Volume 15 Number 23 1987 Nucleic Acids Research

Structure of the catalytic core of the Tetrahymena ribozyme as indicated by reactive abbreviated forms of the molecule

Gerald F.Joyce and Tan Inoue

The Salk Institute for Biological Studies, PO Box 85800, San Diego, CA 92138, USA

Received July 28, 1987; Revised and Accepted October 21, 1987

ABSTRACT

The precursor rRNA of Tetrahymena thermophila contains a group I intervening sequence (IVS) that catalyzes its own excision to yield mature rRNA. The excised IVS catalyzes a number of cleavage/ligation reactions that are analogous to the transesterification reactions of splicing. We examined the behavior of a variety of 3'-truncated forms of the IVS and found several abbreviated molecules that retained catalytic activity. The reactivity of these molecules indicates that the site at which cleavage/ligation occurs lies in close proximity to all of the conserved sequence elements within the catalytic core of the IVS.

INTRODUCTION

The precursor rRNA of Tetrahymena thermophila contains a group I intervening sequence (IVS) of 413 nucleotides that is able to catalyze its own excision to yield mature rRNA (1). This self-splicing reaction is a two step process. The first step involves nucleophilic attack by guanosine at the 5' splice site, following the sequence CUCU. Guanosine becomes joined to the 5' end of the IVS and the 5' exon is released. The second step involves attack by the $CUCU_{OH}$ terminus of the 5' exon at the 3' splice site, following a guanosine residue. The 5' exon becomes joined to the 5' end of the 3' exon and the IVS is released (2,3).

A variety of other IVS-mediated reactions conform to this basic theme; the attack by a guanosine-terminated oligomer at a phosphodiester bond following a sequence of pyrimidines or the attack by an oligopyrimidine-terminated oligomer at a phosphodiester bond following guanosine (4-6). In the simplest version of this theme, the IVS has been shown to catalyze sequence-specific transesterification reactions between G_{OH} and CUN to yield GN and CU_{OH} and between CU_{OH} and GN to yield CUN and G_{OH}; N = A,C,G,U (7).

After the IVS has been excised from Tetrahymena pre-rRNA it exists as a linear molecule with a terminal guanosine 3'-OH. The G_{OH} can attack a phosphodiester bond following a sequence of pyrimidines located near the 5' end of the molecule to produce a circular species (3). The preferred cyclization site is between nucleotides 15 and 16 following the sequence UUU, and a secondary site exists between nucleotides 19 and 20 following the sequence CCU (8). An RNA molecule consisting

of the 5' exon joined to the IVS (5' exon-IVS) cyclizes at the 5' splice site following the sequence CUCU as well as at the 15-16 and 19-20 sites (9). A different type of cyclization reaction, in which a terminal uridine 3'-OH attacks a phosphodiester bond following guanosine, has been observed using subintronic linear RNA derived from a group I intron in yeast mitochondria (10). Both types of cyclization reaction, involving either attack by guanosine or attack by a sequence of pyrimidines, conform to the known sequence specificity of the IVS.

The IVS of Tetrahymena rRNA is classified as a group I intron, based on the presence of several short, highly-conserved sequence elements that result in a characteristic local secondary structure (11-13). These conserved elements, which are designated (in order from 5' to 3') IGS, 9R', A, B, 9L, 9R, and 2, are thought to form short regions of duplex structure; IGS pairing with the 3' end of the 5' exon, 9R' pairing with 9R, A pairing with B, and 9L pairing with 2 (for review, see 14). Their location within the Tetrahymena IVS is shown schematically in Fig. 1. Although the function of the conserved regions is not fully understood, site-directed mutagenesis studies have suggested that they play an important role in the catalytic activity of the molecule (15-16).

We prepared an ensemble of 3'-truncated forms of the IVS, ranging in length from the full IVS to a shortened version terminating approximately 35 nucleotides upstream from conserved element 9L. We incubated these materials under cyclization conditions and selected those molecules that were able to cyclize. This allowed us to assess the importance of the conserved regions in the cyclization reaction as we compared the reactivity of terminal residues located prior to and beyond sequence

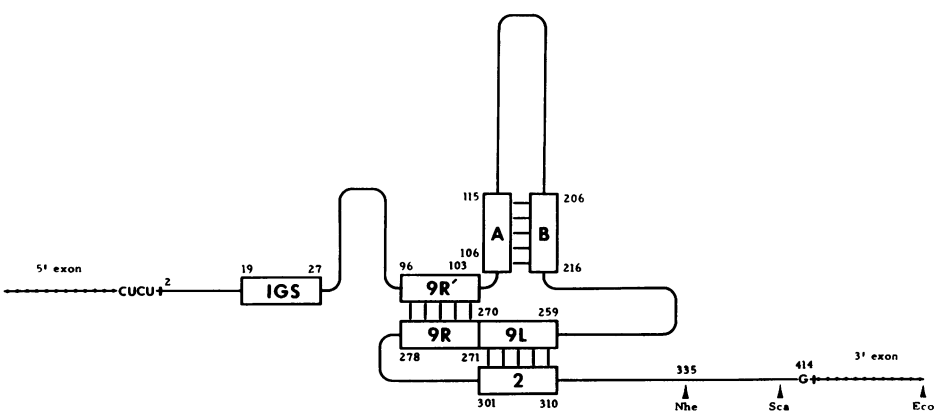


Figure 1: Diagrammatic representation of Tetrahymena pre-rRNA. The conserved sequence elements IGS, 9R', A, B, 9R, 9L, and 2 are shown as boxes, and the proposed base pairing interactions between these elements are shown as hatched lines. In addition, the IGS pairs with the 3' end of the 5' exon. The diagram is not drawn exactly to scale; nucleotide positions within the IVS are indicated by numbers in small font. The base pairing interactions between conserved elements involve only partial duplex structures.

elements 9L, 9R, and 2. It also allowed us to map features of the IVS tertiary structure based on the proximity of reactive 3' termini and their corresponding cyclization sites. In particular, we wished to examine the structural relationship between the conserved sequence elements and the site at which the cleavage/ligation reaction occurs.

MATERIALS AND METHODS

Nucleotides and Enzymes

Unlabeled dinucleotides, nucleoside triphosphates, deoxynucleoside triphosphates, and dideoxynucleoside triphosphates were purchased from Sigma and labeled nucleoside triphosphates from ICN Radiochemicals. Restriction enzymes and Bal31 exonuclease were from New England Biolabs, calf intestine phosphatase from Boehringer, T4 polynucleotide kinase from US Biochemical, RNase T1 and U2 from Pharmacia, RNase A and PhyM from BRL, and reverse transcriptase from Life Sciences. T7 RNA polymerase was prepared as previously described (17), and purified according to a procedure originally developed for SP6 RNA polymerase (18). Plasmid pT7-TT1A3, which contains a 534 base pair fragment of Tetrahymena rDNA (19), was provided by T.R. Cech.

Preparation of 3'-Truncated Forms of the IVS

pT7-TT1A3 was cut with ScaI at a site five nucleotides prior to the end of the IVS segment, and then digested for 5, 7, or 10 min using Bal31 exonuclease (10 units per µg of DNA). The digested materials were pooled to give a broad distribution of truncated templates, which was then used to direct the synthesis of an ensemble of run-off transcripts. The transcription mix contained 125µg of DNA, 1 mM each of the four NTPs (including 400 µCi of [^{3}H]UTP), 15 mM $MgCl_2$, 10 mM dithiothreitol, 2 mM spermidine, 40 mM Tris (pH 7.5), and 12,500 units of T7 polymerase in a volume of 5 ml that was incubated at 37°C for 2 hr. The non-radioactive white precipitate that formed during transcription was removed by centifugation, and T7 polymerase was extracted with phenol. The transcription products were purified by ethanol precipitation and then separated by electrophoresis in a 5% polyacrylamide / 8M urea gel. The gel was cut into fractions corresponding to roughly equal intervals of length over the size range of interest. The RNA contained in each fraction was eluted, twice precipitated in ethanol, and chromatographed on Sephadex G-50.

Selection of Cyclization Products

Each fraction of 3'-truncated forms of the IVS was incubated with 10 mM $MgCl_2$ and 50 mM EPPS (pH 7.5) in a volume of 250 µl at 50°C for 20 min in order to promote cyclization (3). The RNA materials were precipitated in ethanol and then separated by electrophoresis in a 5% polyacrylamide / 8M urea gel. The concentration of Tris/borate in the electrophoresis buffer was reduced to 40 mM (from the usual 90 mM) in order to retard the migration of circular species (20). The slowly-migrating bands (Fig. 2) were visualized by ethidium bromide fluorescence under UV light. The bands were cut from the gel, eluted, twice precipitated in ethanol, and chromatographed on Sephadex G-50.

Sequencing of Cyclization Products

The circular products were sequenced by primer extension analysis using reverse transcriptase in the presence of dideoxynucleotides (21). 1.0 pmol of [5'-^{32}P]-labeled synthetic DNA primer, complementary to positions 67-83 of the IVS, was annealed with 0.3 pmol of circular RNA by incubation at 65°C for 5 min. The primer-extended cDNA products were analyzed on a 10% polyacrylamide / 8M urea sequencing gel.

Reverse Cyclization Using CU or GTP

0.2 pmol of circular RNA was incubated with 2 pmol (2 μCi/pmol) of either [5'-^{32}P]pCpU or [α-^{32}P]GTP in the presence of 100 mM $(NH_4)_2SO_4$, 50 mM $MgCl_2$, and 30 mM Tris (pH 7.5) in a volume of 10 μl at 42°C for 1h in order to promote circle reopening (4). The RNA materials were twice precipitated in ethanol, and in the case of the GTP reactions, further purified by chromatography on Sephadex G-50. The ^{32}P-labeled products were analyzed on a 5% polyacrylamide / 8M urea gel.

Sequencing of Reverse Cyclization Products

The [5'-^{32}P]-labeled linear products resulting from reverse cyclization using [5'-^{32}P]pCpU were sequenced by the enzymatic method of Donis-Keller et al. (22). Partial digestions with ribonucleases T1, A, U2, and PhyM were performed according to the specifications of the supplier, and the products were analyzed on a 20% polyacrylamide / 8M urea sequencing gel. In some cases the circular RNAs were linearized by incubation with unlabeled CpU, and then sequenced by primer extension analysis using reverse transcriptase (21).

RESULTS

Selection of Reaction Products

Plasmid pT7-TT1A3 contains the coding region for a substantial portion of the Tetrahymena pre-rRNA, including 33 nucleotides of the 5' exon, the entire IVS, and 88 nucleotides of the 3' exon, inserted 15 nucleotides downstream from a T7 promoter (19). The plasmid was linearized at a site that lies five nucleotides prior to the 3' end of the IVS segment and then digested briefly with Bal31 exonuclease to produce a distribution of truncated templates. These templates were used to direct the synthesis of an ensemble of run-off transcripts whose 3' termini span the last 185 nucleotides of the IVS. During transcription, some RNA molecules become truncated at their 5' end as well due to attack by GTP at the 5' splice site or 15-16 site (23). The transcripts were separated electrophoretically in a denaturing polyacrylamide gel, and were collected into fractions corresponding to roughly equal intervals of length.

Samples from each fraction were incubated under conditions known to promote cyclization of the intact IVS (10 mM $MgCl_2$, pH 7.5, 42°) (3). Pre- and post-incubation samples were compared side by side after electrophoresis in a denaturing polyacrylamide gel, demonstrating the appearance of several slowly- migrating bands in many of the post-incubation lanes (Fig. 2). These bands appeared

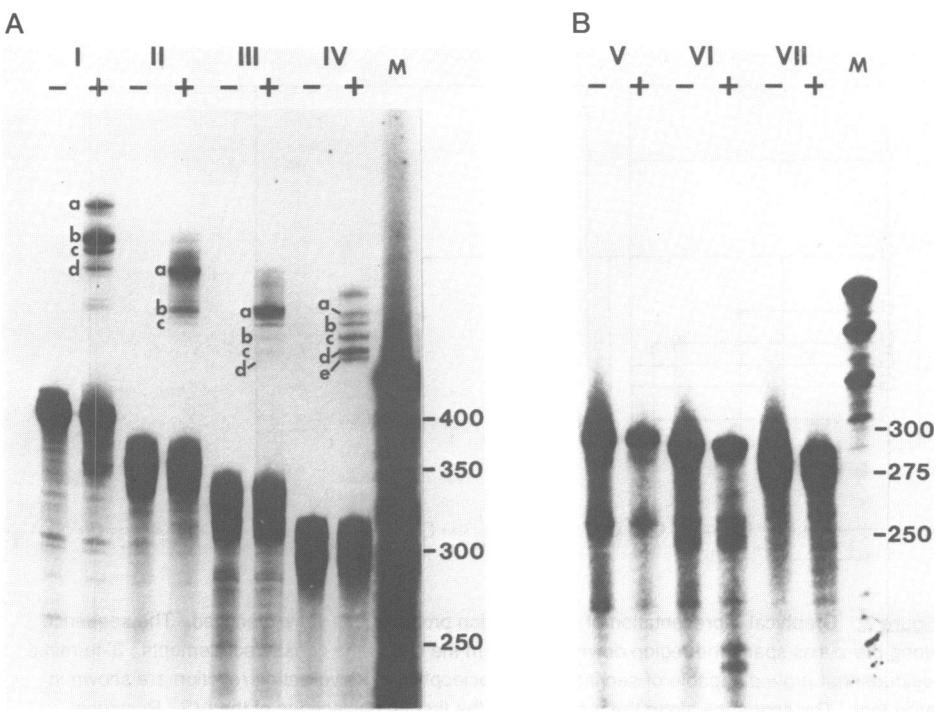


Figure 2: Cyclization reactions of 5'-truncated forms of the IVS. Successive fractions of uniformly-labeled linear RNAs were gel-purified and either not incubated (-) or incubated (+) in the presence of 10 mM $MgCl_2$ and 50 mM EPPS (pH 7.5) at 42° C for 1h. The materials were separated by electrophoresis in a 5% polyacrylamide / 8M urea gel run in 90 mM Tris/borate buffer. The gels were dried and autoradiographed. A) Fractions I-IV, demonstrating the appearance of slowly-migrating bands in the post-incubation lanes. Sixteen of these bands, as labeled in the figure, were cut from the gel and used for further study. B) Fractions V-VII, demonstrating the absence of slowly-migrating bands in reactions containing starting materials with a length less than 310 nucleotides.

in only those fractions containing starting materials with a length greater than approximately 310 nucleotides.

We selected sixteen of the most prominent slowly-migrating bands, as denoted in Fig. 2, for elution, purification, and further study. As an approximation, bands Ia-d were produced from linear molecules having a length of 380-410 nucleotides, bands IIa-c from molecules of length 340-385, bands IIIa-d from molecules of length 315-345, and bands IVa-e from molecules of length 295-320. All of these bands migrate faster than the circular products (of length 399 and 395) that are obtained from cyclization of the intact IVS. Band IVe is the fastest migrating of the group, and to the extent that electrophoretic mobility correlates with size, represents the smallest product that was detected in our survey.

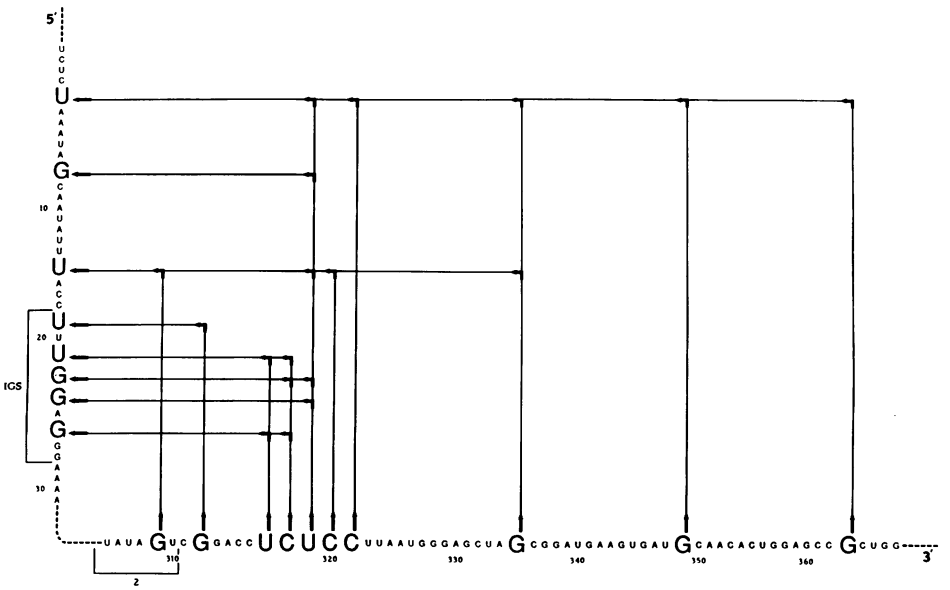


Figure 3: Graphical representation of all cyclization products that were identified. The sequence along the x-axis spans the region downstream from the last of the conserved elements. 3'-terminal residues that proved capable of serving as the nucleophile in a cyclization reaction are shown in large font. The sequence along the y-axis spans the first 30 nucleotides of the IVS. Residues immediately preceding a phosphodiester bond that proved capable of serving as a cyclization site are shown in large font. The arrowed lines connecting a particular 3'-terminal residue to a particular phosphodiester bond near the 5' end of the molecule represent individual circular species that were identified in our study. See Table 1 for information correlating the labeled bands in Fig. 2 to specific circular products in Fig. 3.

Identification of Cyclization Products

We first sought to establish that the collected reaction products are indeed circular molecules rather than some oligomeric linear species. It has been shown, for example, that the linear form of the intact IVS can be converted to dimeric and even trimeric products when incubated under reaction conditions similar to those that we employed (24). However, this oligomerization reaction is known to be absolutely dependent on a specific denaturation/renaturation procedure that presumably allows intermolecular associations to occur. Our truncated linear starting materials were never subject to this type of denaturation/renaturation regime. Furthermore, our reaction products, which appeared as slowly-migrating bands in a polyacrylamide gel, were found to migrate even slower when the concentration of Tris/borate in the electrophoresis buffer was reduced from 90 mM to 40 mM (data not shown), as is characteristic of circular molecules (20). A third indication of the circular nature of these materials came from their use as templates for reverse transcription. In each case, elongation was

Table 1. Identity of Abbreviated Circular Forms of the IVS

Band	Nucleophile	Cyclization Site	Size	Circle Opening	
				CU	GTP
Ia	G_{363}	U_{-1}	362	+	+
Ic	G_{349}	U_{-1}	348	+	+
IIa	G_{335}	U_{-1}	334	+	+
IIb	C_{321}	U_{-1}	320	+	–
	G_{335}	U_{15}	320		
IIIb	U_{319}	U_{15}	304	–	–
	C_{320}	U_{15}	305		
IIIc	U_{317}	U_{21}	296	–	–
	C_{318}	U_{21}	297		
	C_{318}	G_{22}	296		
	U_{319}	G_{23}	296		
IIId	G_{309}	U_{15}	294	–	–
	U_{317}	G_{25}	292		
	C_{318}	G_{25}	293		
IVa	U_{319}	U_{-1}	318	+	+
IVb	U_{319}	G_{7}	312	+	–
IVc	U_{319}	U_{15}	304	+	–
IVd	U_{319}	G_{22}	297	–	+
	U_{319}	G_{23}	296		
IVe	G_{312}	U_{19}	293	–	–
	U_{317}	G_{25}	292		
	C_{318}	G_{25}	293		

The bands described in this table correspond to the slowly-migrating bands as denoted in Fig. 2. The nucleophile is the 3'-terminal residue of the linear starting material. The cyclization site is the residue that precedes the phosphodiester bond at which cyclization occurs. U_{-1} is the 3'-terminal residue of the 5' exon. Circle opening reactions, using either CU or GTP as the nucleophile, were performed as described in the legend to Fig. 4. For bands that contain more than one circular species, the results in the Table describe the aggregate behavior, corresponding to the data presented in Fig. 4a & b.

found to stop at a position corresponding to the full length of a monomeric RNA template, but not at any other position.

We determined the sequence of each circular reaction product by performing primer extension analysis using reverse transcriptase in the presence of dideoxynucleotides (21). A synthetic DNA primer was hybridized to positions 67 through 83 of the circular RNA and then extended through the cyclization junction. This allowed us to identify both the 3' terminus of the linear starting material and the site at which cyclization had taken place. In some cases the sequence diverged to give two or even three superimposed sequence ladders, reflecting the fact that some of the collected bands contained two or more circular species of nearly equal length. However, since the sequence of the entire Tetrahymena IVS is already known (25), it was in most cases possible to decipher the sequence of each circular component without ambiguity.

Figure 3 provides a graphical representation of all cyclization products that were identified in our study. A variety of 3'-truncated forms of the IVS, ending with either guanosine 3'-OH or oligopyrimidine 3'-OH, were found to be capable of cyclizing. The shortest molecules that cyclized end with a guanosine residue inside or just beyond the last of the conserved sequence elements (at positions 309 and 312). There is a long oligopyrimidine sequence at positions 315-323 that provided many of the active 3' termini, and a series of individual guanosine residues at positions 335, 349, and 363 that also demonstrated activity.

All of the cyclization sites were after guanosine or a sequence of pyrimidines located near the 5' end of the IVS. These included such well-established sites as the 5' splice site, the 15-16 site, and the 19-20 site (3,9), as well as the more unusual 21-22 site that had previously been demonstrated only in a mutant form of the IVS that carries an eleven nucleotide deletion near its 5' end (26). Of the eighteen circular molecules that we identified, six appear to have been formed by guanosine attack after oligopyrimidine, six by oligopyrimidine attack after guanosine, and six by oligopyrimidine attack after oligopyrimidine.

Activity of Abbreviated Forms of the IVS

One disadvantage of the selection technique that we employed is that it provides us with only limited knowledge of the events that occurred during the cyclization reaction. We cannot be certain, for example, whether a given circular species was produced by a single-step unimolecular event or a more complicated series of reactions. The linear starting materials were fractionated according to length prior to incubation in order to reduce the complexity of each reaction mixture. However, we cannot exclude the possibility that intermolecular associations may have played a role in the formation of some of the circular products. In order to address this problem, we studied the isolated circular molecules to determine whether they have intrinsic catalytic activity.

The circular form of the intact IVS undergoes specific cleavage by CU_{OH} at its ligation junction ("reverse cyclization"; 4), producing a linear molecule with CU joined to the newly-created 5' end. This is another variant of the familiar reaction involving attack by an oligopyrimidine 3'-OH at a phosphodiester bond following guanosine. One could imagine the converse reverse cyclization reaction involving attack by guanosine after a sequence of pyrimidines. We tested each of the isolated circular species for circle opening activity, using either [5'-^{32}P]pCpU or [α-^{32}P]GTP as the nucleophile. Circles Ia, Ic, IIa, IIb, IVa, IVb, and IVc were all cleaved by [5'-^{32}P]pCpU to give 5'-labeled linear products of appropriate length (Fig. 4a). Only circles Ia, Ic, and IIa were cleaved efficiently by [α-^{32}P]GTP, though circles IVa and IVd showed a trace of activity (Fig. 4b). The efficiency of circle opening varied tremendously, but in all cases was less than that observed for the circular form of the intact IVS.

One of the 3'-truncated forms of the IVS that demonstrated apparent cyclization activity can also be produced by transcription of pT7-TT1A3 that has been cut with NheI (referred to as pT7-TT1A3(NheI)). When this transcript is incubated under cyclization conditions it produces two

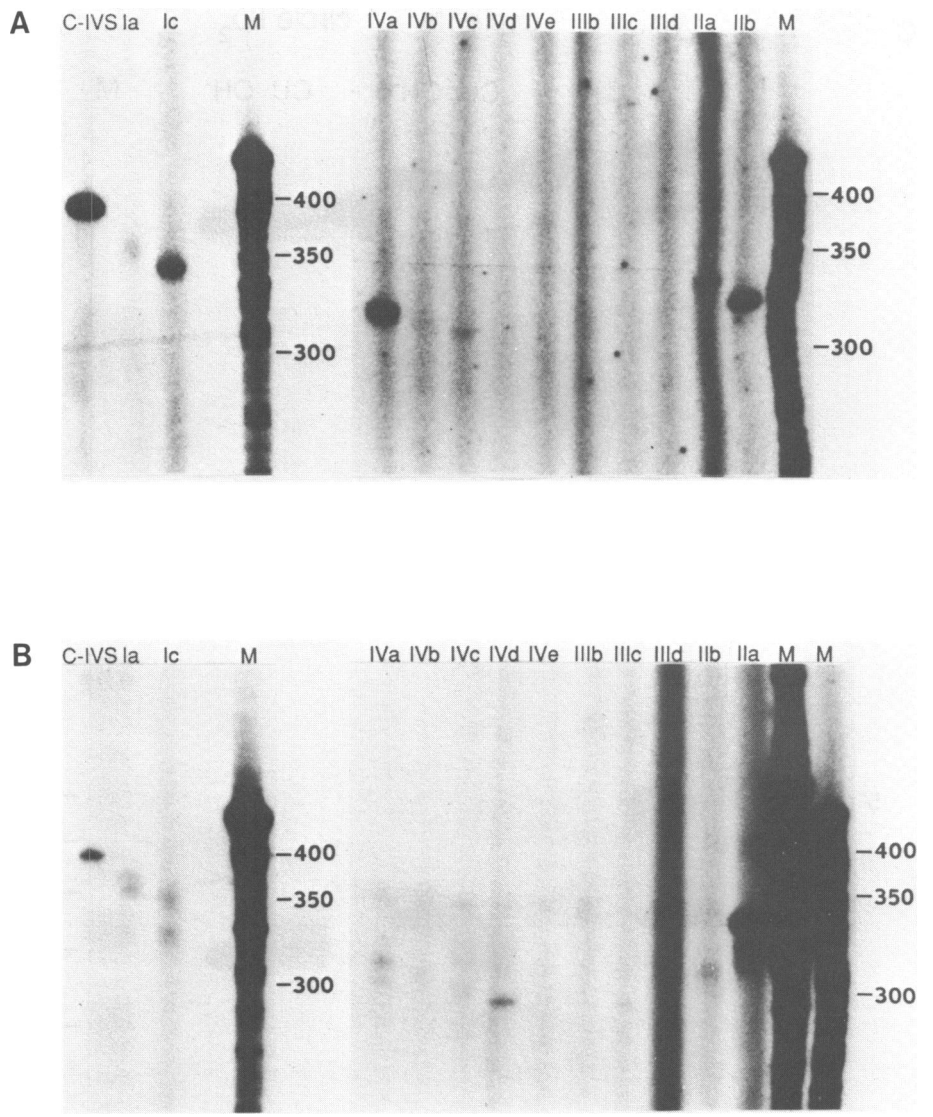
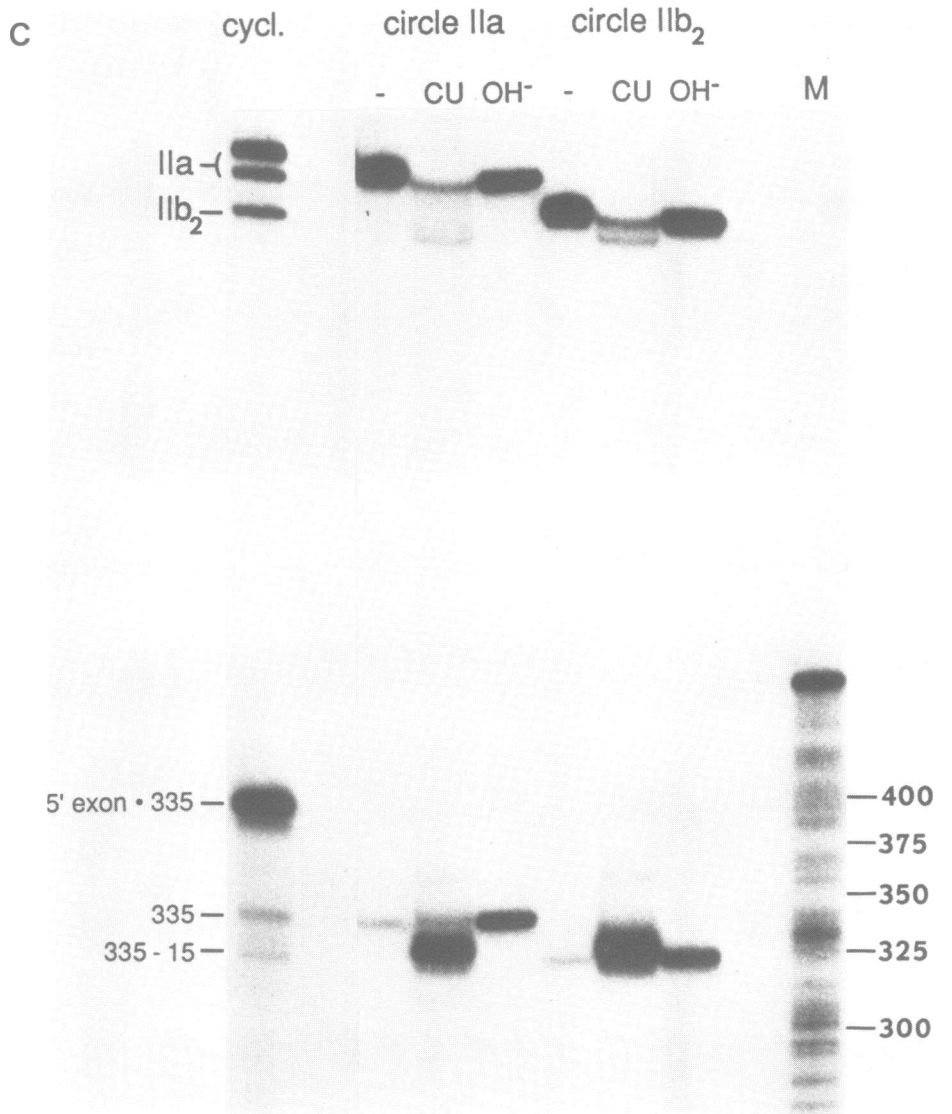
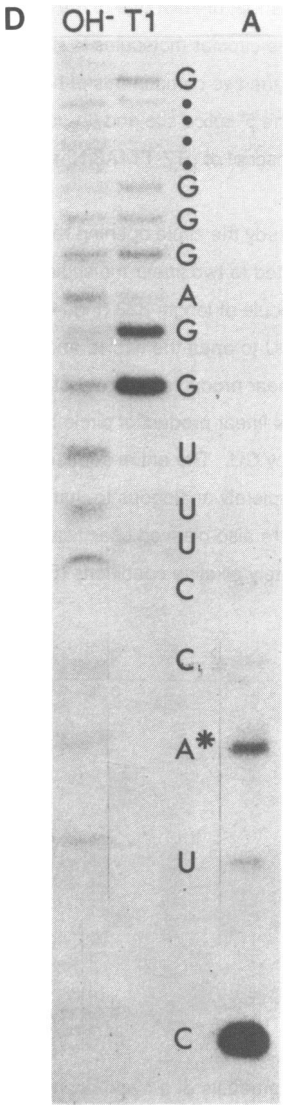


Figure 4: Activity of abbreviated forms of the IVS in circle opening reactions. Individual circular molecules were incubated with either A) [5'-^{32}P]pCpU or B) [α-^{32}P]GTP. The reaction mixture contained 0.3 pmol of circular RNA, 2 pmol (2 μCi/pmol) of CU or GTP, 100 mM $(NH_4)_2SO_4$, 50 mM $MgCl_2$ and 30 mM Tris (pH 7.5) in a volume of 10 μl, which was incubated at 42° C for 1h. The materials were purified by electrophoresis in a 5% polyacrylamide / 8M urea gel run in 40 mM Tris/borate buffer. Autoradiograms are shown. The band(s) in a given lane correspond to the 5' end-labeled linear products of the circle opening reaction. In the majority of lanes, however, no such bands are present. C) Uniformly ^{32}P-labeled RNA, transcribed from pT7-TT1A3(NheI), was incubated



in 10 mM $MgCl_2$ and 50 mM EPPS (pH 7.5) at 42° C for 1h (lane cycl). The resulting circular products, which were identified as IIa and IIb_2 (data not shown), were gel purified and either not incubated (lanes –), or incubated with 5 mM CU, 100 mM $(NH_4)_2SO_4$, 50 mM $MgCl_2$, and 30 mM Tris (pH 7.5) at 42° C for 1h (lanes CU), or with 200 mM NaCl, 10 mM $MgCl_2$, and 50 mM CHES (pH 9.0) at 42° C for 1h (lanes OH^-). The materials were separated by electrophoresis in a 5% polyacrylamide / 8M urea gel run in 40 mM Tris/borate buffer, an autoradiogram of which is shown. Circle IIa appears as a double band in the cyclization lane, although RNA obtained from each of these bands was found to have the same sequence (data not shown). 5' exon·335 is the transcript of pT7-TT1A3(NheI), 335



is the direct product of circle IIa cleavage, and 335-15 is the product of circle IIb_2 cleavage. D) Circle IIb_2 was incubated with [5'-^{32}P]pCpU and the resulting 5' end-labeled linear products were purified by electrophoresis, as described in A. The materials were then partially hydrolyzed by OH-, RNase T1 (specific to G), and RNase A (specific to C and U), and analyzed by electrophoresis in a 20% polyacrylamide / 8M urea sequencing gel, an autoradiogram of which is shown. The sequence listed in the figure is consistent with CU-dependent cleavage at the ligation junction. Cleavage occurred to a very minor extent at the 19-20 site, producing a second sequence ladder which is most evident at the position indicated by the asterisk.

circular molecules. This demonstrates that at least one of the 3'-truncated forms of the IVS has intrinsic cyclization activity. One of the circular molecules is identical to the sole component of band IIa and the other is identical to one of the two components of band IIb (henceforth referred to as IIb_2). IIa is formed by the attack of G_{335} at the 5' splice site and IIb_2 is formed by the attack of G_{335} at the 15-16 site. Beginning with the linear transcript of pT7-TT1A3(NheI), it is possible to synthesize circles IIa and IIb_2 directly.

We used circles IIa and IIb_2 to study the circle opening reaction in some detail. Using CU_{OH} as the nucleophile, circle IIa was converted to two linear molecules of length 336 and 322 and circle IIb_2 was converted to a single linear molecule of length 322 (Fig. 4c). The identity of these compounds was established by using [5'-^{32}P]pCpU to open the circles and then performing enzymatic sequence analysis (22) on the 5' end-labeled linear products. As expected, both IIa and IIb_2 are cleaved by CU at their ligation junction (Fig. 4d). The linear product of circle IIa cleavage then recyclizes to give circle IIb_2, which in turn is reopened by CU. The entire sequence of reactions beginning with the RNA transcript of pT7-TT1A3(NheI) is completely analogous to that which has already been described for 5'exon-IVS (9). Circles IIa and IIb_2 are also cleaved specifically by hydroxyl ion when incubated in the presence of $MgCl_2$ under moderately alkaline conditions (Fig. 4c). This too is an expected result

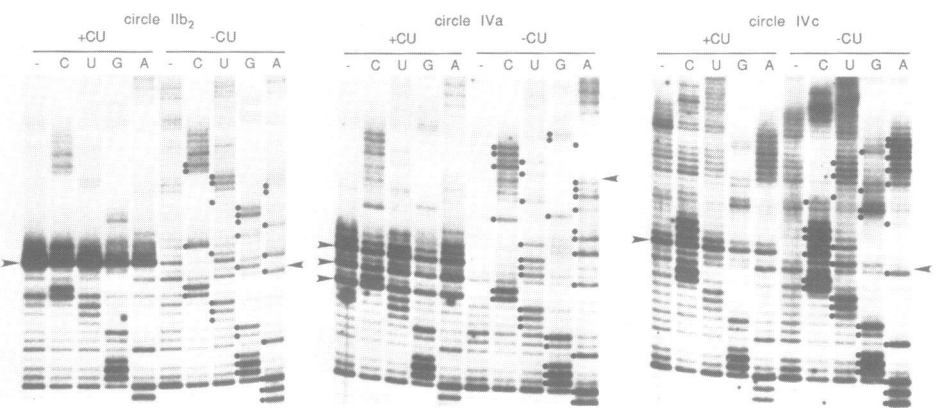


Figure 5: Sequence analysis of the products of a circle opening reaction by the primer extension method. Circles IIb_2, IVa, and IVc were incubated with 100 mM $(NH_4)_2SO_4$, 50 mM $MgCl_2$, and 30 mM Tris (pH 7.5) at 42° C for 1h in either the presence (+CU) or absence (-CU) of 5 mM CU. The resulting materials were purified by ethanol precipitation. A deoxynucleotide complementary to positions 67 to 84 of the IVS was hybridized to the RNA and extended by reverse transcriptase. Lane –, primer extension in the absence of dideoxynucleotides. Lanes C, U, G, and A, primer extension reactions containing ddGTP, ddATP, ddCTP, or ddTTP, respectively. Sequence analysis involved electrophoresis in a 10% polyacrylamide / 8M urea sequencing gel, autoradiograms of which are shown. The sequence of each circular molecule is indicated by dots within the -CU lanes. An arrow to the right of the -CU lanes marks the position of the ligation junction. Arrows to the left of the +CU lanes indicate sites at which CU-dependent cleavage was observed.

based on analogy to the known behavior of the corresponding circular forms of the intact IVS (27,9).

Among the various cyclization products are several species that appear to have been formed by the attack of an oligopyrimidine 3'-OH at a phosphodiester bond following a sequence of pyrimidines (Fig. 3). Circles IVa and IVc, for example, appear to have been formed by the attack of an oligopyrimidine 3'-OH at the 5' splice site and 15-16 cyclization site, respectively. Both of these molecules were cleaved by [5'-^{32}P]pCpU during the circle opening reaction (see above). We opened these circles using unlabeled CpU, and sequenced the resulting linear products by primer extension analysis using reverse transcriptase. Circles IVa, IVc, and (as a control) IIb_2 were incubated in both the presence and absence of CpU. By comparing the two sequence ladders obtained for each circular molecule, it was possible to determine the site(s) at which CU-dependent cleavage had occurred (Fig. 5).

As expected, circle IIb_2 was cleaved by CpU at its ligation junction, producing a characteristic termination point on the sequence ladder. We were surprised to find that circles IVa and IVc were cleaved by CpU at sites following a pyrimidine residue. IVa was cleaved at position 15 (following UUU), position 17 (following UAC), and position 19 (following CCU), although it was not cleaved at the ligation junction (following UCU). IVc was cleaved at position 15, which is its ligation junction, but was not cleaved at positions 17 and 19. Examination of the sequence ladders reveals that none of these sites is particularly susceptible to cleavage in the absence of CpU.

DISCUSSION

Reactivity of Abbreviated Forms of the IVS

Our survey of a large number of 3'-truncated forms of the Tetrahymena IVS demonstrates that several truncated molecules, ending with either guanosine 3'-OH or oligopyrimidine 3'-OH, are capable of cyclizing under physiological conditions. We cannot be certain about the relative abundance of individual species within the population of linear starting materials, since there may be biases in the Bal31 digest procedure and in the transcription process that cause some 3' termini to be either under- or overrepresented. Furthermore, we cannot exclude the possibility that some reactions are prevented from occurring by inappropriate local secondary structure near the 3' end of the molecule. It is nonetheless striking that the only termini that were found to be capable of cyclizing are those that conform to the known sequence-specificity of the IVS (7). Likewise, the only cyclization sites that were observed are those that follow guanosine or a sequence of pyrimidines.

Given the nature of our selection technique, the question can be raised as to whether a particular circular molecule was produced by an intra- or intermolecularly catalyzed reaction. Several of the collected circles were found to catalyze circle opening, using CpU or GTP as the nucleophile. We take this as an indication that the linear forms of these molecules are able to catalyze the forward cyclization reaction. The shortest molecule that cyclizes and then continues to exhibit catalytic activity ends at position 319 following a sequence of pyrimidine residues. Linear molecules, ending with an

oligopyrimidine at position 317 or 318 or ending with a guanosine at position 309 or 312, produced the smallest circular species that were detected in our survey. However, none of these circles exhibited circle opening activity, suggesting that they were formed in an intermolecularly catalyzed reaction.

It is clear from our survey that any truncated molecule which lacks one or more of the conserved sequence elements is incapable of cyclizing. In fact, for truncated molecules whose 3' terminus lies inside or very close to the last of the conserved elements, the circles that do form are without catalytic activity in the circle opening reaction. This is in agreement with phylogenetic data which suggests the importance of the conserved elements based on their presence in all group I introns (13,14). It also supports the results of mutagenesis studies which have shown that conserved elements 9L, 9R, and 2 are essential for the catalytic activity of the IVS (28,15,29,16).

One particularly intriguing finding in our study is that an abbreviated circular form of the Tetrahymena IVS can catalyze an apparent pyrimidine exchange reaction that results in its conversion to a linear molecule. This reaction occurs with good efficiency compared to circle opening involving the circular form of the intact IVS. We found four examples of circle opening via pyrimidine exchange, all of which involved attack by CpU at a phosphodiester bond following a pyrimidine residue. In addition there appear to be six instances of forward cyclization via pyrimidine exchange, involving attack by the oligopyrimidine terminus of a 3'-truncated linear form of the IVS. These reactions are distinct from the oligo(C) disproportionation reaction described by Zaug and Cech (6), which involves the formation of a covalent intermediate between cytidine and the 3' terminal guanosine of the IVS.

Pyrimidine exchange is a transesterification reaction involving attack by the ribose moiety of a nucleotide, resulting in the cleavage of a particular phosphodiester bond and the formation of a new phosphodiester linkage. Mechanistically, all of the IVS-mediated reactions are the same. Ignoring the base sequence of the substrate, there is no difference between pyrimidine exchange and the various splicing-related reactions. In IVS-catalyzed cleavage reactions, free hydroxyl ion rather than ribose 3'-OH is used as the nucleophile, but the result is the same. The target phosphodiester bond is cleaved, and products that have a 5'-phosphate and 3'-OH are obtained (27). It is interesting to note that a chemically equivalent hydrolysis reaction, one that results in products having a 5'-phosphate and 3'-OH, has been observed at the 5' splice site in group II introns (30) and at the cleavage site of pre-tRNA in reactions catalyzed by the RNA component of RNase P (31). This suggests that at the level of the transition state there may be no difference between the reactions catalyzed by group I introns, group II introns, and RNase P.

Catalytic Core of the IVS

There is a set of conserved sequence elements, 9R/9R', A/B, 9L/2 (13,14), whose presence appears to be essential for the catalytic activity of the IVS (see previous section). Examination of the conserved base pairing interactions among these elements and of their position within the primary

structure of the IVS (Fig. 1) demonstrates that all six elements are clustered together within a common region of the molecule. The pairings 9R/9R' and A/B are separated by only two nucleotides between 9R' and A, and the pairings 9R/9R' and 9L/2 are separated by only a phosphodiester bond between 9R and 9L.

There is evidence that the conserved elements are clustered together within the active form of the IVS. Secondary structure analysis of the circular form of the IVS, performed in situ, has suggested that the active form of the IVS contains a large fold so as to allow complementary pairing between conserved elements 9R (positions 271-278) and 9R' (positions 96-103) (32). Site-directed mutagenesis studies using compensatory mutants have demonstrated that 9R/9R' pairing is essential for the catalytic activity of the IVS (16).

The question that we wish to address concerns the relationship between the cluster of conserved elements and the site at which transesterification occurs. Our survey of cyclization activity in truncated forms of the IVS led to the identification of eighteen circular molecules. Based on the location of the cyclization junction, each of these molecules provides a link in a proximity mapping of the tertiary structure of the IVS (Fig. 3). For those molecules that exhibit transesterification activity in a circle opening reaction, the location of the transesterification site is taken as providing structural information about the active form of the IVS. Among the reactive abbreviated circular molecules are two species that contain a linkage bewteen position 319 and either the 5' splice site (circle IVa) or the 15-16 site (circle IVc). Position 319 lies just nine nucleotides away from conserved element 2, indicating that the cluster of conserved elements can lie in close proximity to the site at which transesterification occurs. In the self-splicing reactions of Tetrahymena pre-rRNA and in analogous IVS-catalyzed transesterification reactions the oligopyrimidine, acting as substrate or nucleophile, is held in position by complementary pairing to an "internal guide sequence" (IGS) (33,34). Thus, this supports the idea that the nucleophile, the transesterification site (the electrophilic phosphodiester bond), the IGS, and the cluster of conserved elements 9R/9R', A/B, and 9L/2 all lie in close proximity within the active conformation of the IVS.

ACKNOWLEDGEMENTS

We thank J.M. Burke, P.S. Kay, and L.E. Orgel for helpful comments. This work was supported by grants from the National Institutes of Health (GM35755) and the Alfried Krupp von Bohlen und Halbach-Siftung. G.F.J. is a Merck Fellow of the Life Sciences Research Foundation.

REFERENCES

1. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982) Cell **31**, 147-157.
2. Zaug, A.J. and Cech, T.R. (1982) Nucl. Acids Res. **10**, 2823-2838.
3. Zaug, A.J., Grabowski, P.J., and Cech. T.R. (1983) Nature **301**, 578-583.
4. Sullivan, F.X. and Cech, T.R. (1985) Cell **42**, 639-648.

5. Inoue, T., Sullivan, F.X., and Cech, T.R. (1985) Cell **43**, 431-437.
6. Zaug, A.J. and Cech, T.R. (1986) Science **231**, 470-475.
7. Kay, P.S. and Inoue, T. (1987) Nature **327**, 343-346.
8. Zaug, A.J., Kent, J.R., and Cech, T.R. (1985) Biochemistry **24**, 6211-6218.
9. Inoue, T., Sullivan, F.X., and Cech, T.R. (1986) J. Mol. Biol. **189**, 143-165.
10. Tabak, H.F., Van der Horst, G., Kamps, A.M.J.E., and Arnberg, A.C. (1987) Cell **48**, 101-110.
11. Davies, R.W., Waring, R.B., Ray, J.A., Brown, T.A., and Scazzocchio, C. (1982) Nature **300**, 719-724.
12. Michel, F., Jacquier, A., and Dujon, B. (1982) Biochemie **64**, 867-881.
13. Michel, F. and Dujon, B. (1983) EMBO J. **2**, 33-38.
14. Waring, R.B. and Davies, R.W. (1984) Gene **28**, 277-291.
15. Burke, J.M., Irvine, K.D., Kaneko, K.J., Kerker, B.J., Oettgen, A.B., Tierney, W.M., Williamson, C.L., Zaug, A.J., and Cech, T.R. (1986) Cell **45**, 167-176.
16. Williamson, C.L., Tierney, W.M., Kerker, B.J., and Burke, J.M. (1987) J. Biol. Chem., in press.
17. Davanloo, P., Rosenberg, A.H., Dunn, J.J., and Studier, F.W. (1984) Proc. Natl. Acad. Sci. USA **81**, 2035-2039.
18. Butler, E.T. and Chamberlin, M.J. (1982) J. Biol. Chem. **257**, 5772-5778.
19. Zaug, A.J., Been, M.D., and Cech, T.R. (1986) Nature **324**, 429-433.
20. Grabowski, P.J., Brehm, S.L., Zaug, A.J., Kruger, K., and Cech, T.R. (1983) In Hamer, D.H. and Rosenburg, M.J. (eds), Gene Expression, Alan R. Liss, New York, Vol. 8, pp.327-342.
21. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467.
22. Donis-Keller, H., Maxam, A.M., and Gilbert, W. (1977) Nucl. Acids. Res. **4**, 2527-2538.
23. Tanner, N.K. and Cech, T.R. (1987) Biochemistry **26**, 3330-3340.
24. Zaug, A.J. and Cech, T.R. (1985) Science **229**, 1060-1064.
25. Kan, N.C. and Gall, J.G. (1982) Nucl. Acids. Res. **10**, 2809-2822.
26. Been, M.D. and Cech, T.R. (1985) Nucl. Acids Res. **13**, 8389-8408.
27. Zaug, A.J., Kent, J.R., and Cech, T.R. (1984) Science **224**, 574-578.
28. Waring, R.B., Ray, J.A., Edwards, S.W., Scazzocchio, C., and Davies, R.W. (1985) Cell **40**, 371-380.
29. Szostak, J.W. (1986) Nature **322**, 83-86.
30. Jacquier, A. and Rosbash, M. (1986)Science **234**, 1099-1104.
31. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) Cell **35**, 849-857.
32. Inoue, T. and Cech, T.R. (1985) Proc. Natl. Acad. Sci. USA **82**, 648-652.
33. Waring, R.B., Towner, P., Minter, S.J., and Davies, R.W. (1986) Nature **321**, 133-139.
34. Been, M.D. and Cech, T.R. (1986) Cell **47**, 207-216.