pre-mRNAs due to their different pyrimidine tract sequences. However, given the consistency of our results, it is likely that the distance range is close to what we have observed with this pre-mRNA.

Uncoupling two recognition steps for AG selection. Further insight into the mechanism for AG selection came from comparing the data obtained with the duplicated and single AGs. The single-AG data show that a BPS-to-AG distance between

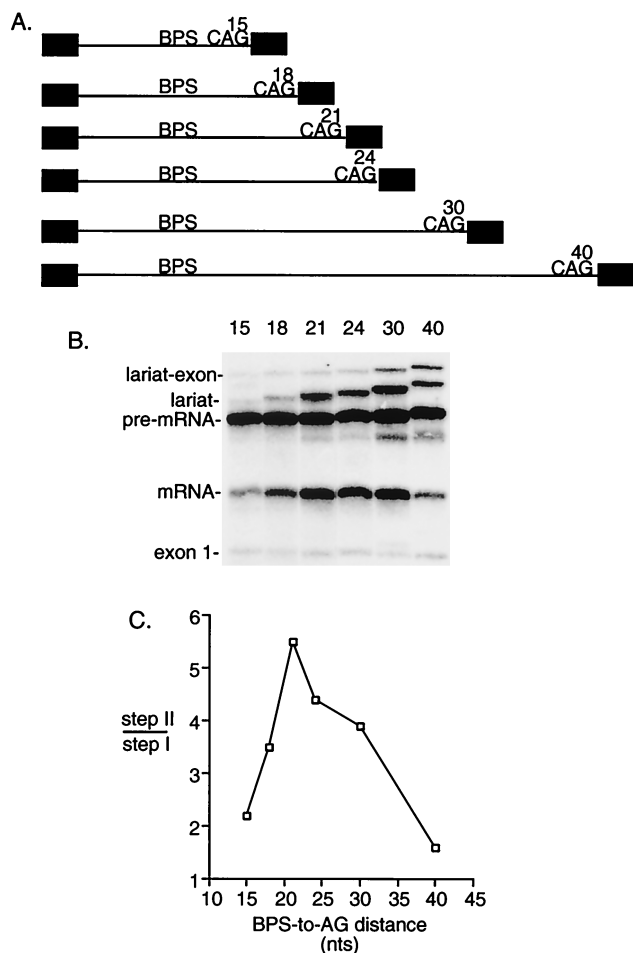


FIG. 3. A BPS-to-AG distance of 19 to 23 nt is optimal for step II. (A) Schematic of pre-mRNAs containing single AGs located at increasing distances from the BPS. (B) Pre-mRNAs were spliced for 45 min and analyzed on a 15% denaturing polyacrylamide gel. (C) Step II efficiency was calculated as the ratio of spliced products (mRNA and lariat) to splicing intermediates (lariat-exon and exon 1).

19 and 23 nt is optimal. Yet with the duplicated AGs, a proximal AG located closer to the BPS was used even when there was a distal AG more optimally located. For example, at a BPS-to-AG distance of 15 nt, the AG at 15 was selected over an AG at 21 nt (Fig. 1). Indeed, the location of the proximal AG, rather than the absolute distance between the BPS and distal AG, determines whether the distal AG is used (Fig. 1 and 2). Specifically, at any given BPS-to-AG distance, distal AGs must be located less than 6 nt from the proximal AG in order to be used. Thus, these data suggest that the proximal AG is recognized, even when it is not used.

To further test this possibility, we asked whether inserting a third AG upstream of two competing AGs alters the use of the two original AGs (see schematic in Fig. 4A). As shown in Fig. 4A (lane 1), the distal AG at 21 nt is used slightly more efficiently than the proximal AG at 18 nt. When a third AG is inserted into this pre-mRNA, 15 nt from the BPS, this AG is not used (Fig. 4A, lane 2). Importantly, however, the presence of the AG at 15 nt alters the pattern of AG selection, as the AG at 18 nt is now exclusively selected (Fig. 4A, lane 2).

The observation that AG selection can be switched depending on the presence or absence of the AG at 15 nt suggests that the AG at 15 nt is recognized even though it is not used. To further test this possibility, the AG at 15 nt was replaced with either a GG or a GA. Significantly, neither of these purine dinucleotides caused a switch to the AG at 18 nt. We conclude that the AG at 15 nt is specifically recognized even when it is not used. Thus, our assay using multiple AGs suggests that there is an initial step in AG recognition that can be uncoupled from the final selection of the AG.

DISCUSSION

Previous studies showed that the first AG downstream of the BPS is normally used as the 3' splice site (reviewed in references 2 and 26). This observation is compatible with several models for 3' splice site selection. In scanning models, the first AG is used because a 5'-to-3' linear search is conducted. Our

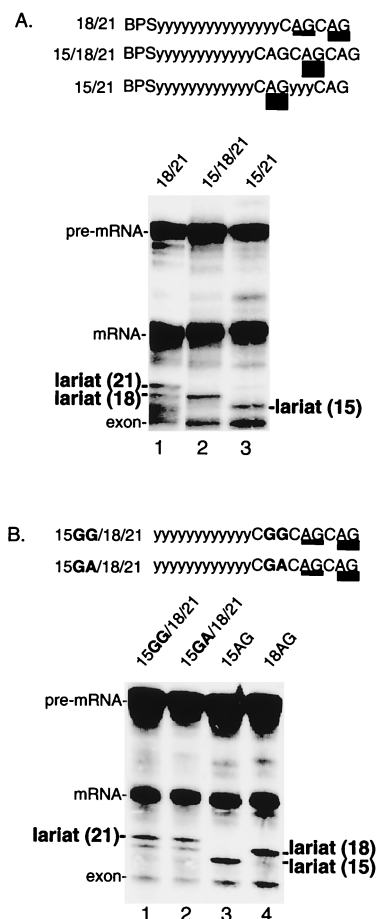


FIG. 4. A proximal AG determines whether a distal AG is selected. (A) An AG at 15 nt determines selection of distal AGs. Splicing reactions were performed using pre-mRNAs containing duplicate AGs at 18 and 21 nt from the BPS (lane 1) or containing a third AG inserted at 15 nt from the BPS (lane 2). Lane 3 shows splicing of pre-mRNA 15/21, which serves as a marker for the lariat. (B) Other purine dinucleotides at 15 nt do not affect selection of distal AGs. Pre-mRNAs are identical to that shown in panel A (lane 2) except that GG (lane 1) or GA (lane 2) replaces the AG at 15 nt. Pre-mRNAs 15AG and 18AG are markers for the lariats (lanes 3 and 4).

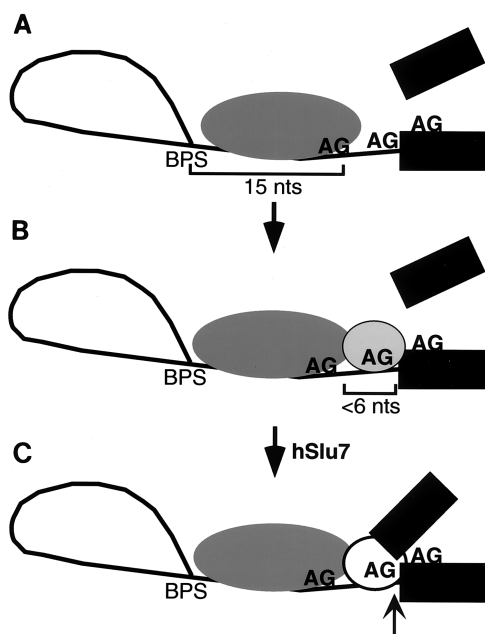


FIG. 5. Model for AG selection. (A) The proximal AG positions the splicing machinery downstream of the BPS. Recognition of this positioning AG is flexible, as this AG can be located outside the optimal distance for step II (in this example, 15 nt). (B) The AG used in the transesterification reaction is selected in a second step. This transesterification AG must be located within a narrow distance (<6 nt) downstream of the positioning AG. If a second AG is not present within this window, the positioning AG itself is used (but inefficiently). In the absence of hSlu7, the transesterification AG is sequestered, and step II is blocked at this location. (C) In the presence of hSlu7, the sequestered transesterification AG is exposed, and catalysis of step II (arrow) occurs.

analysis does not support a simple scanning model. We show that AGs located closer to the BPS than 19 to 23 nt are not efficiently used, most likely due to steric hindrance from a factor bound at the BPS. Yet, the efficiency of step II also begins to decrease when AGs are located immediately beyond this distance. Thus, assuming that scanning does not become inefficient at the exact point at which it is expected to start (i.e., where steric hindrance is relieved), the AG is not likely to be selected by simple scanning. Our data do not address the previously suggested possibility that scanning is involved in selection of AGs when they are located farther than the optimal distance from the BPS (4, 9, 15, 21, 22). A simple BPS-to-AG measuring mechanism also does not explain our data because this model predicts that step II should always occur at an optimally located AG if one is present. However, our data show that AGs optimally located 19 to 23 nt from the BPS are not always selected if an upstream AG is present (see below).

An assay using duplicate AGs uncouples multiple AG recognition steps. We have used pre-mRNAs containing duplicated AGs as an assay to gain insight into the mechanism of AG selection. Our results show that an AG located closer than the optimal BPS-to-AG distance (19 to 23 nt) is recognized even when it is not used. Indeed, the proximal AG determines whether a distal AG will be selected. Thus, these data reveal two distinct AG recognition events. One possible model to explain our observations, shown in Fig. 5, is that the initial

recognition of the proximal AG functions to position the splicing machinery on the RNA downstream of the BPS. The recognition of this positioning AG is relatively flexible, as it can occur when an AG is suboptimally located, only 15 nt from the BPS (Fig. 5A). After this initial positioning of the splicing machinery, selection of the AG that actually functions in the transesterification reaction (the transesterification AG) occurs (Fig. 5B). Our data show that selection of the transesterification AG is more stringent than recognition of the positioning AG. Specifically, the transesterification AG must be located within a narrow range of distance (<6 nt) downstream of the positioning AG. Such a multistep mechanism, which could involve a flexible ruler followed by a stringent ruler, may accommodate the variability in the location of 3' splice sites while still ensuring high fidelity.

In naturally occurring pre-mRNAs, closely clustered AGs are not always present at the 3' splice site. Nevertheless, our assay using the duplicated AGs uncouples two distinct recognition events that most likely occur when only a single AG is present. It is possible that the two events involve more than one factor or a single factor that undergoes a conformational change. The only proteins known so far to affect AG selection during step II are Slu7 and hSlu7 (in yeast and humans, respectively) (6, 10). We analyzed AG selection in extracts lacking hSlu7 using all of the pre-mRNAs containing the duplicated AGs (6; our unpublished observations). In all cases, the AG that was used in normal extracts was specifically not used in the absence of hSlu7 (whereas incorrect AGs were used). Thus, the hSlu7 study shows that the transesterification AG is selected but is specifically sequestered prior to hSlu7 function. hSlu7 is then required for use of this AG in catalytic step II (Fig. 5C).

One candidate for a factor involved in recognition of the positioning AG is U2AF35, which recognizes the AG early in metazoan spliceosome assembly (17, 27, 30). However, no changes in AG selection were observed in extracts lacking U2AF35, arguing against a role for this protein (K. Chua, R. Reed, and R. Gaur, unpublished data). Previously, it was shown that recognition of a downstream AG was necessary for step I but that a proximal AG was used for step II (14, 29). In this case, it is likely that the first recognition event involved U2AF35. In contrast to these studies, in our constructs the duplicated AGs both have effects on step II.

A candidate factor for recognition of the positioning and/or the transesterification AG is the highly conserved U5 snRNP protein Prp8. In yeast, Prp8 can be UV cross-linked to the YAG, and Prp8 mutants that suppress step II defects caused by mutations in the YAG (7, 20, 23–25, 28) have been identified. These and other observations have led to the proposal that Prp8 may play a role in positioning the AG for step II (7, 20). This possibility is compatible with the observation that hPrp8 can be cross-linked on both sides of the AG at or near the time of step II in metazoans (4). Slu7 is also synthetically lethal with Prp8 in yeast (11). Thus, it will be interesting to determine how these factors function in the different events involved in AG selection.

ACKNOWLEDGMENTS

We are grateful to Z. Zhou and O. Gozani for critical comments on the manuscript.

This work was supported by NIH grant GM 43375-11 to R.R.

REFERENCES

1. Anderson, K., and M. J. Moore. 1997. Bimolecular exon ligation by the human spliceosome. *Science* **276**:1712–1716.
2. Burge, C. B., T. H. Tuschl, and P. A. Sharp. 1999. Splicing of precursors to mRNA by the spliceosomes, p. 525–560. *In* R. F. Gesteland, T. R. Cech, and J. F. Atkins (ed.), *The RNA world*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
3. Chen, S., K. Anderson, and M. J. Moore. 2000. Evidence for a linear search in bimolecular 3' splice site AG selection. *Proc. Natl. Acad. Sci. USA* **97**:593–598.
4. Chiara, M. D., O. Gozani, M. Bennett, P. Champion-Arnaud, L. Palandjian, and R. Reed. 1996. Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. *Mol. Cell. Biol.* **16**:3317–3326.
5. Chiara, M. D., L. Palandjian, R. Feld Kramer, and R. Reed. 1997. Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *EMBO J.* **16**:4746–4759.
6. Chua, K., and R. Reed. 1999. The RNA splicing factor hSlu7 is required for correct 3' splice site choice. *Nature* **402**:207–210.
7. Collins, C. A., and C. Guthrie. 1999. Allele-specific genetic interactions between Prp8 and RBA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes Dev.* **13**:1970–1982.
8. Deshler, J. O., and J. J. Rossi. 1991. Unexpected point mutations activate cryptic 3' splice sites by perturbing a natural secondary structure within a yeast intron. *Genes Dev.* **5**:1252–1263.
9. Dix, I., C. S. Russell, R. T. O'Keefe, A. J. Newman, and J. D. Beggs. 1998. Protein-RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae* RNA **4**:1239–1250. (Author's correction, **4**:1675–1686.)
10. Frank, D., and C. Guthrie. 1992. An essential splicing factor, SLU7, mediates 3' splice site choice in yeast. *Genes Dev.* **6**:2112–2124.
11. Frank, D., B. Patterson, and C. Guthrie. 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.
12. Gozani, O., J. G. Patton, and R. Reed. 1994. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. *EMBO J.* **13**:3356–3367.
13. Halfter, H., and D. Gallwitz. 1988. Impairment of yeast pre-mRNA splicing by potential secondary structure-forming sequences near the conserved branchpoint sequence. *Nucleic Acids Res.* **16**:10413–10423.
14. Krainer, A. R., R. Reed, and T. Maniatis. 1985. Mechanisms of human beta-globin pre-mRNA splicing, p. 353–382. *In* The Robert A. Welch Foundation Conferences on Chemical Research XXIX. *In* Genetic chemistry: the molecular basis of heredity. Welch Foundation, Houston, Tex.
15. Liu, Z. R., B. Lagerbauer, R. Luhrmann, and C. W. Smith. 1997. Crosslinking of the U5 snRNP-specific 116-kDa protein to RNA hairpins that block step 2 of splicing. *RNA* **3**:1207–1219.
16. Luukkonen, B. G., and B. Seraphin. 1997. The role of branchpoint-3' splice site spacing and interaction between intron terminal nucleotides in 3' splice site selection in *Saccharomyces cerevisiae*. *EMBO J.* **16**:779–792.
17. Merendino, L., S. Guth, D. Bilbao, C. Martinez, and J. Valcarcel. 1999. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **402**:838–841.
18. Moore, M. J., C. C. Query, and P. A. Sharp. 1993. Splicing of precursors to messenger RNAs by the spliceosome, p. 303–357. *In* R. F. Gesteland and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
19. Patterson, B., and C. Guthrie. 1991. A U-rich tract enhances usage of an alternative 3' splice site in yeast. *Cell* **64**:181–187.
20. Siatecka, M., J. L. Reyes, and M. M. Konarska. 1999. Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. *Genes Dev.* **13**:1983–1993.
21. Smith, C. W., E. B. Porro, J. G. Patton, and B. Nadal-Ginard. 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature* **342**:243–247.
22. Smith, C. W. J., T. T. Chu, and B. Nadal-Ginard. 1993. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Mol. Cell. Biol.* **13**:4939–4952.
23. Teigelkamp, S., A. J. Newman, and J. D. Beggs. 1995. Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO J.* **14**:2602–2612.
24. Umen, J. G., and C. Guthrie. 1996. Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3' splice site selection. *Genetics* **143**:723–739.
25. Umen, J. G., and C. Guthrie. 1995. Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* **1**:584–597.
26. Umen, J. G., and C. Guthrie. 1995. The second catalytic step of pre-mRNA splicing. *RNA* **1**:869–885.
27. Wu, S., C. M. Romfo, T. W. Nilsen, and M. R. Green. 1999. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **402**:832–835.
28. Wyatt, J. R., E. J. Sontheimer, and J. A. Steitz. 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.
29. Zhuang, Y., and A. M. Weiner. 1990. The conserved dinucleotide AG of the 3' splice site may be recognized twice during in vitro splicing of mammalian mRNA precursors. *Gene* **90**:263–269.
30. Zorio, D. A., and T. Blumenthal. 1999. Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* **402**:835–838.