Intracellular folding of the *Tetrahymena* group I intron depends on exon sequence and promoter choice

SUJATHA P. KODUVAYUR\(^1\) and SARAH A. WOODSON\(^2\)

\(^1\)Program in Molecular and Cell Biology, University of Maryland, College Park, Maryland 20742, USA
\(^2\)Department of Biophysics, Johns Hopkins University, 3400 N. Charles St., Baltimore, Maryland 21218, USA

ABSTRACT

The *Tetrahymena* group I intron splices 20- to 50-fold faster in *Tetrahymena* than in vitro, implying that the intron rapidly adopts its active conformation in the cell. The importance of cotranscriptional folding and the contribution of the rRNA exons to the stability of the active pre-RNA structure were investigated by comparing the activity of minimal pre-RNAs expressed in *Escherichia coli*. Pre-RNAs containing exons derived from *E. coli* 23 S rRNA were three to four times more active than the wild-type *Tetrahymena* pre-RNA. *E. coli* transcripts of the chimeric *E. coli* pre-RNA were two to eight times more active than were T7 transcripts. However, the effect of cotranscriptional folding depends on exon sequences. Unexpectedly, the unspliced pre-RNA decays more slowly than predicted from the rate of splicing. This observation is best explained by partitioning of transcripts into active and inactive pools. We propose that the active pool splices within a few seconds, whereas the inactive pool is degraded without appreciable splicing.

Keywords: self-splicing; ribozyme; RNA folding; T7 RNA polymerase

INTRODUCTION

The *Tetrahymena* group I intron splices 20- to 50-fold faster in vivo than it self-splices in vitro (Brehm and Cech 1983; Bass and Cech 1984). The basis for its higher intracellular activity is not understood. Because the intron is spliced as rapidly in *Escherichia coli* as in *Tetrahymena* when inserted near the homologous position in the 23S rRNA of *E. coli*, it is unlikely that a species-specific factor such as a maturase is required in vivo (Zhang et al. 1995b). Instead, intracellular conditions may favor formation of the catalytically active form of the pre-rRNA. Folding in vivo may be enhanced by the structure of the flanking rRNA (Woodson and Cech 1993), favorable ionic conditions (Heilman-Miller et al. 2001; Takamoto et al. 2002), nonspecific RNA-binding proteins (Herschlag et al. 1994; Zhang et al. 1995a; Waldsich et al. 2002), or interactions between the nascent RNA and the polymerase (Pan et al. 1999).

In vitro, the rate of splicing is limited by conformational changes, which are slower than the chemical step of phosphodiester transfer (Herschlag and Cech 1990). Experiments on the *td* group I intron revealed a strong correlation between the stability of the tertiary structure and intracellular activity (Brion et al. 1999). Similarly, the activity of the hairpin ribozyme in yeast is comparable to that under physiological conditions in vitro (Donahue and Fedor 1997). Thus, intracellular activity provides a direct readout of the folding process.

Under physiological conditions in vitro, roughly 90% of the *Tetrahymena* pre-rRNA misfolds (Pan et al. 1997). Because inactive pre-rRNAs refold very slowly in vitro, apparent splicing activity is reduced (Emerick et al. 1996). When the pre-rRNA is expressed in *E. coli*, mutations that cause misfolding in vitro are partially or fully rescued, although splicing becomes more cold sensitive (Nikolcheva and Woodson, 1999). This suggests that intracellular conditions either reduce the probability that the pre-RNA misfolds or allow for rapid refolding of inactive RNAs.

The probability that the pre-RNA folds correctly is likely to depend on the conformation of the surrounding rRNA. Interactions within the flanking rRNA alter the balance between active and inactive conformations of the pre-RNA (Woodson and Cech 1991). An alternative stem-loop in the rRNA [P(-1)] inhibits 5’ splice site recognition by competing with formation of the P1 splice site helix (Fig. 1; Woodson and Cech 1991). This inhibition is balanced by base pairing interactions further upstream that stabilize the ac-
Partition of self-splicing RNA in vivo

FIGURE 1. Plasmids for expression of minimal pre-rRNAs in E. coli. (A) Pre-RNA sequences are transcribed from either the E. coli trc promoter or phage T7 promoter. A rnb or T7 transcription terminator was placed downstream of the cloned insert. Transcripts include 57 or 30 nt of vector sequence upstream of the insert. Ec-Eco, 986 nt; 

RESULTS and DISCUSSION

Comparison of pre-RNAs

Four expression plasmids for pre-RNA were constructed to compare the effects of the transcribing polymerase and the exon sequence on self-splicing (Fig. 1A). Two plasmids were based on pET vectors, for expression of the downstream pre-RNA by T7 RNA polymerase in BL21 cells. The other two contained the E. coli lactose-inducible trc promoter for expression by E. coli RNA polymerase.

Two of the four plasmids contained the sequences of the natural T. thermophila pre-rRNA (Tt). In the other two, the intron was inserted after U1926 of E. coli 23 S rRNA (Ec), and the intron guide sequence was changed to maintain complementarity with the E. coli 5’ exon (Zhang et al. 1995b). The minimal E. coli pre-rRNA was previously found to splice at a rate of 30 min⁻¹ in E. coli cells grown at 42°C (Zhang et al. 1995b), which is comparable to splicing activity in Tetrahymena (Brehm and Cech 1983).

In all cases, the pre-RNAs contained ~250 nucleotides (nt) from domain IV of the large subunit rRNA (Fig. 1B), which correspond to the minimal rRNA exons needed for optimal self-splicing in vitro (Woodson 1992). Because this region of the rRNA is highly conserved, the Tetrahymena and E. coli rRNA exons are expected to adopt similar secondary (and tertiary) structures (Noller et al. 1981; Cunnane et al. 2002). However, differences in primary sequence slightly alter the equilibrium between alternative conformations (Fig. 1C).

Splicing in E. coli

The rate of splicing in E. coli was determined from the steady-state ratio of pre-RNA to spliced intron and the decay rate of free intron, as described previously (Brehm and Cech 1983; Zhang et al. 1995b):

\[
\begin{align*}
\text{pre} & \rightarrow \text{IVS} \\
k_{sp} & = k_{d,IVS} \frac{[\text{IVS}]_{SS}}{[\text{pre}]_{SS}}
\end{align*}
\]

In cells transformed with the plasmids diagrammed in Figure 1, induction with IPTG resulted in good expression of plasmid-encoded pre-RNA (Fig. 2A). Levels of the pre-RNA and spliced linear intron (L IVS) were quantified by primer extension and reached a plateau after 90 min.

To measure the half-life of the spliced intron RNA, transcription by E. coli RNAP was inhibited with rifampicin after a 90-min induction of plasmid-encoded RNA. The amount of intron RNA remaining at various times after addition of rifampicin was measured by primer extension and fit to a first-order rate equation (Fig. 2B). The observed decay rates were 0.08 and 0.18 min⁻¹, respectively, for wild-type (Tt-Eco) and Ec (Ec-Eco) introns at 37°C (Table 1). Because the primer anneals to the 5’ end of the intron, cyclization products missing the first 15 or 19 nt of the intron (Zaug et al. 1983) were not detected in our assay. Consequently, the decay rates reported here reflect the combined loss of full-length intron via cyclization and nucleolytic degradation.

www.rnajournal.org 1527
Polymerase and exon effects

Comparison of the splicing activity of the four transcripts diagrammed in Figure 1 showed that both the polymerase used to transcribe the pre-RNA and the rRNA exon sequence affect intracellular processing (Table 1). At 37°C, E. coli RNAP transcripts spliced two- to threefold faster than T7 transcripts. The pre-RNAs with E. coli rRNA exons (Ec) spliced three- to fourfold faster than pre-RNAs with Tetrahymena rRNA exons (Tt). These effects were cumulative, so that the pre-RNA containing E. coli rRNA exons transcribed by E. coli RNAP (Ec-Eco, \( k_{obs} = 23 \text{ min}^{-1} \)) spliced about eight times faster than the least active pre-RNA (Tt-T7, 3 \text{ min}^{-1} ).

The Ec and Tt pre-RNAs used in these studies had comparable self-splicing activity in vitro, after a brief heat/cool renaturation (data not shown). Hence, differences in intracellular splicing activity likely reflect their ability to fold into the reactive conformation, rather than a change in intrinsic catalytic activity. These results show that the transcribing polymerase or exons have reproducible but subtle effects on pre-RNA folding in vivo.

Temperature dependence of in vivo splicing

To determine whether exon sequence or polymerase choice affect the fraction of misfolded transcripts, the temperature dependence of intracellular splicing activity was compared for the four plasmids in Figure 1. If the pre-RNA is kinetically trapped in misfolded conformers, the overall folding rate should increase with temperature or denaturant because the inactive RNA refolds more easily (Pan and Sosnick 1997; Pan et al. 1997; Rook et al. 1998). If the native structure is unstable, splicing should decrease with temperature (Brion et al. 1999).

Splicing rates were measured at 25°C, 30°C, 37°C, and 42°C for the Ec and Tt pre-RNAs (Fig. 3A,C). All of the pre-RNAs spliced more rapidly with increasing temperature (Fig. 3C), and the activation energies for T7 and E. coli

![Graph](image-url)

**FIGURE 2.** Group 1 splicing in E. coli. (A) Picomoles of intron or pre-RNA per microgram total RNA after induction of plasmid expression (pTN380) with 0.05 mM IPTG at 42°C. RNA was detected by primer extension as previously described (Zhang et al. 1995b). The steady-state ratio of intron to precursor was calculated 90 min after induction (Table 1). (B) Decay of spliced linear intron RNA at 42°C. Rifampicin was added 60 min after induction with IPTG. Full circles indicate wild-type (wt) intron core; open squares, C260G core.

**TABLE 1.** Splicing of the Tetrahymena pre-RNA in E. coli

<table>
<thead>
<tr>
<th>Pre-RNA</th>
<th>Rate constants, 37°C (min(^{-1}))</th>
<th>Intron decay(^{b})</th>
<th>Pre-RNA decay(^{a})</th>
<th>E. kcal/mol/K</th>
<th>( \Phi^{c} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type core</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ec-T7</td>
<td>12.0</td>
<td>—</td>
<td>—</td>
<td>18</td>
<td>0.98</td>
</tr>
<tr>
<td>Tt-T7</td>
<td>3.0</td>
<td>—</td>
<td>—</td>
<td>31</td>
<td>0.82</td>
</tr>
<tr>
<td>Ec-Eco</td>
<td>23.0</td>
<td>0.08</td>
<td>0.26</td>
<td>16</td>
<td>0.99</td>
</tr>
<tr>
<td>Tt-Eco</td>
<td>9</td>
<td>0.18</td>
<td>0.64</td>
<td>ND(^{d})</td>
<td>0.93</td>
</tr>
<tr>
<td>C260G core</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ec-T7</td>
<td>8.0</td>
<td>—</td>
<td>—</td>
<td>34</td>
<td>0.91</td>
</tr>
<tr>
<td>Tt-T7</td>
<td>0.40</td>
<td>—</td>
<td>—</td>
<td>40</td>
<td>0.24</td>
</tr>
<tr>
<td>Ec-Eco</td>
<td>9.0</td>
<td>0.24</td>
<td>0.75</td>
<td>20</td>
<td>0.92</td>
</tr>
<tr>
<td>Tt-Eco</td>
<td>0.05</td>
<td>0.3</td>
<td>1.30</td>
<td>14</td>
<td>0.04</td>
</tr>
</tbody>
</table>

\(^{a}\)In vivo splicing rates were calculated from primer extension assays according to Equation 1. Ec indicates E. coli rRNA exons; Tt, natural Tetrahymena exons; T7, T7 RNAP; and Eco, E. coli RNAP. The steady-state ratios of intron to pre-RNA were from the mean of three trials for the wild-type pre-RNA and four trials for pre-RNA with mutant intron core. These varied 5%–10%, except for Ec-T7 and Ec-Eco which were ±25%. These rates agree within error with previously reported data for splicing in E. coli (Zhang et al. 1995b). Splicing rates at 25°C, 30°C, and 42°C were obtained in the same manner (Fig. 3) to obtain an activation energy (\( E_{a} \)).

\(^{b}\)Decay rates at 37°C (observed) were calculated from primer extension analysis conducted on 0.25 to 0.50 μg of total RNA as in Figure 2. For the intron, decay rates ranged from 0.02 to 0.3 min\(^{-1}\) (wild type) and 0.07 to 0.5 min\(^{-1}\) (C260G) between 25°C and 42°C. For the wild-type intron, decay rates varied 10%–25% between two and three trials.

\(^{c}\)Partition factor \( \Phi \) at 37°C was calculated from Equation 4.

\(^{d}\)ND, the activation energy for Tt-Eco transcript was not determined as the splicing rate at 42°C may be influenced by rapid turnover of the pre-RNA at this temperature.
transcripts of the Ec pre-RNA were similar ($E_a = 18$ and $16 \text{kcal} \cdot \text{mole}^{-1} \cdot \text{K}^{-1}$; Table 1). Thus, the choice of polymerase made little difference in the energy barrier for splicing of the wild-type RNA.

Splicing of the natural Tt pre-RNA, however, was more cold sensitive than was splicing of the chimeric Ec pre-RNA. Although the Tt pre-RNAs were less active than Ec pre-RNAs at all temperatures, this gap widened to $\sim 20$-fold at $25^\circ \text{C}$ (Fig. 3C). The greater cold sensitivity of the natural Tt pre-RNA also suggested either that a higher fraction of transcripts are partitioned into misfolded structures or that Tt intermediates refold less readily than Ec intermediates.

Splicing of misfolding mutant pre-RNA

To further determine whether the $E. coli$ rRNA exons increase the propensity of the pre-RNA to fold correctly, we mutated C260 to G in the intron core. This mutation disrupts the triple helix at the junction of P4 and P6, and increases misfolding of the $Tetrahymena$ ribozyme (Zarrinkar and Williamson 1996). In $E. coli$, this mutation caused splicing to become more cold sensitive, particularly in the absence of the flanking rRNA, consistent with increased misfolding in vivo (Nikolcheva and Woodson 1999).

Splicing rates were measured for C260G pre-RNAs between $25^\circ \text{C}$ and $42^\circ \text{C}$, and activation energies were calculated for each transcript (Fig. 3B,D). As expected, they were less active and more cold sensitive than were their wild-type counterparts, consistent with greater misfolding of the intron core. When comparing the effects of the exons or polymerases, the relative trends were the same as for the analogous wild-type pre-RNAs, except slightly stronger. Surprisingly, the Tt pre-RNAs containing the C260G mutation were $200$-fold or $20$-fold lower than their wild-type counterparts at $42^\circ \text{C}$ (Fig. 3D). These results suggested that not only are the Tt pre-RNAs more likely to misfold than those with $E. coli$ rRNA exons, but the misfolded RNAs are rapidly turned over in the cell.

Precursor decay in $E. coli$ at $37^\circ \text{C}$

Although the Tt-Eco:C260G pre-RNA spliced less than the other pre-RNAs, the unspliced pre-RNA did not accumulate, suggesting rapid degradation of the unspliced RNA. To assess this possibility, the half-lives of $E. coli$ RNAP transcripts were compared at $37^\circ \text{C}$ (Fig. 2C). After inhibition of transcription with rifampicin, the pre-RNA disappears at an observed rate that is the sum of the rates for splicing ($k_{sp}$) and degradation ($k_{d,pre}$):

$$\frac{d[\text{pre}]}{dt} = k_{d,obs}[\text{pre}];$$

$$k_{d,obs} = k_p + k_{d,pre}.$$

If these rates are comparable to each other, the splicing rate can be determined from the difference between the observed half-lives of self-splicing and non-self-splicing precursors (Donahue and Fedor 1997; Donahue et al. 2000). Alternatively, the intrinsic decay rate should be equal to the difference between the overall rate of disappearance and the rate of splicing.

Unexpectedly, the observed decay rate was slower than the splicing rate, for all the wild-type and C260G pre-RNAs (Table 1). Although some pre-RNA may be consumed rapidly after rifampicin is added, the remainder had a longer...
Partitioning into active and inactive populations

The results presented here are inconsistent with a single pool of pre-RNA that splices at a constant rate. Instead, we propose that nascent transcripts are partitioned into active (preN) and inactive (preI) populations (Scheme II). The active RNA splices rapidly, whereas the inactive RNA is either degraded or occasionally refolds into the active form. That splicing determines the observed half-life of the pre-RNA is inconsistent with the data, as those pre-RNAs that produced the least spliced product decayed the fastest.

\[
\Phi \quad \text{pre}_N \xrightarrow{k_{sp}} \text{IVS} \xrightarrow{k_{d,IVS}} \text{pre}_I
\]

Assuming that most of the active pre-rRNA is converted to product, the ratio of spliced intron (IVS) to pre-RNA at steady state can be expressed as

\[
\frac{[\text{IVS}]}{[\text{pre}_N] + [\text{pre}_I]} = \frac{k_{sp}(k_f + \Phi k_{d,pre})}{k_{d,IVS}(k_f + \Phi k_{d,pre} + k_{d,pre}(1 - \Phi))} \tag{3}
\]

where \( \Phi \) is the fraction of active RNA; \([\text{pre}_N]\) and \([\text{pre}_I]\) are the steady-state concentrations of active and misfolded pre-RNA, respectively; and \(k_{d,pre}\) represent decay of preI. If all of the pre-RNA folds correctly (\(\Phi = 1\)), then Equation 3 reduces to Equation 1.

If splicing is rapid compared with refolding of inactive pre-RNA (\(k_{sp} \gg k_f\)), then

\[
\frac{[\text{IVS}]}{[\text{pre}_N] + [\text{pre}_I]} = \frac{[\text{IVS}]}{[\text{pre}_I]} = \frac{k_{d,pre} \Phi}{k_{d,IVS}(1 - \Phi)} \tag{4}
\]

Under these conditions, the yield of spliced product depends only on the partition factor \(\Phi\), because inactive transcripts have no chance to refold. The approximations in Equation 4 lead to estimates of in vivo partitioning that range from \(\Phi = 0.99\) for the most active pre-RNAs (Ec-Eco and Ec-T7) to \(\Phi = 0.04\) for the least active pre-RNA (Tt-Eco:C260G). Very similar values for \(\Phi\) are obtained by assuming that the inactive pre-RNAs can refold at rates similar to those obtained from in vitro experiments (\(k_f = 0.14\) min\(^{-1}\) at 30°C or \(0.2\) min\(^{-1}\) at 37°C; C260G \(k_f = 0.01\) min\(^{-1}\) at 30°C; Pan and Woodson 1998). In contrast, assuming that most of the inactive pre-RNA has a chance to refold and splice (\(k_f \gg k_{d,pre}\)) overestimates the yield of spliced products, particularly for the C260G mutants. Thus, the data are best explained by assuming that the unspliced pre-RNA represents a pool of inactive transcripts that refold slowly, if at all.

The highest intracellular splicing rates reported here (20–40 min\(^{-1}\)) are the same within error as that of the full-length pre-rRNA (Zhang et al. 1996b) and probably represent maximum activity. Unlike the minimal pre-RNAs, the unspliced 23 S rRNA is not degraded rapidly (Nikolcheva and Woodson 1997). Nonetheless, the levels of unspliced pre-23S rRNA are <1%, suggesting very few transcripts become trapped in inactive structures. This can be achieved either by preventing misfolding during transcription or by very rapid refolding of misfolded transcripts by chaperones. Further experiments are needed to distinguish between these possibilities.

Summary

Our results reported show that the intracellular splicing activity of the Tetrahymena group I intron depends on the sequence of the flanking rRNA, and to a lesser extent, the choice of polymerase. Although both the Tetrahymena and E. coli rRNA can form the alternative secondary structures shown in Figure 1C, only in the Tt pre-RNAs is formation of P(-1) directly linked to misfolding of the intron core through an interaction between the internal guide sequence and the 5′ side of P3 (Pan and Woodson 1998). Mutations in the guide sequence of the Ec-IVS abolish this interaction, explaining why the chimeric Ec pre-RNAs misfold less in vivo and are less sensitive to mutations in the intron core.

The choice of polymerase has a small influence on intracellular splicing activity of the group I pre-RNAs, as E. coli transcripts splice two- to threefold faster than do T7 transcripts at 37°C. Such differences could be due to site-specific pausing of the transcription elongation complex, as previously observed for RNase P (Pan et al. 1999). This effect depends on the sequence of the transcript and may reflect different pause sites on the two pre-RNA templates.

In summary, the sequence of upstream RNA, the transcription complex, and the balance between alternative conformations help determine the ultimate fate of a transcript.
Importantly, our results are most consistent with rapid partitioning between active and inactive forms, with the inactive pool contributing little to spliced product. Thus, intracellular factors that increase the yield of active pre-RNA presumably act soon after the RNA is transcribed.

MATERIALS AND METHODS

Plasmids

Plasmid pTN112 contains the EC variant of the Tetrahymena IVS inserted after U1926 in E. coli 23S rRNA (Zhang et al. 1995b) and has been described (Nikolcheva and Woodson 1997). Pre-rRNA sequences of pTN112 and pTN112:C260G were amplified by PCR with primers that introduce the BamHI recognition site at both ends of the insert (upstream primer, 5′-GCAGGGATCCGTAATCCGACTGTTTAATAA; downstream primer, 5′-GAAGCTGCAACTGTTTATT; downstream primer, 5′-GAAGCTGCAACTGTTTATT; downstream primer, 5′-GAAGCTGCAACTGTTTATT). The resulting eight plasmids contain the Tetrahymena precursor RNA template under these conditions. Intracellular splicing rates were calculated from Equation 1.

Intron and precursor decay rates

Transformed cells were grown as above at 25°C, 30°C, 37°C, or 42°C, for 90–120 min after induction of transcription with IPTG to allow expression of plasmid RNA to reach steady state. Transcription was stopped by the addition of 0.8 mg/mL rifampicin, and 20 mL aliquots of culture were removed at various times. Total RNA was extracted and analyzed by primer extension as described above. The fraction of spliced IVS RNA remaining was plotted versus time and fit to a first order rate equation. Precursor decay rates were determined in a similar manner, except cells were harvested at shorter intervals (1–10 and 30 min).

ACKNOWLEDGMENTS

We thank Robert Moss for technical assistance. This work was supported by a grant from the National Institutes of Health (GM46686).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

Received May 19, 2004; accepted July 15, 2004.

REFERENCES


Emerick, V.L., Pan, J., and Woodson, S.A. 1996. Analysis of rate-determining conformational changes during self-splicing of the

www.rnajournal.org
Intracellular folding of the Tetrahymena group I intron depends on exon sequence and promoter choice
SUJATHA P. KODUVAYUR and SARAH A. WOODSON
RNA 2004 10: 1526-1532

References
This article cites 29 articles, 8 of which can be accessed free at:
http://rnajournal.cshlp.org/content/10/10/1526.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.