MOLECULAR AND CELLULAR BIOLOGY, Feb. 1993, p. 1137–1145
0270-7306/93/021137-09$02.00/0

Vol. 13, No. 2

An Alternative Helix in the 26S rRNA Promotes Excision and Integration of the *Tetrahymena* Intervening Sequence

SARAH A. WOODSON* AND VICTORIA L. EMERICK

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20740-2021

Received 1 September 1992/Returned for modification 9 November 1992/Accepted 23 November 1992

A highly conserved ribosomal stem-loop immediately upstream of the *Tetrahymena* splice junction can inhibit both forward and reverse self-splicing by competing with base pairing between the 5′ exon and the guide sequence of the intervening sequence. Formation of this unproductive hairpin is preferred in precursor RNAs with short exons and results in a lower rate of splicing. Inhibition of self-splicing is not observed in longer precursors, suggesting that additional interactions in the extended exons can influence the equilibrium between the productive and unproductive hairpins at the 5′ splice site. An alternative pairing upstream of the 5′ splice site has been identified and is proposed to stabilize the active conformer of the pre-rRNA. Nucleotide changes that alter the ability to form this additional helix were made, and the self-splicing rates were compared. Precursors in which the proposed stem is stabilized splice more rapidly than the wild type, whereas RNAs that contain a base mismatch splice more slowly. The ability of DNA oligomers to bind the RNA, as detected by RNase H digestion, correlates with the predicted secondary structure of the RNA. We also show that a 236-nucleotide RNA containing the natural splice junction is a substrate for intervening sequence integration. As in the forward reaction, reverse splicing is enhanced in ligated exon substrates in which the alternative rRNA pairing is more stable.

There is growing evidence that splice site selection during pre-RNA processing goes beyond recognition of simple sequences. For example, in splicing of pre-mRNAs, it has become clear that splice site selection not only involves the presence of consensus splice site sequences but also is modulated by the differential binding of both constitutive and specific splicing factors to particular pre-mRNAs (14, 23). In addition to regulation through binding of protein factors, there are several examples showing that formation of higher-order structure in the RNA regulates the use of alternative processing sites (6, 9, 10, 22). In protein-dependent systems, the effects of RNA conformation may be transmitted through RNA-protein interactions. In protein-independent systems, activation of a particular processing site can be directly attributed to the structure of the RNA.

We have been investigating the way in which exon RNA structure can influence 5′ splice site recognition during self-splicing of *Tetrahymena* pre-rRNA. In splicing of group I introns, addition of GTP to the 5′ splice site during the first step of the reaction requires base pairing between nucleotides in the 5′ exon with the internal guide sequence (IGS) of the intervening sequence (IVS) (5). Base pairing between the IGS and nucleotides preceding the splice junction is also required for reverse splicing (34). This helical stem is conserved among group I introns and is commonly termed P1. In *Tetrahymena thermophila*, a stable rRNA stem-loop immediately upstream of the 5′ splice site, which we have designated P(−1), involves the same nucleotides that base pair with the IGS during splicing. Consequently, formation of P(−1) directly competes with formation of P1 and can inhibit activation of the 5′ splice site. In short precursor RNAs, P(−1) is preferred, resulting in a 20-fold decrease in the observed rate of self-splicing relative to transcripts that do not contain the sequences of P(−1) (35). In a similar manner, the presence of P(−1) in the ligated exons inhibits integration of the IVS via reverse splicing (35).

We have recently shown that potential inhibition due to P(−1) is apparently relieved in longer *Tetrahymena* pre-RNAs. Longer precursors of natural sequence self-splice more rapidly (32) and at rates similar to that of pre-rRNA isolated from *Tetrahymena* nuclei (2). As many as 146 nucleotides (nt) of the 5′ exon and 86 nt of the 3′ exon are required for optimal splicing in vitro. This effect depends on the specific sequences of the rRNA, since point mutations within this region can result in a 50-fold decrease in activity. Interestingly, the necessary 5′ and 3′ exon sequences correspond to most of the highly conserved domain IV of the 26S rRNA (27). The ribosomal exon sequences are expected to contain many stable and specific RNA-RNA interactions, even in the absence of protein. Our results imply that interactions between the exons and nucleotides near the 5′ splice site alter the equilibrium between P(−1) and P1 and thus modulate the observed rate of self-splicing.

One simple mechanism by which additional exon sequences might destabilize P(−1), and promote formation of P1, is by providing an additional sequence that is complementary to the 5′ strand of P(−1). As drawn on the left side of Fig. 1, formation of the P(−1) stem-loop sequesters the 5′ exon from binding to the guide sequence of the IVS, and this conformer does not lead to spliced products. On the other hand, the presence of an alternative pairing partner for the upstream strand of P(−1) would stabilize active precursors containing P1, as shown on the right. In this model, two duplexes, P1 and a proposed PX pairing, replace P(−1) in the pre-RNA. Following excision of the IVS, the rRNA could refold to adopt the phylogenetically conserved P(−1) stem-loop.

We have used computer-generated secondary structures for the exon sequences to identify a possible alternative PX pairing within the 5′ exon. The requirement for this alternative helix in forward and reverse self-splicing was tested by introducing nucleotide changes that either destabilize or

* Corresponding author.

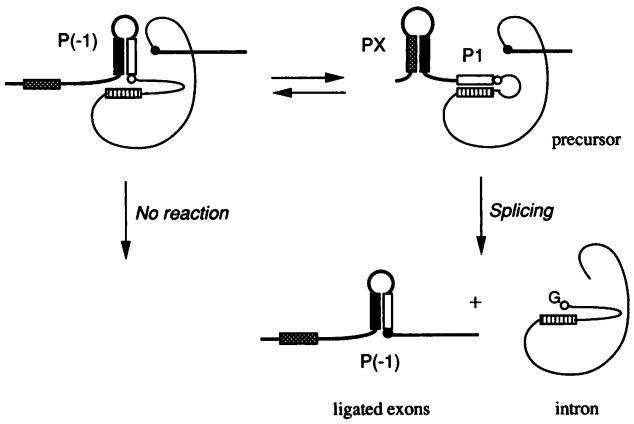


FIG. 1. Model for alternative pairings in the *Tetrahymena* pre-rRNA. A precursor RNA can adopt two alternative conformations, as shown at the top: either the inactive structure containing P(−1) (left) or the active structure containing P1 (right). P(−1) involves nucleotides in the 5′ exon (open box) that also pair with the IGS (striped box) in P1. The active conformer is proposed to be additionally stabilized in long pre-RNAs by the presence of another sequence (cross-hatched box) that can form an alternative helix, PX, with the upstream strand of P(−1) (solid box). Thin lines, IVS; thick lines, exons; open circle, 5′ splice site phosphodiester; solid circle, 3′ splice site phosphodiester.

restore the ability to form this pairing in pre-RNAs. In accordance with the model, nucleotide changes that destabilize the proposed pairing result in lower rates of splicing, while changes that retain or enhance the ability to form PX should result in greater activity. Differences in the binding of RNA and DNA oligonucleotides to mutated precursor RNAs provide additional evidence for the role of this RNA duplex in stabilizing active pre-RNA conformers.

MATERIALS AND METHODS

Computation of RNA secondary structures. The Macintosh version of the computer program MULFOLD was the generous gift of Michael Zuker and is described elsewhere (19, 39). Secondary structures were computed for 246 nt of *Tetrahymena* pre-rRNA sequence, beginning 147 nt upstream of the 5′ splice site, including the first 26 nt of the IVS (all of P1) and the last nucleotide of the IVS, G414, and ending 70 nt downstream of the 3′ splice site. The remaining nucleotides of the IVS were not included in the calculation, via an open excision option in the program. To obtain secondary structures that did not contain the P(−1) stem-loop, computations were carried out with a double-prohibition restraint whereby formation of base pairs between the strands of P(−1) was assigned an additional free-energy penalty. Between 5 and 10 suboptimal structures were examined for each run.

Plasmid DNA and mutagenesis. The 1,299-nt wild-type precursor RNA is encoded by plasmid pJK43-T7 (29, 32). The precursor RNA includes *Tetrahymena* rRNA sequences from the *Hin*dIII site at position 6749 in the rDNA (11) to the *Eco*RI site at position 8047. The intron begins at position 7010 in the rDNA, or position 2261 in 26S rRNA. The 1.3-kb *Hin*dIII-*Eco*RI fragment of JK43-T7 was ligated into the polylinker region of pTZ18U (Bio-Rad), and this plasmid was used to prepare single-stranded DNA as described previously (32).

Oligonucleotide-directed mutagenesis was carried out by the method of Kunkel et al. (20). Plasmid DNA containing the desired mutations was identified by dideoxy sequencing, using Sequenase 2.0 (U.S. Biochemical). The mutated 1.3-kb *Hin*dIII-*Eco*RI fragment was recloned into pTZ19U for expression of pre-RNA. Precursors U2218G and U2218C are encoded by plasmids JKU2855G and JKU2855C, respectively. Other mutations introduced are 26C, −6G, and −16C. Triple and quadruple base changes were not obtained in a single mutagenesis reaction. The mutation at position 26 of the IVS, JK43G26C, was introduced first, and this plasmid was used to prepare single-stranded template DNA for subsequent rounds of mutagenesis.

Plasmid SW012LE was constructed by deleting the IVS sequences from pSW012 (32) by single-strand mutagenesis. This plasmid gives rise to a 244-nt transcript that includes 26S rRNA sequences from positions 2116 to 2346 and 13 nt of plasmid sequence at the 5′ end. The sequence of the deleted plasmid DNA was confirmed by dideoxy sequencing. This plasmid contains an f1 phage origin and was used to prepare single-stranded DNA. Plasmids pSW012LE:U2218C and pSW012LE:U2218G were obtained from oligonucleotide-directed mutagenesis of pSW012 as described above.

Self-splicing reactions. Uniformly radiolabeled precursor RNAs were prepared by in vitro T7 RNA polymerase (7) transcription of linearized plasmid DNA in the presence of [α-^{32}P]ATP (New England Nuclear) and isolated from 8 M urea–4% polyacrylamide gels as described elsewhere (32, 35). Splicing reactions were carried out in 100 mM ammonium sulfate–50 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.5)–5 mM $MgCl_2$–100 μM GTP at 30°C as previously described (32). Prior to the addition of GTP, the RNA was heated to 95°C for 1 min and cooled rapidly in the presence of splicing buffer (31). Products were separated on 4% polyacrylamide gels, and the radioactivity in each lane of the gel was quantitated with a Molecular Dynamics PhosphorImager.

The initial observed rates of splicing were determined from linear fits to $\ln(1 - f_{sp})$ versus time, where f_{sp} is the fraction of spliced products, as deduced by the fraction of ligated exon RNA in each lane. The data were treated in the same manner as previously (32). Typically, at least half of the RNA reacts during the initial linear portion of the curve, and greater than 90% of the precursor RNA has reacted after 4 h of incubation. Linear fits to individual data sets are normally precise, but observed rates can vary 10 to 20% between preparations of an RNA. Reported values reflect several (two to five) isolations and rate determinations for a given precursor.

***trans*-splicing reactions.** Splicing reactions were carried out as described above, but with the addition of 0, 0.1, 1, or 10 μM unlabeled 5′ exon RNA and 0.5 mM GTP. The 8-nt 5′ exon RNA, 5′-rGGCUCUCU, was prepared as described previously (34). After 5 min of incubation at 30°C, an equal volume of buffer containing 10 M urea was added, and the samples were electrophoresed on an 8 M urea–4% polyacrylamide gel.

Preparation of 5′-end-labeled RNA. Nonradioactive RNA was prepared by T7 RNA polymerase transcription of 2.5 μg of linear plasmid DNA in a 0.5-ml volume as described previously (36, 38), and the RNA was precipitated with ethanol. Unincorporated nucleoside triphosphates (Pharmacia) were removed by passing the mixture over a 5-ml Sephadex G-50 column (Pharmacia). The RNA was then 5′ end labeled with T4 polynucleotide kinase (U.S. Biochemical) and [γ-^{32}P]ATP (New England Nuclear) after treatment with calf intestinal phosphatase (Promega) and extraction

with phenol and chloroform. The labeled precursor or ligated exon RNA was isolated from a 4% polyacrylamide gel, soaked from the gel matrix, and precipitated with ethanol twice to remove contaminating urea and salts.

RNase H digestion. Radiolabeled precursor RNA in 20 mM Tris-HCl (pH 7.5) was incubated at 95°C for 1 min and then rapidly cooled to room temperature in the presence of 50 mM NaCl and 5 mM $MgCl_2$ (final concentrations). This renatured RNA mixture was diluted fivefold into RNase H digestion reaction mixtures containing 50 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 100 mM NaCl, 0.1 mM dithiothreitol, 10 μg of bovine serum albumin per ml, and 0, 1, 10, or 100 μM deoxyoligonucleotide (8). The reaction was started with the addition of 0.4 U of RNase H (U.S. Biochemical), incubated 5 min at 10 or 30°C, and stopped with the addition of gel loading buffer containing 10 M urea. Reactions were also initiated with the addition of deoxyoligonucleotide and gave identical results. Incubation in the absence of DNA but the presence of RNase H buffer resulted in some hydrolysis of the precursor at the 5′ and 3′ splice sites that could not be entirely avoided; these species were not present, however, in the gel-purified starting material. Products were electrophoresed on 4% polyacrylamide gels with 5′-end-labeled ϕX/*Hae*III DNA markers (New England Biolabs).

Reverse splicing reactions. Ligated exon RNA was prepared by T7 RNA polymerase transcription of linear pSW012LE (wild type), pSW012LE:U2218C, or pSW012LE:U2218G plasmid DNA and radiolabeled as described above. Reverse splicing reactions were carried out in 100 mM ammonium sulfate–50 mM HEPES (pH 7.5) with 5′-end-labeled ligated exon RNA and linear IVS as previously (34) except that the concentration of ligated exon substrate was approximately 0.5 μM. Following the addition of 25 mM $MgCl_2$, samples were incubated at 42°C for up to 2 h, quenched with the addition of 10 M urea, and electrophoresed on a 4% polyacrylamide gel. The extent of reaction was quantitated with a Molecular Dynamics PhosphorImager.

RESULTS

Identification of an alternative base pairing in the 5′ exon. The suboptimal RNA folding program MULFOLD (18, 19, 39) was used to search for exon sequences that are complementary to the upstream strand of P(−1). A series of secondary structures was generated for *Tetrahymena* pre-rRNA sequences beginning 147 nt upstream of the IVS and ending 70 nt downstream of the 3′ splice site. The first 26 nt of the IVS (all of P1) were included in the calculation. The remainder of the IVS was excluded from the computation in order to reduce the complexity of the structures. The exon sequences used in the computer simulation correspond to those shown to be necessary for efficient self-splicing (32).

If the folding simulation was carried out such that P(−1) was not permitted to form, the 5′ strand of P(−1) (solid box in Fig. 1) was paired with several different sequences within the exons. One candidate for an alternative pairing partner was the sequence 5′-UCUUUA, beginning 75 nt upstream of the 5′ splice site. The location of this sequence in the phylogenetically predicted secondary structure for this region of the 26S rRNA (16, 27) is depicted by a cross-hatched box in Fig. 2. The two strands of P(−1) are symbolized by solid and open boxes, as in Fig. 1. Preliminary data from nuclease protection studies on in vitro-transcribed pre-RNAs are largely consistent with the phylogenetic secondary structure (33). In the phylogenetic folding, the cross-hatched sequences are part of a weak, but conserved,

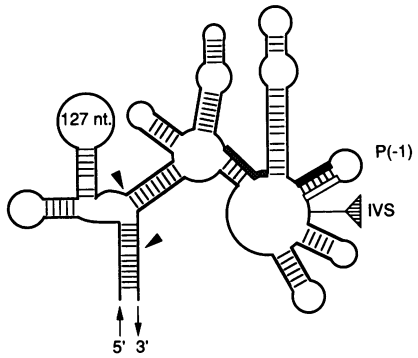
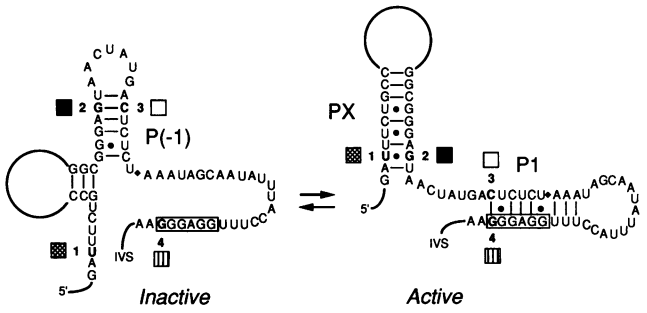


FIG. 2. Phylogenetically predicted secondary structure for domain IV of large-subunit rRNAs (adapted from reference 16). Heavy lines represent the RNA chain; thin lines represent base pairs. The *Tetrahymena* splice junction is marked IVS. Sequences that form the P(−1) stem-loop are keyed with solid and open boxes as in Fig. 1. The cross-hatched box denotes sequences complementary to the solid strand of P(−1). Arrowheads denote the limits of the exon sequences required for optimal self-splicing activity (32) and correspond to the substrate used in the reverse splicing experiments shown in Fig. 7. The exons of the long precursors used in the forward splicing experiments extend well past the 3′ limit of the figure.

long-range pairing that may or may not occur in vitro. In the alternative structure, formation of PX involves hydrogen bonding of bases in the solid box and the cross-hatched box. The cross-hatched sequence is adjacent to another conserved G · C stem that brings it close to the base of P(−1) in the folded RNA. The PX stem can be viewed as an extension of this duplex and is thus expected to be stabilized by continuous base stacking with the neighboring helix.

Self-splicing of precursor RNAs with base changes in PX. To experimentally test whether this proposed pairing in the rRNA has any effect on self-splicing of the *Tetrahymena* IVS, base changes that either stabilize or destabilize the PX pairing relative to the wild type were made in a 1.3-kb precursor RNA transcribed from plasmid pJK43-T7 (29, 32). A uridine 79 nt upstream of the 5′ splice site, U2218, was changed to either a G or a C. This position was also selected to avoid gross disruption of the other weak long-range pairing (Fig. 2). A partial sequence of the 5′ exon is shown in Fig. 3, and U2218 corresponds to site 1 in that drawing. U2218 lies within the cross-hatched box in Fig. 2 and should not directly alter the stability of P1 or P(−1). In precursor U2218G, the wild-type U · G pair is changed to a G · G mismatch, destabilizing PX. According to our simple model, this mutation should result in a decreased rate of splicing. In precursor U2218C, PX is strengthened by the substitution of a C · G pair for the U · G wobble pair. This precursor is predicted to splice more readily than the wild type. The free energies of formation at 37°C for the PX duplexes were estimated according to Freier et al. (12) and are approximately −7.5 kcal (1 kcal = 4.184 kJ)mol for the wild type, −10.1 kcal/mol for U2218C, and −6.3 kcal/mol for U2218G.

The precursor RNAs were incubated in the presence of 100 μM GTP at 30°C under standard splicing conditions (35), and the data are plotted in Fig. 4A. Initial observed rates of splicing are listed in Fig. 3. As expected, precursor U2218G, in which the PX pairing is disrupted, splices more slowly than the wild type (0.11 min^{-1} versus 0.39 min^{-1}). U2218C pre-RNA, in which PX is more stable than the wild type, splices more rapidly at early times (0.75 min^{-1}). Only 55% of



RNA (1 2 3 4)	PX	Splicing
U G C G (wt)	+	0.39 min^{-1}
C G C G	++	0.75
G G C G	–	0.11
G C G C	++	0.46
C C G C	–	0.08
U C G C	–	0.12
G U G C	+	0.42
U U G C	–	0.26

FIG. 3. Mutations that alter the stability of PX in RNA precursors. Alternative secondary structures of the *Tetrahymena* pre-rRNA, including the proposed PX pairing, are diagrammed at the top. Base changes were made at four sites, indicated by the numbers 1 to 4 and by bold type in the sequence. The locations of the sequences are keyed by the squares as in Fig. 1. The base at each of sites 1 to 4 is listed under the first column for each precursor. Note that in the conformer on the left, sites 2 and 3 pair with each other, whereas in the structure on the right, sites 1 and 2 and sites 3 and 4 pair with each other. Nucleotide changes were introduced such that the complementarity of P(−1) and P1 was maintained. The PX pairing was either destabilized with a base mismatch, indicated by a minus sign in the second column, maintained with a G · U pair (+), or further stabilized by a G · C pair (++). The observed rates of splicing at initial times are listed in the third column and were measured as described in Materials and Methods. Rates can vary 10 to 20% among preparations of a given RNA. wt, wild type.

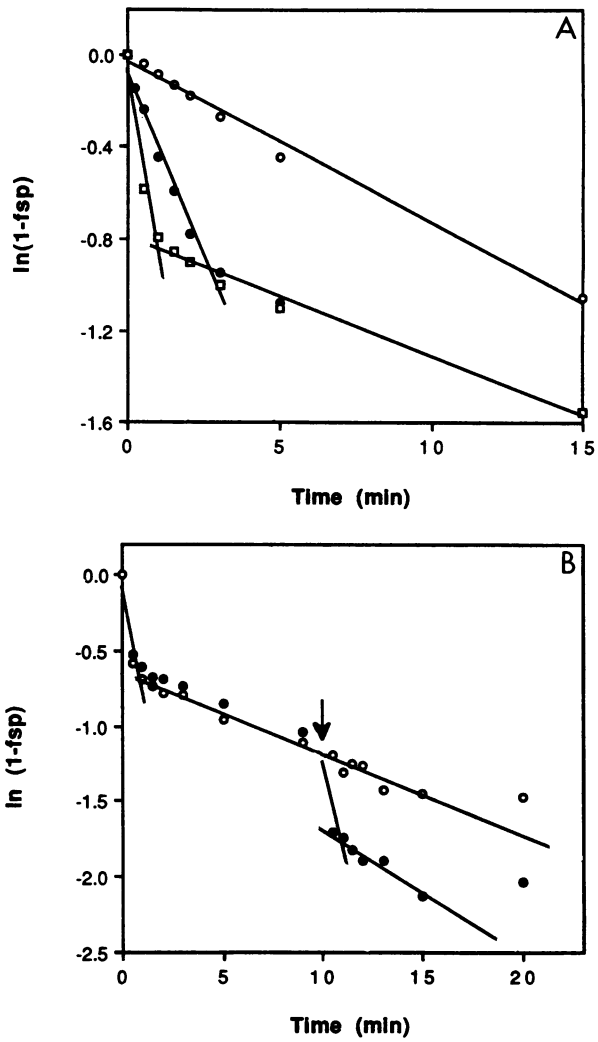


FIG. 4. Self-splicing of precursor RNAs. (A) Precursor RNA was incubated at 30°C in the presence of 100 μM GTP under standard splicing conditions as described in Materials and Methods. The $\ln(1 - f_{sp})$, where f_{sp} is the fraction of spliced products, is plotted versus time (minutes). Closed circles, wild-type RNA; open circles, U2218G; squares, U2218C. Lines indicate linear fits to the data. In the case of U2218C, data at early and late times were fit independently. (B) Two duplicate splicing reactions were carried out on U2218C pre-RNA as in panel A. After 9.5 min, one reaction mixture (closed circles) was heated for 30 s to 95°C then returned to 30°C. The other reaction (open circles) remained at 30°C throughout. Lines represent linear fits to the data points as shown. The fraction of spliced product at 10 min was interpolated from the data represented by open circles.

the precursor RNA, however, reacts during this initial burst. The remainder of the RNA splices much more slowly, at a rate similar to that of precursor U2218G. The observation of two distinct splicing rates for the C mutant suggested that there are two populations of RNA molecules that do not interconvert rapidly at 30°C.

To test whether the rapidly splicing and less active populations of precursor U2218C are conformationally related, duplicate splicing reactions were prepared. Following our standard protocol, the RNA was heated to 95°C and rapidly cooled in the presence of 5 mM $MgCl_2$ prior to the addition of GTP. After the splicing reactions reached the slow phase, one reaction mixture was again heated to 95°C for 30 s and rapidly cooled. As can be seen in Fig. 4B, the remaining precursor RNA again gave rise to a highly reactive component, in approximately the same proportion as in the original mixture. These results suggest that the U2218C precursor is found in two kinetically distinct conformations that are in slow exchange at 30°C.

Compensatory base changes in PX restore splicing activity. To further test whether these alternative helices influence 5′ splice site activation, a series of base changes in the 5′ exon and IGS (Fig. 3) was made. Initially, we wished to test whether a compensatory mutation in the other strand of PX could restore the self-splicing activity of U2218G to wild-type levels. To maintain the ability to form P(−1) and P1, two additional base changes were made in the 5′ exon and the IGS (sites 3 and 4 in Fig. 3). The resulting quadruple mutant, designated GCGC in Fig. 3, has the ability to form all three double helices, just like the wild-type and U2218C pre-RNAs. This precursor splices more rapidly than precursor U2218G and at a rate (0.46 min^{-1}) similar to that of the wild type at initial times. As in the case of U2218C (CGCG), only half of the GCGC RNA reacts with this rate constant, whereas the remaining portion of the starting material splices more slowly.

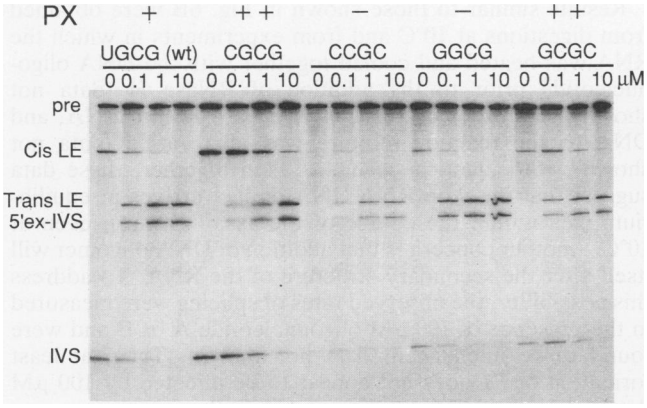


FIG. 5. Competition between intermolecular and intramolecular splicing. Uniformly radiolabeled precursor RNA (20 nM) was incubated 5 min at 30°C in the presence of 0.5 mM GTP and unlabeled 5′ exon oligomer 5′-rGGCUCUCU. The micromolar concentration of RNA oligomer added to the reaction is listed above each lane. Sequence changes introduced into the pre-RNA are indicated above the lanes and are listed in the same manner as in Fig. 3. The ability of each precursor to form PX is indicated with a minus or plus, also as in Fig. 3. The reaction products were fractionated on a 4% polyacrylamide gel. pre, precursor; Cis LE, *cis*-spliced ligated exons; Trans LE, *trans*-spliced ligated exons; 5′ex-IVS, 5′ exon-IVS RNA; IVS, excised intron. Cis LE and the IVS are the products of intramolecular splicing. Trans LE and 5′ex-IVS are the products of intermolecular splicing.

To ensure that the observed rates are not due to the base changes at positions 2, 3, and 4 per se, the nucleotide at position 1 in the GCGC quadruple mutant was changed to a C, to give CCGC. As expected, this precursor splices slowly, at a rate of 0.08 min^{-1}. Thus, the ability to self-splice rapidly does not correlate strongly with the particular nucleotides at each position but rather appears to depend on the relative stability of PX. This situation is similar to the previous observation that the sequence of P1 can be altered without loss of activity as long as base complementarity is maintained (3, 26). Another set of experiments in which the U · G pair of PX was changed to G · U was carried out in a similar fashion. As before, this precursor reacted at a rate (0.42 min^{-1}) similar to that of the wild type, whereas precursors containing U · C or U · U mismatches spliced more slowly (Fig. 3).

Competition between *trans* and *cis* self-splicing. The relative ability of a precursor RNA to undergo inter- and intramolecular splicing reactions has been used previously to assay the equilibrium between P(−1) and P1 (35). Precursors that preferentially form P(−1), and are thus normally unreactive in intramolecular splicing, can readily bind an exogenous 5′ exon RNA and undergo intermolecular exon ligation, or *trans* splicing (17, 35). Conversely, in precursors in which P1 is more stable than P(−1), the IGS is not available to base pair with the free 5′ exon RNA, and the *cis*-spliced products predominate.

This approach was used to assess the equilibrium between P(−1) and P1 in pre-RNAs in which the stability of PX has been altered. Uniformly radiolabeled precursor RNAs were incubated in the presence of 0.5 mM GTP and various concentrations of an 8-nt free 5′ exon RNA of sequence 5′-rGGCUCUCU. The results are shown in Fig. 5. The sequence of each pre-RNA is represented above the lanes in the same manner as in Fig. 3, and the relative stability of PX is indicated by a plus or minus.

In agreement with our model, precursors that can form a stable PX duplex, such as U2218C (CGCG), undergo intramolecular splicing more readily than intermolecular splicing, even with 1 μM exogenous 5′ exon RNA. On the other hand, pre-RNAs in which PX is disrupted, such as U2218G (GGCG), *cis* splice poorly but do give rise to intermolecularly spliced products at 0.1 μM oligomer. Thus, in these precursors, P(−1) appears to be preferred over P1, resulting in a lower rate of G addition to the 5′ splice site and an increase in the availability of the IGS to bind an exogenous RNA. The *trans*-splicing experiments, moreover, demonstrate that the reduced activity of these precursors is not due to misfolding of the IVS itself and that the 3′ splice site can still be activated for exon ligation.

Probing alternative structures with RNase H digestion. Another means of measuring changes in RNA secondary structure is to compare the extent to which deoxyoligonucleotides bind the RNA, using RNase H to detect RNA-DNA hybrids. This method has been used to distinguish between alternative secondary structures (1) and to investigate tRNA-rRNA interactions (24). As diagrammed in Fig. 6A, two DNA probes were used to assay the structure of wild-type and mutant precursor RNAs. Ideally, the length of the probes should be such that the DNA-RNA helix is adequate to permit partial digestion by RNase H but is itself not stable enough to alter the secondary structure of the RNA. Oligonucleotide A is complementary to the upstream strand of PX and is expected to bind only the inactive conformer. Precursors that can form a stable PX duplex should be cleaved to a lesser extent in the presence of a given concentration of DNA than are precursors in which PX is disrupted.

5′-end-labeled wild-type, U2218G, and U2218C pre-RNAs were subjected to RNase H cleavage in the presence of 0, 1, 10, or 100 μM oligonucleotide A at 30°C (Fig. 6B). As expected, precursor U2218C, in which the PX pairing should be most stable, is more resistant to cleavage than is the wild type. Less expected is the observation that the wild-type and U2218G precursors are digested to similar extents. One possible complication is that a different DNA probe was used for each precursor, since oligonucleotide A anneals to position 2218. The observed differences in cleavage, however, cannot easily be explained by the one-nucleotide difference in the sequence of the hybrid duplexes, since U2218C is cleaved to a lesser extent than the wild type yet contains one more G · C pair. Thus, at least in the case of the 2218C RNA, stabilization of the PX pairing appears to result in a lower extent of digestion, even at 100 μM oligonucleotide.

Oligonucleotide B is of the same sequence as the 5′ exon (open box) and is complementary to both the IGS (striped box in Fig. 6A) and the upstream half of P(−1) (solid box). This oligonucleotide can bind to conformers containing P1 as well as those containing P(−1), but at different sites along the pre-RNA. The lengths of the expected RNase H products differ by 20 nt and are easily separated on a 4% polyacrylamide gel. As shown in Fig. 6C, the wild-type and U2218C precursors are cleaved at both sites to similar extents. In the case of U2218G, in which PX is destabilized, only the downstream IGS site is degraded by RNase H. These results are consistent with the notion that U2218G precursors predominantly form P(−1), leaving the IGS unpaired and susceptible to RNase H cleavage. The wild-type and U2218C precursors, which are more active, should form P1 more readily and thus make sequences of P(−1) more available for binding with the DNA probe.

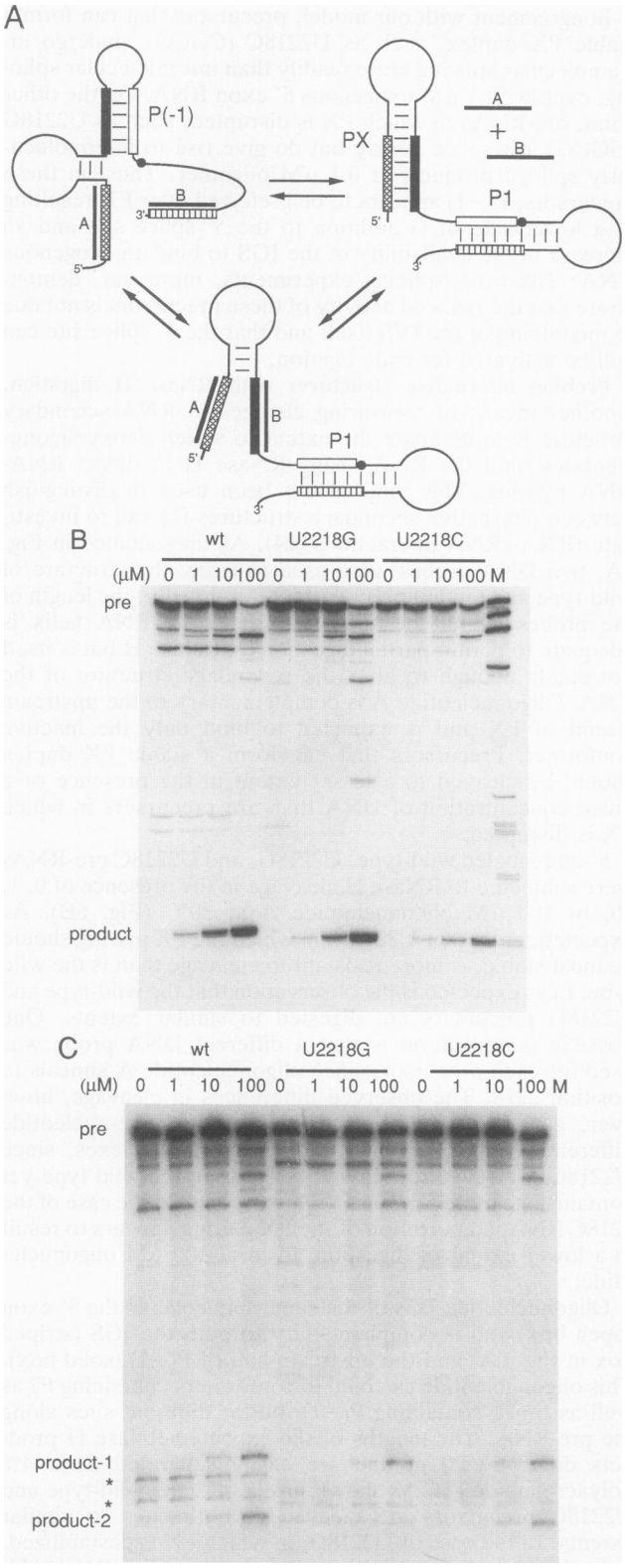


FIG. 6. RNase H digestion of precursor RNAs. (A) Binding of DNA oligonucleotides to alternative secondary structures of the pre-RNA. Sequences of the precursor are represented as in Fig. 1; DNA oligomers are represented by short, thick lines. Oligonucleotide A is complementary to sequences of PX (cross-hatched box) and should preferentially bind the conformer on the left, corresponding to a loss of PX. Oligonucleotide B has the same sequence as the 5′ exon (open box) and is complementary to both the IGS (striped box) and the upstream half of P(−1) (solid box). In the absence of P1, oligonucleotide B can pair with the IGS as shown on the left. In the absence of P(−1) and PX, oligomer B can also pair with nucleotides in the 5′ exon, as shown in the lower middle. (B) 5′-^{32}P-labeled precursor RNAs (0.15 μM) were digested with RNase H in the presence of a DNA oligonucleotide complementary to the upstream sequence of PX, as diagrammed in panel A. The wild-type (wt) precursor was digested with oligonucleotide 5′-dGCAG AAATC; precursor U2218G was digested with oligonucleotide 5′-dGGAGAACTC; precursor U2218C was digested with oligonucleotide 5′-dGGAGAAGTC. The micromolar concentration of DNA oligomer in each reaction is indicated above the lanes. Reaction products were electrophoresed on a 4% polyacrylamide gel. pre, precursor; product, the 5′-labeled digestion product. The lengths of the expected 5′ nucleolytic fragments are 194 nt for the wild type and 189 nt for the mutants. Other minor products are due to hydrolysis of the pre-RNA at the 3′ splice site and hybridization of oligomer to other sites at high DNA concentrations. Lane M contains ϕX/*Hae*III DNA markers. (C) 5′-end-labeled precursor RNAs were digested with RNase H in the presence of oligonucleotide B, 5′-dCUCUCU, as in panel B. pre, precursor; product-1, 287-nt RNA arising from digestion in the IGS; product-2, 267-nt fragment resulting from cleavage in the 5′ exon. Bands marked with asterisks result from attack of G414 at the 5′ splice site and at nt 15 of the IVS following hydrolysis of the 3′ splice site. Lane M is as in panel B.

Results similar to those shown in Fig. 6B were obtained from digestions at 10°C and from experiments in which the RNA was heated and cooled together with the DNA oligonucleotide prior to the addition of RNase H (data not shown). The order of addition of RNase H, pre-RNA, and DNA to the reaction mixture was also varied (data not shown), with identical results. Taken together, these data suggest that the 9-bp RNA-DNA duplex arrives at equilibrium well within the course of the assay (5 min), even at 10°C. Another concern is that addition of DNA oligomer will itself alter the secondary structure of the RNA. To address this possibility, the observed rates of splicing were measured in the presence of 100 μM oligonucleotide A or B and were found to be unchanged (data not shown). Thus, at least formation of P1 does not appear to be affected by 100 μM deoxyoligonucleotide.

Reverse self-splicing is enhanced by the ability to form PX. Since the ability to form an alternative secondary structure in the rRNA appears to promote splicing of the *Tetrahymena* IVS, we wished to test whether this conformational change can also affect the reverse splicing reaction. The formation of P(−1) in ligated exon substrates has been shown to inhibit integration of the IVS at the splice junction (35). Earlier attempts to observe reverse splicing by using an 88-nt transcript of rRNA sequence were also unsuccessful (37). Since that time, additional work on the forward splicing reaction has demonstrated that at least 145 nt nucleotides 5′ and 86 nt 3′ of the IVS must be present to ensure optimal splicing in vitro. It seemed reasonable that structures that destabilize P(−1) in the forward reaction might also enable recognition of the splice junction in the reverse reaction. Thus, we anticipated that RNAs containing all of these exon sequences would be better substrates for IVS integration.

Reverse splicing reactions were carried out with a 244-nt ligated exon RNA. This substrate RNA corresponds to the ligated exon product of a 657-nt precursor that was previously found to self-splice efficiently (32) and includes 146 nt of rRNA sequence upstream and 86 nt downstream of the

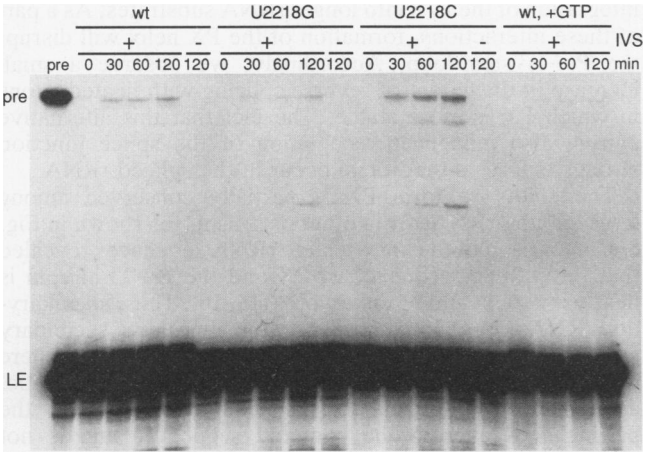


FIG. 7. Integration of the IVS into a 244-nt ligated exon RNA. One micromolar linear IVS RNA was incubated with 0.5 μM 5′-^{32}P-labeled ligated exon RNA at 42°C under standard reverse splicing conditions. The length of the substrate corresponds to the solid arrowheads in Fig. 2. Aliquots were removed from the reaction at the time (minutes) indicated above each lane and electrophoresed on a 4% polyacrylamide gel. The sequence of the ligated exon substrate is indicated above the lanes. −IVS indicates identical reactions carried out in the absence of IVS RNA; +GTP indicates a reaction containing wild-type (wt) substrate and 100 μM GTP. pre, position of authentic precursor RNA and of the expected reverse spliced product; LE, position of the ligated exon substrate. The extents of reaction at 2 h are approximately 0.1% for the wild type, 0.08% for U2218G, and 0.3% for U2218C.

splice junction. The extents of these sequences are denoted by the solid arrowheads in Fig. 2. 5′-^{32}P-labeled substrate RNA was incubated with unlabeled linear IVS RNA under standard reverse splicing conditions (34). As shown in Fig. 7, a labeled product that comigrates with authentic precursor RNA in an adjacent lane appears with time. As expected for a product of reverse splicing, appearance of this band depends on the addition of IVS RNA to the reaction and is sensitive to the presence of GTP.

As a next step, we wished to test whether integration of the IVS was influenced by the ability to form a stable PX helix. Reverse splicing reactions were also carried out with U2218C and U2218G ligated exon substrates. As can be seen in Fig. 7, the extent of product formation with the U2218C substrate, in which PX is more stable, is slightly greater than with the wild-type substrate. Similarly, U2218G ligated exons reverse splice slightly less readily than the wild type. Thus, under these conditions, the stability of PX appears to enhance the ability of the IVS RNA to integrate into a longer ribosomal transcript. Reactions were also carried out at 30 and 42°C in the presence of 50 mM $MgCl_2$ (data not shown). At the higher magnesium concentration, the differences in reactivity between the ligated exon substrates disappear.

DISCUSSION

An alternative helix in the 5′ exon increases the rate of self-splicing. The data presented above strongly support the idea that an alternative base pairing in the rRNA increases the rate of self-splicing in *Tetrahymena* precursors of natural sequence. Formation of this alternative helix, which we have termed PX, increases recognition of the 5′ splice site by shifting the equilibrium between the competing P(−1) and P1 hairpins in favor of the active P1-containing conformer. Base changes that decrease the stability of PX result in lower rates of splicing, whereas changes that maintain or increase the stability of PX result in higher initial rates of reaction. The ability to compensate for base changes suggests that this effect is largely due to secondary structure rather than to a specific nucleotide. A mutation that is proposed to stabilize the PX helix results in less RNase H digestion in the presence of oligonucleotide A, providing additional experimental evidence for the PX pairing.

Our interpretation that the observed rates of splicing result from changes in the equilibrium between P(−1) and P1 is supported by the absence of IVS-3′ exon intermediate, even with saturating (0.5 mM) GTP. Thus, exon ligation does not appear to be rate limiting in these experiments. The qualitative results from the intermolecular splicing experiments argue that the 3′ splice site is still functional in the mutant precursors. The extent of intramolecular versus intermolecular splicing is taken to reflect the proportion of precursors containing P1 and PX, as opposed to P(−1). Indeed, the ratio of intramolecular to intermolecular spliced products correlates with the predicted stability of the PX helix. Somewhat surprisingly, the wild-type precursor undergoes both intramolecular and intermolecular splicing quite readily. Furthermore, incubation in the presence of oligonucleotide B and RNase H results in a similar extent of cleavage in the IGS and P(−1). In the presence of oligonucleotide A, we observed little difference in the digestion of the wild-type and U2218G precursors, even though the wild-type RNA clearly splices more rapidly. These results suggest that formation of PX may be transient and that the equilibrium constant for the alternative structures in the wild type is near one.

Apparent differences in the splicing data and RNase H experiments for the wild-type precursor may arise from the nature of the assays. Spliced product can result from even brief formation of P1, whether in *cis* or in *trans*. The rate of splicing is directly sensitive to P(−1) and P1 but is only indirectly affected by formation of PX. Protection from RNase H digestion, on the other hand, requires that each RNA molecule remain folded and inhibit formation of RNA-DNA hybrid throughout much of the incubation period. Cleavage of upstream sequences with oligonucleotide A should be directly sensitive to formation of the PX pairing but could also be affected by other secondary structures, such as the long-range pairing shown in Fig. 2. If the wild-type RNA adopts both conformers in roughly equal amounts, and if the RNA can isomerize between the two secondary structure models, this might account for the observation that half of the RNA population *cis* splices rapidly while much of the RNA can also *trans* splice, and could account for our RNase H results.

A requirement for the RNA sequence 75 to 80 nt upstream of the 5′ splice site is consistent with earlier results from exon deletion studies (32). In those experiments, a precursor containing a 120-nt 5′ exon and a 624-nt 3′ exon still spliced at a rate similar to that of the wild type, but a precursor in which the 5′ exon was shortened to 55 nt spliced fourfold slower (0.12 min^{-1}). The ability to form a PX helix, however, may be only one part of the mechanism by which exon sequences influence splicing of the *Tetrahymena* IVS. For instance, our simple model does not account for the large changes in splicing efficiency arising from deletions and point mutations in the 3′ exon (32). One explanation is that these mutations eliminate interactions within the rRNA domain that maintain proper folding of the exon sequences. This, in turn, might influence the equilibrium between the

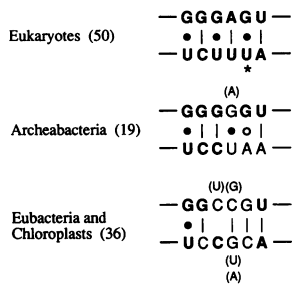


FIG. 8. Consensus PX pairings in eukaryotic, archaebacterial, and eubacterial rRNAs. Known sequences of large-subunit rRNAs were compared (15, 16), and rough consensus sequences for the PX pairing are as shown. The number of sequences in each category is listed. Chloroplast sequences are included with eubacteria. The ability to form PX is poorly conserved among mitochondrial rRNAs. Invariant nucleotides are shown in large bold type, conserved nucleotides are shown in large type, and less frequent variations are shown in small type. Watson-Crick base pairs are indicated with a line, U · G wobble pairs are marked with a solid dot, and G · A pairs are marked with an open circle. The position marked with an asterisk is a C in *Euglena gracilis*.

alternative helices 5′ of the IVS. Another possibility is that 3′ exon sequences affect 3′ splice site activation. If activation of the two splice sites is coupled, changes in the 3′ exon could also lower the reactivity of the 5′ splice site. Finally, mutations introduced at position 2218 may alter the stability of another long-range pairing in the 26S rRNA. Such an interaction could have a direct or indirect effect on the equilibrium between P(−1) and PX. Although this pairing is phylogenetically conserved, we have no experimental evidence yet that it forms in vitro.

Mechanism of conformational change in RNA precursors. An interesting question is what the mechanism and rate of interconversion between the alternative secondary structures might be. Although the results seem to indicate a change in the equilibrium between the two conformers, our data do not allow us to assess whether this equilibrium is rapid with respect to the observed rate of splicing, or whether refolding of the RNA occurs on a time scale similar to that of splicing. Juxtaposition of the upstream strand (cross-hatched box in Fig. 6A) of PX with the base of P(−1) might facilitate interconversion of the helices by strand displacement. Newly formed PX base pairs would be stabilized by coaxial stacking of the adjacent G · C stem. A similar mechanism of branch migration has been proposed for the *L. collosoma* spliced leader RNA, in which the time scale of helix exchange is less than 1 s (21). The rate of helix exchange is expected to be rapid, because refolding does not require complete duplex dissociation. On the other hand, the strongly biphasic reaction profile for precursor U2218C indicates that about half of the RNA molecules must undergo some slow conversion at 30°C before forming spliced products. The nature of this slow transition has not been identified yet, and it may or may not be related to P(−1) and PX.

Integration of the IVS is facilitated by rRNA structure. The extended exon sequences required for efficient forward splicing can also promote recognition of the splice junction during IVS integration. In contrast to previous results (34, 35), we were able to observe reverse splicing into a 244-nt ribosomal transcript, corresponding to the ligated exon product of the smallest efficiently spliced precursor (32). Apparently, RNA interactions that relieve potential inhibition of splicing by P(−1) in the pre-RNA can also promote integration of the IVS into longer rRNA substrates. As a part of these interactions, formation of the PX helix will disrupt the P(−1) stem-loop. Accordingly, we observe a small increase in the extent of reverse splicing with ligated exons in which PX is more stable. The fact that this alternative pairing also influences recognition of the splice junction strongly suggests that it can occur in the spliced rRNA.

The ability to form PX is partially conserved among large-subunit rRNAs from other organisms, as shown in Fig. 8. Comparison of the known large rRNA sequences revealed that the primary sequence of PX and the P(−1) hairpin is nearly invariant among eukaryotes (15, 16). Thus, all eukaryotic rRNAs could potentially form the alternative secondary structure proposed for the *Tetrahymena* precursor. There are some compensatory changes in PX between eukaryotic, archaebacterial, and eubacterial sequences, although the pairing is weaker in the bacterial sequences and is not conserved in mitochondrial rRNAs. Nonetheless, conservation of the PX sequences among eukaryotes, along with the results from the reverse splicing experiments, suggests that a transient switch between PX and P(−1) may be a normal feature of eukaryotic rRNA. Alternatively, the ability of the *Tetrahymena* sequence to form PX may represent a special adaptation to a group I intron. In vivo, the equilibrium between these alternative RNA structures is likely to also depend on protein binding.

The known group I introns in rRNA genes are clustered in a few locations, all of which correspond to highly conserved regions that are thought to interact with tRNA or other rRNAs (13, 28). It has been proposed that group I introns were at one time transposed to new locations by reverse splicing at the RNA level (4, 30). If this is the case, the current locations of introns in rRNAs should correspond to regions of the rRNA that are most accessible. Many group I introns, however, are found inserted next to secondary structures that could interfere with splice site recognition (35). We suggest that formation of alternative RNA structures, either by RNA-RNA or RNA-protein interactions, could be one method of relieving inhibition of splicing in ribosomal transcripts. In addition, RNA structures or RNA-protein interactions (25) that facilitate forward splicing can be expected to promote the reverse reaction. Thus, the position of introns in modern rRNA may reflect regions that are particularly susceptible to conformational rearrangement.

ACKNOWLEDGMENTS

We thank Michael Zuker for the gift of MULFOLD and advice on the RNA secondary structure calculations. We also thank Robin Gutell for comments on the phylogenetic sequence comparisons and communication of unpublished observations. Finally, we acknowledge Art Zaug for many helpful discussions and Dan Celander for a critical reading of the manuscript.

This work was supported by Public Health Service grant GM46686 from the National Institutes of Health.

REFERENCES

1. **Altuvia, S., Kornitzer, D., Teff, D., and A. B. Oppenheim.** 1989. Alternative mRNA structures of the *c*III gene of bacteriophage λ determine the rate of its translation initiation. J. Mol. Biol. **210**:265–280.
2. **Bass, B. L., and T. R. Cech.** 1984. Specific interaction between the self-splicing RNA of Tetrahymena and its guanosine substrate: implications for biological catalysis by RNA. Nature (London) **308**:820–826.
3. **Been, M. D., and T. R. Cech.** 1986. One binding site determines sequence specificity of Tetrahymena pre-rRNA self-splicing,

trans-splicing and RNA enzyme activity. Cell **47**:207–216.
4. **Cech, T. R.** 1985. Self-splicing RNA: implications for evolution. Int. Rev. Cytol. **93**:3–22.
5. **Cech, T. R.** 1990. Self-splicing of group I introns. Annu. Rev. Biochem. **59**:543–568.
6. **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.
7. **Davanloo, P., A. H. Rosenberg, J. J. Dunn, and F. W. Studier.** 1984. Cloning and expression of the gene for bacteriophage T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **81**:2035–2039.
8. **Donis-Keller, H.** 1979. Site-specific enzymatic cleavage of RNA. Nucleic Acids Res. **7**:179–192.
9. **D'Orval, B. C., Y. D. 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.
10. **Eng, F. J., and J. R. Warner.** 1991. Structural basis for the regulation of splicing of a yeast messenger RNA. Cell **65**:797–804.
11. **Engberg, J., and H. Nielsen.** 1990. Complete sequence of the extrachromosomal rRNA molecule from the ciliate Tetrahymena thermophila strain B1868VII. Nucleic Acids Res. **18**: 6915–6919.
12. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **85**:9373–9377.
13. **Gerbi, S. A., R. L. Gourse, and C. G. Clark.** 1982. Conserved regions within ribosomal DNA: locations and some possible functions, p. 351–386. *In* H. Busch and L. Rothblum (ed.), The cell nucleus, vol. 10. Academic Press, New York.
14. **Green, M. R.** 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7**:559–599.
15. **Gutell, R. R. (University of Colorado).** Personal communication.
16. **Gutell, R. R., M. N. Schnare, and M. W. Gray.** 1990. A compilation of large subunit (23S-like) ribosomal RNA sequences presented in a secondary structure format. Nucleic Acids Res. **18**(Suppl.):2319–2330.
17. **Inoue, T., F. X. Sullivan, and T. R. Cech.** 1985. Intermolecular exon ligation of the rRNA precursor of Tetrahymena: oligonucleotides can function as 5′ exons. Cell **43**:431–437.
18. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Improved predictions of secondary structures for RNA. Proc. Natl. Acad. Sci. USA **86**:7706–7710.
19. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Predicting optimal and suboptimal secondary structure for RNA. Methods Enzymol. **183**:281–306.
20. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
21. **LeCuyer, K. A., and D. M. Crothers.** Biochemistry, in press.
22. **Libri, D., A. Piseri, and M. Y. Fiszman.** 1991. Tissue-specific splicing in vivo of the β-tropomyosin gene: dependence on an RNA secondary structure. Science **252**:1842–1845.
23. **Maniatis, T.** 1991. Mechanisms of alternative pre-mRNA splicing. Science **251**:33–34.
24. **Marconi, R. T., and W. E. Hill.** 1989. Evidence for a tRNA/rRNA interaction site within the peptidyltransferase center of the Escherichia coli ribosome. Biochemistry **28**:893–899.
25. **Mohr, G., and A. M. Lambowitz.** 1991. Integration of a group I intron into a ribosomal RNA sequence promoted by a tyrosyl-tRNA synthetase. Nature (London) **354**:164–167.
26. **Murphy, F. L., and T. R. Cech.** 1989. Alteration of substrate specificity for the endoribonucleolytic cleavage of RNA by the Tetrahymena ribozyme. Proc. Natl. Acad. Sci. USA **86**:9218–9222.
27. **Noller, H. F.** 1984. Structure of ribosomal RNA. Annu. Rev. Biochem. **53**:119–162.
28. **Noller, H. F.** 1991. Ribosomal RNA and translation. Annu. Rev. Biochem. **60**:191–227.
29. **Price, J. V., G. L. Kieft, J. R. Kent, E. L. Sievers, and T. R. Cech.** 1985. Sequence requirements for self-splicing of the *Tetrahymena thermophila* pre-ribosomal RNA. Nucleic Acids Res. **13**:1871–1889.
30. **Sharp, P. A.** 1985. On the origin of RNA splicing and introns. Cell **42**:397–400.
31. **Walstrum, S., and O. C. Uhlenbeck.** 1989. Gel purification traps the self-splicing RNA of Tetrahymena in a less active conformation. Biochemistry **29**:10573–10576.
32. **Woodson, S. A.** 1992. Exon sequences distant from the splice junction are required for efficient self-splicing of the Tetrahymena IVS. Nucleic Acids Res. **20**:4027–4032.
33. **Woodson, S. A.** Unpublished data.
34. **Woodson, S. A., and T. R. Cech.** 1989. Reverse self-splicing of the Tetrahymena group I intron: implication for the directionality of splicing and for intron transposition. Cell **57**:335–345.
35. **Woodson, S. A., and T. R. Cech.** 1991. Alternative secondary structures in the 5′ exon affect both forward and reverse self-splicing of the Tetrahymena intervening sequence RNA. Biochemistry **30**:2042–2050.
36. **Zaug, A. J., M. D. Been, and T. R. Cech.** 1986. The Tetrahymena ribozyme acts like an RNA restriction endonuclease. Nature (London) **34**:429–433.
37. **Zaug, A. J., and T. R. Cech.** 1985. Oligomerization of intervening sequence RNA molecules in the absence of proteins. Science **229**:1060–1064.
38. **Zaug, A. J., and T. R. Cech.** 1986. The intervening sequence RNA of Tetrahymena is an enzyme. Science **231**:470–475.
39. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244**:48–52.