The EMBO Journal vol.12 no.9 pp.3599-3605, 1993

Novel RNA polymerization reaction catalyzed by a group I ribozyme

Bharat M.Chowrira[1], Alfredo Berzal-Herranz[2] and John M.Burke[3]

Department of Microbiology and Molecular Genetics, Markey Center for Molecular Genetics, Stafford Hall, The University of Vermont, Burlington, VT 05405, USA

[1]Present address: Ribozyme Pharmaceuticals Inc., Boulder, CO 80301, USA
[2]Present address: Centro de Investigaciones Biologicos (CSIC), 28006 Madrid, Spain

[3]Corresponding author

Communicated by P.A.Sharp

We have converted a bacterial tRNA precursor containing a 205 nt self-splicing group I intron into a RNA enzyme that catalyzes polymerization of an external RNA substrate. The reaction involves transesterification steps analogous to both the forward and reverse exon ligation steps of group I splicing; as such it depends entirely on 3′ splice site reactions. The RNA substrate is a 20 nt analogue of the ligated exons (E1·E2), whose 3′ end resembles the 3′ terminus of the intron RNA enzyme (IVS). The splice junction of the substrate is attacked by the 3′ end of the intron, then the molecule displaces the original 3′ terminal guanosine so that the new 3′ terminus is brought into the active site and used as the attacking nucleophile in the next reaction. Polymerization occurs via a series of covalent enzyme-linked intermediates of the structure IVS·$(E2)_n$, where n = 1 to ≥18. The 5′ exon accumulates during the course of the reaction and can attack the covalent intermediates to produce elongation products of structure E1·$(E2)_n$, regenerating the intron RNA enzyme in unchanged form. In this manner, the enzyme converts 20 nt oligoribonucleotides into polyribonucleotides up to at least 180 nt by 10 nt increments. These results have significant implications for the evolution of RNA-based self-replicating systems.

Key words: ribozyme/RNA/RNA polymerization

Introduction

In addition to their role as carriers of genetic information, catalytic RNA molecules, or ribozymes, exhibit a range of activities centering on the cleavage and formation of phosphodiester bonds in nucleic acids (Cech *et al.*, 1992; Symons, 1992). The discovery of the catalytic properties of RNA was closely followed by proposals that RNA self-replication may have formed the biochemical basis for the prebiotic evolution of self-replicating systems (Pace and Marsh, 1985; Sharp, 1985; Cech, 1986; Darnell and Doolittle, 1986; Orgel, 1986; Joyce, 1989). Indeed, speculation along these lines preceded the discovery of catalysis by RNA (Woese, 1967; Crick, 1968; Orgel, 1968).

The concept of RNA self-replication has been supported by several observations concerning reactions catalyzed by group I intron RNAs. The ability of a catalytic RNA molecule to mediate the assembly of larger species was first demonstrated by the finding that the *Tetrahymena* group I intron RNA can react in an intermolecular manner to form intron oligomers (Zaug and Cech, 1985). However, the relevance of this reaction to RNA self-replication is uncertain, since it involves the multimerization of very large pre-formed (~140 kDa) polymers. The *Tetrahymena* intron RNA also catalyzes a disproportionation reaction using an oligocytidine substrate, using steps analogous to intron cyclization and exon ligation (Zaug and Cech, 1986).

The 5′ splice site reactions catalyzed by group I introns have been used to catalyze the addition of nucleotides (Kay and Inoue, 1987; Been and Cech, 1988; Bartel *et al.*, 1991) and oligonucleotides (Szostak, 1986; Doudna and Szostak, 1989; Doudna *et al.*, 1991, 1993; Green and Szostak, 1992) to the 3′ end of RNA 'primers'. Reaction schemes used for these experiments have been based on reversal of the first step of splicing, i.e. attack of an exon 1 (E1) analogue on a G·N linkage (where N is a mono- or oligonucleotide) to produce E1·N and guanosine. Template-dependent addition of mononucleotides is relatively inefficient, resulting in the addition of only 1–5 nucleotides to the 3′ end of the primer. Assembly of oligonucleotides by this mechanism can generate significantly longer polymers (Doudna *et al.*, 1991); recently, it has been shown that a series of oligomers can be assembled into RNA products by a series of sequential steps (Doudna *et al.*, 1991; Green and Szostak, 1992). For all of these reactions, sequence specificity results primarily from helix P1, an interaction that involves association of sequences at the 5′ splice site with an internal or external guide sequence (Been and Cech, 1986; Garriga *et al.*, 1986; Waring *et al.*, 1986).

We wished to examine details of reactions at the 3′ splice site that are catalyzed by group I ribozymes. At 205 nt, the $tRNA^{Ile}$ intron from the bacterium *Azoarcus* is the smallest known naturally occurring group I intron for which self-splicing activity has been demonstrated (Reinhold-Hurek and Shub, 1992). The secondary structure model for this intron (Figure 1) exhibits all characteristic structural elements of group I introns. The 5′ splice site follows a U·G wobble pair within a short stem (P1). The 3′ splice site follows a characteristic G and is selected using two short helices, P9.0 and P10. As for all group I introns, a sequence near the 5′ end of the intron termed the internal guide sequence (IGS, Davies *et al.*, 1982) serves to align both the 5′ and 3′ splice sites. In the case of the *Azoarcus* intron, these sequences are overlapping and the resulting helices (P1 and P10) are predicted to have only very modest stability. Stacking of either or both of these helices on the anticodon stem of the tRNA precursor might confer additional stability in pre-tRNA splicing.

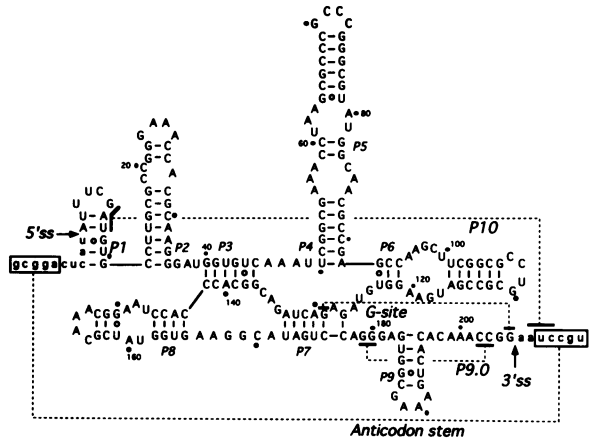

Fig. 1. Secondary structure of the $tRNA^{Ile}$ group I intron derived from *Azoarcus* sp. (Reinhold-Hurek and Shub, 1992). Upper case, intron sequences; lower case, exon sequences. Arrows indicate the 5′ and 3′ splice sites. P1–P10 are phylogenetically conserved structural elements of group I introns (Burke *et al.*, 1987). Nucleotides 6–11 constitute the internal guide sequence (IGS). G-site, guanosine binding site in catalytic core of the intron (Michel *et al.*, 1989; Been and Perrotta, 1991; Yarus *et al.*, 1991).

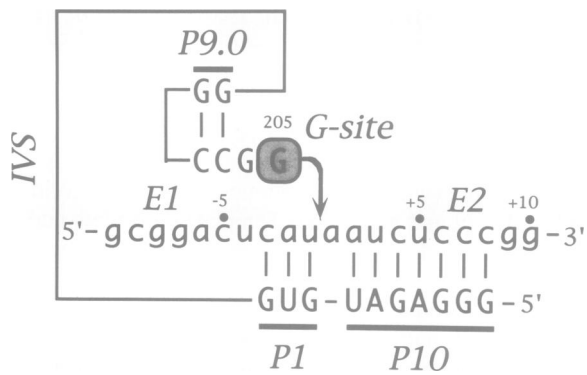

Fig. 2. Alignment of modified *Azoarcus* intron with the ligated exons before reverse exon ligation. IVS, intervening sequence (intron); E1, exon 1; E2, exon 2. Nucleotides corresponding to the unmodified intron sequence (Figure 1) are numbered 1–205. Exon numbering is relative to the splice junction; nucleotides in exon 1 have negative numbers, while those in exon 2 have positive numbers.

For analysis of 3′ spice site reactions involving the *Azoarcus* intron, we chose to examine the reverse exon ligation reaction (Figure 2), previously demonstrated for the *Tetrahymena* rRNA intron (Woodson and Cech, 1989, 1991). This reaction begins with excised intron (IVS) and ligated exons (E1·E2) and proceeds via attack of the terminal guanosine of the intron (G_{205}) on the splice junction. The expected reaction products are the intron linked to exon 2 (IVS·E2) and free exon 1. These experiments have led to the discovery of a novel RNA polymerization reaction catalyzed by the group I ribozyme, the ribozyme elongates an external 20 nt oligoribonucleotide into polyribonucleotides up to more than 180 nt in length.

Results

RNA-catalyzed RNA polymerization

A transcriptional template for the modified intron RNA was prepared using polymerase chain reaction (PCR) methods as described in Materials and methods. An *in vitro* transcript of this template yielded the modified IVS RNA terminating with G_{205}. A 20 nt analogue of the ligated exons (E1·E2) was synthesized using solid phase methods. This molecule includes 10 nt of exon 1 derived from the tRNA sequence and a 10 nt exon 2 analogue. Because of the expected instability of the 3 bp wild type P10 structure, we modified the 3′ exon sequence to stabilize P10 and destabilize the anticodon stem. Thus, the exon 2 analogue retains four bases of the tRNA exon proximal to the splice junction, four bases of unrelated sequence and two bases of tRNA sequence at its 3′ end. The predicted P10 structure of the complex between the modified intron and synthetic ligated exons includes seven base pairs and exon 2 ends with an unpaired GG sequence (Figure 2).

When partially 3′ end-labeled (only 0.8% of the 3′ ends are pCp-labeled) ligated exon RNA (E1·E2) was incubated with unlabeled modified *Azoarcus* intron RNA (IVS), a prominent series of radiolabeled species of higher molecular weight than the input material was obtained (Figure 3A). The details of this apparent RNA polymerization reaction were investigated.

At all time points, high molecular weight RNA was observed. This RNA corresponds in size to the predicted product of the reaction, IVS·E2. A series of bands intermediate in size between E1·E2 and IVS·E2 was also observed. These species appear as a distribution with regular spacing. Comparison with a RNA hydrolysis ladder indicated that bands occurred at 10 nt intervals (data not shown). The size distribution of these intermediate species increases with time, such that after 90 min, RNA products >180 nt could be detected. Retention of the radioactive label in these RNA species suggested that they contained E2 at their 3′ end.

Closer inspection of the IVS-sized band on appropriate gels revealed that it is not the homogeneous species IVS·E2 (Figure 3B). Rather, it too is a series of bands larger in size than the input IVS RNA. The size distribution of the series of larger bands also increases with time.

The polymerization reaction was repeated using 5′ end-labeled E1·E2 (Figure 3C). As for the 3′ end-labeled E1·E2, a 10 nt ladder of bands was observed, indicating that these reaction products probably contained exon 1 on their 5′ end. Two significant differences were noted. First, RNA species equal to or greater than the size of the IVS were not observed. However, a high molecular weight species was obtained when 5′ end-labeled E1·E2 was used (Figure 3C). RNA sequencing showed that this species does not contain the full intron RNA. Instead, it results from attack of E1 on a site within the intron, resulting in a 214 nt product (Been and Cech, 1988; data not shown). Second, a prominent band 10 nt shorter than E1·E2 was observed. RNA sequence analysis showed that this species was the free 5′ exon, E1 (data not shown).

The polymerization products of intermediate mobility were characterized by enzymatic RNA sequencing (Figure 4). To obtain starting material for sequencing, a reaction was carried out using 3′ end-labeled E1·E2, and the reaction products of sizes corresponding to 30 and 40 nt were purified and sequenced. Results showed that the 30 nt species had the structure E1·E2·E2, containing a tandem duplication of the 10 nt 3′ exon. The 40 nt species was E1·E2·E2·E2. These results are consistent with the 10 nt spacing observed in the polymerization products and indicate that the structure

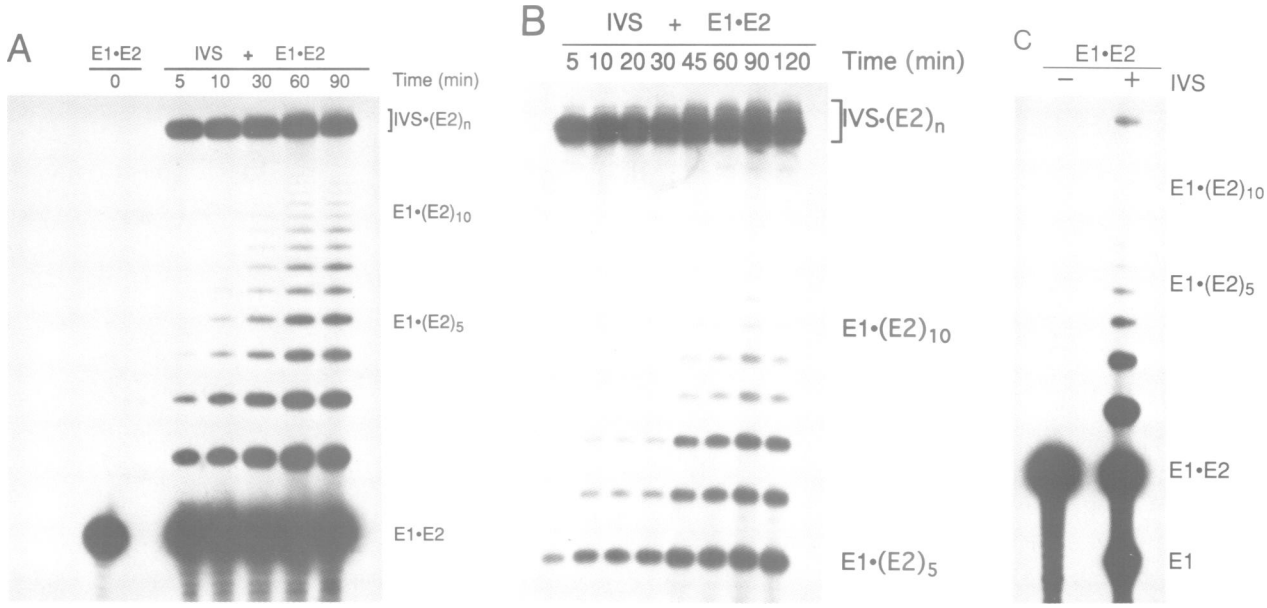

Fig. 3. (**A**) Time course of RNA polymerization. E1·E2, ligated exons; IVS·$(E2)_n$, multimers of exon 2 linked to the 3′ end of the modified intron RNA; E1·$(E2)_5$ and E1·$(E2)_{10}$, exon 1 covalently linked to five or 10 tandem repeats of exon 2, respectively. Autoradiogram of a 10% denaturing polyacrylamide gel is shown. (**B**) The products of a RNA polymerization time course, similar to the one described in panel A, were resolved by extended electrophoresis on an 8% denaturing polyacrylamide gel to resolve high molecular weight reaction intermediates. Polymerization products shorter than E1·$(E2)_5$ were allowed to run off the gel. (**C**) RNA polymerization reaction using 5′ end-labeled E1·E2. Nomenclature is as described in panel A. Minus and plus signs indicate the absence and presence of unlabeled intron RNA.

of the series of polymerization products is E1·$(E2)_n$. Thus, the intron RNA is acting to polymerize exon 2 on the 3′ end of exon 1.

Polymerization requirements

Conditions for the RNA polymerization reaction were characterized (Figure 5). The reaction increased with temperature to an optimum at approximately 50°C (Figure 5A). As the concentration of ligated exons increased, the reaction rate saturated at ~10 μM E1·E2 (Figure 5B). We estimate that K_m for the ligated exons is ~5 μM using these RNA constructs. The polymerization reaction requires Mg^{2+} and efficient reactions were observed at Mg^{2+} concentrations >20 mM (data not shown).

Transesterification reactions catalyzed by group I introns typically involve the 3′-hydroxyl group of G attacking a phosphodiester to the 3′ side of U (as for the first step of splicing) or the 3′-hydroxyl group of U attacking a phosphodiester 3′ of G (as for exon ligation; Cech *et al.*, 1981; Inoue *et al.*, 1985; Kay and Inoue, 1987). Therefore, we predicted that the G at the 3′ end of E1·E2 (G_{+10}) was likely to be important in the polymerization mechanism. Substitution of G_{+10} with A blocked polymerization, but did not inhibit formation of IVS·E2 (Figure 5C). We therefore conclude that G_{+10} is essential for the polymerization activity, but is not important for the reverse exon ligation reaction.

Model for polymerization mechanism

We noted that the four bases at the 3′ end of exon 2 are identical to those at the 3′ end of the IVS (CCGG). In the exon ligation step of splicing, those positions in the IVS participate in two interactions known to be important for aligning the 3′ splice site with the catalytic core. The CC sequence forms part of P9.0 (Burke, 1989; Burke *et al.*, 1990; Michel *et al.*, 1990) and the terminal G interacts with the G-site (Figures 1 and 2). Because of this similarity, we

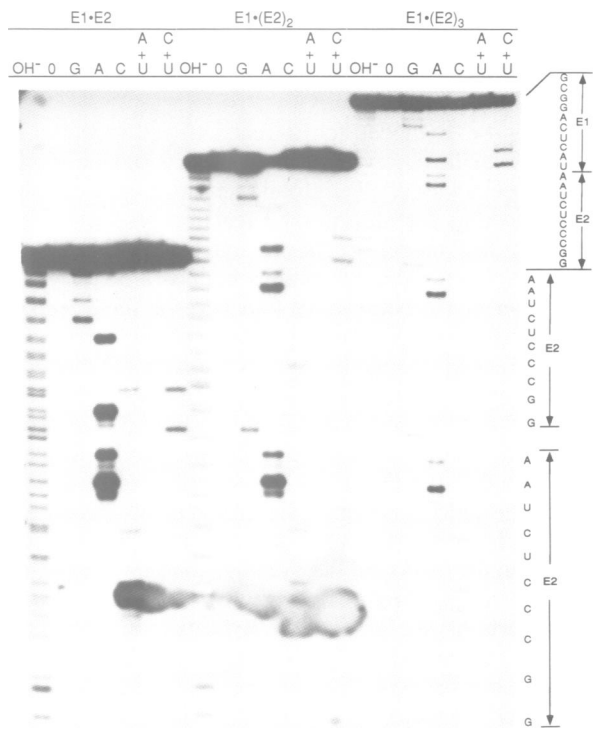

Fig. 4. Sequence analysis of the RNA polymerization reaction products. Enzymatic RNA sequencing was carried out on reaction products obtained in a reaction using 5′ end-labeled E1·E2. Autoradiogram of a 20% denaturing polyacrylamide gel is shown. Derived sequences are indicated. OH^-, partial alkaline hydrolysis ladder. 0, incubation of RNA in the absence of any nucleases or alkali.

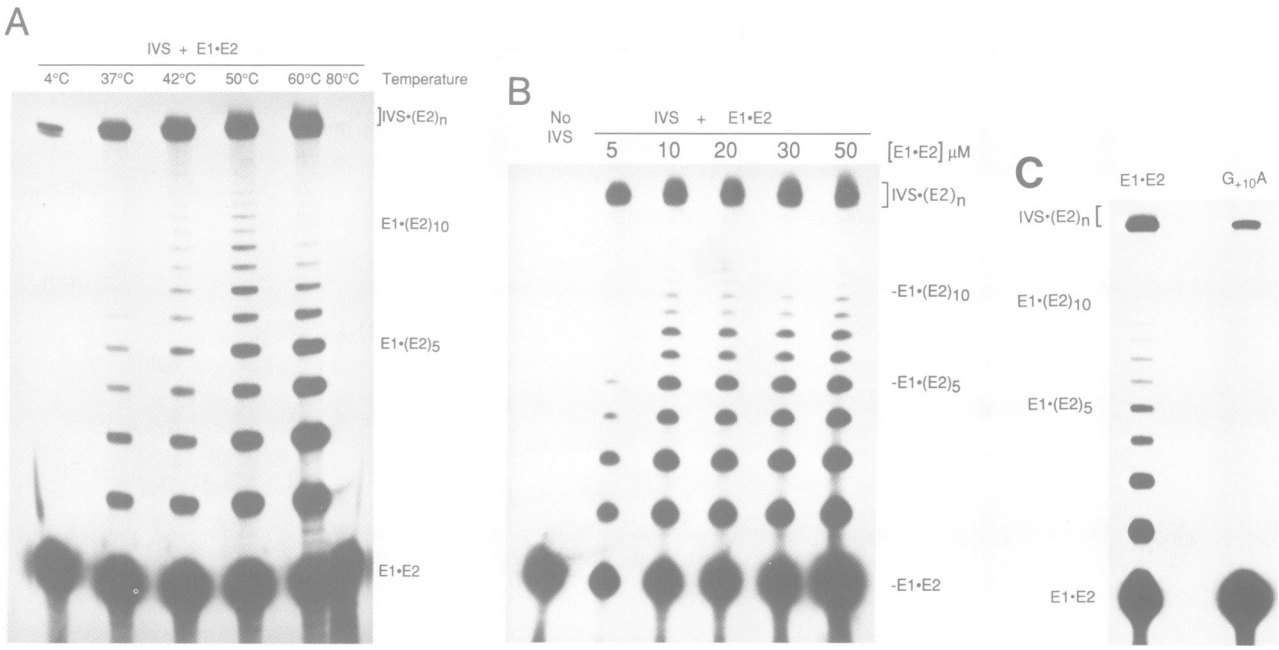

Fig. 5. (**A**) Temperature dependence of the polymerization reaction. Reaction and electrophoresis conditions are as described in Materials and methods except that reaction temperatures were varied from 4 to 80°C. Reactions were stopped after 60 min. (**B**) Dependence of the polymerization reaction on E1·E2 concentrations. The reaction conditions were as described in Materials and methods. Reactions were terminated after 30 min. (**C**) The 3′ terminal G of E1·E2 is essential for polymerization, but not reverse exon ligation. The G at the 3′ end of E2 (position +10) was changed to A. Reaction conditions are as described in Materials and methods, and the reaction was stopped after 30 min.

postulated that C_{+7} and C_{+8} in E2 could form an alternative P9.0 structure that would bring the 3′ end of exon 2 into the active site of the intron.

We propose a model for a reaction pathway that is consistent with the observed reaction products and the known reaction specificities of other group I introns (Figure 6). This model acccounts for all observed reaction products. Two features of this model are novel and essential for polymerization. First, exon 2 functions both as a 3′ exon and an intron terminus, because the four bases at the 3′ end of exon 2 are identical to those at the 3′ end of the intron RNA (CCGG). Second, the model invokes bidirectional displacement. The enzyme must move RNA through its active site, first in a 5′ to 3′ direction and then in a 3′ to 5′ direction, in order for the observed reaction products to be formed.

In our model, the IVS RNA initiates polymerization by attacking the splice junction of the ligated exons, producing IVS·E2 and free E1 (Figure 7, steps 1 and 2) in a reaction analogous to the reverse of the exon ligation step of splicing. This initial step is then followed by 5′ to 3′ displacement (step 3), such that the 3′ end of IVS·E2 is brought into the active site of the ribozyme to serve as the attacking nucleophile for the next reaction. Next, IVS·E2 binds and attacks a second molecule of E1·E2, yielding IVS·E2·E2 and E1 (steps 4 and 5). Further cycles of 5′ to 3′ displacement and reaction result in the production of a series of molecules of structure IVS·$(E2)_n$, which represent covalent enzyme-linked intermediates of the polymerization reaction. In this model, each round of displacement and reaction increases the size of the IVS by 10 nt, accounting for the ladder of bands shown in Figure 3B. In each round of reaction, an additional E2 sequence is added to the 3′ end of the IVS RNA and a molecule of E1 is released.

After production of the molecule IVS·E2·E2, displacement in the 3′ to 5′ direction aligns the authentic 3′ splice

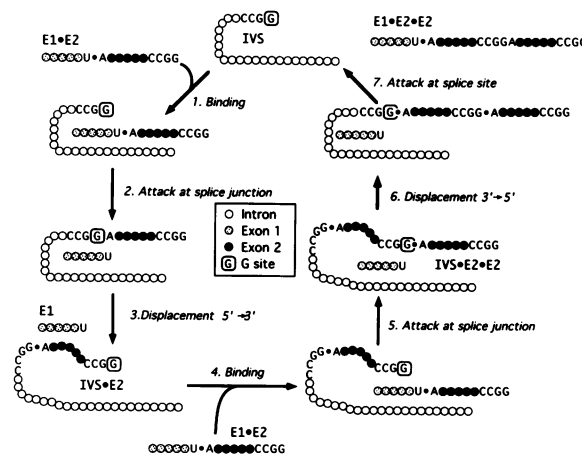

Fig. 6. Reaction mechanism for RNA polymerization catalyzed by the modified *Azoarcus* group I ribozyme.

site (i.e the IVS·E2 junction) with the active site (step 6). Attack of E1 on the 3′ splice site (step 7), a reaction equivalent to the exon ligation step of splicing, generates the elongation product E1·E2·E2 and regenerates the enzyme in unchanged form. For this series of steps, the net reaction is 2 E1·E2 → E1 + E1·E2·E2. In a similar fashion, 3′ to 5′ displacement of a molecule of structure IVS·$(E2)_n$ followed by E1 attack would be expected to result in products with structure E1·$(E2)_m$ and IVS·$(E2)_p$, where m + p = n, if all potential 3′ splice sites were reactive.

Analysis of polymerization mechanism

The proposed mechanism makes four testable predictions. First, free exon 1 accumulates during the course of the polymerization reaction. Second, the series of high molecular

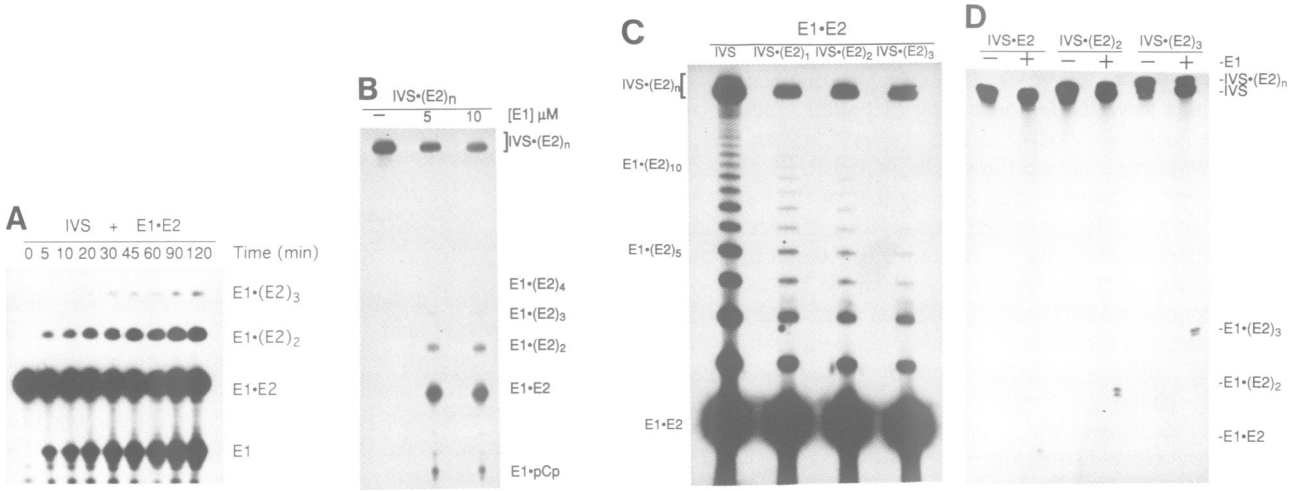

Fig. 7. Analysis of polymerization mechanism. (**A**) Accumulation of free E1 during polymerization. 5′ end-labeled E1·E2 was incubated in a standard polymerization assay. Polymerization reaction products longer than E1·$(E2)_3$ are not shown. (**B**) IVS·$(E2)_n$ reacts with free exon 1 to yield E1·$(E2)_n$. Bands corresponding to 3′ end-labeled IVS·$(E2)_n$ were isolated from a preparative scale polymerization reaction. This mixture of higher molecular weight RNA molecules was incubated in the absence (−) or presence of unlabeled E1 at the indicated concentrations in a standard polymerization reaction. The band corresponding to E1·pCp results from attack of E1 on the phosphodieser linkage between G_{+10} of exon 2 and pCp, used to 3′ end-label the RNA. (**C**) Synthetic covalent intermediates IVS·$(E2)_{1-3}$ react with ligated exons to yield IVS·$(E2)_{n+1}$ and polymerization products. Synthetic versions of the putative enzyme-linked covalent intermediates of the polymerization reaction [IVS·$(E2)_1$, IVS·$(E2)_2$, IVS·$(E2)_3$] were synthesized by transcription of DNA templates containing 1, 2 and 3 exon sequences at the 3′ end of the modified intron. Transcriptional templates were generated by PCR methods as described in Materials and methods. The synthetic putative intermediates were incubated with 3′ end-labeled E1·E2 as described in Figure 3A. (**D**) Synthetic covalent intermediates can react with free E1 to yield IVS and E1·E2. Uniformly labeled IVS·$(E2)_1$, IVS·$(E2)_2$ and IVS·$(E2)_3$ were incubated in the absence (−) and presence (+) of 20 μM synthetic unlabeled exon 1 in a standard polymerization reaction for 30 min. Autoradiogram of a 10% gel is shown. Heterogeneity of reaction products is due to addition of a nontemplated nucleotide during transcription by T7 RNA polymerase (Milligan and Uhlenbeck, 1989).

weight species shown in Figure 3B represent covalent enzyme-linked intermediates of structure IVS·$(E2)_n$. Third, the covalent intermediates can react with ligated exons to add further exon 2 units to the intermediate, in the reaction IVS·$(E2)_n$ + E1·E2 → IVS$(E2)_{n+1}$ + E1. Fourth, the covalent intermediates can react with free exon 1 to produce polymerization products, in the reaction E1 + IVS·$(E2)_n$ → E1·$(E2)_m$ + IVS·$(E2)_p$, where m + p = n. Each of these predictions was experimentally tested and confirmed.

First, accumulation of exon 1 was demonstrated by conducting a polymerization reaction using 5′ end-labeled ligated exons (Figure 7A). As polymerization proceeded, a species corresponding to the expected mobility of E1 was observed to accumulate. The identity of this species as E1 was confirmed by direct RNA sequencing (data not shown), supporting steps 1 and 2 of the proposed reaction scheme.

Second, the high molecular weight species, consisting of approximately six major bands obtained in the reaction, were gel-purified and tested for reaction with synthetic exon 1 (Figure 7B). Four reaction products were observed. The major product was E1·E2, resulting from attack of exon 1 at the last bond formed. Two elongation products were observed, E1·$(E2)_2$ and E1·$(E2)_3$. These products result from attack of exon 1 at sites to the 5′ side of the last bond formed. Formation of these elongation products requires displacement of the 3′ end of the covalent intermediate in a 3′ to 5′ direction relative to the enzyme's active site. Interestingly, a reaction product shorter than E1·E2 was also obtained. Sequencing showed that this species was E1·pCp, formed by reaction of E1 at the 3′ terminus of the RNA containing the radioactive label. Formation of this product requires displacement of the 3′ end of the covalent intermediate in a 5′ to 3′ direction relative to the active site (step 3), and demonstrates that P10 is not absolutely required for exon ligation in this system.

The identity of the enzyme-linked covalent intermediates was confirmed by directly synthesizing the putative intermediates IVS·$(E2)_1$, IVS·$(E2)_2$ and IVS·$(E2)_3$, and demonstrating that they react in the same way as the authentic intermediates obtained from the polymerization reaction. First, the three synthetic intermediates were incubated with 3′ end-labeled ligated exons, in order to test their activity in attacking the splice junction (Figure 7C). In each case, radioactive label was transferred from E1·E2 to the unlabeled intermediates and a series of polymerization products was obtained. Therefore, the three synthetic intermediates can bind and react with E1·E2, displacement in a 3′ to 5′ direction, and react with free exon 1. Second, the synthetic intermediates were tested for their ability to react with synthetic exon 1 (Figure 7D). When uniformly labeled IVS·E2 was incubated with synthetic E1, both of the predicted products, IVS and E1·E2, were formed. Similar reactions were carried out with synthetic IVS·$(E2)_2$ and IVS·$(E2)_3$. The major reaction products were E1·$(E2)_2$ and E1·$(E2)_3$, respectively. These experiments confirmed that exon 1 can attack the covalent enzyme-linked intermediates to form products of the structure E1·$(E2)_n$. Reaction site preference was observed for both IVS·$(E2)_2$ and IVS·$(E2)_3$. IVS·$(E2)_2$ reacts to produce E1·$(E2)_2$ as the major product and E1·E2 as a minor product. Similarly, IVS·$(E2)_3$ reacts to produce E1·$(E2)_3$ as the major product and E1·$(E2)_2$ as a minor product. Therefore we observed a marked preference for cleavage at the 3′ splice site closest to the catalytic core of the intron under the conditions used, which involved renaturation of the gel-purified synthetic intermediates prior to the reaction.

Discussion

Scenarios for RNA self-replication play a prominent role in discussions of prebiotic evolution. Our results make two contributions towards demonstrating that such speculation is biochemically reasonable. First, we have shown that a RNA enzyme can catalyze the synthesis of a RNA product of a size approximately equivalent to that of the catalyst. Second, we have demonstrated that a RNA molecule can achieve displacement, that is, movement of the active site relative to the substrate used for a repetitive reaction.

The reactions used in the polymerization pathway are those of the reverse and forward exon ligation reactions of the well-established splicing pathway used by RNA molecules containing group I introns. As such, the reaction is completely dependent on 3′ splice site reactions and the system is therefore distinct from previously described group I systems that depend entirely on 5′ splice site reactions (Doudna and Szostak, 1989; Bartel *et al.*, 1991; Doudna *et al.*, 1991, 1993; Green and Szostak, 1992).

A critical structural feature of the system described here is that the 3′ end of exon 2 is bifunctional and can also function as the 3′ end of the intron because the four bases at the 3′ end of exon 2 are identical to those at the 3′ end of the intron RNA (CCGG). Following intron attack at the splice junction E1·E2 to produce IVS·E2, the 3′ end of the IVS bound at the active site dissociates and then is replaced by the 3′ end of E2, which serves as the nucleophile for the next round of reaction. Although our experiments do not directly address the issue of processivity, this movement of the active site of the enzyme relative to the 3′ end of the molecule represents a displacement step that is similar, in some respects, to 'translocation'.

In this system, displacement requires that the 3′ splice site alignment structure (involving P9.0, P10 and the G site structure) be stable enough to permit reactivity, but unstable enough that the reactive structure can disassemble and be reassembled using new sequences in the displacement event. We do not know whether disassembly of the 3′ splice site alignment structure is a passive process or if disassembly is actively triggered by a conformational change that occurs during reverse exon ligation or another step, for example dissociation of P1 as exon 1 leaves.

The generation of reaction products of the structure $E1\cdot(E2)_n$ requires bidirectional displacement. The active site must move in a 5′ to 3′ direction relative to the 3′ splice site as covalent intermediates of the structure $IVS\cdot(E2)_n$ are elongated to $IVS\cdot(E2)_{n+1}$. Subsequently, displacement must occur in the 3′ to 5′ direction before attack of exon 1 to produce the elongation products $E1\cdot(E2)_{n+1}$. These results can be explained most simply by a model in which the direction of displacement is stochastic. Thus, when the 3′ splice site alignment structure disassembles it can refold so as to reassemble the same structure (no displacement), or refold using alternative 3′ splice sites downstream (5′ to 3′ displacement) or upstream (3′ to 5′ displacement) of the previous reaction site. Our results using the synthetic covalent intermediates $IVS\cdot(E2)_1$, $IVS\cdot(E2)_2$ and $IVS\cdot(E2)_3$ show a preference for exon 1 attack at the 3′ splice site closest to the catalytic core. However, these data were obtained using intermediates that were renatured following denaturing gel electrophoresis and may not necessarily reflect the RNA folding and relative site utilization in intermediates as they are formed during the course of the polymerization reaction.

The synthetic mechanism and reaction products described here resemble those of telomerase during the synthesis of repetitive sequences on the 3′ ends of eukaryotic chromosomes (Blackburn, 1991). In both systems, the 3′ end of the chain to be elongated is bound by an RNA template and elongated through repeats of several nucleotides. Following the addition of one oligomeric repeat, the newly synthesized 3′ end dissociates and rebinds so that the next repeat can be added.

The reactions of the RNA enzyme described here differ from telomerase and traditional polymerases in that an internal polynucleotide template is used. The feasibility of utilization of an external template (external guide sequence) by a group I ribozyme in catalysis of 5′ splice site reactions has been demonstrated (Szostak, 1986; Doudna and Szostak, 1989; Doudna *et al.*, 1991; Green and Szostak, 1992; Doudna *et al.*, 1993). We expect that use of an external RNA template will also prove to be possible for the 3′ splice site reactions described here.

Although the reaction that we observe is reasonably efficient, it should be noted that we have not optimized the reaction by systematic manipulation of the stability of the relevant RNA structures, nor have we varied the length and sequence of exon 2. We expect that it may be possible to adapt this reaction to polymerize exon 2 in slightly smaller or significantly longer units, or to use a mixture of exon 2 analogues with different sequences in the same reaction. The use of multiple external guide sequences to polymerize a variety of exon sequences in a complex mixture of oligoribonucleotides may begin to approach the recapitulation of a protein-free primordial RNA replication system in the test tube. Our results, coupled with the recent demonstration that nonenzymatic synthesis of oligoribonucleotides can be catalyzed on the surface of clays (Ferris and Ertem, 1992), suggest that prebiotic polyribonucleotide synthesis could conceivably have occurred through the action of RNA enzymes in the assembly of oligoribonucleotides generated by nonenzymatic mechanisms.

Materials and methods

RNA synthesis

Plasmid pBGB1.1 (Reinhold-Hurek and Shub, 1992) was digested with *Sph*I (BRL) to generate a 0.7 kb fragment containing the group I intron from *Azoarcus* sp. This fragment was gel-purified and used as template for PCR. The 5′ primer (5′-AATACGACTCACTATAGGGAAGCGAAAGAGC-AGGGAGATGTGCCTTGCGCC-3′) and 3′ primer (5′-CCGGTTTGT-GTGACTTT-3′) were synthesized using DNA phosphoramidite chemistry. The 5′ primer incorporates a promoter for T7 RNA polymerase. The PCR product serves as a T7 RNA polymerase transcription template for a modified group I intron where exon 1 and five nucleotides at the 5′ end of the intron (positions 1–5) were deleted and 18 nt unrelated sequence was substituted.

The transcription templates used to synthesize $IVS\cdot E2$, $IVS\cdot(E2)_2$ and $IVS\cdot(E2)_3$ were prepared by PCR using three modified versions of the downstream primer described above. These primers were modified in such a way that either one, two or three copies of exon 2 were added to the 3′ end of the RNA product.

E1·E2 and other RNAs containing 20 or fewer nucleotides were synthesized using solid phase methods, deprotected and gel-purified as described by Scaringe *et al.* (1990), and Chowrira and Burke (1991). RNA phosphoramidites were obtained from Chemgenes. Intron RNA was transcribed from a DNA template using T7 RNA polymerase (Epicenter) basically as described by Milligan and Uhlenbeck (1989), and Chowrira and Burke (1991). Transcription products were separated by electrophoresis through a 10% denaturing polyacrylamide gel. RNA was visualized by UV shadowing and eluted as described by Chowrira and Burke (1991).

End-labeling reactions

Substrate RNAs were 5′ end-labeled with [γ-^{32}P]ATP (NEN) and T4 polynucleotide kinase (US Biochemical) at 37°C for 45 min (Chaconas and van de Sande, 1980; Chowrira and Burke, 1991). Labeling of 3′ end of RNA was accomplished by incubating RNA with T4 RNA ligase (Pharmacia) and [5′-^{32}P]pCp (NEN) at 4°C for 120 min (England *et al.*, 1980). Following the end-labeling reaction, RNAs were gel purified as described above.

Polymerization reactions

In polymerization reactions, 0.4 μM unlabeled modified IVS RNA and 20 μM 3′ end-labeled E1·E2 RNA were denatured at 95°C for 1 min and renatured on ice for 1 min in a buffer containing 40 mM Tris–HCl (pH 7.5) and 100 mM ammonium sulfate. Reactions were initiated by preincubating at 50°C and adding $MgCl_2$ to a final concentration of 30 mM. At the indicated times, 10 μl aliquots were withdrawn and the reaction was quenched by adding an equal volume of formamide loading buffer and freezing on crushed dry ice. The samples were resolved on 10% denaturing polyacrylamide gels. The gels were dried and quantified by radioanalytic imaging with a betascope instrument (Betagen).

RNA sequencing

RNA bands corresponding to 3′ end-labeled E1·E2, E1·$(E2)_2$ and E1·$(E2)_3$ were isolated from preparative gels and sequenced using sequence-specific endoribonucleases T1 (G), Phy M (A + U), *Bacillus cereus* (C + U), U2 (A) and Cl3 (C) (Kuchino and Nishimura, 1989). Enzymes were obtained from US Biochemicals. The products were resolved on 20% sequencing gels.

Acknowledgements

We thank David Shub for the plasmid clone containing the *Azoarcus* intron, Johanna Polsenberg for chemical synthesis of DNA and RNA molecules used in this study and Gerald Joyce for comments on the manuscript. A.B.-H. is a long-term EMBO postdoctoral fellow. J.M.B. is a member of the Vermont Cancer Center. This work was supported by NIH grant GM36981.

References

Bartel,D.P., Doudna,J.A., Usman,N. and Szostak,J.W. (1991) *Mol. Cell. Biol.*, **11**, 3390–3394.
Been,M.D. and Cech,T.R. (1986) *Cell*, **47**, 207–216.
Been,M.D. and Cech,T.R. (1988) *Science*, **239**, 1412–1416.
Been,M.D. and Perrota,A.T. (1991) *Science*, **252**, 434–437.
Blackburn,E.H. (1991) *Trends Biochem. Sci.*, **16**, 378–381.
Burke,J.M. (1989) *FEBS Lett.*, **250**, 129–133.
Burke,J.M., Belfort,M., Cech,T.R., Davies,R.W., Schweyen,R.J., Shub,D.A., Szostak,J.W. and Tabak,H.W. (1987) *Nucleic Acids Res.*, **15**, 7217–7221.
Burke,J.M., Esherick,J.S., Burfeind,W.R. and King,J.L. (1990) *Nature*, **344**, 80–82.
Cech,T.R. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 4360–4364.
Cech,T.R., Zaug,A.J. and Grabowski,P.J. (1981) *Cell*, **27**, 487–496.
Cech,T.R., Herschlag,D., Piccirilli,J.A. and Pyle,A.M. (1992) *J. Biol. Chem.*, **267**, 17479–17481.
Chaconas,G. and van de Sande,J.H. (1980) *Methods Enzymol.*, **65**, 75–85.
Chowrira,B.M. and Burke,J.M. (1991) *Biochemistry*, **30**, 8518–8522.
Crick,F.H.C. (1968) *J. Mol. Biol.*, **38**, 367–379.
Darnell,J.E. and Doolittle,W.F. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 1271–1275.
Davies,R.W., Waring,R.B., Ray,J.A., Brown,T.A. and Scazzocchio,C. (1982) *Nature*, **300**, 719–724.
Doudna,J.A. and Szostak,J.W. (1989) *Nature*, **339**, 519–522.
Doudna,J.A., Couture,S. and Szostak,J.W. (1991) *Science*, **251**, 1605–1608.
Doudna,J.A., Usman,N. and Szostak,J.W. (1993) *Biochemistry*, **32**, 2111–2115.
England,T.E., Bruce,A.G. and van de Sande,J.H. (1980) *Methods Enzymol.*, **65**, 65–74.
Ferris,J.P. and Ertem,G. (1992) *Science*, **257**, 1387–1389.
Garriga,G., Lambowitz,A.M., Inoue,T. and Cech,T.R. (1986) *Nature*, **322**, 86–89.
Green,R. and Szostak,J.W. (1992) *Science*, **258**, 1910–1915.
Inoue,T., Sullivan,F.X. and Cech,T.R. (1985) *Cell*, **43**, 431–437.
Joyce,G.F. (1989) *Nature*, **338**, 217–224.
Kay,P.S. and Inoue,T. (1987) *Nature*, **327**, 343–345.
Kuchino,Y. and Nishimura,S. (1989) *Methods Enzymol.*, **180**, 154–158.
Michel,F., Hanna,M., Green,R., Bartel,D.P. and Szostak,J.W. (1989) *Nature*, **342**, 391–395.
Michel,F., Netter,P., Xu,M.-Q. and Shub,D.A. (1990) *Genes Dev.*, **4**, 777–788.
Milligan,J.F. and Uhlenbeck,O.C. (1989) *Methods Enzymol.*, **180**, 51–62.
Orgel,L.E. (1968) *J. Mol. Biol.*, **38**, 381–393.
Orgel,L.E. (1986) *J. Theor. Biol.*, **123**, 127–149.
Pace,N.R. and Marsh,T.L. (1985) *Orig. Life*, **16**, 97–116.
Reinhold-Hurek,B. and Shub,D.A. (1992) *Nature*, **357**, 173–176.
Scaringe,S.A., Francklyn,C. and Usman,N. (1990) *Nucleic Acids Res.*, **18**, 5433–5441.
Sharp,P.A. (1985) *Cell*, **42**, 397–400.
Symons,R.H. (1992) *Annu. Rev. Biochem.*, **61**, 641–661.
Szostak,J.W. (1986) *Nature*, **322**, 83–85.
Waring,R.B., Towner,P., Minter,S.J. and Davies,R.W. (1986) *Nature*, **321**, 133–139.
Woese,C. (1967) *The Genetic Code*. Harper and Row, New York, pp. 179–195.
Woodson,S.A. and Cech,T.R. (1989) *Cell*, **57**, 335–345.
Woodson,S.A. and Cech,T.R. (1991) *Biochemistry*, **30**, 2050–2056.
Yarus,M., Illangesekare,M. and Christian,E.J. (1991) *J. Mol. Biol.*, **222**, 995–1005.
Zaug,A.J. and Cech,T.R. (1985) *Science*, **229**, 1060–1064.
Zaug,A.J. and Cech,T.R. (1986) *Science*, **231**, 470–475.

Received on April 6, 1993; revised on May 21, 1993