RNA (1996), 2:509–522. Cambridge University Press. Printed in the USA.

Intramolecular structure in yeast introns aids the early steps of in vitro spliceosome assembly

BRUNO CHARPENTIER and MICHAEL ROSBASH

Howard Hughes Medical Institute and Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA

ABSTRACT

rp51B pre-mRNA, like many *Saccharomyces cerevisiae* primary transcripts, contains a secondary structure within its intron sequence. The structure is required for optimal in vivo splicing efficiency and includes two complementary regions near the 5′ splice site and the branchpoint (UB1 and DB1, respectively). An intron-containing RNA was probed in vitro with RNase T1 and dimethyl sulfate (DMS), and is folded as expected. We have also examined in vitro splicing of rp51B pre-mRNA, by analyzing the formation of splicing complexes as well as splicing products. Similar analyses were done with mutant pre-mRNAs containing shortened or lengthened UB1/DB1 base pairing regions. Our experiments indicate that the secondary structure acts at an early step of spliceosome assembly to aid the formation of U1 snRNP-containing commitment complexes. pre-mRNAs were probed with DMS in vivo and the folding takes place inside cells. The effect of the different UB1/DB1 interactions on in vivo splicing efficiency was also analyzed. The results are consistent with the idea that the intramolecular interaction takes place prior to or at the beginning of spliceosome assembly.

Keywords: intron secondary structure; pre-mRNA splicing; *S. cerevisiae*

INTRODUCTION

Gene expression in eukaryotes requires the removal of intervening sequences (splicing of introns) from primary transcripts (pre-mRNA). Although most yeast (*Saccharomyces cerevisiae*) genes do not contain introns, splicing of a number of intron-containing pre-mRNAs has been characterized (Rymond & Rosbash, 1992). Yeast pre-mRNA splicing proceeds by the same basic pathway as in mammalian systems and can be divided into two steps. The first step involves nucleophilic attack at the 5′ splice junction of the intron by the 2′-hydroxyl of the branchpoint adenosine. The resultant products are the lariat intermediate and free 5′ exon. The second involves nucleophilic attack at the 3′ splice junction of the intron by the 3′-hydroxyl of the free 5′ exon. The final products are ligated exons and the released lariat-intron (for review see, Rymond & Rosbash, 1992; Moore et al., 1993).

During in vitro splicing, an ordered assembly process is observed during which snRNPs and proteins associate with the substrate to form a mature spliceosome within which the splicing reaction occurs (Moore et al., 1993). The four snRNPs that constitute the spliceosome contain five snRNAs (U1, U2, U5, and U4/U6 snRNAs), which are quite conserved between mammals and yeast (Guthrie, 1991; Steitz, 1992; Moore et al., 1993). The ordered addition of the snRNPs is also conserved. First, U1 snRNP is added, then U2 snRNP, and finally, the U4/U5/U6 tri-snRNP (Green, 1991; Rymond & Rosbash, 1992; Moore et al., 1993). In the yeast in vitro system, the U1 snRNP complex is called a commitment complex, because its formation appears to stably commit the pre-mRNA to the splicing pathway (Legrain et al., 1988; Séraphin & Rosbash, 1989). Formation of the commitment complex depends on RNA–RNA interactions, for example, base pairing takes place between the highly conserved 5′-end of U1 snRNA and the pre-mRNA 5′ splice site sequence GUAUGU (for review see, Rosbash & Séraphin, 1991). Protein–RNA and protein–protein interactions also play important roles, e.g., the MUD2 protein (Mud2p) associates with the pre-mRNA and U1 snRNP during commitment complex formation (Abovich et al., 1994). These experiments and others indicate that the branchpoint region as well as the 5′ splice site is important for early splicing complex formation.

In addition to the short, highly conserved splice site signals, there is evidence that other less-conserved in-

Reprint requests to: Michael Rosbash, Howard Hughes Medical Institute and Department of Biology, Brandeis University, Waltham, Massachusetts 02254, USA; e-mail: rosbash@binah.cc.brandeis.edu.

tron elements can also affect splicing efficiency (Pikielny & Rosbash, 1985; Newman, 1987). For example, there is an intramolecular interaction between segments located near the 5′ splice site and branchpoint regions of the *S. cerevisiae* rp51B intron that is required for efficient in vivo splicing (Goguel & Rosbash, 1993). (These two interacting segments have been called UB1 and DB1 [Libri et al., 1995].) Similar pairing interactions, in similar locations, can be drawn for many large yeast introns (Parker & Patterson, 1987; Rymond & Rosbash, 1992). A more thorough understanding of this interaction in the rp51B intron has been gained recently by using an in vivo randomization-selection approach, where it was shown that base pairing in any of several locations within this segment of the intron promotes splicing efficiency (Libri et al., 1995). Despite this progress, there is no report of the effects of the intramolecular pairing on yeast in vitro splicing, i.e., it is not known at which splicing step the pairing might function. The precise nature of the UB1/DB1 structure is also not known; it has not been examined directly in vitro or in vivo, but only inferred from its effect on splicing.

In this paper, we correct these two deficiencies. The UB1/DB1 paring is assayed directly, and we show that it acts at an early step of spliceosome assembly.

RESULTS

In vitro splicing of rp51B pre-mRNA

To study the effect of the intramolecular base pairing on in vitro splicing, we used a number of substrates (Fig. 1). The wild-type substrate (B) contains the complete 315-nt rp51B intron (plus an innocuous 10-nt sequence carrying a *Sal* I restriction site between positions 110 and 111 [Goguel & Rosbash, 1993]). This RNA and the natural rp51B intron has a noncanonical 5′ splice site (GUACGU). Because a mutation at position four to a canonical U (GUAUGU) has been shown to improve in vivo splicing efficiency (Goguel & Rosbash, 1993), a substrate with this optimal 5′ splice site was examined (B*; Fig. 1A). In addition, we generated a number of substrates that altered the UB1-DB1 pairing region. These include mutations that eliminate the pairing by inverting or modifying the UB1 or DB1 sequence so that no base pairing is possible (mut-UB1i and mut-DB1i, respectively), a mutation that restores strong pairing by inverting UB1 to be complementary to DB1i (mut-UB1iDB1i), a mutation that reduces the consecutive pairing region to 5 base pairs (mut-5), a single nucleotide change that improves pairing to 12 consecutive base pairs (mut-12), and a 3-nt change that extends pairing to 18 consecutive base pairs (mut-18). These mutations were placed in the contexts of the wild-type rp51B 5′ splice site (B) or of the improved rp51B 5′ splice site (B*) and are so designated (e.g., B-5, B*-5). To study their effects on in vitro splicing, the wild-type or mutant pre-mRNA substrates were added to standard yeast extract splicing reactions.

The results with the wild-type rp51B pre-mRNA substrate are shown in Figure 2A (lane 2). No detectable splicing intermediates or products were detected. pre-mRNA from the efficiently spliced rp51A gene, the other gene encoding ribosomal protein 51 with a 398-nt intron, served as a positive control (Fig. 2A, lane 1). The relatively poor splicing of rp51B pre-mRNA parallels what has been shown previously in vivo, namely, rp51B splicing is significantly less efficient than rp51A splicing (Abovich, 1988; Goguel & Rosbash, 1993). Those in vivo experiments showed further that replacement of the noncanonical rp51B 5′ splice site (GUACGU) with the consensus sequence (GUAUGU), present in rp51A and most other pre-mRNAs, improved splicing efficiency (Goguel & Rosbash, 1993). This change (to rp51B*), also enhanced in vitro splicing (Fig. 2A, lane 4).

Effect of UB1-DB1 mutations on in vitro splicing efficiency

Interestingly, a more stable UB1-DB1 stem also improved in vitro splicing efficiency of wild-type B pre-mRNA with its noncanonical 5′ splice site (B-UB1iDB1i) (Fig. 2, lane 3); inversion of either UB1 or DB1 alone had no positive effect (Fig. 2A, lanes 5 and 6). To further analyze the relationship between intramolecular base pairing and in vitro splicing efficiency, additional mutants of the UB1-DB1 stem were tested in both 5′ splice site contexts (Fig. 2B). Both mut-12 and mut-18 improved substantially splicing efficiency in the B context (compare lane 7 with lanes 8 and 9). Mut-12 was especially informative, because it involves only a single base pair change from the wild-type sequence. In the B* context, a positive effect of the mut-12 and mut-18 mutations was not obvious (compare lane 10 with lanes 13 and 14). However, we observed a prominent decrease in splicing efficiency with a reduction in pairing potential (B*-DB1i, B*-5). This suggests that, in this natural long-intron context, wild-type 5′ splice site and branchpoint consensus sequences are not sufficient for optimal in vitro splicing efficiency. The pairing effects are consistent with the previous in vivo results (Goguel & Rosbash, 1993). Because they impact both lariat species, pairing contributes to splicing prior to the first catalytic step, 5′ splice site cleavage and lariat formation.

The stem acts at an early step in spliceosome formation

Splicing in vitro is preceded by formation of spliceosomal complexes, which can be assayed by native gel analysis (Konarska & Sharp, 1986; Pikielny et al., 1986; Séraphin & Rosbash, 1991). In yeast extracts, splice-

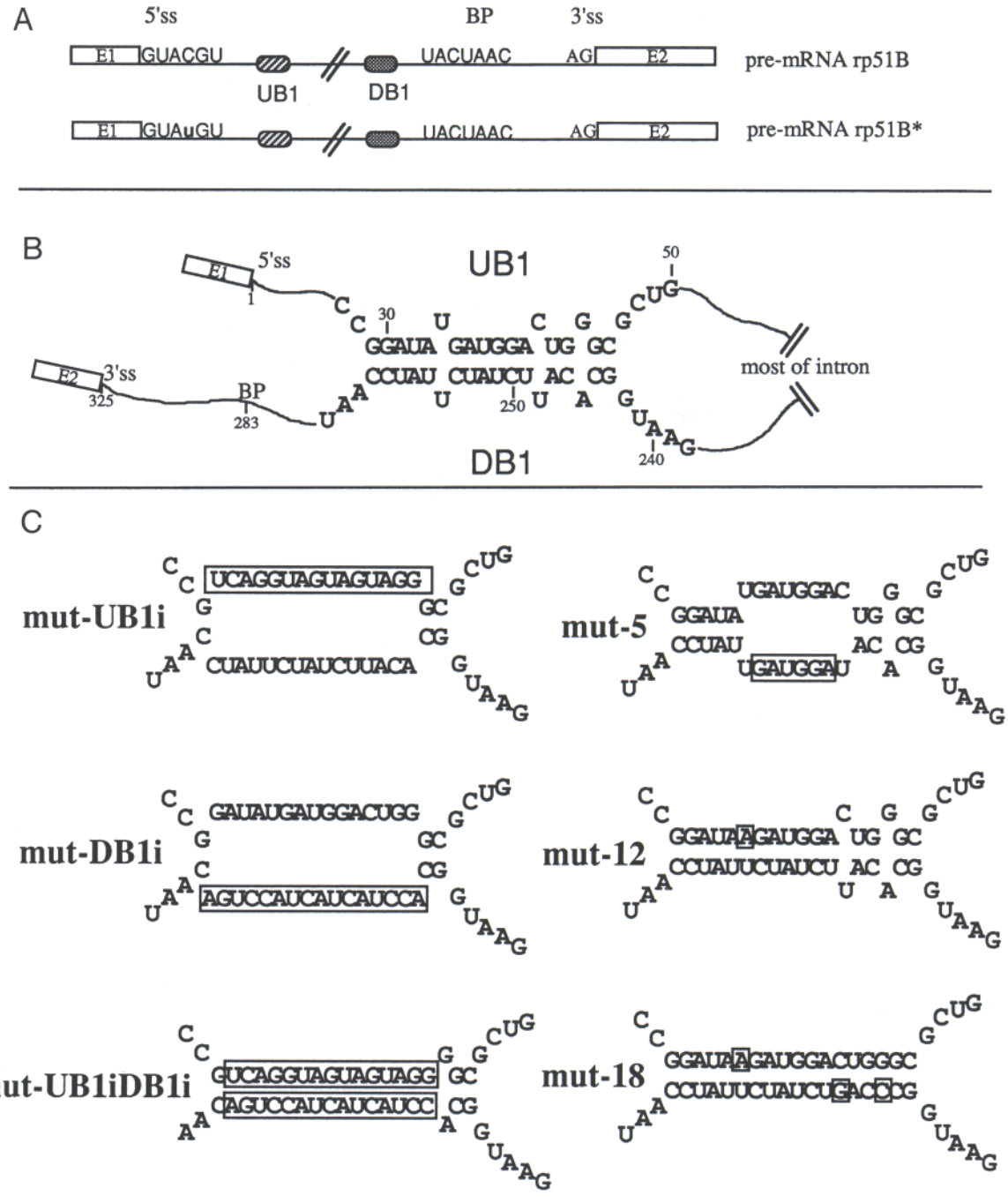

FIGURE 1. **A:** Schematic representation of the rp51B pre-mRNA used in this study. The intron is represented by a thin line and the exons (E1 and E2) by open boxes. The conserved *cis*-elements that define an intron are indicated: 5′ splice site (5′ss), branchpoint (BP), and 3′ splice site (3′ss). The two regions of complementary, UB1 and DB1, are indicated. The uracil, introduced by directed mutagenesis in place of a cytosine, is indicated in lower case in the rp51B* pre-mRNA. **B:** Representation of the putative interaction between UB1 and DB1 regions. Numbering starts from the first nucleotide of the intron (nt 1). The sequence reported here is the correct genomic sequence of rp51B (based on sequencing of many clones) and is different from that studied in Libri et al. (1995). The A between positions 33 and 34 is not present in the correct sequence. Its presence promoted a bulge (U34) outside of the stem rather than a mismatch between U36 and U255. **C:** Nucleotide sequences and putative secondary structures of the different mutants used in this study. The nucleotides introduced by directed-mutagenesis (see the Materials and methods) are boxed.

osome formation (operationally defined here as U2 snRNP-containing splicing complexes) is preceded by commitment complex formation (U1 snRNP binding). We compared the different pre-mRNA substrates for the formation of commitment complexes and spliceosomes by analyzing complex formation with and without oligonucleotide-mediated RNase H digestion of U2 snRNA, respectively.

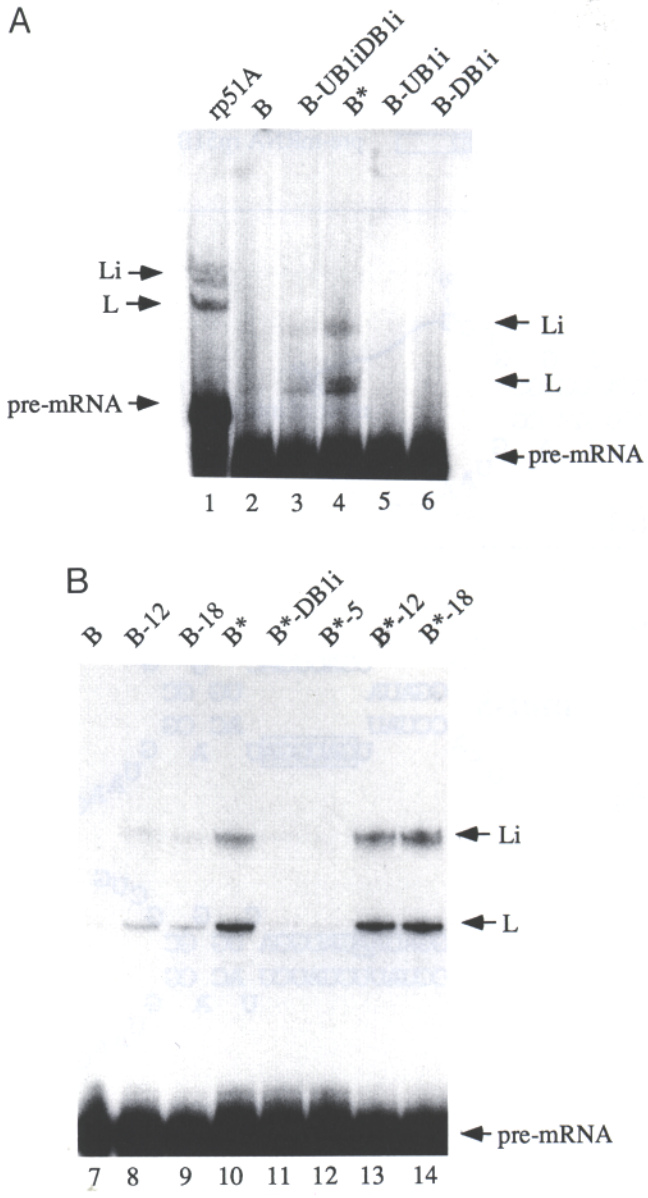

FIGURE 2. A,B: In vitro splicing of the different substrates. ^{32}P-labeled pre-mRNAs used as substrates are indicated on top of each lane. They were incubated under standard conditions (see the Materials and methods) for 30 min at 25 °C with a yeast extract. RNA products were analyzed on a 5% denaturing polyacrylamide gel. Arrows indicate positions of pre-mRNA, lariat intermediate (Li), and lariat product (L) species.

We were unable to detect significant amounts of either complex with any of the B substrates (Fig. 3A, lane 1 and data not shown), which correlates with their undetectable or low levels of splicing products. With the B* substrate, however, spliceosomes were observed (Fig. 3A, lane 2). In this optimal 5′ss context, inversion of the UB1 or DB1 sequences eliminates apparent spliceosome formation (Fig. 3A, lanes 3 and 4), but the pairing-positive combination (B*-UB1iDB1i) (lane 5 in Fig. 3A) restores complex formation. Consistent with the splicing assay (Fig. 2), less complexes were observed with less pairing (the B*-5 substrate; Fig. 3B), and comparable or even larger amounts of spliceosomes observed with increased pairing (the B*-12 and B*-18 substrates). Parallel results were obtained for B* commitment complex formation: less complexes with the B*-5, B*-UB1i, B*-DB1i substrates and restoration of commitment complex formation with increased pairing (the B*-UB1iDB1i, B*-12, and B*-18 substrates; Fig. 3C,D). The results suggest that the differences in splicing efficiency are caused by differences in complex formation at or before commitment complex formation.

To verify the pairing effect on commitment complex formation, we also compared three of the substrates with an indirect competition assay (Séraphin & Rosbash, 1991) (Fig. 4); we measured commitment complex formation with radioactive B*-18 pre-mRNA in the presence of increasing amounts of nonradioactive substrates, either B*-18, B*-5, or B*. There was no striking difference in this assay between B* and the increased pairing of B*-18. But the B*-5 substrate was a worse competitor, consistent with the notion that the UB1-DB1 pairing makes a contribution to commitment complex formation.

In vitro structure probing

An effect on early complex formation suggests that the UB1-DB1 pairing might affect the overall folding of the intron prior to factor recognition. If so, the in vitro secondary structure of the pairing region might be more or less as drawn in Figure 1B. To test this prediction, we subjected three of the substrates to the single-strand probing reagents RNase T1 (Fig. 5A,B) and dimethyl sulfate (DMS) (Fig. 5C,D). The substrates (B*, B*-5, and B*-18) were incubated in the same buffer and conditions that allow the formation of commitment complexes and spliceosomes in extract. The sites of RNase hydrolysis or DMS methylation were analyzed by primer extension with reverse transcriptase, as described in the Materials and methods.

The results indicate that a UB1-DB1 pairing is indeed formed in solution and that the structure of this region is close to the one drawn (Figs. 1B, 5F). First, only in case of the B*-18 substrate, with its predicted more stable stem, do strong cDNA pauses occur just prior to or within the pairing region (arrows in Fig. 5A,B,C,D). These cDNA stops are a low percentage of the total extension products (data not shown; compare band intensity above UB1 or DB1 in Fig. 5A,B,C,D); therefore, B*-18 cDNA stops beyond UB1 or DB1 can be compared with the stops in the B* or B*-5 reactions. Second, the guanines, adenines, and cytosines of UB1 and DB1 are more protected than nucleotides outside these regions, indicating that they are engaged in interactions. Third, the UB1 region of B*-18 is completely protected compared to the same region in B*, which

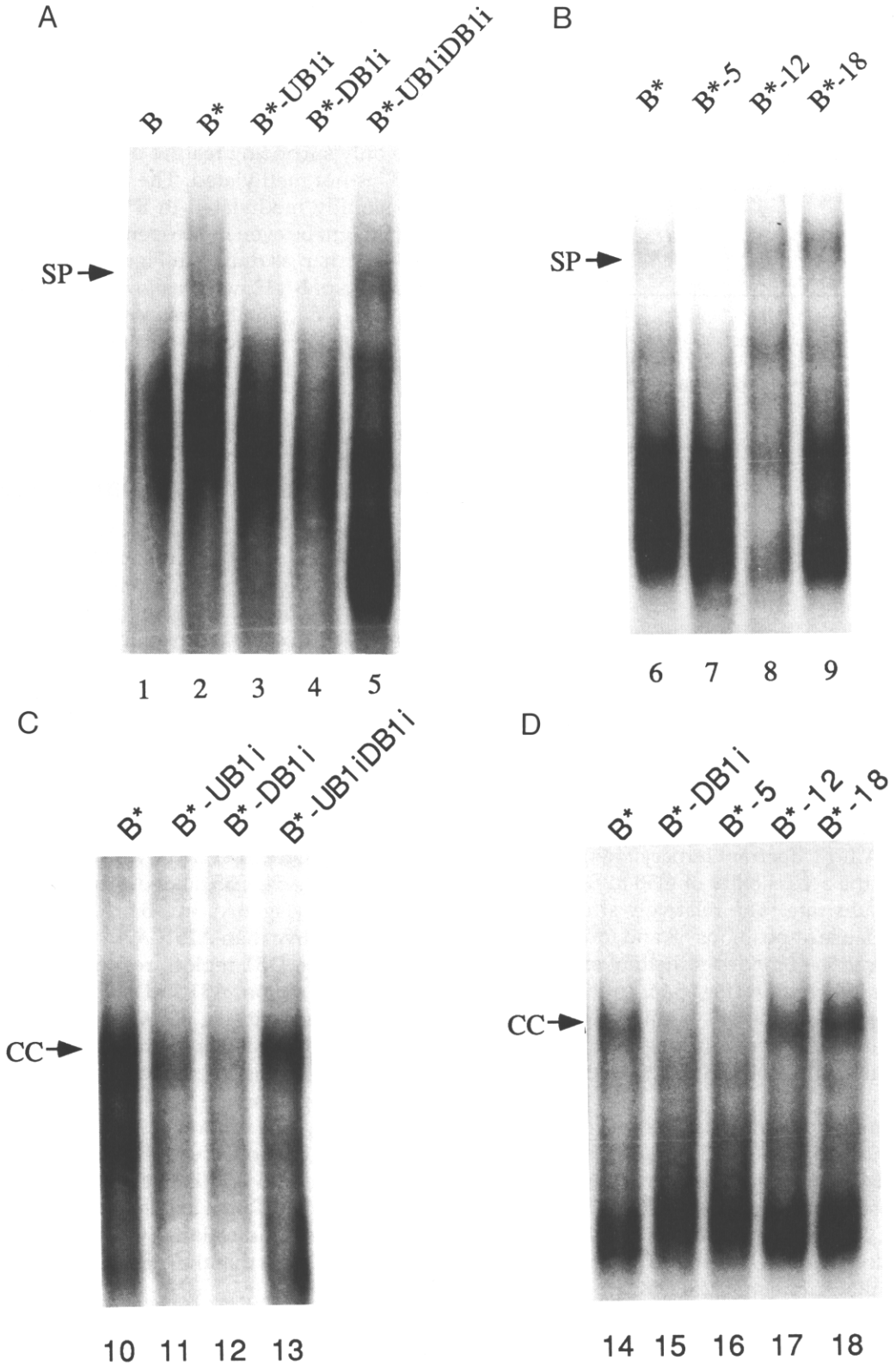

FIGURE 3. Effect of the mutations in the UB1/DB1 base pairing on the spliceosome assembly and commitment complex formation. The complexes formed with two different extracts were resolved by electrophoresis in nondenaturing conditions (Pikielny et al., 1986; Séraphin & Rosbash, 1991). The different substrates used are indicated on top of each lane. **A,B:** Complexes formed after incubation of the ^{32}P-labeled substrate with a complete extract. Arrow indicates the position of the spliceosome (SP). **C,D:** Complexes formed in the presence of an extract for which U2 snRNP was inactivated with oligonucleotide-directed RNaseH digestion (Legrain et al., 1988). Arrow indicates the position of the commitment complex (CC). In these conditions, nonspecific complexes are formed when B is used as the substrate. These complexes mask the visualization of the real commitment complex (data not shown).

TABLE 1. Effect of the UB1-DB1 mutations on the splicing efficiency in vivo.

	WT strain			MUD2 ko strain		
pre-mRNA	Copper[a] resistance	M[b]	P/M[c]	Copper[a] resistance	M[b]	P/M[c]
1 B-UB1i	0.8	ND	<0.2	ND		
2 B-DB1i	0.6/0.7	ND	<0.1	ND		
3 B-5	0.5/0.6	0.2	8.4	<0.1	0.05	21
4 B	2.0	1.2	1	0.6/0.7	0.23	3.5
5 B-12	2.0/2.1	1.2	0.8	1	0.4	2.3
6 B-UB1iDB1i	2.0/2.1	ND	1.1/1.2	ND		
7 B-18	2.0/2.1	1.25	ND	1.2	0.47	ND
8 B*-UB1i	1.8	ND	<0.2	ND		
9 B*-DB1i	1.9/2.0	ND	0.2	ND		
10 B*-5	1.9	0.86	1.6	0.1/0.2	0.06	16.2
11 B*	2.1/2.2	1.1	0.7	1.2	0.44	2.3
12 B*-12	2.1/2.2	1.05	0.7	1.2/1.3	0.48	2
13 B*-UB1iDB1i	2.1/2.2	ND	1.2	ND		
14 B*-18	2.1/2.2	1.1	ND	1.2/1.3	0.47	ND
15 A	>2.2	ND		1.4/1.5	ND	

[a] Yeast Y59ΔCUP1Δrp51B::URA (WT strain lanes) and Y59ΔCUP1Δrp51B::URAΔMUD2::ADE (MUD2 ko strain lanes) were transformed with a plasmid containing a reporter cup gene (JH21 series) expressing a CUP1 protein that allows transformants to grow on copper-containing media in a dose-dependent manner (Hamer et al., 1985; Lesser & Guthrie, 1993). Each plasmid expressed pre-mRNAs rp51B or rp51B* carrying mutations described in Figure 1C or pre-mRNA rp51A. The efficiency of splicing was analyzed by growth on copper-containing plates (Stutz & Rosbash, 1994; Libri et al., 1995) as described in the Materials and methods. Values reported in this table obtained from at least two experiments, represent the maximum copper concentration (mM) at which the different transformants can grow.

[b] Amounts of messenger RNA (M) encoding the CUP1 gene were quantified from a reverse transcriptase gel. Values represent the relative amount of messenger to the U1 snRNA.

[c] Ratio between the pre-messenger RNA (P) and the messenger RNA (M). Primer extension products were quantified and normalized to the U1 snRNA internal control. Values represent an average of the P/M ratios obtained from two independent experiments. The B ratio was set arbitrarily to 1 and the others were expressed accordingly.

taining media and is a sensitive indicator of splicing efficiency (Hamer et al., 1985; Lesser & Guthrie, 1993; Stutz & Rosbash, 1994; Libri et al., 1995). The cells transformed with a plasmid expressing the rp51A intron grow at a higher copper concentration than those expressing rp51B or rp51B* introns, consistent with expectations (Goguel & Rosbash, 1993), and confirming our in vitro splicing reactions (Fig. 2A,B). In a wild-type strain background, the pairing mutations that decrease the stability of the stem have a significant effect on copper resistance of B-containing strains, i.e., mut-UB1i, mut-DB1i, and mut-5 substrates (B-UB1i, B-DB1i, and B-5) grow less well than the controls. In contrast, eliminating or reducing the pairing in the B* splice site context had little or no effect (B*-UB1i, B*-DB1i, and B*-5, respectively). The concentrations of copper at which each strain grows correlate well with relative levels of mRNA; these are reduced by a factor of 6 in B-5 compared to B, but almost the same in B*-5 compared to B. The ratios of pre-mRNA and mRNA give an independent estimate of splicing efficiency (Pikielny & Rosbash, 1985), and they confirm the copper data (Table 1). All these in vivo observations parallel those made in vitro.

Because the inessential MUD2 gene affects commitment complex formation (Abovich et al., 1994), and because our in vitro experiments indicated an early role of the pairing in a spliceosome assembly, we suspected that use of a MUD2-deletion background (ΔMUD2) might be informative. Indeed, the copper resistance of B and B* was reduced, and the pairing mutant effects were more pronounced in this strain. Eliminating or reducing the pairing in the context of the B 5′ splice site (B-UB1i, B-DB1i, or B-5) decreased the copper resistance in the ΔMUD2 strain much more dramatically than what was observed in a wild-type background (Table 1, compare rows 1, 2, 3, and 4 of the ΔMUD2 strain with the WT strain). Significantly, strengthening the pairing (B-UB1iDB1i, B-12, and B-18) increased the copper resistance relative to the wild-type (B) substrate in this strain. In the B* context, decreasing the pairing also had a pronounced effect, namely, the copper resistance of the B*-UB1i, B*-DB1i, and B*-5 constructs were substantially reduced as compared to the B* control. However, increasing the pairing had little effect; presumably the natural pairing in the context of a consensus 5′ splice site is sufficient for optimal efficiency. All these observations were confirmed by direct mRNA and pre-mRNA analyses.

DISCUSSION

The results of the experiments presented in this report provide new information concerning the role that in-

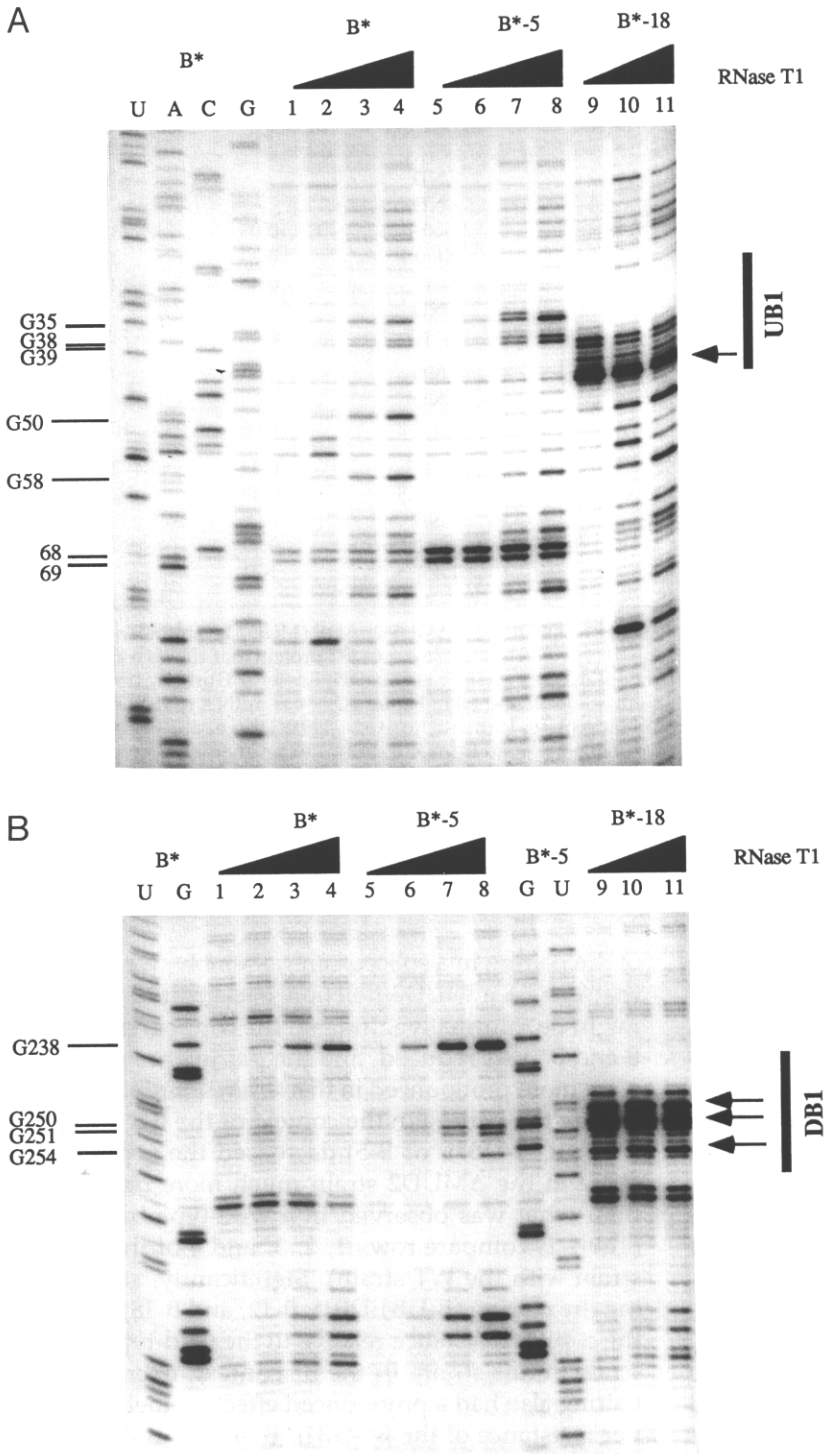

FIGURE 5. Structural analysis of the UB1/DB1 interaction. Autoradiograms of gels obtained in the nuclease (RNase T1) (**A,B**) and chemical (DMS) (**C,D**) probing experiments. Pre-mRNAs were incubated in the conditions described in the Materials and methods. The identity of the pre-mRNAs used in each experiment is indicated above each lane. The enzymatic hydrolysis was done with 0.1 (lanes 2, 6, 9), 0.25 (lanes 3, 7, 10), or 0.5 U (lanes 4, 8, 11) of RNase T1, which cleaves the RNA backbone 3′ to a single-stranded guanine. The DMS methylation (+) of adenine (N-1) and cytosine (N-3) was achieved with a 1:2.5 (v/v) dilution of the DMS stock solution in 95% ethanol. Positions of cleavages and methylations were identified by primer extension with reverse transcriptase, 5′-end labeled oligonucleotide DT2091 as the primer for UB1 region analysis, and 5′-end labeled oligonucleotide DT2770 for DB1 region analysis. Synthesized cDNAs were fractionated by electrophoresis on a 6% sequencing gel. The cDNAs fractionation of nontreated pre-mRNA (−) was used as a control (lanes 1, 5, a, c, e, g, i, m). No nuclease controls were also made for B*-18: cDNA fractionation profiles are comparable to lanes e and m. The A, C, G, U lanes refer to RNA sequence: cDNAs were generated by dideoxynucleotide sequencing of B* or B*-5 pre-mRNAs. Arrows along the right side of each gel indicate strong stops of the reverse transcriptase elongation when pre-mRNA B*-18 is used as the template (see details in text). Positions of sequences implicated in the UB1 or DB1 interactions are delimited along each gel. Positions of guanines, adenines, and cytosines very accessible to probes are indicated by a long dash. Position of adenines poorly accessible are indicated by a shorter dash. **E:** Densitograms of gels shown in A and C. Only profiles of the UB1 region of B* (grey line) and B*-5 pre-mRNAs (dark line), which has the same nucleotide sequence for the different mutants, is shown. Lanes 4, 8, b, and d were analyzed using the Molecular Analysis software. Peaks labeled with Stop represents a pause of the reverse transcriptase during cDNA synthesis. **F:** Accessibility of the probes to the UB1/DB1 interaction region. Reactivity of the reagents on the three pre-mRNAs analyzed are represented by symbols. Accessibility is based on visual estimation and profiles analysis of autoradiograph band intensities from at least two experiments. (*Continues on facing page.*)

tramolecular secondary structure plays in the splicing of the yeast rp51B pre-mRNA. This interaction, between complementary regions near the 5′ splice site and branchpoint regions, is required for optimal in vivo splicing (Goguel & Rosbash, 1993; Libri et al., 1995). Because several and perhaps most yeast genes with large introns contain these regions of potential complementarity at approximately the same location (Parker & Patterson, 1987; Rymond & Rosbash, 1992), the contribution to yeast splicing efficiency is probably of general significance. A reasonable hypothesis is that the structure might serve to reduce the effective distance between the two ends of the intron so that long yeast introns would functionally resemble short introns, gen-

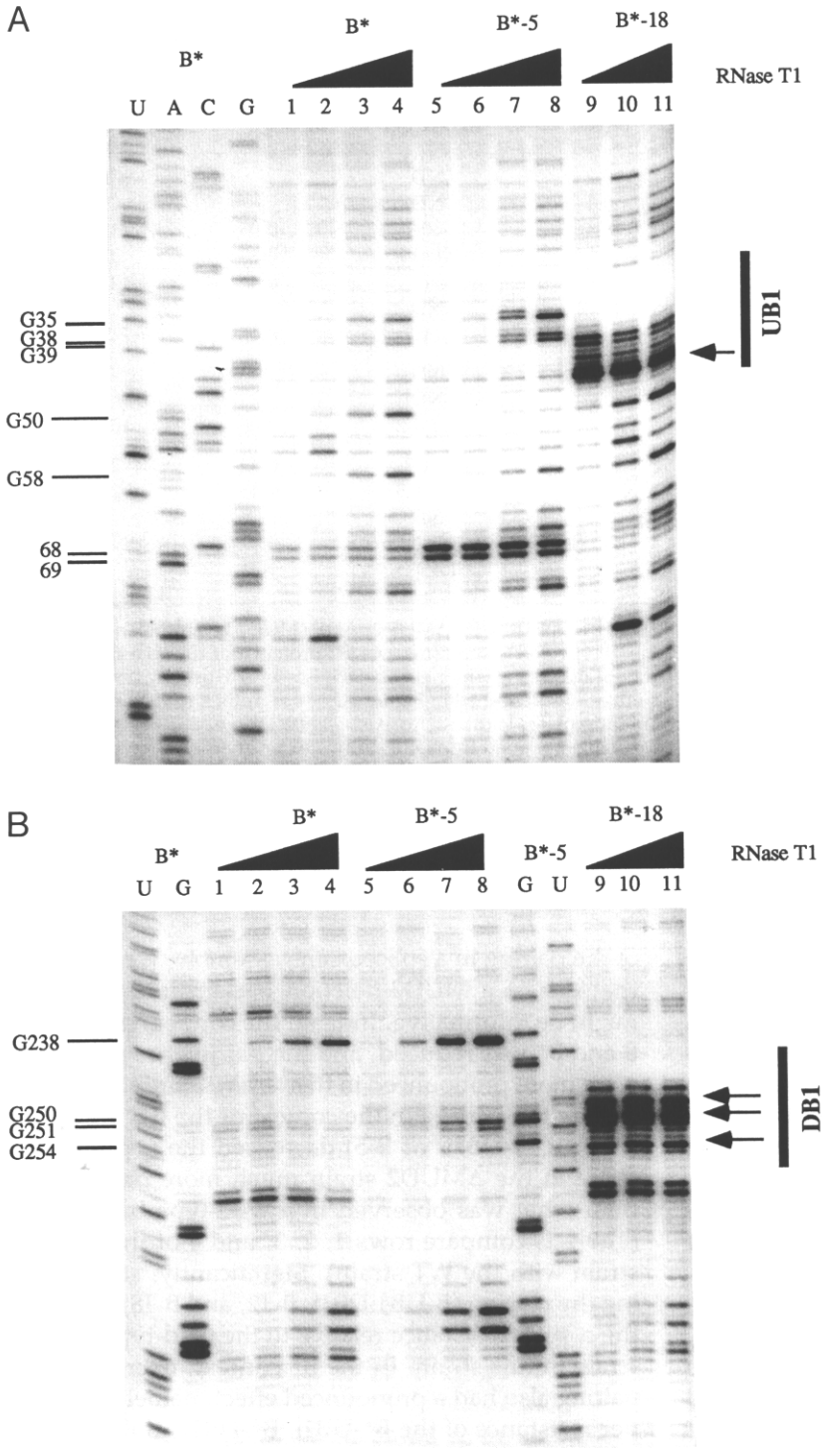

FIGURE 5. Structural analysis of the UB1/DB1 interaction. Autoradiograms of gels obtained in the nuclease (RNase T1) (**A,B**) and chemical (DMS) (**C,D**) probing experiments. Pre-mRNAs were incubated in the conditions described in the Materials and methods. The identity of the pre-mRNAs used in each experiment is indicated above each lane. The enzymatic hydrolysis was done with 0.1 (lanes 2, 6, 9), 0.25 (lanes 3, 7, 10), or 0.5 U (lanes 4, 8, 11) of RNase T1, which cleaves the RNA backbone 3′ to a single-stranded guanine. The DMS methylation (+) of adenine (N-1) and cytosine (N-3) was achieved with a 1:2.5 (v/v) dilution of the DMS stock solution in 95% ethanol. Positions of cleavages and methylations were identified by primer extension with reverse transcriptase, 5′-end labeled oligonucleotide DT2091 as the primer for UB1 region analysis, and 5′-end labeled oligonucleotide DT2770 for DB1 region analysis. Synthesized cDNAs were fractionated by electrophoresis on a 6% sequencing gel. The cDNAs fractionation of nontreated pre-mRNA (−) was used as a control (lanes 1, 5, a, c, e, g, i, m). No nuclease controls were also made for B*-18: cDNA fractionation profiles are comparable to lanes e and m. The A, C, G, U lanes refer to RNA sequence: cDNAs were generated by dideoxynucleotide sequencing of B* or B*-5 pre-mRNAs. Arrows along the right side of each gel indicate strong stops of the reverse transcriptase elongation when pre-mRNA B*-18 is used as the template (see details in text). Positions of sequences implicated in the UB1 or DB1 interactions are delimited along each gel. Positions of guanines, adenines, and cytosines very accessible to probes are indicated by a long dash. Position of adenines poorly accessible are indicated by a shorter dash. **E:** Densitograms of gels shown in A and C. Only profiles of the UB1 region of B* (grey line) and B*-5 pre-mRNAs (dark line), which has the same nucleotide sequence for the different mutants, is shown. Lanes 4, 8, b, and d were analyzed using the Molecular Analysis software. Peaks labeled with Stop represents a pause of the reverse transcriptase during cDNA synthesis. **F:** Accessibility of the probes to the UB1/DB1 interaction region. Reactivity of the reagents on the three pre-mRNAs analyzed are represented by symbols. Accessibility is based on visual estimation and profiles analysis of autoradiograph band intensities from at least two experiments. (*Continues on facing page.*)

tramolecular secondary structure plays in the splicing of the yeast rp51B pre-mRNA. This interaction, between complementary regions near the 5′ splice site and branchpoint regions, is required for optimal in vivo splicing (Goguel & Rosbash, 1993; Libri et al., 1995). Because several and perhaps most yeast genes with large introns contain these regions of potential complementarity at approximately the same location (Parker & Patterson, 1987; Rymond & Rosbash, 1992), the contribution to yeast splicing efficiency is probably of general significance. A reasonable hypothesis is that the structure might serve to reduce the effective distance between the two ends of the intron so that long yeast introns would functionally resemble short introns, gen-

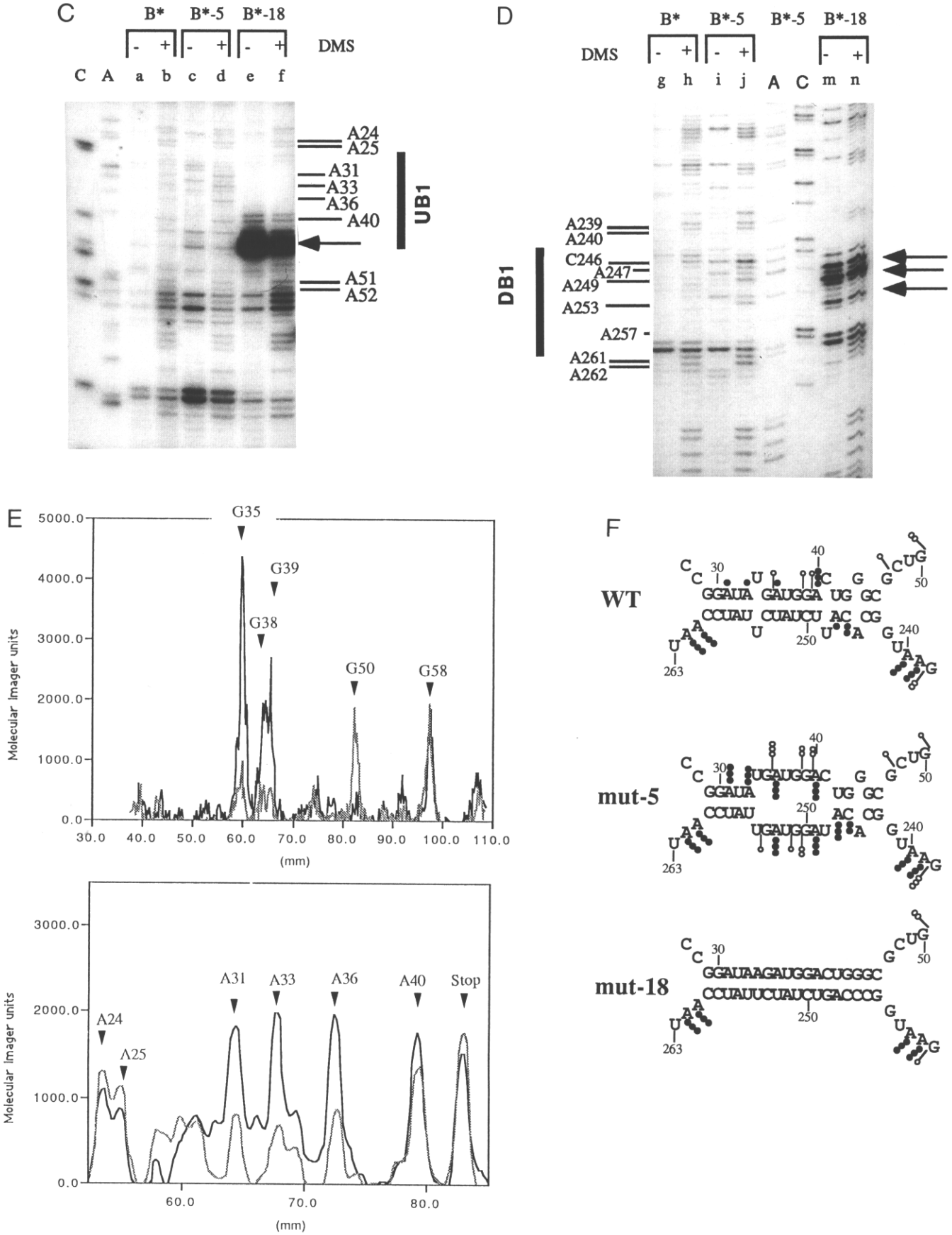

FIGURE 5. *Continued.*

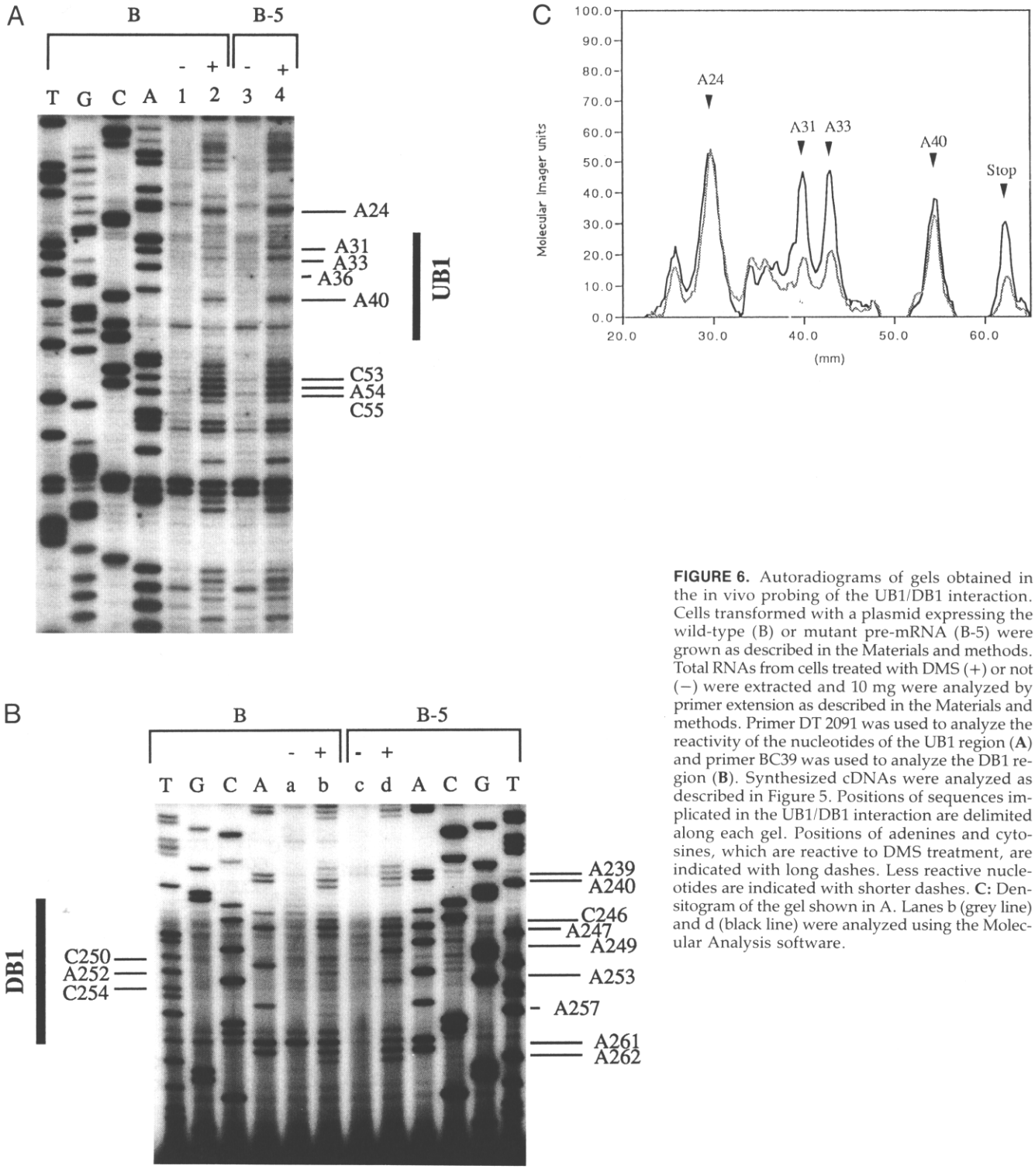

FIGURE 6. Autoradiograms of gels obtained in the in vivo probing of the UB1/DB1 interaction. Cells transformed with a plasmid expressing the wild-type (B) or mutant pre-mRNA (B-5) were grown as described in the Materials and methods. Total RNAs from cells treated with DMS (+) or not (−) were extracted and 10 mg were analyzed by primer extension as described in the Materials and methods. Primer DT 2091 was used to analyze the reactivity of the nucleotides of the UB1 region (**A**) and primer BC39 was used to analyze the DB1 region (**B**). Synthesized cDNAs were analyzed as described in Figure 5. Positions of sequences implicated in the UB1/DB1 interaction are delimited along each gel. Positions of adenines and cytosines, which are reactive to DMS treatment, are indicated with long dashes. Less reactive nucleotides are indicated with shorter dashes. **C:** Densitogram of the gel shown in A. Lanes b (grey line) and d (black line) were analyzed using the Molecular Analysis software.

erally less than 100 nt in length (Rymond & Rosbash, 1992). Short introns might experience significantly reduced folding difficulties, i.e., key *cis*-acting regions might be more accessible to *trans*-acting factors in short introns. Alternatively or in addition, these *cis*-acting regions, e.g., the 5′ splice site and branchpoint regions, may interact with each other more easily in short introns than in long introns.

The intramolecular pairing could aid at any of three different splicing steps. It could aid in the spliceosome assembly steps that precede the first catalytic step, it could facilitate the first catalytic step when the branch

site attacks the 5′ splice site, or it could contribute to the second catalytic step when the 5′ and 3′ exons undergo ligation. Our in vitro experiments suggest that the first possibility is correct, at least in part.

Importantly, splicing of the full-length large rp51B intron was very inefficient, i.e., it was usually impossible to detect products or intermediates of the splicing reaction (Fig. 2). Previous in vivo studies had also shown that the rp51B intron is spliced less efficiently than the rp51A intron (Goguel & Rosbash, 1993; Libri et al., 1995). This was due in part to the noncanonical 5′ splice site sequence GUACGU (Goguel & Rosbash, 1993), and replacing this sequence with the canonical 5′ splice site sequence improved in vitro splicing efficiency (Fig. 2A,B). Analysis of splicing complexes indicated that the differences between B and B* are likely due to differences in the formation of commitment complexes, i.e., proper and stable U1 snRNP complexes. The contribution of the canonical U at position 4 to the commitment complex formation is enigmatic because it is probably not base paired with the 5′ end of U1 snRNA (Rosbash & Séraphin, 1991; Moore et al., 1993).

Even more interesting was the qualitatively similar contribution made by the intramolecular pairing. At this time, it is not certain whether the observed effects on complex formation reflect kinetic or thermodynamic differences. However, we observed no obvious stability differences during chase incubation conditions between the commitment complexes containing rp51A, B*, B*-12, and B*-18 pre-mRNAs. Even the low level of complex formed with the B*-5 substrate was similarly stable (data not shown). Previous studies showed that the commitment complexes formed with the Δ2 substrate, the short efficiently spliced version of rp51A pre-mRNA, were also very stable (Séraphin & Rosbash, 1991). Consequently, we favor the notion that a stable stem contributes to complex formation rather than just complex stability.

In addition to U1 snRNP and the intramolecular base pairing, the MUD2 protein is a third identified contributor to commitment complex formation. Mud2p is a commitment complex component, and there is genetic as well as biochemical evidence that Mud2p interacts directly or indirectly with both U1 snRNP and the branchpoint region (Abovich et al., 1994). This adds to the substantial in vitro evidence that early steps in yeast splicing complex formation benefit from interactions between these two intronic regions (Séraphin & Rosbash, 1991). A similar role is proposed for the related mammalian splicing factor U2AF, which binds to the polypyrimidine tract and is a component of mammalian E complexes (Zamore & Green, 1989; Bennett et al., 1992; Michaud & Reed, 1993).

In vitro splicing with long yeast introns generates a single commitment complex (Fig. 3), which does not correspond to either of the two well-characterized complexes observed with short intron substrates. But a number of arguments indicate that it probably corresponds to the more complete Mud2p-containing complex (present data; H.V. Colot, unpubl. results). This view also supports the more optimistic interpretation of the in vivo splicing assays (Table 1). The UB1-DB1 stem and MUD2p contribute in synergistic fashion to commitment complex formation. However, the in vivo nature of the observations makes it possible that the effects are indirect, i.e., that the absence of Mud2p does not impact the same biochemical step as the UB1-DB1 pairing.

Our structure probing indicates that the stem forms in vitro. The folding as drawn in Figure 1 is almost certainly inexact, but the reactivity profiles suggest that most molecules are folded in something that resembles the UB1-DB1 interaction. In addition, mutations in the UB1-DB1 region affect the structure in the predicted fashion: mutations that open the structure increase reactivity, whereas stabilizing mutations have the opposite effect (Fig. 5F).

There are, however, a number of unanswered structural questions. First, what is the in vitro structure of the entire intron? Second, what is the structure in splicing extract, i.e., what is the effect of extract proteins? We have resisted addressing the first question without being able to answer the second. But we have experienced considerable difficulty in analyzing the UB1-DB1 structure in extract; the RNA ladders are much less sharp, due almost certainly to low but persistent levels of RNA degradation (data not shown). Examining the effect of extract proteins has an additional complication: the observed intron structure is derived from a mixture of at least two forms, commitment complex and RNP (i.e., RNA–protein complexes without U1 snRNP). Despite these difficulties, no prominent differences were observed between UB1-DB1 folding in buffer and in extract (data not shown).

Formation of the UB1-DB1 stem also occurs in vivo (Fig. 6). However, this structure probing is even more subject to the "mixture of forms" criticism than the in vitro experiments. For example, examination of the steady-state pre-mRNA population may focus on molecules that have escaped from the splicing pathway, molecules that may even be cytoplasmic (Legrain & Rosbash, 1989). Because spliceosome assembly probably occurs coincident with transcription (Wuarin & Schibler, 1994), it may be the structure of nascent pre-mRNA that is most relevant to the early steps in spliceosome formation. Although this kind of study has yet to be undertaken, the experiments presented here and elsewhere (e.g., Goguel & Rosbash, 1993; Libri et al., 1995) predict that pre-mRNA folding contributes to early steps of in vivo spliceosome assembly. Apparently, *S. cerevisiae* has only low levels of HnRNP-like proteins, as measured with a nucleic acid annealing activity assay (G. Dreyfuss, pers. comm.). Intramolecu-

lar pre-mRNA structure might play a more prominent role in the absence of extensive chaperone activity. Stated differently, mammalian HnRNP proteins may promote the pre-mRNA *trans*-acting factor interactions in a manner that replaces the function of yeast pre-mRNA structure.

In any case, we favor the notion that the correct structure maintains one or more key *cis*-acting regions of yeast pre-mRNA in a favorable configuration for interaction with the *trans*-acting factors that do the initial recognition of the pre-mRNA substrate. An alternative possibility is that the structure favors interactions between the two ends of the intron, i.e., the 5′ splice site and branchpoint regions. Because the latter occurs at very early steps of spliceosome assembly (Séraphin & Rosbash, 1991; Abovich et al., 1994), the distinction between these two possibilities may be more apparent than real.

MATERIALS AND METHODS

Strains

Escherichia coli strain TG1 was used for plasmid propagation, strain CJ236 (*dut*$^-$, *ung*$^-$, F′) was used for oligonucleotide-mediated mutagenesis. Yeast strain Y59ΔCUP1Δrp51B::URA (*MATa*, *leu2*-3, *leu2*-112, *ura3*-52, *trp1*-289, *arg4*, *ade2*, ΔCUP1) and Y59ΔCUP1Δrp51B::URAΔMUD2::ADE (D. Libri & N. Abovich, unpubl. results) were used for copper growth assays.

Site-directed mutagenesis and plasmid constructions

Plasmids were constructed according to standard procedures (Sambrook et al., 1989). pTZ19R recombinant vectors pTZ18, pTZB/B, pTZB*/B, pTZB/Bi, and pTZBi/Bi have been described previously (Goguel & Rosbash, 1993). Mutant pTZB*/Bi plasmid was obtained by cloning a *Sal* I–*Sac* I fragment from pTZB/Bi into pTZB*/B. pTZB/B and pTZB*/B were used as template for oligonucleotide single-strand mutagenesis (Kunkel et al., 1987). Each mutation on the DB1 region was done on pTZB/B, whereas mutations on UB1 region were done on pTZB/B as well as on pTZB*/B. The oligonucleotides used are listed in Table 2. pTZBmut-5 was obtained after directed mutagenesis with oligonucleotide DT2756. Mutant pTZB*mut-5 was obtained after cloning the pTZBmut-5 *Sal* I–*Sac* I fragment into pTZB*/B. pTZBmut-12 and pTZB*mut-12 were obtained directly by using oligonucleotide DT2757 in a directed mutagenesis reaction. pTZBmut-18 and pTZB*mut-18 were obtained after cloning the pTZB/B *Sal* I–*Sac* I fragment, which was mutagenized with oligonucleotide DT2758, into pTZBmut-12 and pTZB*mut-12, respectively. pTZB*iB and pTZB*iBi were obtained after cloning a *Pst* I–*Sal* I PCR fragment obtained from pTZBiBi with oligonucleotides BC44 (5′-ATGCCTGCAGGTCGAGAAAAAAGCAGATAAAAATGGTATGTACCACGAGAT-3′) and BC45 (5′-AGAATACCCTGCTGGTGGTCGACCGTCCCCTCGTTCCCA-3′) into pTZB/B and pTZB/Bi, respectively. Yeast vectors (JH21 series) were prepared according to (Libri et al., 1995). Briefly, the wild-type rp51A and rp51B and mutant intron sequences were cloned by PCR and inserted between the *Bam*H I and *Sma* I sites of a rp51B-CUP1 fusion construct (Libri et al., 1995).

Copper growth assay

All the DNA constructs (JH21 series) were transformed into the copper-sensitive strains Y59ΔCUP1Δrp51B::URA and Y59ΔCUP1Δrp51B::URAΔMUD2::ADE according to standard procedures (Ito et al., 1983). The transformants were grown to saturation and spotted on Leu$^-$ plates containing increasing concentrations of copper and were subsequently grown for 4 days at 30 °C.

In vitro complex assembly, native gel electrophoresis, and splicing reactions

Yeast extracts used in the different experiments were prepared as described previously (Liao et al., 1992). ^{32}P-labeled wild-type rp51A and rp51B, and mutant pre-mRNAs were prepared according to (Pikielny et al., 1986), except that PCR fragments were used for the in vitro T7 transcription reaction. The pTZ18, pTZB/B, pTZB*/B, and pTZmut series of plasmids were used as substrate for PCR reactions. Oligonucleotide DT2139 (5′-CTTTGGATAGTAGAATTCAATCAAAGCC-3′), which carries the T7 promoter of pTZ19R, was used as the upstream primer of the amplification reaction, whereas oligonucleotide DT2140 (5′-GCTCTAATACGACTCACTATAGGGAAAGCTTGC-3′), which binds the 5′-end region of exon 2 of rp51A and rp51B genes, was used as the downstream primer. Cold pre-mRNAs used for the competition assays were also synthesized by in vitro T7 transcription in conditions described by (Séraphin & Rosbash, 1991). In vitro assembly complexes analysis by native gels and splicing reactions were done as described (Séraphin & Rosbash, 1991; Liao et al., 1992). In the competition experiments, increasing amounts of cold pre-mRNA were mixed with the ^{32}P-labeled sub-

TABLE 2. Oligonucleotides used for site-directed mutagenesis of rp51B intron.

Name	Function	Sequence (5′-3′)[a]
DT 2756	Mutation in pTZB/B DB1	CTTTCGAATGGCACATAGGTAGTTATCCAATGGTCTTG
DT 2757	Mutation in pTZB/B and pTZB*/B UB1	AGCCGGATAAGATGGACTGG
DT 2758	Mutation in pTZB/B DB1	GCTTTCGAATGGCCCAGTCTATCTTATCCA

[a] Mutations introduced are underlined.

strate before the addition of the yeast extract, as described (Séraphin & Rosbash, 1991).

In vitro RNase and chemical probing and primer extension analysis of modified RNA

Prior to enzymatic cleavage with RNase T1 or chemical modification with DMS, cold pre-mRNA (100 ng), synthesized as described above (Séraphin & Rosbash, 1991), was incubated for 10 min at 25 °C in 20 μL (for RNase T1 hydrolysis) or 100 μL (for DMS methylation) of 2.5 mM $MgCl_2$, 3% PEG 8000, 60 mM potassium phosphate, pH 7, buffer. tRNA (10 μg) was added for the DMS modification. The RNA was digested by addition of 1 μL of RNASE T1 (0.1, 0.25, or 0.5 U/μL), followed by incubation for 5 min. The reaction was stopped by phenol extraction and ethanol precipitation. The RNA was methylated for 5 min by the addition of 1 μL of a 1:2.5 (v/v) dilution of DMS stock in 95% ethanol. The reaction was quenched by adding 50 μL of a stop solution (1 M Tris-HCl, pH 7.5, 2 M 2-mercaptoethanol, 0.1 mM EDTA) and the samples were placed on ice prior to ethanol precipitation. The RNA pellets were resuspended in 18 μL of sterile H_2O. Modified nucleotides and RNase cleavage sites were detected by primer extension using two different oligonucleotides as the primer, DT2091 (5′-TCATAAATAAGAATACCCTGCTGG-3′) or DT2770 (5′-GTCTTGGTTCTAACTCTACCC-3′), complementary to positions 134–111 in the intron and to the beginning of the 3′ exon, respectively. ^{32}P end-labeled primer (5 ng) was annealed to one-fourth (4.5 μL) of the modified RNA and extended by reverse transcriptase as described (Pikielny & Rosbash, 1985). The cDNAs were resolved in denaturing 6% sequencing gels. To prepare sequencing ladders of the RNA, 25 ng of untreated RNA was used as the template and dideoxynucleosides triphosphate were added (dNTP:ddNTP = 1/5) to the extension reaction.

In vivo DMS probing

The yeast strain Y59ΔCUP1Δrp51B::URAΔMUD2::ADE were transformed with plasmids JH21B or JH21B-5. Cells were grown overnight at 30 °C in Leu^- Ade^- medium containing glucose as the carbon source. The day after, the cells were harvested, washed in water, and resuspended in YM-1 medium containing glucose to an OD_{600} of 0.3. Then, the cells were grown 2 h before DMS treatment that was performed according to Zavanelli and Ares (1991) with some modifications. Ten milliliters of cells were incubated with or without 200 μL of a 1:4 (v/v) dilution of DMS in 95% ethanol. The reaction was quenched with 5 mL of 0.6 M of 2-mercaptoethanol and 5 mL of water-saturated isoamyl alcohol. Cells were pelleted and washed once with 5 mL of 0.6 M 2-mercaptoethanol. Total RNA was extracted as described previously (Pikielny & Rosbash, 1985). Ten micrograms of total RNA were analyzed by primer extension as described above using two oligonucleotides primers. The modifications in the UB1 region were analyzed with the same oligonucleotide (DT2091) that was used for the in vitro probing of UB1. The oligonucleotide used for the in vitro probing of the DB1 region could not be employed for that experiment: because it binds in the 3′ exon, it allows the analysis of pre-mRNAs as well as mRNAs. To avoid the background of the cDNAs produced from the mRNA template, we used oligonucleotide BC39 (5′-GCAATAATCAAATGAGACAAC-3′), which binds between the UACUAAC region and the 3′ splice site (positions 314–294 in the intron).

ACKNOWLEDGMENTS

We are grateful to members of the Rosbash laboratory for helpful suggestions during the course of these experiments, and to M. Moore, D. Libri, and H.V. Colot for critical comments on the manuscript. H.V. Colot is thanked for extract preparation. We thank Ed Dougherty for help in preparing the figures and L.A. Monaghan for secretarial assistance. This work was supported in part by a grant from National Institutes of Health to M.R. (GM23549).

Received September 8, 1995; returned for revision October 17, 1995; revised manuscript received April 29, 1996

REFERENCES

Abovich N. 1988. Regulation of yeast ribosomal protein gene expression [thesis]. Waltham, Massachusetts: Brandeis University.

Abovich N, Liao XC, 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.

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

Goguel V, Rosbash M. 1993. Splice site choice and splicing efficiency are positively influenced by intramolecular base-pairing in yeast. *Cell 72*:893–901.

Green MR. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu Rev Cell Biol 7*:559–599.

Guthrie C. 1991. Messenger RNA splicing in yeast: Clues to why the spliceosome is a ribonucleoprotein. *Science 253*:157–163.

Hamer DH, Thiele DJ, Lemontt JE. 1985. Function and autoregulation of yeast copperthionein. *Science 228*:685–690.

Ito H, Fukuda Y, Murata K, Kimura A. 1983. Transformation of intact yeast cells treated with alkali cations. *J Bacteriol 153*:163–168.

Konarska MM, Sharp PA. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell 46*:845–855.

Kunkel TA, Roberts JD, Zakour RA. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol 154*:367–382.

Legrain P, Rosbash M. 1989. Some *cis*- and *trans*-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell 57*:573–583.

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

Lesser CF, Guthrie C. 1993. Mutational analysis of pre-mRNA splicing in *Saccharomyces cerevisiae* using a sensitive new reporter gene, CUP1. *Genetics 133*(4):851–863.

Liao XC, Colot HV, Wang Y, Rosbash M. 1992. Requirements for U2 snRNP addition to yeast pre-mRNA. *Nucleic Acids Res 20*:4237–4245.

Libri D, Stutz F, McCarthy T, Rosbash M. 1995. RNA structural patterns and splicing: Molecular basis for an RNA-based enhancer. *RNA 1*:425–436.

Michaud S, 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 MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNAs by the spliceosome. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–357.

Newman A. 1987. Specific accessory sequences in *Saccharomyces cerevisiae* introns control assembly of pre-mRNAs into spliceosomes. *EMBO J 6*:3833–3839.

Parker R, Patterson B. 1987. Architecture of fungal introns: Implications for spliceosome assembly. In: Dudock B, Inouye M, eds. *Molecular biology of RNA: New perspectives*. New York: Academic Press. pp 133–149.

Pikielny CW, Rosbash M. 1985. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell 41*:119–126.

Pikielny CW, Rymond BC, Rosbash M. 1986. Electrophoresis of ribonucleoproteins reveals an ordered pathway of yeast splicing complexes. *Nature 324*:341–345.

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

Rymond BC, Rosbash M. 1992. Yeast pre-mRNA splicing. In: Jones EW, Pringle JR, Broach JR, eds. *The molecular and cellular biology of the yeast Saccharomyces: Gene expression*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 143–192.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning, a laboratory manual, 2nd ed*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

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

Séraphin B, Rosbash M. 1991. The yeast branchpoint sequence is not required for the formation of a stable U1 snRNP–pre-mRNA complex and is recognized in the absence of U2 snRNA. *EMBO J 10*: 1209–1216.

Steitz JA. 1992. Splicing takes a Holliday. *Science 257*:888–889.

Stutz F, Rosbash M. 1994. A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *EMBO J 13*:4096–4104.

Wuarin J, Schibler U. 1994. Physical isolation of nascent RNA chains transcribed by RNA polymerase II: Evidence for cotranscriptional splicing. *Mol Cell Biol 14*:7219–7225.

Zamore PD, Green MR. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc Natl Acad Sci USA 86*:9243–9247.

Zavanelli MI, Ares MJ. 1991. Efficient association of U2 snRNPs with premessenger RNA requires an essential U2 RNA structural element. *Genes & Dev 5*:2521–2533.