Nucleic Acids Research, Vol. 20, No. 20 5357–5364

Group I permuted intron–exon (PIE) sequences self-splice to produce circular exons

M.Puttaraju and Michael D.Been*
Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

Received July 16, 1992; Revised and Accepted September 4, 1992

ABSTRACT

Circularly permuted group I intron precursor RNAs, containing end-to-end fused exons which interrupt half-intron sequences, were generated and tested for self-splicing activity. An autocatalytic RNA can form when the primary order of essential intron sequence elements, splice sites, and exons are permuted in this manner. Covalent attachment of guanosine to the 5′ half-intron product, and accurate exon ligation indicated that the mechanism and specificity of splicing were not altered. However, because the exons were fused and the order of the splice sites reversed, splicing released the fused-exon as a circle. With this arrangement of splice sites, circular exon production was a prediction of the group I splicing mechanism. Circular RNAs have properties that would make them attractive for certain studies of RNA structure and function. Reversal of splice site sequences in a context that allows splicing, such as those generated by circularly permuted group I introns, could be used to generate short defined sequences of circular RNA *in vitro* and perhaps *in vivo*.

INTRODUCTION

Group I introns are characterized by a linear array of conserved sequences and structural features, and are excised by two successive transesterifications in a well-defined splicing pathway (1, 2). The reaction is initiated with a nucleophilic attack by the 3′ hydroxyl group of guanosine (G) on the phosphorous at the 5′ splice site, resulting in cleavage of the precursor RNA. A free 3′ hydroxyl group is generated on the end of the 5′ exon, and the G is ligated to the 5′ end of the intron. Release of the intron and ligation of the exons occur in the second step by attack of the 5′ exon's 3′ hydroxyl group on the 3′ splice site phosphorous.

A number of group I introns are autocatalytic and the normal splicing reaction is considered to be essentially intramolecular, even though each step actually involves reactions between two molecules which are not covalently joined. So called *trans*-splicing reactions, reactions involving the combining of separate intron and exon sequences or fragmented precursor sequences (3–7), have been useful for *in vitro* studies of the splicing mechanism. The intermolecular reactions demonstrate that a linear array of sequences and structural elements in a single transcript is not required for the formation of the active tertiary structure. Also, because the function of splice sites is not determined by their relative positions (5′ or 3′), but by splice site sequence and structural features which recognize, bind, and activate those sites (8–12), it is possible to reverse the order of the splice sites and generate circles of RNA (13). With the *Tetrahymena* group I intron, it was shown, by Price *et al.* (13), that removal of the normal 5′ splice site can result in activation of a cryptic 5′ splice site in the 3′ exon (3′ to the normal 3′ splice site) and, as predicted, the production of a small circle of 3′ exon sequences.

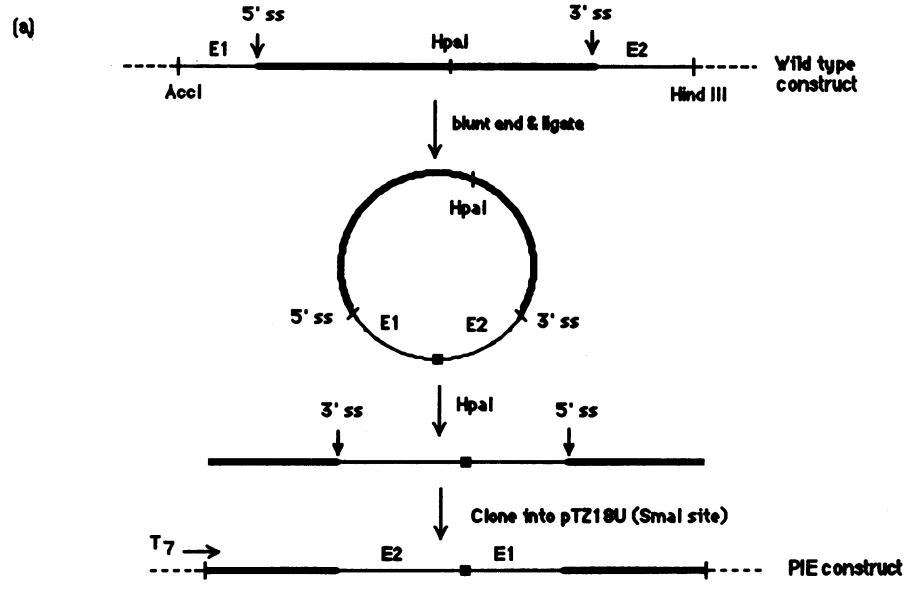
A general method for making RNA circles, especially if the process was autocatalytic and occurred during transcription *in vitro* or *in vivo*, could be useful. It was not clear whether the method of circle production discovered by Price *et al.* (13) would be generally useful for making RNA circles because it would appear to rely on only the P1 duplex of the internal guide sequence (IGS) (14, 11) to bring the cryptic 5′ splice site into service. By reversing large sections of the intron as well as the order of the two splice sites, it may be possible to utilize the entire secondary and tertiary structure of the intron to stabilize the required interactions and facilitate circular exon generation. We have examined the question of whether circularly permuted forms of group I introns, which are interrupted by exon sequence, would fold into an active structure capable of splicing *via* the normal pathway. If the permuted intron-exon (PIE) sequence is active, the prediction is that a circular form of the ligated exon would be generated. It was found that, with four different group I PIE sequences, splicing resulted in the generation of circular exon sequences.

MATERIALS AND METHODS

Enzymes, reagents and chemicals

T7 RNA polymerase was prepared from an over-expressing clone provided by W.Studier (15). *Eco*RI was from P.Modrich (Duke University). RNase A, PEI–cellulose plates, cellulose acetate strips and DEAE–cellulose paper were provided by D.A.Steege (Duke University). Modified T7 DNA polymerase, T4 DNA polymerase, T4 DNA ligase, T4 polynucleotide kinase, reverse transcriptase, and other restriction endonucleases were purchased

* To whom correspondence should be addressed



(b)

GGAAUUCGAG CUCGGUACCC aacaacagau aacuuacagc uaaucggaag
gugcagagac ucgacgggag cuacccuaac gucaagacga ggguaaagau
gagaguccaa uucucaaagc caauaggcag uagcgaaagc ugcgggagaa
g^AAAAUCCGU UGACCUUAAA CGGUCGUGUG GGUUCAAGUC CCUCCACCCC
CAUGAAGAGA AAAGCAAGUU CUUGUAAGUU UAACCAUAAA UUUGCCAGCC
ACGGCAAGCU AGACGCUACG GACUU*aaaua auugagccuu agagaagaaa
uucuuuaagu ggauacucuc aaacucaggg aaaccuaaau cuagcgauag
acaaggcaau ccugagccaa gccgaaguag uaauuaguaa guuGGGGAUC

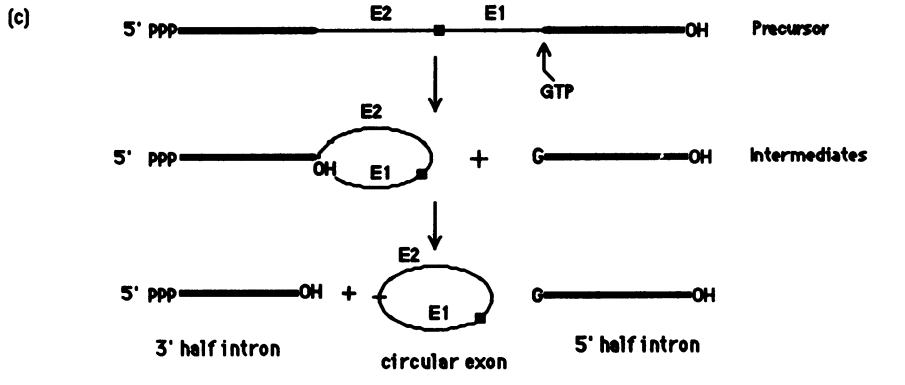


Figure 1. (a) Construction of the plasmid, pPR100. The thin line represents exon sequences, with the heavier dot indicating the approximate position at which they were fused, the thick line represents intron sequences and the dashed lines are vector sequences. E1 and E2 are the 5′ exon and the 3′ exon, respectively. See text for details. (b) Nucleotide sequence of T7 RNA polymerase transcript from *Bam*HI cut pPR100. Lower case letters, intron derived sequences; upper case letters, exon and vector derived sequences. Vector derived sequences are underlined. The 5′ and 3′ splice sites are indicated by * and A ^, respectively. (c) Proposed splicing pathway based on the normal group I splicing mechanism.

from US Biochemicals. RNase T_1 was obtained from Calbiochem. Nucleoside triphosphates were purchased from Pharmacia. ^{32}P-labeled nucleotides were purchased from ICN. The plasmid pAtRNA-1, containing the sequence of the group I intron with tRNA exons from *Anabaena* PCC7120, was constructed by A.Zaug (University of Colorado) from pXAb107 (D. Shub, SUNY Albany).

Construction of PIE plasmids

Plasmid DNA containing the *Anabaena* pre-tRNA group I intron (16), pAtRNA-1, was cleaved with *Acc*I and *Hin*dIII in the 5′ and 3′ exon sequences, respectively. The ends of the 370 bp fragment containing the intron, exons, and flanking sequences were filled using T4 DNA polymerase and then ligated together under dilute conditions to circularize the fragments. It was then

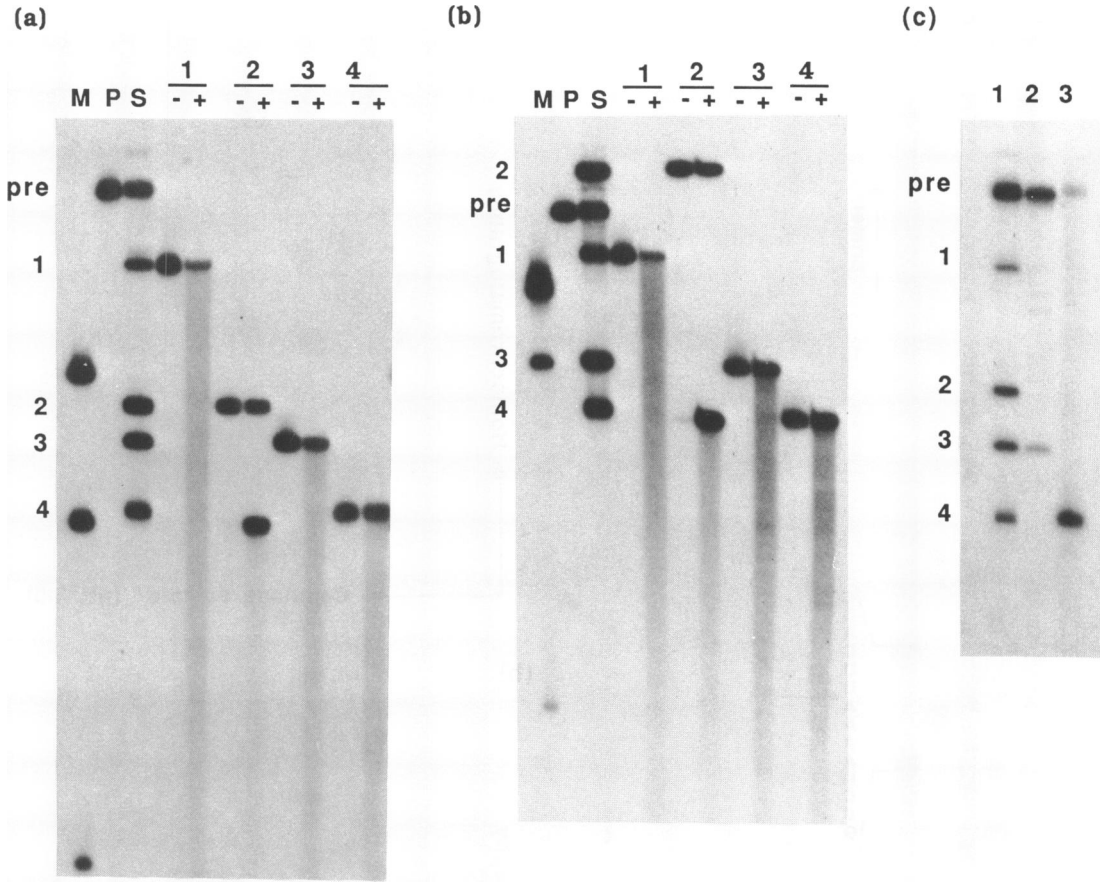


Figure 2. (a and b) Splicing of PR100 PIE sequence and identification of a circular RNA product. Gel purified precursor (pre) received either no further treatment (lane P) or was incubated for 60 min in 15 mM $MgCl_2$, 0.2 mM GTP, 40 mM Hepes pH 7.5, at 37°C (lane S). The reaction was stopped with EDTA. Gel purified splicing products (bands 1 to 4) were treated (+) or not treated (−) to nicking conditions (25 mM $NaHCO_3$ (pH 9.0), 1 mM EDTA at 90°C for 3 min). Formamide with tracking dyes was added to each sample and the products were analyzed by electrophoresis on 6% (a), and 12% (b) polyacrylamide gels containing 7 M urea. Lane M, Marker RNAs (181, 127, 54 nt) from a self-cleavage reaction of an HDV genomic RNA containing fragment (30, 31). (c) Identification of product bands 3 and 4. Lane 1, splicing of [^{32}P]body-labeled precursor with cold guanosine. Lane 2, splicing of [5′^{32}P]end-labeled precursor with cold guanosine. Lane 3, splicing of unlabeled precursor with [α-^{32}P]GTP.

cleaved with *Hpa*I to reopen the circle, ligated to *Sma*I cleaved pTZ18U, and transformed into *E.coli* strain JM83. Miniprep plasmid DNA was isolated from overnight cultures by the boiling method (17) and the DNA sequenced by primer extension with modified T7 DNA polymerase (Sequenase) (18) and dideoxynucleotide chain terminators (19). Following retransformation, plasmid DNA containing the permuted sequence (pPR100) was prepared by boiling lysis and purified by CsCl equilibrium density centrifugation in the presence of ethidium bromide (17), and resequenced.

Plasmid DNA pBFSN1 (20), containing a modified form of the *Tetrahymena* group I intron, was cut with *Eco*RI and *Hin*dIII in the vector sequences which flank the exons. The excised fragment containing the intron (413 basepairs) and flanking sequences (~60 basepairs) was circularized as described above. The circle was then cut with one of three restriction endonucleases (*Bgl*II, *Nsi*I, *Stu*I) to regenerate linear fragments containing the permuted sequence. These DNA fragments were ligated into the multicloning site of pTZ18U, and transformants containing the desired orientation of the insert were used to prepare plasmid DNA. In pBFSN1, the restriction sites *Stu*I, *Bgl*II, and *Nsi*I, are found in sequences corresponding to intron loop structures at the ends of P5, P6 and P8, respectively (20). The plasmids, pDUS1, pDUB1, and pDUN1 contain sequences permuted from the *Stu*I, *Bgl*II, and *Nsi*I sites, respectively.

Preparation of precursor RNA

Plasmid DNA from pDUN1, pDUB1 and pDUS1 (*Tetrahymena* derived), and pAtRNA-1 and pPR100 (*Anabaena*) were linearized with restriction endonuclease *Hin*dIII or *Bam*HI. The products were extracted once with buffered phenol and once with chloroform, and the DNA was recovered by ethanol precipitation. Transcription was carried out in 40 mM Tris−HCl (pH 7.5), 10 mM $MgCl_2$, 2 mM spermidine, 5 mM dithiothreitol, ribonucleoside triphosphates at 1 mM each, 0.1 mg/ml linear plasmid DNA, and 50 units of T7 RNA polymerase/mg of DNA. For body-labeled RNA, an [α-^{32}P]NTP was included in the reaction at 0.5 μCi/μl. After 60 minutes at 37°C, EDTA was added to 25 mM, formamide to 50%, and the RNA was fractionated by electrophoresis on a 6% polyacrylamide gel containing 7 M urea. RNA bands were located by UV shadowing or autoradiography. The RNA was eluted from the gel into 0.1% SDS, 10 mM EDTA and recovered by ethanol precipitation.

5′-End labeling

Unlabeled gel purified pre-RNA (1 μg) was dephosphorylated using calf intestinal alkaline phosphatase (1 unit) in 10 mM

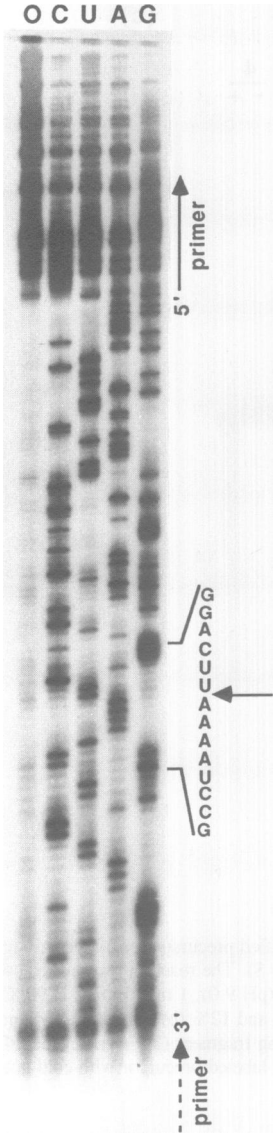


Figure 3. Primer extension analysis. Product-band 2 RNA was gel purified on a 6%, followed by a 12% polyacrylamide gel. An end-labeled oligonucleotide primer ([5'-^{32}P]pATGGGGGTGGAGGGACTTGAACC) was annealed to 3'-exon sequence and extended in the presence of dideoxynucleotides using reverse transcriptase (32). The extension products were fractionated on a 6% sequencing gel. Arrow indicates the ligation junction. Lane 0 was a control lane (no dideoxynucleotide). Lanes C, U, A, and G contained ddGTP, ddATP, ddTTP and ddCTP, respectively.

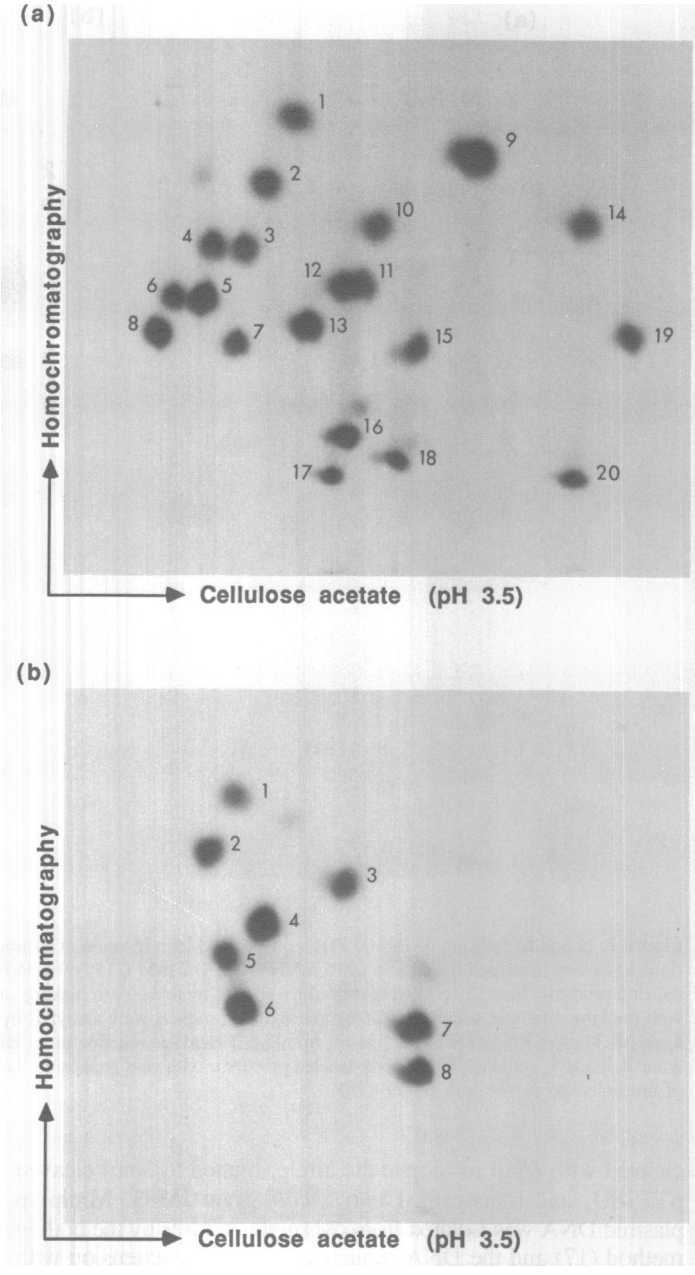


Figure 4. Autoradiograms of a ribonuclease T_1 fingerprint of circular RNAs. The gel purified circular RNAs from *Anabaena* (**a**) and *Tetrahymena* (**b**) PIE precursors were digested with RNase T_1 and the oligonucleotides were fractionated by electrophoresis in the first dimension and homochromatography in the second dimension. The spots were identified by their characteristic mobility (21, 22) and confirmed by secondary digestion with RNase A. Numbering corresponds to the oligonucleotide sequences in Table I.

Tris−HCl (pH 8.0), 1 mM EDTA. After 30 min at 37°C, the RNA was extracted once with phenol and once with chloroform and recovered by ethanol precipitation. The dephosphorylated RNA was labeled in a 20 μl reaction containing 50 mM Tris−HCl (pH 8.9), 10 mM $MgCl_2$, 5 mM EDTA, 22 pmole of [γ-^{32}P]ATP and 10 units of polynucleotide kinase. Incubation was on ice (0−4°C) for 2 hr and the reaction was stopped by addition of EDTA to 25 mM. The end-labeled intact pre-RNA was gel purified as described above.

Fingerprinting

Gel purified RNA (labeled using [α-^{32}P]GTP) was digested with ribonuclease T_1 in 10 mM Tris−HCl (pH 8.0) and 1 mM EDTA at 37°C for 30 min. Fractionation of oligonucleotides: the first dimension was electrophoresis on cellulose acetate strips (55×3 cm) at pH 3.5 and the second dimension was homochromatography on DEAE thin layer plates using homomix C (21, 22). Initial identification of spots was made by correlation of the predicted oligonucleotide sequence with characteristic mobilities. Identities were confirmed by secondary analysis of the oligonucleotides as described by Barrell (21).

Analysis of splicing reactions

Products were fractionated by electrophoresis on 6% or 12% polyacrylamide gels containing 7 M urea, 50 mM Tris−Borate (pH 8.3), and 0.5 mM EDTA. Following electrophoresis, 6%

gels were dried on filter paper and an autoradiogram prepared. The 12% gels were not dried. To quantify results from gels, bands were excised and quantified by measuring Cherenkov scintillation. To determine a rate constant and end point, the fraction of precursor reacted (F) was calculated for each time point (t). A first-order rate constant (k) and end point (c) were obtained by fitting the data to $F = c(1-e^{-kt})$, using a non-linear least squares fitting program (Systat). The half-life ($t_{1/2}$) of the precursor was calculated from $t_{1/2} = 0.693/k$.

RESULTS

Construction of a group I PIE from the *Anabaena* pre-tRNA intron

The general method used for making the permuted sequence is illustrated for the *Anabaena* group I intron (16) (Figure 1a). A DNA fragment containing the wild-type intron sequence, together with short 5′ and 3′ exon sequence and associated vector sequence, was excised from the parent plasmid (pAtRNA-1), circularized with DNA ligase, and relinearized by cutting at a unique site within the intron sequence. In this form, the ends of the two exon sequences have been fused but the splice site sequences remain intact. In the *Anabaena* sequence there is a *Hpa*I site in the sequence forming the P6 (16, 11) of the intron structure. The resulting DNA fragment contained a permuted version of the original sequence which began and ended within P6 of the intron, and contained 124 nucleotides (nt) of fused-exon plus vector sequence between the two splice sites; it was ligated into the multicloning site in a T7 promoter containing plasmid to generate pPR100. A 400 nt long *Bam*HI-runoff T7 RNA polymerase transcript from pPR100, the *Anabaena* PIE sequence (Figure 1b), was used for these studies.

Products of splicing of the *Anabaena* PIE sequence

Gel purified *Anabaena* PIE sequence reacted under group I intron splicing conditions to generate four major products (Figure 2a, lane S). There was no reaction in the absence of $MgCl_2$ and very little product was formed in the absence of guanosine (data not shown). The hypothesized mechanism for PIE splicing (Figure 1c) would result in 3 to 4 products: a 5′ half-intron, a 3′ half-intron, a circular form of the exon, and possibly a 3′ half-intron–exon intermediate. To identify which product, if any could be the predicted circular form, each product was gel purified and then rerun on a 6% polyacrylamide gel after treating the isolated RNA to a brief nicking treatment (90°C, pH 9.0, 3 minutes). Three of the bands generated a smear of random sized fragments indicating that the original product was probably linear (Figure 2a and 2b, lanes 1, 3, 4). One of the products (band 2), however, showed a discrete shift in mobility (lane 2). In this case, the new faster migrating band ran slightly ahead of band 4 and is also a linear form as indicated by a smear of fragments below it. When the same material is run on a 12% polyacrylamide gel under denaturing conditions, there is a significant shift in the mobility of band 2; it now migrates slower than the precursor (Figure 2b, lane 2). However, the nicked form of band 2 again migrates close to the position of band 4. The above data are consistent with the original form of band 2 being non-linear and the second form being linear.

The origin of the other products, tentatively identified by relative mobility, were confirmed by the following experiments. The 5′ end of an unlabeled PIE precursor was dephosphorylated and then labeled with ^{32}P. Upon splicing, a labeled product co-migrates with band 3 (Figure 2c, lane 2); this would identify band 3 as 3′ half-intron sequence (151 nt). When unlabeled precursor is spliced in the presence of ^{32}P-labeled GTP, a labeled product is generated that co-migrates with band 4 (lane 3). Thus, band 4 is presumably labeled by covalent addition of the exogenous G at the 5′ end of the intron during cleavage at the 5′ splice site; this would identify band 4 as the 5′ half-intron sequence (125 nt). Band 1 has been identified as the exon containing intermediate (275 nt) (Figure 1c) by primer extension (data not shown). Gel purified RNA from band 1 showed no further activity under splicing conditions, which is consistent with the premise that both intron halves are required for the reaction.

Table I. Sequences of oligonucleotides produced by RNase T1 and RNase A digestion of PR100 and *Tetrahymena* circular RNA.

Spot No	Products from digestions with RNase T1	RNase A
(Fig. 4A)		
1	G	G
2	AG	AG
3	AAG	AAG
4	ACG	AC
5	CAAG	AAG
6	CCAG	AG
7	AAAAG	AAAAG
8	CCACGp§	AC,G
9	UGp	U,G
10	UCG	C
11	UAAG	AAG
12*	CU–AG	AG
13	CUACGp	AC,G
14	UUG	U
15	UUCAAG	AAG
16	ACCUUAAACGp	AAAC,G
17	UCCCUCCACCCCCAUG	AU
18†	ACUU–AAAAUCCG	C
19	UUCUUG	U
20	UUUAACCUAAAUUUG	U
(Fig. 4B)		
1	G	G
2	AG	AG
3	UAG	AG
4	CUCGp	C,G
5	CCUCG	C
6	UCACAAG and CCAAAUG	AAG,AAAAU
7*	UAACUAUG and CU–AAUUCG	AU,C
8†	ACUCUCU–UAAGp	AAG,G

*Oligonucleotides containing the fused sequence.
†Oligonucleotides containing the spliced junction.
§Oligonucleotides ending in p were followed by an additional G and thus contained an extra labeled phosphate, this was confirmed by Cherenkov counting of the eluted material.

The exon sequences are circular

Unlabeled band 2 was isolated by preparative gel electrophoresis and sequenced by primer extension with a [5′-^{32}P]end-labeled primer. Despite a number of stops in the sequencing gel, nearly the complete sequence from a position near the 3′ end of the primer to the 5′ end of the primer can be determined (Figure 3). These data also identify the sequence of the splice-site junction as accurate.

As an alternative method for identifying the entire circular RNA sequence, band 2 was analyzed by 2-dimensional fingerprinting (Figure 4a). RNA, labeled during transcription with [α-^{32}P]GTP, was digested with RNase T_1 and separated in the first dimension by electrophoresis on cellulose acetate (pH 3.5)

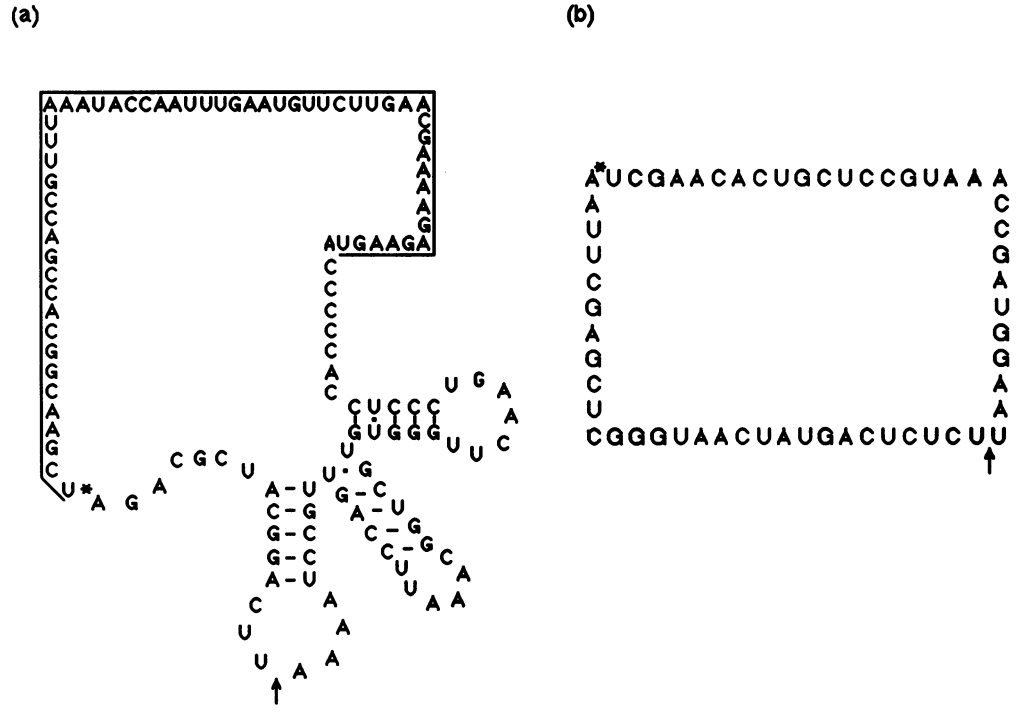


Figure 5. (a) Circular representation of the exon RNA sequences from the *Anabaena* PIE sequence (PR100). The secondary structure is not known, however, for convenience the exon sequences derived from the tRNA sequence are drawn in the classical fashion. Vector sequence is underlined. (b) Circular exon sequence from the *Tetrahymena* PIE sequences (DUS1, DUB1, and DUN1). For both (a) and (b), 5′→3′ is counter-clockwise, the arrows indicate the spliced junction, and asterisks indicate the location of the fused exon position.

and in the second dimension by homochromatography on PEI TLC plates (22) (Figure 4a). Each spot was tentatively identified by comparison to known standard patterns (provided by D.A.Steege), then each spot was eluted, quantified, and subjected to secondary analysis to confirm its identity (Table 1). Spot 18 (Figure 4a) is the ligated exon containing oligonucleotide and spot 12 is the oligonucleotide generated by fusion of the exons during the PIE plasmid construction. An extra spot (AAAAUCCG) which would have been generated by the linear form of the exon sequence was not detected.

Thus, by two independent sequencing methods, the band 2 product is a circular form of the exon sequence (Figure 5a), and there is no evidence for unusual linkages or branched structures.

PIE splicing is less efficient than the normal splicing reaction

A time course for the splicing reaction of the *Anabaena* sequences indicated that, at 37°C, the rate of splicing of the PR100 PIE sequence was slower than that of the original sequence (Figure 6a). The bands were excised and the data were quantified. For the normal intron sequence, splicing proceeds with a $t_{1/2}$ of approximately 2 min and it reacts to near completion (87%), whereas for the PIE sequence under the same conditions, the $t_{1/2}$ is 20 min and the extent of the reaction is about 60%. At 42°C the rate of PR100 splicing increases ($t_{1/2}$ = 8 min) and the extent of the reaction approaches 90%, whereas at 30°C both the rate ($t_{1/2}$ = 30 min) and extent of splicing (28%) decrease (data not shown). The temperature dependence on the extent of splicing suggests that the PIE RNA may not fold into the active structure as readily at lower temperatures, or it may have assumed a less active confirmation during RNA isolation (23). A brief incubation at 80°C prior to addition of GTP did not enhance the rate of splicing at 37°C (data not shown).

Splicing of *Tetrahymena* PIE sequences

Three different PIE sequence-containing plasmids were constructed from a modified form of the *Tetrahymena* pre-rRNA group I intron. Methods similar to those described for generating the *Anabaena* PIE plasmid were used; a DNA fragment from pBFSN1 (20) was excised, circularized and reopened at each of 3 unique sites within the intron sequence. A natural *Bgl*II site, and engineered *Nsi*I and *Stu*I sites (20) allowed the production of PIE sequences permuted at P6, P5, and P8 (11), respectively. With each of these, splicing of ^{32}P-labeled *Hin*dIII runoffs (Figure 6b, lanes N, B, and S) generated a more complex set of products than did the *Anabaena* PIE. However, a common fragment migrating more slowly than the size expected for a linear form of the exon sequence (58 nt) was observed (Figure 6b). This was gel purified, pooled, and shown to behave like a single circular species when run on different percentage gels under denaturing conditions (data not shown). Nicking with alkali (Figure 6b, '+' lane) also indicated that it was probably a circle. The slower migrating form of the RNA was fingerprinted as previously described. Spot 8 (Figure 4b) corresponds to the oligonucleotide predicted to contain the sequence of the spliced junction, and one of the oligonucleotides co-migrating in spot 7 corresponds to the fused exon sequence (Table I and Figure 5b). These results are consistent with generation of a circular form of the exon.

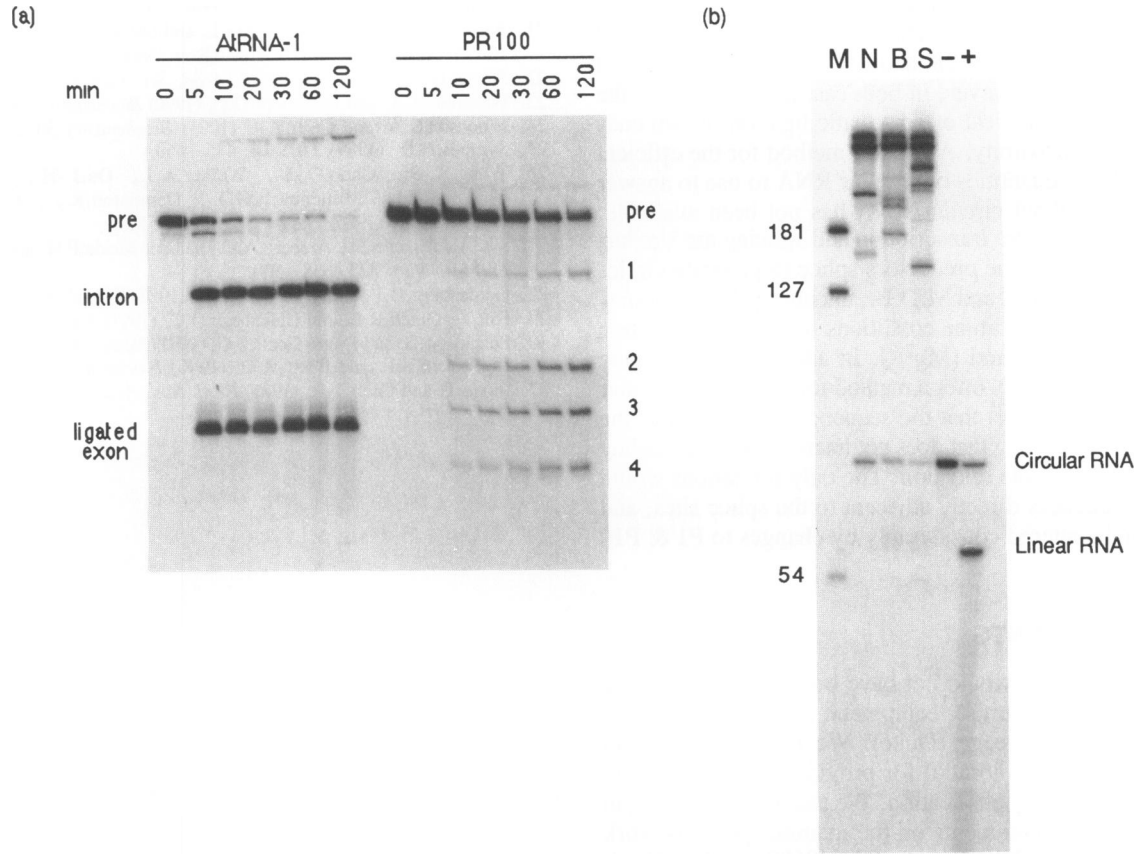


Figure 6. (a) Splicing time courses of precursor RNAs from pAtRNA-1 and pPR100. Precursor RNAs, body-labeled with ^{32}P, were incubated in 40 mM Hepes (pH 7.5), 200 mM NaCl, 0.05 mM EDTA, 20 mM $MgCl_2$ and 0.2 mM GTP. The RNA was not subjected to a heat denaturation step (23). The initial reaction volume was 80 μl; 10 μl was stopped at the indicated time points by adding formamide stop mix containing EDTA and tracking dyes. The products were fractionated on a 6% polyacrylamide gel under denaturing conditions. An autoradiogram of the gel is shown. (b) Splicing and exon circularization of *Tetrahymena* PIE sequences. The reaction conditions for splicing were similar to those described above, except the concentration of $MgCl_2$ was increased to 50 mM. The isolated circular RNA was nicked as described for Figure 2. The products were separated on a 12% denaturing polyacrylamide gel. Lane M, marker lane with sizes in nt. Lanes N, B, and S are splicing products of PIE RNAs transcribed from pDUN1, pDUB1 and pDUS1. Lane −, untreated RNA. Lane +, RNA nicked with alkali treatment.

DISCUSSION

The order of the well defined bench-mark elements in group I introns (P, Q, R, and S) is: E1-5′SS-P-Q-R-S-3′SS-E2 (11), where E1 and E2 are the 5′ and 3′ exons, and 5′SS and 3′SS are the splice sites. We constructed and tested permuted versions of the *Tetrahymena* intron represented by: (5′)-Q-R-S-3′SS-E2-E1-5′SS-P, (5′)-R-S-3′SS-E2-E1-5′SS-P-Q, and (5′)-S-3′SS-E2-E1-5′SS-P-Q-R, and one version of a permuted *Anabaena* intron also represented by (5′)-R-S-3′SS-E2-E1-5′SS-P-Q. In each case circular exons were generated under splicing conditions *in vitro*, indicating that folding of a single transcript containing a linear array of sequence elements into a functional tertiary structure is not dependent on a unique linear arrangement of those sequences. However, these sequences were only circularly permuted; correct folding of randomly permuted sequence elements would certainly provide larger obstacles to formation of the active structure.

The *Anabaena* PIE sequence, generating four major identifiable products, gave cleaner results than the *Tetrahymena* PIE sequences which generated a number of product bands, only some of which have been identified (M.D.B., unpublished results). Even so, the *Tetrahymena* PIE exons are excised to a large degree as circles. A direct comparison of the reactions with *Tetrahymena* and *Anabaena* PIE sequences is difficult because, not only do the introns differ in size and sequence, but so do the fused exon sequences. Even subtle variations in exon sequences have been shown to greatly influence splicing activities (24). Any number of these differences could affect the efficiency of splicing and circle formation. One interesting difference, however, is that the *Anabaena* intron is located in the anticodon of a pre-tRNA and it is possible that the anticodon stem, which is intact in the PIE sequence and excised circle (Figure 5a), facilitates the alignment of exons and perhaps stabilizes the overall correct structure of the intron. PIE sequences with a shorter (17 nt) fused-exon sequence, containing essentially only the anticodon stem and loop, are very active ($t_{1/2}$ < 5 min at 37°C), but the circular product contained very little label and was not recovered (M.P. and M.D.B., unpublished results).

Circular RNAs are a common feature among a number of small pathogenic plant RNAs (viroids and virosoids) (25). In humans, the genome of hepatitis delta virus is an RNA circle (26, 27). It is possible that circles, in addition to their advantages as templates for replication, have certain properties advantageous

to these pathogens such as resistance to cellular nucleases, or stabilization of secondary or tertiary structures. Circular and circularly permuted proteins (28) and RNAs (29) have been used in structural studies. However, in both cases, generation of the circle depended on chemical or enzymatic ligation of two ends that were in close proximity. A general method for the efficient production of large quantities of circular RNA to use to answer specific questions about circular RNA has not been available. The conditions for *in vitro* transcription and splicing are similar, and since 20 to 50% of the precursors splice to generate circles during transcription (M.P. and M.D.B., unpublished data), it may be possible to further optimize conditions for circle production by varying temperature and [Mg^{2+}]. In addition, self-splicing group I PIE sequences may offer a method for generating circular RNAs *in vivo*. We suggest that the sequence that is used as the fused-exon can be anything that does not interfere with the folding of the intron or splice site selection. The only limitations would appear to be sequences directly adjacent to the splice sites, and even these can be altered considerably by changes to P1 & P10 (8, 9, 12).

ACKNOWLEDGEMENTS

The fingerprint analysis would not have been done without the knowledge, technical expertise, equipment, and generous hands-on efforts of Dr D.A.Steege (Duke). We thank A.Zaug and T.Cech (University of Colorado) for providing pAtRNA-1 and suggestions prior to their publication. We thank S.P.Rosenstein and A.T.Perrotta for comments on the manuscript. This work was supported by an NIH grant (GM-40689) and a North Carolina Biotechnology Academic Research Initiation Grant. This project was initiated by M.D.B. while supported by a Junior Faculty Research Award from the American Cancer Society (JFRA-233).

REFERENCES

1. Cech,T.R. and Bass,B.L. (1986) *Annu. Rev. Biochem.* **55**, 599–629.
2. Cech,T.R. (1990) *Annu. Rev. Biochem.* **59**, 543–568.
3. Sullivan,F.X. and Cech,T.R. (1985) *Cell* **42**, 639–648.
4. Inoue,T., Sullivan,F.X. and Cech,T.R. (1986) *J. Mol. Biol.* **189**, 143–165.
5. Zaug,A.J. and Cech,T.R. (1986) *Science* **231**, 470–475.
6. Zaug,A.J., Been,M.D. and Cech,T.R. (1986) *Nature* **324**, 429–433.
7. Szostak,J.W. (1986) *Nature* **322**, 83–86.
8. Waring,R.B., Towner,P., Minter,S.J. and Davies,R.W. (1986) *Nature* **321**, 133–139.
9. Been,M.D. and Cech,T.R. (1986) *Cell* **47**, 207–216.
10. Michel,F., Hanna,M., Green,R., Bartel,D.P. and Szostak,J.W. (1989) *Nature* **342**, 391–395.
11. Burke,J.M., Belfort,M., Cech,T.R., Davies,R.W., Schweyen,R.J., Shub,D.A., Szostak,J.W. and Tabak,H.F. (1987) *Nucleic Acids Res.* **15**, 7217–7221.
12. Suh,E.R. and Waring,R.B. (1990) *Mol. Cell. Biol.* **10**, 2960–2965.
13. Price,V.J, Engberg,J. and Cech,T.R. (1987) *J. Mol. Biol.* **196**, 49–60.
14. Davies,R.W., Waring,J.R., Brown,T.A. and Scazzocchio,C. (1982) *Nature* **300**, 719–724.
15. Davanloo,P., Rosenberg,A.H., Dunn,J.J. and Studier,F.W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2035–2039.
16. Xu,M.-Q., Scott,D., Kathe,S.D., Goodrich-Blair,H., Nierzwicki-Bauer,S.A. and Shub,D.A. (1990) *Science* **250**, 1566–1570.
17. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. Cold Spring Harbor.
18. Tabor,S. and Richardson,C.C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4767–4771.
19. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
20. Been,M.D. and Perrotta,A.T. (1990) *Science* **252**, 434–437.
21. Barrell,B.G. (1971) in Cantoni,G.L. and Davies,D.R. (eds), *Procedures in Nucleic Acid Research*. Harper and Row, New York. Vol. 2, pp. 751–779.
22. Steege,D.A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4163–4167.
23. Walstrum,A.A. and Uhlenbeck,O.C. (1990) *Biochemistry* **29**, 10573–10576.
24. Woodson,S A. and Cech,T.R. (1991) *Biochemistry* **30**, 2042–2050.
25. Symons,R.H. (1989) *TIBS* **14**, 445–450.
26. Wang,K.-S., Choo,Q.-L., Weiner,A.J., Ou,J.-H., Najarian,R.C., Thayer,R.M., Mullenbach,G.T., Denniston,K.J., Gerin,J.L. and Houghton,M. (1986) *Nature* **323**, 508–514.
27. Kos,A., Dijkema,R., Arnberg,A.C., van der Meide,P.H. and Schellekens,H. (1986) *Nature* **323**, 558–560.
28. Goldenberg,D.P. and Creighton,T.E. (1983) *J. Mol. Biol.* **165**, 407–413.
29. Pan,T., Gutell,R.R. and Uhlenbeck,O.C. (1991) *Science* **254**, 1361–1364.
30. Belinky,M.G. and Dinter-Gottlieb,G. (1991) *Nucleic Acids Res.* **19**, 559–564.
31. Rosenstein,S.P. and Been,M.D. (1990) *Biochemistry* **29**, 8011–8016.
32. Inoue,T. and Cech,T.R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 648–652.