

Short Artificial Hairpins Sequester Splicing Signals and Inhibit Yeast Pre-mRNA Splicing

VALÉRIE GOGUEL,^{1†} YUE WANG,² AND MICHAEL ROSBASH^{1*}

Howard Hughes Medical Institute, Department of Biology,¹ and Department of Biochemistry,²
Brandeis University, Waltham, Massachusetts 02254

Received 10 May 1993/Returned for modification 21 June 1993/Accepted 23 August 1993

To examine the stability of yeast (*Saccharomyces cerevisiae*) pre-mRNA structures, we inserted a series of small sequence elements that generated potential RNA hairpins at the 5' splice site and branch point regions. We analyzed spliceosome assembly and splicing in vitro as well as splicing and nuclear pre-mRNA retention in vivo. Surprisingly, the inhibition of in vivo splicing approximately paralleled that of in vitro splicing. Even a 6-nucleotide hairpin could be shown to inhibit splicing, and a 15-nucleotide hairpin gave rise to almost complete inhibition. The in vitro results indicate that hairpins that sequester the 5' splice site have a major effect on the early steps of spliceosome assembly, including U1 small nuclear ribonucleoprotein binding. The in vivo experiments lead to comparable conclusions as the sequestering hairpins apparently result in the transport of pre-mRNA to the cytoplasm. The observations are compared with previous data from both yeast and mammalian systems and suggest an important effect of pre-mRNA structure on in vivo splicing.

Pre-mRNA splicing is a ubiquitous feature of gene expression in eukaryotes. The pre-mRNA *cis*-acting sequence elements (i.e., the 5' splice site, branch point, and 3' splice site regions) are initially recognized by *trans*-acting factors. These include small nuclear ribonucleoproteins (snRNPs) and other protein factors that add to the pre-mRNA in an ordered pathway to form the spliceosome, within which the two steps of the splicing reaction take place. U1, U2, U5, and U6 small nuclear RNAs are known to interact with the pre-mRNA directly; it is generally thought that nuclear pre-mRNA splicing, like group I and group II self-splicing, is likely to be an RNA-catalyzed process (2, 30, 41).

In higher eukaryotes, splice site sequences are not well conserved, and they are probably insufficient to define splice site partner assignment perfectly. In addition, splice site context (adjacent sequence information) has been shown to have a major influence on splice site selection in vitro as well as in vivo (1, 13, 21, 23, 28, 31, 34). Several studies have shown that long-range context effects also influence splice site choice (14, 44).

In the yeast *Saccharomyces cerevisiae*, nuclear introns are relatively rare (less than 5% of yeast genes have introns) and short (50 to 500 nucleotides). With one exception (29), there is only a single intron per gene. Yeast 5' splice site and branch point consensus sequences (GUAPyGU and UACU AAC, respectively) are also highly conserved. Despite these features that might seem adequate for the selection of splice sites, there is evidence that context effects (i.e., nonconserved sequences) also influence splicing efficiency and splice site selection in the yeast system (16, 32, 33).

A significant fraction of context effects may be due to natural pre-mRNA structures. In *S. cerevisiae*, as Eng and Warner have shown, the regulation and inefficient splicing of the gene encoding ribosomal protein RPL32 depend on the formation of a stem-loop structure at the 5' splice site (8). A natural pre-mRNA structure is also required for the proper

excision of the actin gene's intron in the related budding yeast *Kluyveromyces lactis*; formation of an RNA hairpin sequesters a proximal 3' splice site and is required to activate the correct, more distal 3' splice site (5). Long-range context effects influence the splicing efficiency of at least two yeast genes, *CYH2* and *RP51B*, and are likely mediated by intramolecular pairing within the intron (16, 32). In higher eukaryotic systems, a role of natural pre-mRNA structure in constitutive splicing is not established, but its importance in alternative splicing is well documented (e.g., references 3, 4, 6, 27, and 45).

In both yeast and mammalian systems, several groups have investigated the effects of artificial stem-loop structures on pre-mRNA splicing (e.g., references 12, 18, and 46). One of these mammalian studies suggested that such effects might not be large in natural pre-mRNA settings (9), and others indicated that the effects on in vivo splicing can be substantially weaker than those on in vitro splicing (10, 39, 40), raising the possibility that other processes compete with the formation and/or stability of these intramolecular RNA duplexes in vivo (see reference 10 for an elaboration of these possible competing processes).

Because there is no existing study that systematically evaluates the effect of intramolecular structure on yeast splicing, we decided to analyze the role of local pre-mRNA secondary structure on in vitro as well as in vivo yeast splicing. The results proved surprisingly easy to interpret and suggest a strong effect of pre-mRNA structure on spliceosome assembly, both in vitro and in vivo.

MATERIALS AND METHODS

Plasmid constructions. Plasmids were constructed according to standard protocols (35) and propagated in *Escherichia coli* TG1. Vectors for in vitro studies are derived from the vector pTZΔ2-3, which contains a short version of the *RP51A* intron under the control of the T7 promoter (38). Vectors for in vivo studies are derived from HZ18 (42) and pLGAcc/NdeAcc (24), which contain the *RP51A* intron and a short synthetic intron, respectively, inserted with some

* Corresponding author. Electronic mail address: BITNET% "ROSBASH@BRANDEIS".

† Present address: Laboratoire de Génétique Moléculaire, Ecole Normale Supérieure, 75230 Paris Cedex 05, France.

5'-6HP	5'-9HP	5'-12HP	5'-15HP	BP-15HP
5' A U G GUAUGU ^U A CAUACA _U A A U G 3'	5' A A U GUAUGU ^A U CCAUACA ^A U A C U 3'	5' A A U GUAUGU ^U AAU ^U GG CCAUACA ^U U ^A U ^A A A A 3'	5' A A U GUAUGU ^U AAU ^A U ^G GG ^A C CCAUACA ^U U ^A U ^A ACC ^A U A A G 3'	5' U U G CUUGUUGUAUACUG ^A U GAA _U CAAUCAU ^A U ^G AC ^U A U U G 3'
Acc-6HP NdeAcc-6HP	Acc-9HP NdeAcc-9HP	Acc-12HP NdeAcc-12HP	Acc-15HP NdeAcc-15HP	Acc-BP15HP NdeAcc-BP15HP
5' A U G GUAUGU ^U A CAUACA _U A A U G 3'	5' A A U GUAUGU ^A U CCAUACA ^A U G G U 3'	5' A A U GUAUGU ^U AAU ^U GG CCAUACA ^U U ^A U ^A U A A 3'	5' A A U GUAUGU ^U AAU ^A U ^G GG ^U U CCAUACA ^U U ^A U ^A ACC ^A U C G U 3'	5' U U U GGCCUGUUGUAUAG ^U GA CCGG _U CAAUCAU ^U AUC ^U A U U U 3'

FIG. 1. Description of the different hairpins. Inserted nucleotides are represented in boldface. Boxes show 5' splice sites and branch point consensus sequences.

flanking exonic sequences into the pLGSD5 *lacZ* reporter gene (17).

Insertions into the pTZΔ2-3, pTZAcc, and pTZNdeAcc 5' splice site or branch point regions were made by single-strand oligonucleotide-directed mutagenesis (22). pTZΔ2-3 hairpin-containing 5' splice sites were subcloned in the HZ18 background. 5' splice site duplications were made by the insertion of the various 75-bp *XhoI-SalI* fragments containing the modified 5' splice sites into HZ18 to create 5'-wt/5'-6HP, 5'-wt/5'-9HP, 5'-wt/5'-12HP, and 5'-wt/5'-15HP and into HZ12 (5'II mutation; GUAUGU→GUAUaU [20]) to create 5'II/5'-6HP, 5'II/5'-9HP, 5'II/5'-12HP, and 5'II/5'-15HP. pTZAcc and pTZNdeAcc *BamHI* fragments were subcloned into pLGSD5 in order to create the Acc and NdeAcc hairpin-containing series. In vitro and in vivo constructs were all sequenced by the chain termination method of Sanger et al. (36) with Sequenase version 2.0 (U.S. Biochemical).

In vitro complex assembly and splicing. Yeast DBY745 (*MATa ade1-100 leu2-112 ura3-52*) miniextracts were prepared by the glass bead procedure as described previously (37). Radioactively labeled pre-mRNA substrates were synthesized as described previously (38), and in vitro reactions were carried out either under standard conditions (37) to study RNA products and spliceosome formation or with U2 snRNP-inactivated miniextracts (26) to study commitment complex formation. Native gel analysis was as described previously (38), as was the analysis of splicing products (38).

β-Galactosidase assays. DNAs were used to transform yeast cells by the LiCl method (19). The strain used was the wild-type strain XL8 (*MATa leu2-3 leu2-112 ura3-52 trp1-289 arg4 ade2 snr19::leu2 pXL8* [25]). After induction of the

reporter gene for 3 h in 2% galactose, the β-galactosidase assays were performed as described previously (24). β-Galactosidase values are reported as a percentage of the wild-type construct. Because some *cis* competition constructs show a decrease in overall splicing efficiency, β-galactosidase values do not always reflect perfectly the relative use of the two 5' splice sites (e.g., reference 15). Results are expressed as the average of a minimum of eight values from four independent experiments.

Primer extension analyses. RNA extraction and primer extension were carried out according to published procedures (33). The primer used, RB1, was complementary to the beginning of the second exon of the *RP51A* gene (42). Extension products were analyzed on 6% polyacrylamide denaturing gels. Short extensions were also performed with 0.025 mM ddGTP instead of 0.125 mM dGTP in order to get a 36-nucleotide fragment with *RP51A* mRNAs and a 40-nucleotide fragment with *RP51A* pre-mRNAs (33). These products were analyzed on 12% polyacrylamide denaturing gels.

RESULTS

Hairpin structure inhibits pre-mRNA splicing in vitro. To analyze the effects of local secondary structure on in vitro pre-mRNA splicing, oligonucleotides complementary to either 5' splice site or branch point sequences were introduced into a shortened version of the yeast *RP51A* intron (WT-Δ2 [38]). The putative hairpin structures of these are shown in Fig. 1. The size of the stems varied from 6 to 15 nucleotides, and that of the loops varied from 4 to 6 nucleotides (Fig. 1). The hairpin-containing constructs are designated according

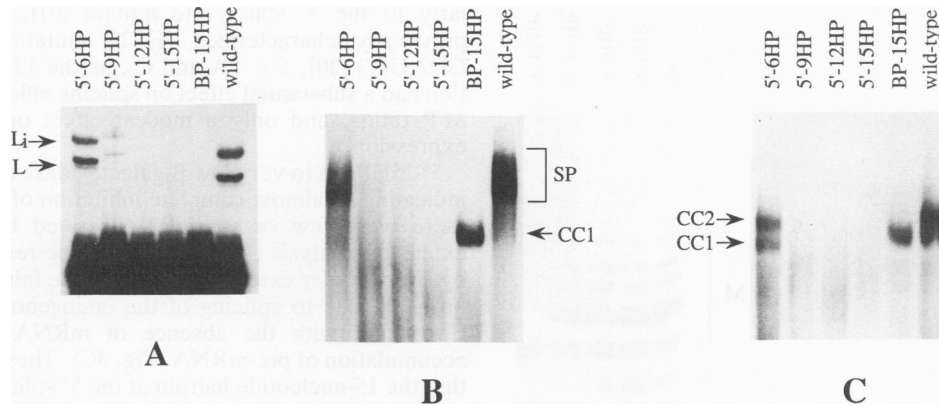


FIG. 2. In vitro splicing of the hairpin-containing substrates. Radioactively labeled pre-mRNA substrates were incubated under standard conditions (see Materials and Methods) with wild-type miniextracts (A and B) or after U2 snRNP inactivation with oligonucleotide-directed RNase H digestion (C). The wild-type $\Delta 2$ pre-mRNA was used as a control. (A) RNA products were analyzed on a 6% denaturing polyacrylamide gel. Arrows indicate positions of lariat intermediate (Li) and lariat intron (L) species. (B) Spliceosomes were analyzed on a native gel (see Materials and Methods). Positions of spliceosomes (SP) and the lower band of commitment complex (CC1) are indicated. (C) Commitment complexes were analyzed on a native gel. Arrows indicate the positions of the commitment complex bands CC1 and CC2.

to the location and stem size of the hairpins. Radioactively labeled pre-mRNA substrates were synthesized and incubated in yeast extract under standard conditions for in vitro splicing.

Splicing of the 5'-6HP substrate gave rise to significant levels of splicing intermediates and products, approximately half of what was generated by the control wild-type construct (Fig. 2A). The 5'-9HP substrate gave low levels of intermediates and products (Fig. 2A), indicating that splicing was strongly affected. The 5'-12HP substrate as well as the 5'-15HP substrate inhibited splicing essentially completely (Fig. 2A). The construct with a 15-nucleotide insertion at the branch point region (BP-15HP, Fig. 1) gave results similar to those obtained with 5'-15HP, i.e., no detectable products (Fig. 2A). In all cases, substrates with insertions of the reverse sequences produced the same levels of splicing intermediates and products as those generated by the wild-type substrate, indicating that this modest expansion of intron size was not itself inhibitory (data not shown). We conclude that secondary structures which sequester yeast splice sites inhibit in vitro splicing. The longer the size of the stem is, the stronger the inhibition is.

Assembly of spliceosomes and commitment complex formation are blocked in hairpin-containing introns. In both *S. cerevisiae* and higher eukaryotes, in vitro splicing is preceded by a required spliceosome assembly process during which the splicing factors add to the pre-mRNA in an ordered fashion. The binding of U1 snRNP occurs first and gives rise to a stable substrate-U1 snRNP complex (commitment complex) that apparently commits the substrate to the in vitro splicing pathway. Then, U2 snRNP is added and is followed by the addition of the U4/U5/U6 triple snRNP. To test whether the observed splicing inhibition could be accounted for by an inhibition of splicing complex formation, we assayed the ability of the hairpin-containing substrates to undergo U2 snRNP and U1 snRNP addition.

Similar to its effect on splicing, 5'-6HP gave rise to approximately half of the spliceosomes generated by the wild-type substrate (Fig. 2B; for simplicity, all U2 snRNP-containing complexes are referred to as spliceosomes). The 5'-9HP substrate led to a barely detectable level of spliceosome formation whereas the 5'-12HP and the 5'-15HP substrates led to no detectable complex formation (Fig. 2B).

We observed nearly identical results when U1 snRNP addition (commitment complex formation) was assayed subsequent to U2 snRNP inactivation by oligonucleotide-directed RNase H digestion. The wild-type substrate gave rise to the two characteristic commitment complex bands, CC1 and CC2 (Fig. 2C) (37). In the case of the 5'-6HP substrate, the intensity of these two bands was approximately half that from the wild-type construct. The 5'-9HP substrate generated very little complex (Fig. 2C; visible on the autoradiogram but not on the photograph). For the 5'-12HP and 5'-15HP substrates, there was no detectable commitment complex formation (Fig. 2C). These results parallel the ones described above for splicing products and spliceosome formation, suggesting that splicing inhibition results from an effect on stable commitment complex formation.

The BP-15HP construct gave rise to completely different results, as we observed in both analyses a single high-intensity band corresponding to CC1 (Fig. 2B and C). This result is consistent with previous data indicating that a functional branch point region as well as a functional 5' splice site region is required for formation of the upper commitment complex band (CC2) and for U2 snRNP addition but that only a functional 5' splice site is required for basal U1 snRNP addition (CC1). We conclude that the BP-15HP structure prevented branch point recognition, a step required for CC2 formation and subsequent U2 snRNP addition. The inhibition of CC2 formation is sufficient to account for the splicing inhibition observed for the BP-15HP substrate.

In vivo splicing of hairpin-containing introns. In order to study the effect of these local RNA structures on in vivo splicing, we inserted the sequences described above into the well-studied reporter construct HZ18 (Fig. 1). This gene contains a wild-type *RP51A* intron inserted into the *lacZ* gene so that only precise excision of the intron generates β -galactosidase expression (42). In vivo splicing was assayed with enzyme activity and with primer extension analyses. The latter method was used to compare relative mRNA levels as well as to estimate mRNA/pre-mRNA (M/P) ratios, which reflect relative splicing efficiency more accurately than mere comparisons of enzyme activity or mRNA levels (33). To determine M/P ratios, we also performed short primer extension with ddGTP to generate only a single

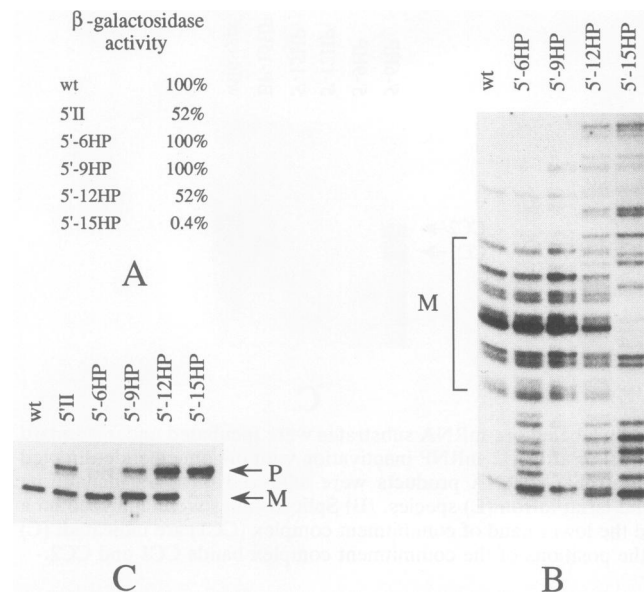


FIG. 3. In vivo splicing of the 5' hairpin-containing *RP51A* constructs. (A) β -Galactosidase analysis. β -Galactosidase activities were assayed in wild-type cells transformed by the different constructs and are reported as a percentage of the β -galactosidase expression generated by the wild-type *RP51A* intron construct. (B) RNA analysis by primer extension. After extraction, RNAs were analyzed by primer extension with the oligonucleotide RB1 (see Materials and Methods) and electrophoresed on a 6% denaturing polyacrylamide gel. The bracket (M) indicates the positions of cDNAs corresponding to mRNAs (33). (C) mRNA/pre-mRNA analysis. Short primer extensions were performed with ddGTP (see Materials and Methods). Products were analyzed on a 12% denaturing polyacrylamide gel. Arrows indicate the positions for mRNA (M) and pre-mRNA (P) species.

cDNA species from the mRNAs and a single cDNA species from the pre-mRNAs; this procedure also minimizes problems due to pre-mRNA degradation, which complicates measurement of M/P ratios with standard primer extension.

The 6-nucleotide hairpin-containing construct, 5'-6HP (Fig. 1), was not detectably different from the control, i.e., it gave rise to a level of β -galactosidase expression similar to that observed for the wild-type *RP51A* construct (Fig. 3A). This result was confirmed by primer extension analysis: we observed a set of mRNA-derived bands of intensity similar to that generated by the control (Fig. 3B). (These bands correspond to the multiple start sites of the *CYC1* promoter [11, 42]). Furthermore, short primer extension analysis showed an intense band corresponding to mRNA and only a very faint band corresponding to the pre-mRNA species, similar to what was observed with the wild-type RNA (Fig. 3C).

The 5'-9HP construct also generated β -galactosidase expression and mRNA levels similar to those of the control (Fig. 3A and B). However, we observed a significant accumulation of precursor species (Fig. 3C). As previously described (33), an increase in pre-mRNA levels with little or no decrease in mRNA levels is characteristic of a modest decrease in splicing efficiency.

The 5'-12HP construct gave rise to a modest twofold decrease in β -galactosidase activity relative to that of the wild-type *RP51A* construct (Fig. 3A), and short primer extension analysis showed a dramatic accumulation of the pre-mRNA-derived band (Fig. 3C). 5'-12HP behaved simi-

larly to the 5' splice site mutant 5'II, which has been previously characterized (HZ12 mutation; GUAUGU \rightarrow GUAUaU [20]; Fig. 3A and C), as the 12-nucleotide insertion had a substantial effect on splicing efficiency, defined by M/P ratios, and only a modest effect on β -galactosidase expression.

5'-15HP led to very low β -galactosidase activity (Fig. 3A), indicating an almost complete inhibition of splicing. Indeed, there were few or no mRNA-derived bands by primer extension analysis (Fig. 3B). The same result was obtained by short primer extension (Fig. 3C; the faint mRNA band is probably due to splicing of the endogenous *RP51A* gene). Consistent with the absence of mRNA was the strong accumulation of pre-mRNA (Fig. 3C). These results indicate that the 15-nucleotide hairpin at the 5' splice site completely inhibits in vivo splicing or nearly so.

In summary, the in vivo effects of the different structures on mRNA levels or enzyme activity were weaker than their effects on in vitro splicing. For example, the 9-nucleotide hairpin had a dramatic effect on in vitro splicing but led to near-wild-type mRNA levels in vivo. However, the more sensitive M/P ratio assay suggested that the in vivo effects on splicing efficiency might be more similar to those observed in vitro.

Splice site duplications. To verify and extend these observations, we turned to another sensitive assay in which two 5' splice sites are in *cis* competition for a single 3' splice site. To this end, we generated a set of constructs in which the hairpin-containing 5' splice sites were inserted downstream of the *RP51A* 5' splice site. Only mRNAs derived from use of the upstream site can give rise to β -galactosidase expression. Primer extension analyses were also carried out to verify splicing at the upstream 5' splice site and to monitor directly splicing at the downstream 5' splice site. Also, short primer extension analyses were performed with ddGTP in order to assess M/P (i.e., both forms of M/P) ratios.

We previously showed that with a wild-type duplication of the *RP51A* 5' splice site (VG18-18 [15], renamed here for clarity 5'-wt/5'-wt [Fig. 4]), both sites were used to a similar extent, indicating little or no proximity effect in this construct. A 5' duplication with the mutant site 5'II (GUAUaU [20]) adjacent to the wild-type site (either VG18-12 or VG12-18 [15], renamed here for clarity 5'-wt/5'II and 5'II/5'-wt [Fig. 4]) showed a predominant use of the wild-type site, regardless of its position (upstream or downstream). This mutant 5' splice site serves as a benchmark against which the effects of the hairpin constructs can be compared.

The 5'-wt/5'-6HP duplication gave rise to an increase in β -galactosidase activity compared with that of the wild-type *RP51A* 5' splice site duplication (Fig. 4A). However, enzyme activity was somewhat lower than that observed for the wild-type/mutant duplication (Fig. 4A; compare 5'-wt/5'-6HP with 5'-wt/5'II). Primer extension analysis showed the presence of two sets of bands corresponding to splicing occurring at both the upstream and the downstream 5' splice sites (Fig. 4B). Primer extension performed with ddGTP did not show any significant pre-mRNA accumulation (data not shown). These results indicate that the 6-nucleotide hairpin-containing 5' splice site was selected, though not as efficiently as the wild-type site.

The 5'-wt/5'-9HP, 5'-wt/5'-12HP, and 5'-wt/5'-15HP substrates also gave rise to elevated β -galactosidase activities, similar to that of the 5'-wt/5'II construct (Fig. 4A). Primer extension analysis confirmed the predominant use of the wild-type upstream 5' splice site, as we could no longer detect the downstream set of mRNA-derived bands (Fig.

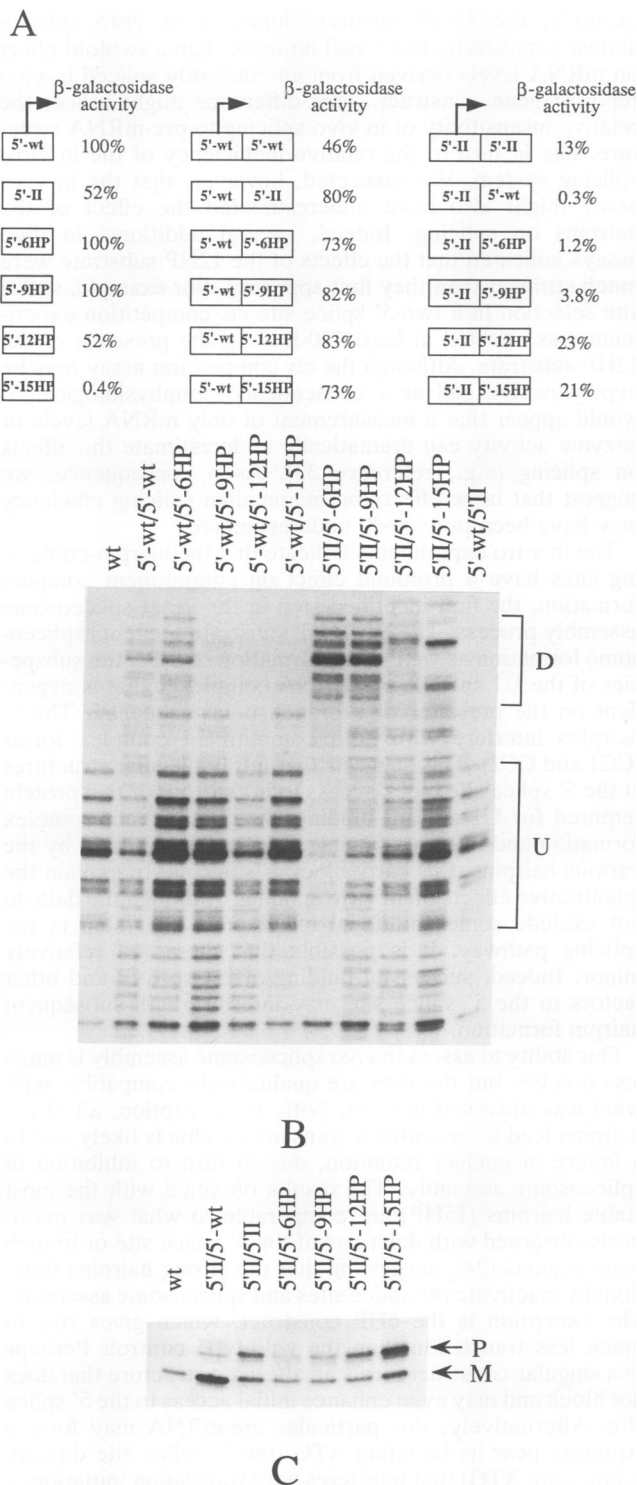


FIG. 4. In vivo splicing of the 5' splice site duplications. (A) β -Galactosidase analysis. Boxes represent the duplicated 5' splice sites. Arrows indicate where splicing should occur to produce mRNAs in frame for β -galactosidase expression. Activities are reported as a percentage of the β -galactosidase expression generated by the wild-type *RP51A* intron construct. (B) RNA analysis by primer extension. After extraction, RNAs were analyzed as described above (Fig. 3B). Brackets indicate positions of cDNAs corresponding to mRNAs resulting from splicing taking place at the upstream (U) and the downstream (D) 5' splice sites (15). (C) mRNA/pre-mRNA analysis. After extraction, RNAs were analyzed

(4B). No substantial pre-mRNA accumulation was observed (data not shown).

To improve the assay's sensitivity, we generated a parallel set of duplication constructs in which the reference (upstream) 5' splice site carried the weak 5' splice site 5'II. Replacement of the wild-type site from 5'II/5'-wt by the 6-nucleotide hairpin-containing 5' splice site (5'II/5'-6HP) led to a fourfold increase in β -galactosidase expression (Fig. 4A; compare 5'II/5'-wt with 5'II/5'-6HP), and 5'II/5'-9HP produced another threefold increase (Fig. 4A). Primer extension confirmed inefficient selection of the upstream mutant 5' splice site and preferred selection of the downstream hairpin-containing sites (Fig. 4B). Insertion of the 12-nucleotide hairpin-containing 5' splice site downstream of the 5'II site led to a large increase in enzyme expression, similar to what was generated by the duplication containing two mutant sites, 5'II/5'II (Fig. 4A). The 5'II/5'-15HP duplication also gave rise to similar levels of β -galactosidase expression (Fig. 4A). For both 5'II/5'-12HP and 5'II/5'-15HP, primer extension analysis confirmed the predominant use of the upstream 5' splice site (Fig. 4B). Short primer extension analysis showed a monotonic increase in pre-mRNA levels and indicated that the 5'-12HP is comparable to the 5'II mutant and that the 5'-15HP is an even worse splice site (Fig. 4C).

Transport of pre-mRNA from the nucleus to the cytoplasm: effect of hairpin-containing splice sites. To test how the hairpins inhibit in vivo pre-mRNA splicing, we created similar hairpins in a pair of in vivo reporter gene constructs, pLGAcc and pLGNdeAcc (24). They both contain an almost identical short inefficient intron inserted in the *lacZ* coding sequences under the control of the galactose-inducible *CYC1* promoter. A 2-nucleotide difference in the two artificial introns creates a frameshift such that the β -galactosidase coding sequence is in frame with the *CYC1* AUG for only spliced mRNA in the case of pLGAcc and for only the pre-mRNA in the case of pLGNdeAcc (24). We generated variants of pLGAcc and pLGNdeAcc in which hairpins of different lengths were introduced at the 5' splice site and branch point regions (Fig. 1). Analyses of these constructs allowed us to assay pre-mRNA as well as mRNA translation.

As previously described (24), pLGAcc is a poor splicing substrate; a fraction of the pre-mRNA from the parallel construct, pLGNdeAcc, is also translated (Fig. 5; wild type). This indicates that some pLGNdeAcc pre-mRNA is transported to the cytoplasm, probably reflecting rather inefficient nuclear retention of the artificial intron. The hairpin-containing Acc constructs showed an increase in splicing inhibition as the size of the hairpin increased (Fig. 5).

Similar results were observed for the pLGNdeAcc constructs: namely, as the size of the hairpin increased, the β -galactosidase values increased (Fig. 5), indicating the presence of more pre-mRNA in the cytoplasm. (Note that the NdeAcc-6HP construct gave rise to a decrease in enzyme activity compared with that of the control; see below.) The data obtained with the 15-nucleotide hairpin constructs are similar to those previously obtained with constructs in which the 5' splice site or the branch point region was deleted, i.e., a complete inhibition of splicing and a high level of pre-mRNA translation were observed, respectively

as described above (Fig. 3C). Arrows indicate the positions of cDNAs corresponding to pre-mRNAs (P) and to mRNAs generated by splicing occurring at either 5' splice site (M).

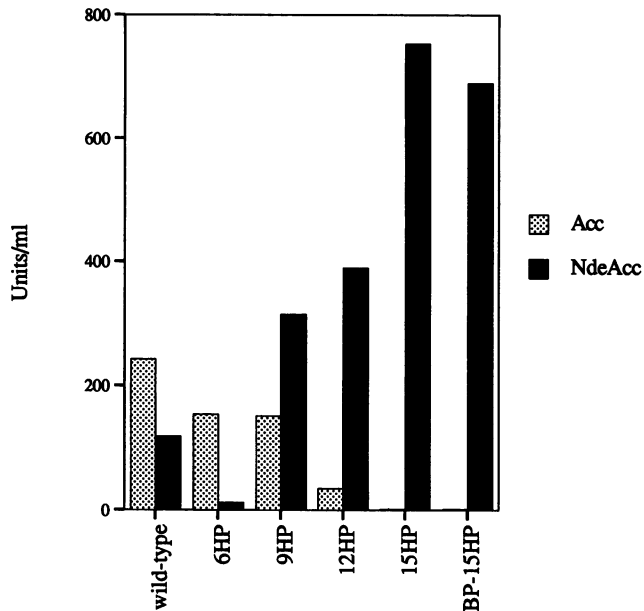


FIG. 5. β -Galactosidase analysis of the pLGAcc- and pLGNdeAcc-derived constructs. The Acc and NdeAcc hairpin-containing series (Fig. 1) were transformed into strain XL8 as described above. β -Galactosidase values were measured as described in Materials and Methods. Note that the hairpins inhibit expression from the Acc series, presumably by inhibiting mRNA formation. In contrast, they increase expression from the NdeAcc series, presumably by decreasing stable complex formation which leads to increased pre-mRNA translation.

(24). The observed effects were not due to the insertions per se, as constructs with insertions of the reverse sequences generated the same levels of β -galactosidase expression as the parent constructs (data not shown). We conclude that the fraction of pre-mRNA transported to the cytoplasm is related to the extent of hairpin structure (except for the 6HP construct; see Discussion), suggesting that hairpins inhibit stable complex formation in vivo, which is in part responsible for the splicing inhibition observed in the hairpin-containing Acc series.

DISCUSSION

Our experiments indicate that pre-mRNA secondary structures that sequester important splicing sequences can have a profound effect on splicing efficiency and splice site selection. Although similar conclusions have already been reached for mammalian systems, the experiments presented in this report are more complete in several respects than most of those already in the literature. The effects of the same hairpin structures on splicing were assayed in vitro as well as in vivo. Stem length was varied in a systematic way; the extent of inhibition, in the main, varied in monotonic fashion from modest to complete. Inhibition of spliceosome assembly was also assayed and found to vary in parallel with splicing inhibition.

In previous reports from mammalian systems, there are indications that in vivo splicing is much less sensitive to pre-mRNA structure than in vitro splicing (39, 40). Our initial attempts to compare the effects of the sequestering hairpins in the two assay systems gave rise to similar results, namely, much bigger effects in vitro than in vivo. For

example, the 12HP substrate inhibited in vitro splicing almost completely, but it had no more than a twofold effect on mRNA levels derived from an efficiently spliced in vivo reporter gene construct. This difference might reflect the relative insensitivity of in vivo splicing to pre-mRNA structure, due in turn to the relative inefficiency of the in vitro splicing system. We suspected, however, that the in vivo assay might also have underestimated the effect of the hairpins on splicing. Indeed, several additional in vivo assays indicated that the effects of the 12HP substrate were much stronger than they first appeared. For example, splice site selection in a two-5' splice site *cis* competition experiment was affected at least 100-fold by the presence of the 12HP substrate. Although the *cis* competition assay may be hypersensitive and as a consequence nonphysiological, it would appear that a measurement of only mRNA levels or enzyme activity can dramatically underestimate the effects on splicing (e.g., reference 33). As a consequence, we suggest that large effects on mammalian splicing efficiency may have been previously underestimated.

The in vitro experiments indicate that the hairpin-containing sites have a profound effect on commitment complex formation, the first identified step in the yeast spliceosome assembly process. The BP-15HP substrate prevents spliceosome formation as well as the formation of CC2, the subspecies of the U1 snRNP-pre-mRNA complexes that is dependent on the presence of a branch point sequence. The 5' hairpins interfere with both commitment complex forms (CC1 and CC2). This suggests that intramolecular structures at the 5' splice site block access to U1 snRNP or to a protein required for U1 snRNP binding. As commitment complex formation and in vitro splicing are affected similarly by the various hairpins, this early block is sufficient to explain the quantitative effects on in vitro splicing. Although the data do not exclude some additional effect on a later step in the splicing pathway, it is possible that these are relatively minor. Indeed, successful binding of U1 snRNP and other factors to the 5' splice site may interfere with subsequent hairpin formation.

Our ability to assess in vivo spliceosome assembly is much less precise, but the data are qualitatively compatible with what was observed in vitro. With one exception, all of the hairpins lead to pre-mRNA translation. This is likely due to a failure in nuclear retention, due in turn to inhibition of spliceosome assembly. The results obtained with the most stable hairpins (15HP) are comparable to what was previously observed with deletions of the 5' splice site or branch point regions (24), suggesting that the strong hairpins functionally inactivate the splice sites and spliceosome assembly. The exception is the 6HP construct, which gives rise to much less translation than the wild-type control. Perhaps this singular construct forms an alternate structure that does not block and may even enhance initial access to the 5' splice site. Alternatively, this particular pre-mRNA may form a structure near its initiating ATG (the 5' splice site directly follows the ATG) that interferes with translation initiation, a consideration that generally complicates quantitative comparisons with the more directly observed effects on splicing.

It is of considerable interest to compare the hairpins' effects on splicing as a function of relative helix stability. The splicing effects can be estimated from the various in vitro and in vivo assays. The best estimates of the 6HP inhibitory effect come from (i) the in vitro splicing and assembly experiments (Fig. 2), in which a twofold decrease in splicing and spliceosome assembly was observed, and from (ii) the splice site duplication experiments (Fig. 4A), in

which a fourfold increase in splice site choice was obtained (from 0.3 to 1.2%) when the competing wild-type 5' splice site was replaced with the less potent 6HP site. The 9HP substrate has a strong effect in vitro and an estimated 10-fold effect from the splice site duplication experiments (Fig. 4A; 0.3 to 3.8%). An estimate of the 12HP effect also comes from the splice site duplication experiments (Fig. 4A), in which the effect was approximately 70-fold (from 0.3% to 23%). Because this was the maximum value in the series, it may represent an underestimate of the hairpin's inhibitory effect. The effect of the 15HP site, assayed more simply in a single intron reporter gene construct, is 100-fold stronger than that of the 12HP site (Fig. 4A; 12HP = 52%; 15HP = 0.4%).

In contrast to the effects on splicing, the absolute helix stabilities are more difficult to evaluate; it is uncertain how the predicted values will compare with T_m or ΔG measurements inside nuclei or in splicing extract, in which RNA-binding proteins and helicases are likely to influence helix stability. To the best of our knowledge, no general measurement of this nature has been undertaken. Equally interesting would be a comparison of these helix stabilities with those previously used in comparable mammalian experiments. Unfortunately, many of the relevant mammalian experiments preceded the general appreciation of the effect of loop size and sequence on helix stability (e.g., reference 43). As a consequence, this parameter was often not well controlled, which makes a comparison with our 6- to 15-helix series difficult. It is even possible that the relatively small loop size in this study contributed to the sizeable effects observed in vivo.

With this variable in mind (the strong effect of loop size on helix stability), it would be interesting to assay more directly the various RNA structures, both in vivo within nuclei and in vitro in splicing extract. As yeast heterogeneous nuclear RNP proteins appear to be less abundant than or different from their mammalian counterparts (7), there might even be significant differences in helix stabilities between yeast and mammalian cells, i.e., a direct comparison of the same series of intramolecular helices in *S. cerevisiae* and mammals (in both cells and extracts) might be revealing. Taken together with recent studies that indicate a positive effect of natural pre-mRNA structures on splicing efficiency (16), a detailed examination of pre-mRNA structure appears to be warranted.

ACKNOWLEDGMENTS

We thank Terri McCarthy for technical assistance and members of the Rosbash laboratory for suggestions as well as comments on the manuscript.

This work was supported by grant GM-23549 from the National Institutes of Health.

V.G. and Y.W. contributed equally to this work.

REFERENCES

- Aebi, M., H. Hornig, R. A. Padgett, J. Reiser, and C. Weissmann. 1986. Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. *Cell* 47:555-565.
- Cech, T. R. 1986. The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* 44:207-210.
- Chebli, K., R. Gattoni, P. Schmitt, G. Hildwein, and J. Stevenin. 1989. The 216-nucleotide intron of the E1A pre-mRNA contains a hairpin structure that permits utilization of unusually distant branch acceptors. *Mol. Cell. Biol.* 9:4852-4861.
- Clouet d'Orval, B., Y. D'Aubenton-Carafa, P. Sirand-Pugnet, M. Gallego, E. Brody, and J. Marie. 1991. RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. *Science* 252:1823-1828.
- 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.
- Domenjoud, L., H. Gallinaro, L. Kister, S. Meyer, and M. Jacob. 1991. Identification of a specific exon sequence that is a major determinant in the selection between a natural and a cryptic 5' splice site. *Mol. Cell. Biol.* 11:4581-4589.
- Dreyfuss, G., M. J. Matunis, S. Piñol-Roma, and C. G. Burd. 1993. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* 62:289-321.
- Eng, F. J., and J. R. Warner. 1991. Structural basis for the regulation of splicing of a yeast messenger RNA. *Cell* 65:797-804.
- Eperon, L. P., J. P. Estibeiro, and I. C. Eperon. 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature (London)* 324:280-282.
- Eperon, L. P., I. R. Graham, A. D. Griffiths, and I. C. Eperon. 1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* 54:393-401.
- Faye, G., D. W. Leung, K. Tatchell, B. D. Hall, and M. Smith. 1981. Deletion mapping of sequences essential for *in vivo* transcription of the iso-1-cytochrome *c* gene. *Proc. Natl. Acad. Sci. USA* 78:2258-2262.
- Fu, X.-Y., and J. L. Manley. 1987. Factors influencing alternative splice site utilization in vivo. *Mol. Cell. Biol.* 7:738-748.
- Furdon, P. J., and R. Kole. 1988. Length of the downstream exon and the substitution of specific sequences affect pre-mRNA splicing in vitro. *Mol. Cell. Biol.* 8:860-866.
- Gallego, M. E., and B. Nadal-Ginard. 1990. Myosin light-chain 1/3 gene alternative splicing: *cis* regulation is based upon a hierarchical compatibility between splice sites. *Mol. Cell. Biol.* 10:2133-2144.
- Goguel, V., X. Liao, B. C. Rymond, and M. Rosbash. 1991. U1 snRNP can influence 3'-splice site selection as well as 5'-splice site selection. *Genes Dev.* 5:1430-1438.
- Goguel, V., and M. Rosbash. 1993. Splice site choice and splicing efficiency are positively influenced by intramolecular base-pairing in yeast. *Cell* 72:893-901.
- Guarente, L., R. R. Yocum, and P. Gifford. 1982. A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. *Proc. Natl. Acad. Sci. USA* 79:7410-7414.
- 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.
- Ito, H., Y. Fukuda, K. Murata, and A. Kimura. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* 153:163-168.
- Jacquier, A., J. R. Rodriguez, and M. Rosbash. 1985. A quantitative analysis of the effects of 5' junction and TACTAAC box mutants and mutant combinations on yeast mRNA splicing. *Cell* 43:423-430.
- Kuhne, T., B. Wieringa, J. Reiser, and C. Weissmann. 1983. Evidence against a scanning model of RNA splicing. *EMBO J.* 2:727-733.
- Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.
- Lear, A. L., L. P. Eperon, I. M. Wheatley, and I. C. Eperon. 1990. Hierarchy for 5' splice preference determined in vivo. *J. Mol. Biol.* 211:103-115.
- Legrain, P., and M. Rosbash. 1989. Some *cis*- and *trans*-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell* 57:573-583.
- Liao, X., L. Kretzner, B. Séraphin, and M. Rosbash. 1990. Universally conserved and yeast-specific U1 snRNA sequences are important but not essential for U1 snRNP function. *Genes Dev.* 4:1766-1774.
- Liao, X. C., H. V. Colot, Y. Wang, and M. Rosbash. 1992. Requirements for U2 snRNP addition to yeast pre-mRNA.

- Nucleic Acids Res. **20**:4237–4245.
27. Libri, D., A. Piseri, and M. Fiszman. 1991. Tissue-specific splicing *in vivo* of the β -tropomyosin gene: dependence on an RNA secondary structure. *Science* **252**:1842–1845.
 28. Mayeda, A., and Y. Ohshima. 1988. Short donor site sequences inserted within the intron of β -globin pre-mRNA serve for splicing *in vitro*. *Mol. Cell. Biol.* **8**:4484–4491.
 29. Miller, A. M. 1984. The yeast *MAT α 1* gene contains two introns. *EMBO J.* **3**:1061–1065.
 30. Moore, M. J., C. C. Query, and P. A. Sharp. 1993. Splicing of precursors to mRNA 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.
 31. Nelson, K. K., and M. R. Green. 1988. Splice site selection and ribonucleoprotein complex assembly during *in vitro* pre-mRNA splicing. *Genes Dev.* **2**:319–329.
 32. Newman, A. 1987. Specific accessory sequences in *Saccharomyces cerevisiae* introns control assembly of pre-mRNAs into spliceosomes. *EMBO J.* **6**:3833–3839.
 33. Pikielny, C. W., and M. Rosbash. 1985. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell* **41**:119–126.
 34. Reed, R., and T. Maniatis. 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell* **46**:681–690.
 35. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 36. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
 37. Séraphin, B., and M. Rosbash. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**:349–358.
 38. Séraphin, B., and M. Rosbash. 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.
 39. Solnick, D. 1985. Alternative splicing caused by RNA secondary structure. *Cell* **43**:667–676.
 40. Solnick, D., and S. I. Lee. 1987. Amount of RNA secondary structure required to induce an alternative splice. *Mol. Cell. Biol.* **7**:3194–3198.
 41. Steitz, J. A. 1992. Splicing takes a holiday. *Science* **257**:888–889.
 42. Teem, J. L., and M. Rosbash. 1983. Expression of a β -galactosidase gene containing the ribosomal protein 51 intron is sensitive to the *ma2* mutation of yeast. *Proc. Natl. Acad. Sci. USA* **80**:4403–4407.
 43. Tuerk, C., P. Gauss, C. Thermes, D. R. Groebe, M. Gayle, N. Guild, N. Stormo, Y. D'Aubenton-Carafa, O. C. Uhlenbeck, I. Tinoco, Jr., E. N. Brody, and L. Gold. 1988. CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes. *Proc. Natl. Acad. Sci. USA* **85**:1364–1368.
 44. Ulfendahl, P. J., J.-P. Kreivi, and G. Akusjärvi. 1989. Role of the branch site/3'-splice site region in adenovirus-2 E1A pre-mRNA alternative splicing: evidence for 5'- and 3'-splice site co-operation. *Nucleic Acids Res.* **17**:925–938.
 45. Watakabe, A., K. Inoue, H. Sakamoto, and Y. Shimura. 1989. A secondary structure at the 3' splice site affects the *in vitro* splicing reaction of mouse immunoglobulin μ chain pre-mRNAs. *Nucleic Acids Res.* **17**:8159–8169.
 46. Yoshimatsu, T., and F. Nagawa. 1989. Control of gene expression by artificial introns in *Saccharomyces cerevisiae*. *Science* **244**:1346–1348.