Efficient Trans-Splicing of a Yeast Mitochondrial RNA Group II Intron Implicates a Strong 5′ Exon–Intron Interaction

ALAIN JACQUIER* AND MICHAEL ROSBASH

The reaction mechanism for self-splicing introns requires the existence of a 5′ exon binding site on the intron. Experimental evidence is now presented consistent with the existence of such a binding site by demonstrating efficient and accurate trans–self-splicing of a yeast mitochondrial group II intron. Partial and complete trans-splicing reactions take place in the absence of branch formation, part of the usual pathway of nuclear splicing and group II self-splicing. In addition to indicating the existence of a 5′ exon binding site on the intron, the results have mechanistic implications for group II self-splicing and perhaps for nuclear splicing as well.

IN THE LAST TWO OR THREE YEARS WE HAVE WITNESSED remarkable progress in our understanding of RNA splicing. While transfer RNA (tRNA) splicing has been understood in detail for some time (*1*), only recently has a relatively clear picture of the other major modes of RNA splicing emerged. Nuclear (pre-messenger RNA) splicing is dependent on relatively few conserved primary and secondary structural features within the intron but apparently requires an extensive set of trans-acting factors that are necessary for proper and efficient splicing. These factors, both proteins and RNA, contribute to assembling the primary transcript into a spliceosome within which intron removal and exon ligation take place. Indeed, the primary sequence signals within the intron (5′ junction, branch point, 3′ junction) serve, at least in part, as recognition signals for proper spliceosome assembly (*2, 3*). Because of the remarkable self-splicing activity that at least some group I and group II introns manifest in vitro (*4–7*), it is almost certain that these molecules are less dependent on an extensive set of trans-acting factors rather than on the substantial primary and secondary structure features that they contain. Indeed, it was initially on the basis of their primary and secondary structure characteristics that group I and group II introns were originally identified and classified (*8–10*), and more recent insight into the two mechanisms supports this division (*11*).

Group I introns, as exemplified by the *Tetrahymena* ribosomal RNA (rRNA) intron studied by Cech and his collaborators, have been proposed to splice via two successive phosphate transfer, transesterification reactions (*5, 12*). The first one is initiated by nucleophilic attack at the 5′ junction by the 3′ OH of a free guanosine nucleotide, which adds to the 5′ end of the intron–3′ exon and liberates the 5′ exon with a 3′ OH. The second transesterification reaction is initiated by nucleophilic attack at the 3′ junction by the 3′ OH of the 5′ exon, which results in exon ligation and liberates the intron.

Group II introns also probably splice by way of two successive phosphate transfer, transesterification reactions. There is, however, one prominent difference between the reaction mechanisms proposed for group I and group II introns. While cleavage at the 5′ junction in group I splicing is due to nucleophilic attack by a free guanosine nucleotide, cleavage at the 5′ junction in group II splicing is almost certainly due to nucleophilic attack by a 2′ OH from within the intron. This creates a lariat intermediate with the 5′ end of the intron attached through a 2′,5′-phosphodiester bond to a residue near the 3′ end of the intron. Subsequent cleavage at the 3′ junction results in exon ligation and liberates the "free" intron in the form of a lariat, the diagnostic experimental observation on which this mechanism is based (*6, 7*). The nature of the initiating nucleophile notwithstanding, the two self-splicing mechanisms appear quite similar as both undergo 5′ junction cleavage first, and subsequently 3′ junction cleavage and exon ligation as a consequence of nucleophilic attack by the 5′ exon. Nuclear pre-mRNA splicing shares the diagnostic features of group II splicing, that is, the same intermediates and products are visible. Indeed, it was on this basis, the prior and more extensive characterization of nuclear lariat intermediates and excised lariats, that the group II mechanism was proposed (*6, 7*).

Thus, it appears that group I, group II, and nuclear pre-mRNA splicing all follow a similar pathway. The first step consists of 5′ junction cleavage; the second consists of 3′ junction cleavage and exon ligation by subsequent nucleophilic attack at the 3′ junction by the 5′ exon. As a result, all three splicing modes also share the feature that the 5′ exon is no longer covalently bound to the intron–3′ exon subsequent to the first step. In the case of nuclear splicing, trans-acting factors in the spliceosome can be invoked to hold the released 5′ exon in place. For example, the 5′ exon might be associated with the U1 small ribonucleoprotein particle (snRNP) at this intermediate stage of the reaction (*13*). In the case of group I splicing, a 5′ exon binding site within the intron probably serves this function. This base pairing interaction, originally proposed as a result of sequence comparisons of different group I introns (*8–10, 14*), has recently received substantial experimental support. In particular, genetic experiments from two groups demonstrate that pairing between the 5′ exon and an intron sequence is required for

The authors are in the Department of Biology, Brandeis University, Waltham, MA 02254.

*Present address: Centre de Genetique Moleculaire du CNRS, G-91190 Gif sur Yvette, France.

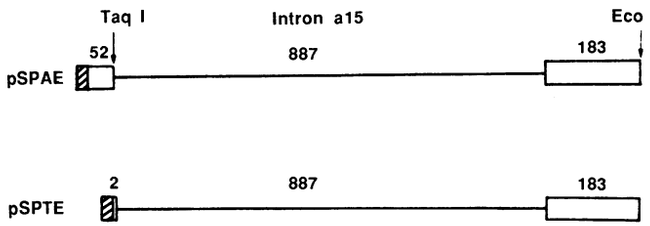

Fig. 1. The sequences transcribed from each of the the two plasmids, pSPAE and pSPTE. The open boxes represent exon sequences, the lines represent intron sequences, and the shaded boxes represent transcribed vector-encoded sequences. These transcribed vector sequences are GATACAAGCTT for pSPAE and GAATACAAGCTTGGGCTGCAGG for pSPTE. The numbers indicate, in nucleotides, the length of the corresponding mitochondrial segments. When pSPAE and pSPTE were digested with Eco RI before in vitro transcription, they generated transcripts corresponding to the complete sequences shown. "Full length precursor" refers to the transcript generated from pSPAE and "intron–3′ exon" to the transcript generated from pSPTE. The DNA sequence of the Taq I site, TCGA, is such that the C is the last nucleotide of the 5′ exon and the G is the first nucleotide of the intron (*19*). When pSPAE was digested with Taq I prior to in vitro transcription, it generated a transcript including the vector-encoded sequences, the 52-nucleotide 5′ exon and the first nucleotide of the intron, referred to as "5′ exon-G." RNA concentrations were measured by optical density and by calculating from specific activities when the RNA's were internally labeled with α-labeled nucleotides.

proper and efficient splicing (*15, 16*). Moreover, an interesting consequence of this pairing is that an oligonucleotide, when complementary to the 5′ exon binding site, behaves as a trans-acting 5′ exon and initiates directly the second transesterification reaction, resulting in 3′ junction cleavage and trans-exon ligation (*17, 18*).

The similarities in proposed reaction mechanism suggest that some pairing might also exist between the 5′ exon and intron in group II splicing. Yet, in the case of group II splicing, a comparable 5′ exon binding site has not been found or even proposed, thereby precluding a genetic test, similar to those cited above and involving site-directed mutagenesis.

To circumvent this difficulty, we designed a set of trans-splicing experiments, somewhat related to those described above, that test for the presence of a 5′ exon binding site within the group II intron. The results of these experiments suggest that, indeed, there exists a strong interaction between the 5′ exon and the intron. Moreover, a number of unexpected findings have implications for the group II reaction mechanism and perhaps for nuclear splicing as well.

Efficient and accurate trans–self-splicing. To test a possible noncovalent interaction between the 5′ exon and the intron–3′ exon, we constructed plasmids from which appropriate transcripts could be synthesized in vitro. A portion of the yeast (*Saccharomyces cerevisiae*) cytochrome oxidase subunit I gene [from the strain D273-10B (*18, 19*)] containing 52 nucleotides of exon 5, the entire fifth intron (intron aI5), and 183 nucleotides of exon 6, was cloned into the SP6 promoter-containing plasmid pSP64 (*20*) (plasmid pSPAE in Fig. 1). We took advantage of a convenient Taq I site that spans the 5′ junction to generate a subclone that deletes all but two nucleotides of the 5′ exon and retains the entire intron and the same 3′ exon (plasmid pSPTE in Fig. 1). Digestion of plasmid pSPAE with Eco RI, and subsequent in vitro transcription with SP6 polymerase, generates a full-length precursor RNA while digestion with Taq I and in vitro transcription generates a major transcript that terminates after the first nucleotide (a guanosine) of the intron (5′ exon-G RNA). The structure of the 5′ exon-G RNA was predicted from the position of the Taq I site and confirmed by size analysis at high resolution and by analysis of the 3′ terminal nucleotide, that is, addition of radioactive pCp, digestion with T2 ribonuclease, and thin-layer chromatography (TLC) (*21*). Digestion of subclone pSPTE with Eco RI and in vitro transcription generates the intron–3′ exon RNA which contains two nucleotides of 5′ exon, the entire intron, and the same 183 nucleotides of the 3′ exon. In all three cases, the transcripts contain short, plasmid-encoded sequences at their 5′ ends (shaded boxes in Fig. 1).

As expected, incubation of the full-length precursor RNA under self-splicing conditions gives rise to the products described by van der Veen *et al.* (*6*) and Peebles *et al.* (*7*), that is, lariat intron, nicked lariat intron (*7*), and ligated exons (Fig. 2A). Incubation of the intron–3′ exon RNA, under identical conditions, manifests no detectable reactivity (Fig. 2B), suggesting that two nucleotides of 5′ exon are not sufficient for self-splicing. The 5′ exon-G molecule is also apparently inert when incubated alone (*21*) (see below). In contrast, incubation of internally labeled intron–3′ exon RNA, in the presence of an excess of nonradioactive 5′ exon-G RNA, gives rise to two products. One is approximately the size of the linear intron and the other is the expected size of the ligated exons (Fig. 2, C and D). These results, and a dependence on both substrates, suggest that the observed products are generated by a trans-splicing reaction.

The simplest trans-splicing scenario suggests that the 5′ exon, derived from the 5′ exon-G molecule, ligates to the 3′ exon from the intron–3′ exon molecule and liberates a linear intron. This linear intron would then retain the two nucleotides of 5′ exon and plasmid-derived sequences at its 5′ end. This is indeed the case as an identical reaction, but with the intron–3′ exon end-labeled at its 5′ end by polynucleotide kinase (and [γ-^{32}P]ATP) rather than internal-

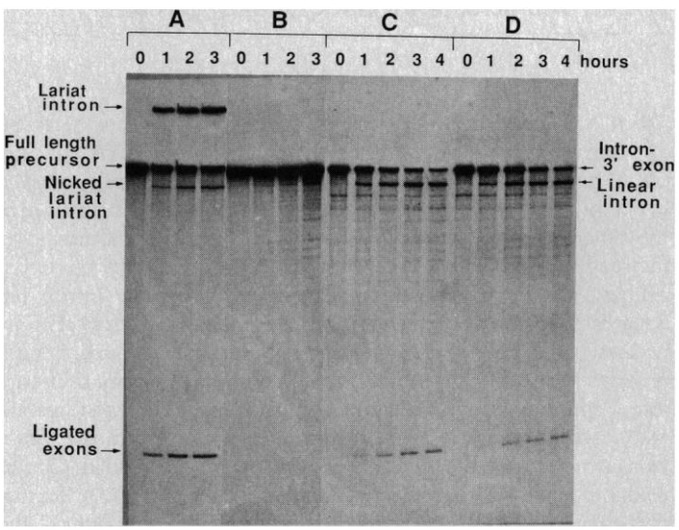

Fig. 2. A comparison of the time course of cis- and trans-splicing. Full-length precursor (panel A) or intron–3′ exon (panels B to D) were synthesized by in vitro transcription from the Eco RI–digested plasmids pSPAE and pSPTE, respectively. 5′ exon was synthesized by in vitro transcription of Taq I–digested pSPAE (Fig. 1). (A) Full-length precursor RNA, synthesized in vitro (*20*), in the presence of [α-^{32}P]UTP (*20*), was incubated at 0.02 μ*M* in 5 m*M* tris-HCl, 5 m*M* $MgCl_2$, 2 m*M* spermidine, and carrier yeast tRNA (0.2 μg/μl) at *p*H 7.7 and 45°C. Samples (1 μl) were taken at 0, 1, 2, and 3 hours and mixed with 3 μl of loading buffer (98 percent formamide, 10 m*M* EDTA, 0.1 percent xylene cyanol FF, 0.1 percent bromphenol blue). The samples were heated for 2 minutes at 90°C and analyzed on a 4 percent acrylamide-urea gel. (B) Intron–3′ exon was synthesized, incubated, and analyzed at the same concentration and under the same conditions as in (A). (C and D) The conditions were the same as in (B) except that 5 μ*M* (C) or 0.5 μ*M* (D) of unlabeled 5′ exon-G transcript was added, and samples were analyzed during the incubation period at 0, 1, 2, 3, and 4 hours. The linear intron generated in these trans-splicing experiments is longer than the nicked lariat intron in the cis-splicing experiment in (A) because of the vector-encoded sequences on the 5′ end of pSPTE (Fig. 1).

ly labeled, gives rise to the same radioactive intron product but no detectable radioactive ligated exons. This result reinforces the conclusion drawn from the inactivity of the intron–3′ exon molecules when incubated alone, namely that this short 5′ exon, containing only two nucleotides of the normal 5′ exon, does not participate in the splicing reactions.

The trans-splicing reaction is very efficient under these conditions, as the time course is almost as fast as the cis reaction (compare Fig. 2, A and C). The trans-reaction shown here takes place efficiently at relatively low concentrations of 5′ exon, namely, 0.5 μ*M* in Fig. 2D. Furthermore, the time course is quite fast at these relatively low concentrations of 5′ exon as a tenfold increase in 5′ exon concentration leads to considerably less than a tenfold increase in the apparent rate (compare Fig. 2C with Fig. 2D).

Products can also be visualized by incubating internally labeled 5′ exon-G RNA with an excess of nonradioactive intron–3′ exon (Fig. 3). As above, no reaction is detectable when the two molecules are incubated separately. The ligated exon product (Fig. 3) is the same size as that visible in Fig. 2, which reinforces the idea that it is the product of a trans-splicing event (because it is labeled by both substrates, the intron–3′ exon in Fig. 2 and the 5′ exon in Fig. 3).

To sequence the junction of the trans-splicing generated ligated exons, we labeled the product at the 5′ end. This was accomplished by incubating a 5′ end-labeled 5′ exon and nonradioactive intron–3′ exon, so that the resultant ligated exons product was 5′ end-labeled. (The gel pattern is identical to that in Fig. 3.) The ligated exons band was purified on a sequencing gel and subjected to standard sequence analysis (*22*). The results (Fig. 4) show that the band indeed represents the ligated exons and that the junction is correct. As the junction contains no extra G between the sequence of the two exons, the 3′ terminal G of the 5′ exon-G RNA must have been removed prior to exon ligation. This indicates that both the first step in splicing, 5′ junction cleavage, and the second step of splicing, 3′ junction cleavage and exon ligation, take place properly in this trans-splicing reaction.

Trans-splicing occurs without branch formation. If the trans-splicing reaction follows the same mechanistic pathway as the normal (cis) group II self-splicing reaction, cleavage at the 5′ junction should be due to nucleophilic attack by the 2′ OH of the normal branch site within the intron (*6, 7*). This should result in transfer of the 3′ terminal G from the 5′ exon-G molecule to the intron, in a 2′,5′-phosphodiester bond at the branch site. Yet, no radioactive intron band is visible at the expected position in Fig. 3 (well above the ligated exons). We believe this negative result is significant because, as there are only 17 G's in the 5′ exon-G molecule, transfer of one G to the intron should result in an intron band 1/16th the intensity of the ligated exons band, a signal easily detectable in the exposure shown.

To search for the missing G, the same splicing reaction was analyzed by TLC (Fig. 5A). Clearly visible is the time-dependent appearance of radioactive GMP (guanosine monophosphate). Similarly, if the 5′ exon-G molecule is end-labeled with [^{32}P]pCp and RNA ligase, and incubated with nonradioactive intron–3′ exon, a larger product, almost certainly pGpCp appears (Fig. 5B). In this last mentioned case, where the molecule is end-labeled, more than 50 percent of the radioactivity chromatographs in this spot after 3 hours of incubation. This correlates reasonably well with the ratio of 5′ exon-G to ligated exons observed after 3 hours of incubation (Fig. 3), suggesting that a major fate of the "extra" nucleotide at the end of the 5′ exon-G RNA is hydrolysis and release rather than transfer to the branch site within the intron.

The 5′ junction is cleaved by hydrolysis in trans-splicing. The previous observation can be explained by either of two hypotheses. Perhaps the 2′ OH of the branch site is the nucleophile as in normal cis splicing, but the branch structure is unstable, resulting in release of the G. Alternatively, 5′ cleavage takes place through nucleophilic attack by hydroxide ion (OH^-) or water, an attack that requires a prior interaction between the 5′ exon-G RNA and the intron–3′ exon. The former hypothesis suggests that 5′ cleavage should not take place when the 5′ exon-G RNA is incubated with lariat-intron, as the branch site is already occupied in a 2′,5′-phosphodiester bond. Thus, if 5′ cleavage does occur under these experimental conditions, the second hypothesis is likely to be correct.

To this end, nonradioactive lariat intron, purified from a prepara-

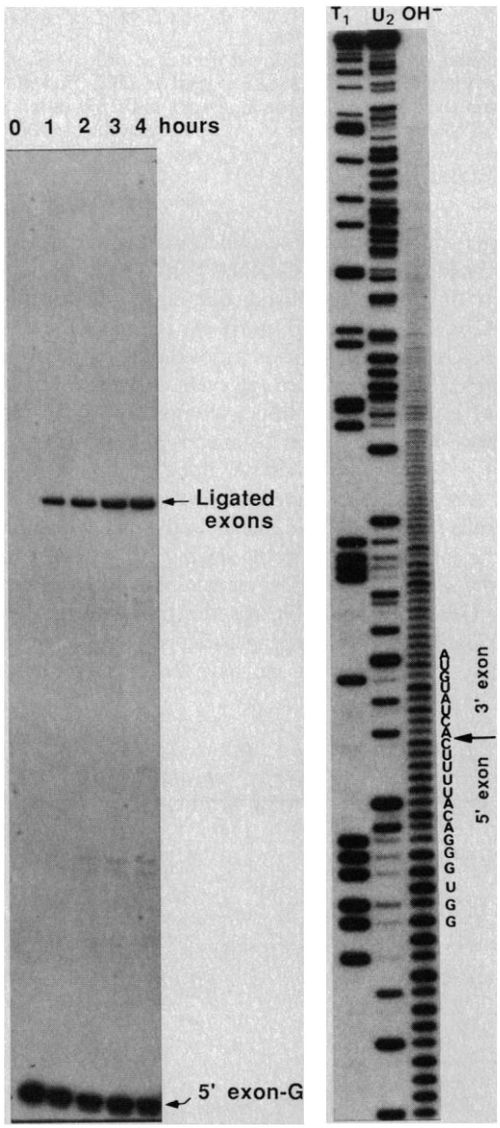

Fig. 3 (left). Trans-splicing with labeled 5′ exon-G. The 5′ exon-G was internally labeled with [α-^{32}P]GTP and incubated at 0.5 μ*M* with 2 μ*M* unlabeled intron-3′ exon transcript, as in Fig. 2. Samples (1 μl) were taken at 0, 1, 2, 3, and 4 hours, mixed with loading buffer (as in Fig. 2), and analyzed on an 8 percent acrylamide-urea gel. Fig. 4 (right). Sequence of the ligated exons generated by trans-splicing. The 5′ exon-G molecule, 5′ end-labeled with polynucleotide kinase and [γ-^{32}P]ATP, was incubated with unlabeled intron–3′ exon for 4 hours (Fig. 3). The ligated exons product was then purified from a 6 percent acrylamide-urea gel and sequenced (*22*). The sequence was analyzed on an 8 percent acrylamide-urea gel. The letters show the sequence as interpreted from the previously known DNA sequence (*19*) since the reactions used did not distinguish between the U and C. The arrow points to the 5′ exon–3′ exon junction. Lane T_1 is partial T_1 digestion; lane U_2 is partial U_2 digestion; and lane OH^- is partial alkali digestion.

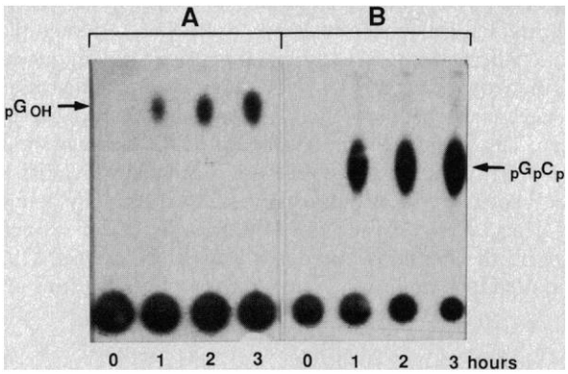

Fig. 5. Analysis of nucleotides cleaved from 5′ exon. (A) Samples from a reaction identical to that in Fig. 3 were spotted on a PEI (polyethyleneimine) cellulose thin-layer plate with chromatography in 0.25*M* sodium phosphate (*p*H 3.5). (B) Same conditions as in (A) except that the 5′ exon-G molecule was labeled at its 3′ end with [^{32}P]pCp and RNA ligase (*37*) instead of internally labeled with [α-^{32}P]GTP.

tive, cis, self-splicing reaction, was incubated with a limiting amount of 5′ end-labeled 5′ exon-G RNA. The results show the time-dependent appearance of a band, one nucleotide shorter than the 5′ exon-G RNA substrate (Fig. 6). As the molecule is 5′-labeled, the conversion must result from removal of the 3′ terminal G, that is, the proper 5′ junction cleavage event. Almost 50 percent of the 5′ exon-G is cleaved in 3 hours, comparable to the TLC assay shown above for the 5′ exon-G, intron–3′ lariat incubation. Incubation of 5′ exon-G with tRNA alone results in no detectable cleavage. Also, and as expected, incubation with radioactive lariat intron results in no detectable change in the migration of the intron, that is, the intron lariat is stable during the reaction. These data support the second hypothesis, namely that nucleophilic attack by water or OH^- is responsible for the trans-splicing, 5′ junction cleavage. The lariat intron appears to act as an enzyme that catalyzes 5′ junction hydrolysis, but we have not analyzed the turnover

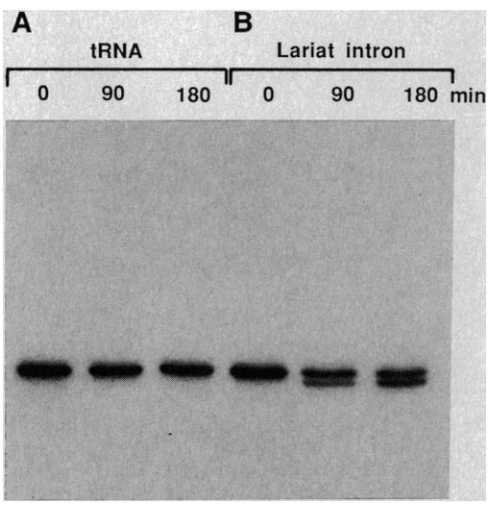

Fig. 6. Cleavage of 5′ exon-G by the intron. The lariat intron, generated by a self-splicing reaction carried out with about 25 μg of unlabeled full-length precursor, was purified on a 4 percent acrylamide gel. At a concentration comparable to the concentration of intron–3′ exon in Fig. 3 (that is, 2 μ*M*), the purified lariat intron was incubated with about 0.2 μ*M* 5′ exon-G that had been 5′ end-labeled with polynucleotide kinase and [γ-^{32}P]ATP. The reaction was incubated under the same conditions as in Fig. 3 for 0, 90, or 180 minutes and analyzed on an 8 percent acrylamide-urea gel (B). In a mock experiment (A), the same 5′ exon-G molecule was mixed with tRNA (0.2 μg/ml) (but no lariat intron) and incubated and analyzed in parallel.

number so this analogy is tenuous at present. The data also suggest that the 5′ exon:intron–3′ exon interaction, necessary for the first step of splicing, is in reality an interaction of the 5′ exon and intron.

The second step occurs efficiently without a branch or lariat. The above data indicate that the two intermediates in the trans-splicing reaction are the 5′ exon and linear intron–3′ exon. The second step of splicing, 3′ cleavage and exon ligation, apparently occurs efficiently in this trans-splicing reaction without any detectable lariat or branch structure. To test directly this assertion, we purified from a preparative incubation (similar to the analytical reaction presented in Fig. 6) 5′ exon molecules, that is, the 5′ exon-G RNA with the G removed. When incubated with nonradioactive, standard intron–3′ exon substrate, ligated exons are efficiently produced (Fig. 7). Since no 5′ junction cleavage takes place under these conditions, the first step is effectively bypassed, and thus no branch is formed. As the second step occurs efficiently, we conclude that the lariat, or branch, plays no obligatory role in 3′ junction cleavage or exon ligation. In fact, the reaction is markedly (about five times) faster with 5′ exon than with 5′ exon-G (Fig. 7), suggesting that 5′ cleavage is rate-determining in the trans-splicing reaction.

The 5′ exon–intron interaction precedes 5′ junction cleavage. The experiments shown above were designed to test for the existence of an interaction between the 5′ exon and the intron during group II self-splicing. Although the two-step pathway for splicing, also applicable to group II molecules, only dictates that a pairing site is needed at an intermediate stage (that is, to "hold" the 5′ exon to the intron–3′ exon), our experiments suggest that a binding site is also important prior to the first step, 5′ junction cleavage.

The efficiency of the trans-splicing, at relatively low concentrations of substrate, suggests that there exists an interaction, or binding, between the 5′ exon and intron–3′ exon. Moreover, the fact that the time course of the reaction does not greatly increase with a tenfold increase in the concentration of the 5′ exon indicates that the trans-splicing reaction is almost zero order with respect to 5′ exon concentration, like the cis reaction (Fig. 2). This suggests that binding, under these trans-splicing conditions, is not rate-limiting and that virtually all the intron–3′ exon transcript has already formed a complex with 5′ exon-G RNA before the first step of the reaction, 5′ junction cleavage. The similarity of the cis and trans rates suggests that the rate-limiting step follows binding and is similar in the cis- and the trans-splicing reactions. Furthermore, the observation that the rate of the trans-splicing reaction significantly increases when the first step is bypassed [when the 5′ exon rather than 5′ exon-G is used (Fig. 7)] suggests that the rate-limiting step occurs subsequent to binding but before 3′ junction cleavage and exon ligation. The fact that the intron–3′ exon does not accumulate, and is, in general, difficult to detect during self-splicing, is also in support of this argument. Perhaps a (normally) intramolecular rearrangement is required and is rate-limiting for 5′ junction cleavage.

In vitro trans-splicing has been observed for nuclear pre-mRNA's (*23, 24*). In these cases, and more generally, no available evidence suggests a 5′ exon binding site on the intron or 3′ exon. It is possible that exclusively intronic sequences are sufficient for 5′ junction recognition and that trans-acting factors in the spliceosome are responsible for maintaining the proper association of 5′ exon and 3′ exon after 5′ cleavage but before 3′ cleavage and exon ligation.

A 5′ exon binding site has been proposed in the group I self-splicing system (*8–10, 14*). Genetic and biochemical experiments indicate that this interaction is important both for an early stage

prior to 5′ cleavage (*15, 16, 25, 26*), and for a later stage in which the 5′ exon is held in place to attack the 3′ junction, thereby accomplishing the second step of the reaction. The demonstration of this latter pairing event is based on a series of elegant, group I, trans-splicing experiments in which oligoribonucleotides play the role of 5′ exon and bypass the first step (*16, 17*). This experimental paradigm is rather similar to the experiment (Fig. 6) in which a 5′ exon, when incubated with intron–3′ exon, induces cleavage and exon ligation, thereby generating a trans-reaction analog of the second splicing step. We have also generated a trans-splicing analog of the first splicing step, by incubating 5′ exon-G RNA with purified lariat intron. All of the above arguments indicate that the G is efficiently and accurately hydrolyzed as a consequence of the interaction between the 5′ exon-G and the intron. This reaction is somewhat similar to a trans-splicing experiment of group I, in which the first step of group I splicing is catalyzed in trans (*27*). Finally, we have generated a complete trans-splicing reaction in which proper 5′ cleavage, proper 3′ cleavage, and proper exon ligation take place. The fact that we have been able to dissect the group II splicing pathway into its two steps reinforces the validity of this pathway, as proposed by van der Veen *et al.* (*6*) and Peebles *et al.* (*7*).

The 5′ exon-G cannot bypass the first step. The ability to separate the pathway into its component parts reveals that the 5′ exon-G, unlike the 5′ exon, is unable to bypass the first step and induce directly 3′ exon cleavage and exon ligation. Similar conclusions were drawn from in vivo experiments with mutant yeast nuclear splicing substrates in which aberrant 5′ cleavage events did not continue through the second step (*28–30*). Thus, in both group II and yeast nuclear splicing, it appears that substantial specificity resides in the second step. With yeast nuclear splicing, however, the aberrant lariat structures generated along with the aberrant 5′ exons are likely to be responsible, at least in part, for the block after the first step of splicing in that other mutations, affecting only the branch structure, have a similar effect (*29–32*).

Role of branch formation. Group II splicing bears a fundamental relation to nuclear splicing as, in both cases, and in contrast to group I splicing, lariat molecules are generated. Indeed, the proposed pathway for group II splicing (*6, 7*) is identical to the well-established pathway for nuclear splicing (*2, 3*). This pathway includes the postulate that the 2′ OH of the branch site is the nucleophile that attacks the 5′ splice junction (*2*). Our finding that the trans-splicing reaction occurs apparently in the absence of branch formation was rather unexpected. It suggests that the accuracy of 5′ cleavage is not determined by attack of the branch site 2′ OH but, rather, that specific activation of the 5′ junction is separable from subsequent attack by the nucleophile, normally the branch site 2′ OH but for the trans-reactions shown here, water or OH^-. This possibility bears great similarity to the splicing of group I introns, based on the 5′ and 3′ splice site hydrolysis that occurs in the *Tetrahymena* pre-rRNA (*33*).

Branch formation may have an important role for the second step. Its proximity to the 3′ junction is consistent with this possibility and, indeed, in nuclear splicing, mutants that affect the branch structure have a profound effect on the rate of the second step (*29–32*). Yet, our trans-splicing experiments also show that branch formation is not required for an efficient second step. These results argue against a major enzymatic role of the branch structure for the second step, suggesting that the nuclear mutants referred to above may manifest an indirect effect on the second step, for example, by way of aberrant interactions with trans-acting factors.

It is likely that nucleophilic attack by the 2′ OH of the branch site in the normal cis reaction is indeed much more efficient than nucleophilic attack by water or OH^-. The aberrant structure of the trans-splicing complex precludes the use of the branch site 2′ OH and allows the use of the water or OH^- as the initiating nucleophile. Nucleophilic attack by water or OH^- in the trans-reaction might be much less efficient than nucleophilic attack by the branch site 2′ OH in the cis reaction, but this difference may be obscured because some other step in splicing, rather than nucleophilic attack at the 5′ junction, might be rate-limiting. The rate-determining step, as manifest in vitro in the absence of proteins, might be different in vivo. The 2′ OH of the branch site might be a more efficient nucleophile because of the structure of the active site; for example, the branch site may be positioned appropriately so that nucleophilic attack is optimized. Alternatively, or in addition, competing nucleophiles such as water or OH^- may normally have very low local concentrations. This latter possibility may be especially important for the 3′ junction where attack by a nucleophile other than the 3′ OH of the 5′ exon would prevent exon ligation. In any case, it is likely that group I, group II, and nuclear splicing all share the feature that 5′ cleavage normally takes place by transesterification rather than hydrolysis.

Recent DNA sequence data from the chloroplast genome of two species, *Nicotiana* (*34, 35*) and *Marchantia* (*36*), indicate that group II trans-splicing may normally occur between the first and second exons of the gene coding for the chloroplast ribosomal protein S12.

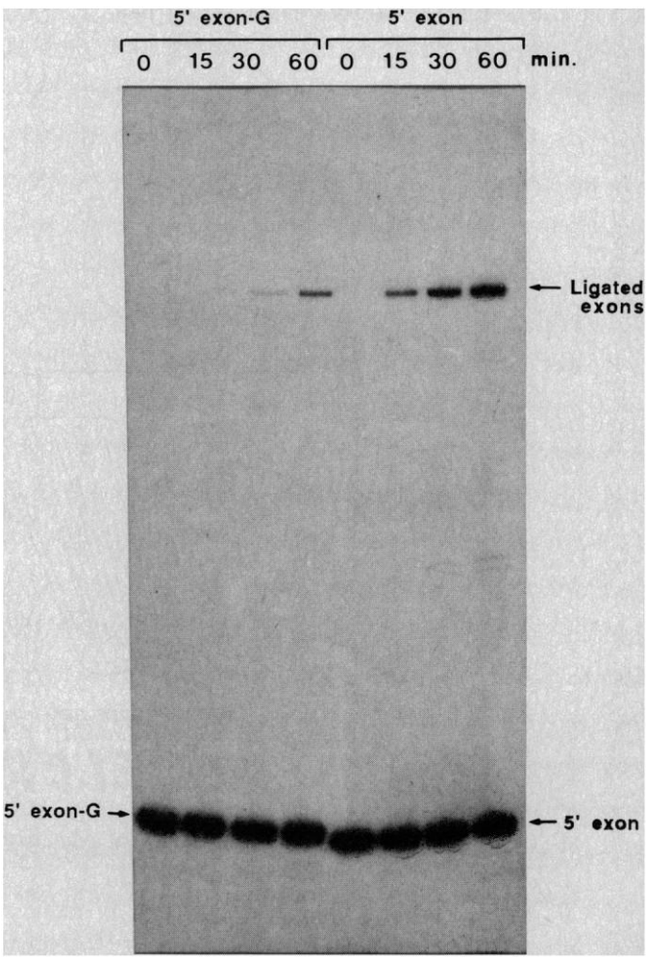

Fig. 7. Comparison of the time course of trans-splicing with 5′ exon and 5′ exon-G. The 5′ end-labeled 5′ exon was purified from a gel similar to the one shown in Fig. 6. It was incubated for 0, 15, 30, and 60 minutes in the presence of unlabeled intron–3′ exon as described in Fig. 3. The same amount of 5′ exon-G molecule (the same number of counts per minute of the labeled 5′ exon-G preparation from which the 5′ exon was derived) was incubated with intron–3′ exon in an identical manner and analyzed in parallel.

Although these data do not speak to the mechanism involved in this in vivo trans-splicing, an interesting possibility is that it occurs through a pathway similar to the one we describe for this in vitro reaction. It would be of particular interest to know the structure of the intron released by this in vivo trans-splicing reaction since it might suggest the nature of the attacking nucleophile involved in the first step. It is even possible that transcription creates a perfect 5′ exon and that only the second step of the trans-splicing event takes place.

We recognize that an interaction between the 5′ exon and intron is implied but not directly demonstrated by these studies. The sequences that interact and the nature of the interaction remain to be defined.

REFERENCES AND NOTES

1. C. L. Greer, C. L. Peebles, P. Gegenheimer, J. Abelson, *Cell* **32**, 537 (1983).
2. For review, R. A. Padgett, P. J. Grabowski, M. M. Konarska, S. Seiler, P. A. Sharp, *Annu. Rev. Biochem.* **55**, 1119 (1986).
3. For review, M. R. Green, *Annu. Rev. Genet.* **20**, 671 (1986).
4. K. Kruger *et al., Cell* **31**, 147 (1982).
5. A. J. Zaug, P. J. Grabowski, T. R. Cech, *Nature (London)* **301**, 578 (1983).
6. R. van der Veen *et al., Cell* **44**, 225 (1986).
7. C. L. Peebles *et al., ibid.*, p. 213.
8. F. Michel, A. Jacquier, B. Dujon, *Biochimie* **64**, 867 (1982).
9. R. W. Davies, R. B. Waring, J. A. Ray, T. A. Brown, C. Scazzocchio, *Nature (London)* **300**, 719 (1982).
10. F. Michel and B. Dujon, *EMBO J.* **2**, 33 (1983).
11. T. R. Cech and B. L. Bass, *Annu. Rev. Biochem.* **55**, 599 (1986).
12. T. R. Cech, A. J. Zaug, P. J. Grabowski, *Cell* **27**, 487 (1981).
13. S. M. Mount, I. Petterson, M. Hinterberger, A. Karmas, J. A. Steitz, *ibid.* **33**, 509 (1983).
14. R. B. Waring and R. W. Davies, *Gene* **28**, 277 (1984).
15. R. B. Waring, P. Towner, S. J. Minter, R. W. Davies, *Nature (London)* **321**, 133 (1986).
16. M. D. Been and T. R. Cech, *Cell*, in press.
17. T. Inoue, F. X. Sullivan, T. R. Cech, *ibid.* **43**, 431 (1985).
18. G. Garriga, A. M. Lambowitz, T. Inoue, T. R. Cech, *Nature (London)* **322**, 86 (1986).
19. S. E. Bonitz, G. Coruzzi, B. E. Thalenfeld, A. Tzagologg, G. Macino, *J. Biol. Chem.* **255**, 11927 (1980).
20. D. A. Melton *et al., Nucleic Acids Res.* **12**, 7035 (1984).
21. A. Jacquier and M. Rosbash, unpublished results.
22. H. Donis-Keller, A. M. Maxam, W. Gilbert, *Nucleic Acids Res.* **4**, 2527 (1977).
23. M. M. Konarska, R. A. Padgett, P. A. Sharp, *Cell* **42**, 165 (1985).
24. D. Solnick, *ibid.*, p. 157.
25. M. B. Been and T. R. Cech., *Nucleic Acids Res.* **13**, 8389 (1985).
26. F. X. Sullivan and T. R. Cech, *Cell* **42**, 639 (1985).
27. J. W. Szostak, *Nature (London)* **322**, 83 (1986).
28. R. Parker and C. Guthrie, *Cell* **41**, 107 (1985).
29. A. Jacquier, J. R. Rodriguez, M. Rosbash, *ibid.* **43**, 423 (1985).
30. L. A. Fouser and J. D. Friesen, *ibid.* **45**, 81 (1986).
31. A. Jacquier and M. Rosbash, *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 5835 (1986).
32. A. J. Newman, R.-J. Lin, S.-C. Cheng, J. Abelson, *Cell* **42**, 335 (1985).
33. T. Inoue, F. X. Sullivan, T. R. Cech, *J. Mol. Biol.* **189**, 143 (1986).
34. H. Fromm, M. Edelman, B. Koller, P. Goloubinoff, E. Galun, *Nucleic Acids Res.* **14**, 883 (1986).
35. K. Torazawa, N. Hayashida, J. Obokata, K. Shinozaki, M. Sugiura, *ibid.*, p. 3143.
36. H. Fukuzawa *et al., FEBS Lett.* **198**, 11 (1986).
37. T. E. England, A. G. Bruce, O. C. Uhlenbeck, *Methods Enzymol.* **65**, 65 (1980).

38. We thank Hildur Colot and Bernard Dujon for comments and our collaborators, especially Claudio Pikielny, for their advice in the course of these experiments. Supported by NIH (GM23549).

19 August 1986; accepted 4 November 1986