The EMBO Journal Vol.18 No.22 pp.6491–6500, 1999

Self-splicing of the *Tetrahymena* intron from mRNA in mammalian cells

Moira Hagen[1] and Thomas R.Cech[2]

Department of Chemistry and Biochemistry and Howard Hughes Medical Institute, University of Colorado, Boulder, CO 80309-0215, USA

[1]Present address: University of Colorado Health Science Center, 4200 East Ninth Avenue, Denver, CO 80262, USA

[2]Corresponding author
e-mail: Thomas.Cech@Colorado.Edu

The *Tetrahymena* pre-rRNA self-splicing intron is shown to function in the unnatural context of an mRNA transcribed by RNA polymerase II in mammalian cells. Mutational analysis supports the conclusion that splicing in cells occurs by the same RNA-catalyzed mechanism established for splicing *in vitro*. Insertion of the intron at five positions spanning the luciferase open reading frame revealed 10-fold differences in accumulation of ligated exons and in luciferase activity; thus, the intron self-splices in many exon contexts, but the context can have a significant effect on activity. In addition, even the best self-splicing constructs, which produced half as much mRNA as did an uninterrupted luciferase gene, gave ~100-fold less luciferase enzyme activity, revealing an unexpected discontinuity between mRNA production and translation in cells. The finding that production of accurately spliced mRNA in cells does not guarantee a corresponding level of protein production is surprising, and may have implications for the development of *trans*-splicing ribozymes as therapeutics.

Keywords: group I intron/mammalian cells/mRNA/ribozyme/RNA splicing/translation

Introduction

The *Tetrahymena* intron is a self-sufficient catalytic unit: the RNA folds into a specific three-dimensional shape that provides an active site for the splicing of the adjacent exons (Cech and Golden, 1999). As such, it might be expected that this intron could be transplanted to a new gene location in a different cell type and would retain its self-splicing ability. Indeed, this assumption has provided the basis for development of *trans*-splicing group I ribozymes to repair defective mRNAs produced in individuals with genetic disorders. Such ribozymes can pair with mRNA sequences that precede a mutation or deletion, cleave off the defective downstream portion of the mRNA and splice on an unmutated RNA segment, thereby regenerating a functional, translatable mRNA (Sullenger and Cech, 1994; Sullenger, 1997). Thus, the questions of efficiency and accuracy of group I intron self-splicing in the mammalian cell milieu has a growing biotechnology interest, and also a basic biology interest—to what extent is a ribozyme activity that has undergone evolutionary selection in one cellular environment transplantable into a different environment?

Previous work provides the expectation that the exon context of a self-splicing intron can affect its activity. Truncating the 5′ or 3′ exon of the *Tetrahymena* pre-rRNA at different locations can have a significant effect on self-splicing *in vitro* (Woodson and Cech, 1991; Woodson, 1992). The last few nucleotides of the 5′ exon must pair with the intronic internal guide sequence (IGS) for splicing; if they instead form a stem–loop with other exon sequences, splicing is inhibited. The same sort of exon structures can inhibit *trans*-cleavage of an RNA target, which requires that the IGS bind intermolecularly to a complementary target sequence (Zaug *et al.*, 1986; Campbell and Cech, 1995). Similarly, secondary and tertiary structure of some target RNAs can interfere with their cleavage by *trans*-cleaving hammerhead ribozymes (Fedor and Uhlenbeck, 1990; Heidenreich and Eckstein, 1992). *In vivo*, proteins that act as RNA chaperones may facilitate RNA catalysis by preventing misfolding of RNAs and resolving RNAs that have misfolded (Coetzee *et al.*, 1994; Herschlag *et al.*, 1994; Herschlag, 1995; Semrad and Schroeder, 1998; Clodi *et al.*, 1999). However, such chaperones can overcome inhibitory exon sequences only if the active structure is reasonably stable relative to the inactive structures.

In predicting whether a foreign cellular environment might support group I ribozyme reactions, the cellular milieu is another consideration. Self-splicing requires only ~1 mM Mg^{2+} and micromolar concentrations of guanosine or GTP, which would seem to be readily available in diverse cells. Yet, activity can be enhanced or inhibited by polyamines, other cations or proteins whose presence and concentrations are difficult to predict. Thus, an experimental approach is useful.

The *Tetrahymena* group I intron inserted into the α-fragment of the β-galactosidase gene undergoes self-splicing in *Escherichia coli* (Price and Cech, 1985; Waring *et al.*, 1985), but neither the rate nor the efficiency of splicing was assessed. A ribozyme constructed from a bacteriophage T4 group I intron was shown to convert active mRNA sequences into a covalently closed circular form in both bacteria and yeast (Ford and Ares, 1994). The yeast RNA species also contained a spliceosomal actin intron; the distribution of RNA products suggested that, for the majority of transcripts, the actin intron was spliced prior to the ribozyme reaction (Ford and Ares, 1994). Thus, the group I reaction appeared to be slow. Finally, the *Tetrahymena* ribozyme was able to *trans*-splice a new 3′ exon sequence onto truncated β-galactosidase mRNA transcripts in *E.coli* (Sullenger and Cech, 1994) and in the cytoplasm of mouse cells (Jones

et al., 1996), and when transfected as RNA it underwent *trans*-splicing onto endogenous β-globin transcripts in erythrocyte precursors derived from sickle cell patients (Lan *et al.*, 1998). Optimization of the efficiency of one of these reactions by increasing the ribozyme:substrate ratio showed that a *trans*-splicing ribozyme can convert as much as 25–50% of targeted RNA in mammalian cells (Jones and Sullenger, 1997). Protein production from RNAs *trans*-spliced in mammalian cells has not been reported. The overall summary of these past studies is that group I ribozymes typically show activity when transplanted into mRNAs in other cell types, that the efficiency of these RNA reactions in cells has been explored in only a few cases and that translation of the reacted mRNA has rarely been assessed.

In the current study we demonstrate and analyze *cis*-splicing of the *Tetrahymena* intron from a luciferase mRNA transcript in three mammalian cell lines. *Cis*-splicing ensures that the ribozyme is in immediate proximity to its target sequence, thereby circumventing a possible rate-limiting step in *trans*-splicing. The luciferase reporter is sensitive and gives a linear response over a wide range. Five main questions are addressed. (i) Does group I intron self-splicing occur in mammalian cells? (ii) How efficient is it? (iii) How greatly does efficiency depend on the site of insertion, i.e. on exon context? (iv) Do *in vitro* self-splicing rates or efficiencies provide reasonable predictions of activity in cells? (v) Is the self-spliced RNA always competent to be translated in cells? We found that all the intron-containing precursor RNAs we constructed gave rise to accurately spliced mRNA in mammalian cells but, unexpectedly, the amount of translation varied depending on the position of intron insertion, and even in the best case was inefficient. Possible explanations for this discrepancy include perturbation by the intron of mRNA surveillance or transport events.

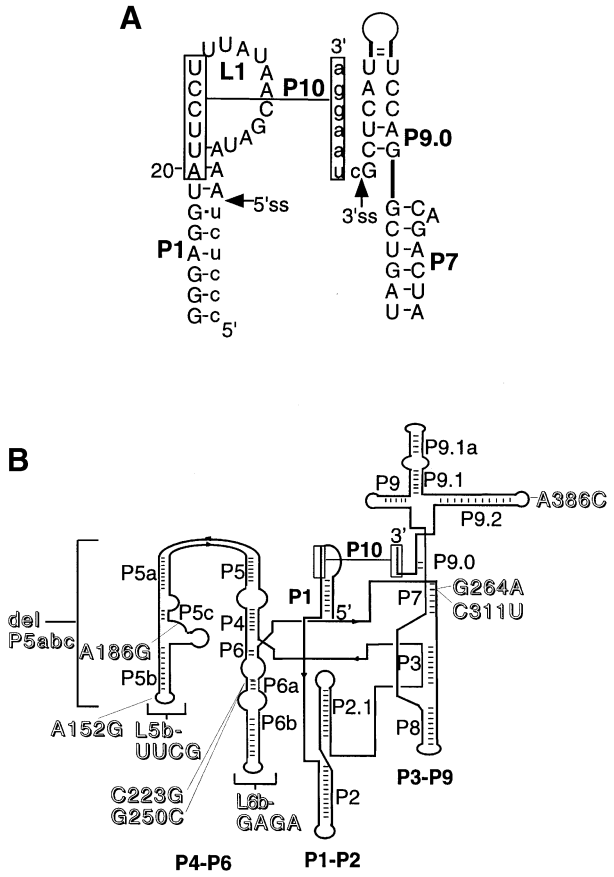

Fig. 1. Secondary structure of the *Tetrahymena* group I intron. (**A**) IGS–exon interactions for an intron (upper case letters) inserted at nucleotide 958 of the luciferase mRNA (lower case letters). Arrows indicate 5′ and 3′ splice sites. In this case, the natural *Tetrahymena* sequences matched the luciferase 5′ exon, allowing formation of paired region P1. The sequence of the other half of the IGS (boxed) was altered to be complementary to the luciferase 3′ exon, allowing formation of P10. (**B**) Schematic representation of the entire intron. P1 and P10 are highlighted, and the positions of mutations used to test for self-splicing in cells are indicated. To prevent crowding in the diagram, straight lines are used to show connectivity of the three structural domains (labeled P4–P6, P1–P2 and P3–P9). Arrowheads on lines indicate 5′→3′ polarity.

Results

Construction of a reporter gene interrupted by the Tetrahymena intron

Position 958 of the firefly luciferase open reading frame (ORF) is preceded by the sequence CCCUCU, which is complementary to the IGS of the *Tetrahymena* intron (GGAGGG; Waring *et al.*, 1983). Thus, the *Tetrahymena* intron inserted at this site would be expected to form a normal P1 (5′ exon–IGS pairing; Figure 1A), including the required G·U wobble pair at the 5′ splice site (Strobel and Cech, 1995), and undergo self-splicing (Been and Cech, 1986; Waring *et al.*, 1986). The P10 (IGS–3′ exon) interaction is not as essential as P1 but nevertheless facilitates exon ligation (Suh and Waring, 1990). Thus, nucleotides 15–20 of the intron sequence were altered to be complementary to nucleotides 960–965 of the luciferase mRNA, allowing a 6 bp P10 interaction (Figure 1A).The intron-containing luciferase gene was cloned into a plasmid vector that had a T7 promoter for *in vitro* transcription and a cytomegalovirus (CMV) promoter for expression in mammalian cells. The gene did not contain any spliceosomal mRNA intron, but did contain a bovine growth hormone gene polyadenylation signal.

To test whether the intron was capable of undergoing self-splicing in the context of these foreign exons, the plasmid DNA was linearized with *Eco*RI (which cleaves eight nucleotides past the luciferase termination codon) and transcribed with phage T7 RNA polymerase. The artificial ~2100 nucleotide transcript underwent guanosine-dependent self-splicing to generate the expected 1650 nucleotide ligated exons and the 414 nucleotide excised intron (data not shown).

A point mutation within the intron (A187G) was missed by automated sequencing but later found in Luc(958*) and all the mutant derivatives described below. [The asterisk in Luc(958*) denotes this mutation.] A187 is involved in a non-Watson–Crick base pair with U135 and therefore has a supporting role in tertiary folding of the P4–P6 domain of the intron (Cate *et al.*, 1996). Any destabilization caused by A187G may help to reveal negative effects of other mutations, a strategy employed purposefully by Jacquier and Michel (1987) and Couture *et al.* (1990).

Table I. Effect of intron mutations on expression of an intron-containing luciferase gene in 293 cells

Mutant	Predicted effect	Luminometer units	% Activity
Luc(958*)	–	42 650 ± 7582	100
G264A	catalytic	76 ± 33	0
C311U	catalytic	115 ± 7	0
del P5abc	structural	153 ± 112	0
L5b-UUCG	structural	188 ± 69	0
A152G	structural	327 ± 70	1
C223G	structural	9781 ± 2168	23
G250C	structural	35 ± 3	0
A186G	structural	795 ± 197	2
A386C	none	23 284 ± 4204	55
L6b-GAGA	none	15 916 ± 3580	37

For comparison, the intronless luciferase gene gave 1×10^8 luminometer units and Luc(958) (without A187G mutation) gave 2×10^6 units.

Self-splicing of a pol II transcript in human cells

Luc(958*) DNA, which contains the luciferase gene interrupted by the *Tetrahymena* intron, was transfected into human 293 cells, an immortal embryonic kidney cell line. After 40 h, the cells were harvested and monitored for luciferase activity. Luciferase activity (Table I) was easily detectable, but was low compared with the 10^8 luminometer units obtained with an analogous plasmid containing an uninterrupted luciferase gene.

The luciferase activity obtained with Luc(958*) could be due to self-splicing of the intron in the 293 cells, as we anticipated. Alternatively, there could perhaps be some splicing-independent pathway of translating luciferase from the intron-containing transcripts. A genetic test was devised to assess self-splicing in cells. Mutations were introduced into the intron that were predicted to: (i) abolish its ribozyme activity; (ii) destabilize its folding and therefore decrease its ribozyme activity; or (iii) change peripheral sequences that are not involved directly in intron tertiary structure or ribozyme activity (Figure 1B and Table I).

The mutations in the guanosine-binding site, G264A and C311U (Michel *et al.*, 1989), greatly reduce *in vitro* self-splicing activity [$(k_{cat}/K_m)^G$ reduced $\geqslant$1000-fold; Legault *et al.*, 1992]. These mutations reduced luciferase activity in cells to approximately background level (Table I). The mutations predicted to destabilize the catalytic core structure had a wide range of effects. The deletion of the entire P5abc subdomain (del P5abc) or replacement of its critical L5b tetraloop (L5b-UUCG) almost eliminated activity in cells. This tetraloop docks into a 'tetraloop receptor' comprising 11 nucleotides in the vicinity of P6a and P6b (Costa and Michel, 1995; Cate *et al.*, 1996). Even a single base mutation of the tetraloop, A152G, which is known to disrupt folding of P4–P6 (Murphy and Cech, 1994), reduced luciferase activity to ~1% of wild-type level. Single base mutations in the tetraloop receptor had unequal effects: C223G retained 23% of wild-type activity, while G250C was as defective as the intron with the entire P5abc element deleted.

Mutations not expected to disrupt the intron tertiary structure were also tested. P9.1 and P9.2 previously had been shown to be dispensable for core catalytic activity, although exon ligation takes place with decreased efficiency when both stem–loops are deleted (Barfod and Cech, 1988). A single base change in the L9.2 hairpin loop, A386C, gave only a small decrease in luciferase activity. The L6b loop does not interact with other portions of the intron (Latham and Cech, 1989) and can be replaced by a much longer helical segment without effect on activity (Nakamura *et al.*, 1995). Changing the L6b sequence to a stable tetraloop, GAGA, resulted in a small decrease in luciferase activity (Table I). Although these two 'negative control' mutations had the most modest effects of all the mutations tested, it is noteworthy that they did result in reproducible 2-fold reductions in luciferase activity. Thus, gene expression in the 293 cells is extremely sensitive to the integrity of the group I intron.

Overall, the mutations had similar effects on luciferase activity in the reporter gene system as they have on self-splicing activity: active site mutations eliminated luciferase activity; mutations that affect structural interactions had intermediate to strong effects; and mutations predicted to be neutral had the smallest effects on luciferase activity. These results support the conclusion that luciferase activity in cells is dependent upon the same self-splicing activity that has been characterized *in vitro*.

Effect of exon context on self-splicing and translation in vitro

Given that luciferase gene expression was very sensitive to the integrity of the intron, we explored whether it was also sensitive to the nature of the flanking exon sequences. The intron was therefore inserted after nucleotides 57, 257, 553 and 1578 of the luciferase ORF, all of which are uracil residues as required for group I intron self-splicing (Figure 2A, top). In each case, the intronic IGS was altered to match the 5′ and 3′ exon sequences, thereby allowing formation of P1 and P10. In addition, a new Luc(958) plasmid was constructed to correct the point mutation within P5abc.

Self-splicing was tested for *in vitro* transcripts of each luciferase gene. The main products generated in each splicing reaction were the 1650 nucleotide ligated exons and the excised intron (Figure 2A). Luc(958) had a slightly smaller excised intron than the other RNAs because it underwent efficient circularization and circle reopening, accumulating a linear form missing 15 nucleotides from its 5′ end (Hagen, 1999). [This pathway has been worked out in detail by Zaug *et al.* (1983, 1984).] The other introns had altered IGS sequences, and thus circularized less efficiently or not at all.

The effect of intron position on the rate of self-splicing was determined (Figure 2B). As is typical for self-splicing kinetics, the reactive fraction of the precursor RNA required two kinetic components to be adequately fitted; another portion (54–76% of the total) remained unreacted even after a 2 h incubation. The different precursor RNAs differed with respect to the fraction of the RNA reacting in the most rapid kinetic phase (f_1), the rate constant for this rapid phase (k_1) and the percentage reaction after 2 h (Table II). However, all precursors had similar splicing activity within a 5-fold range.

We expected that self-spliced RNA should be translated just as effectively as Luc(WT) mRNA, which had never been interrupted by an intron, because they should be

identical in nucleotide sequence. To test this expectation, each RNA construct was self-spliced and translated *in vitro* in rabbit reticulocyte lysates. All of them programmed the synthesis of luciferase proteins, which appeared as a doublet; one band having the mobility expected for luciferase (63 kDa) and another with a slower mobility (Figure 2C). Luc(WT) mRNA gave rise to the same doublet (Figure 2C, lane 6). A transcript interrupted by a mutant intron that could not undergo splicing gave neither of these translation products (Figure 2C, lane 7). The amount of luciferase protein translated from ligated exons was within 3-fold of the amount translated from the same amount of Luc(WT) mRNA (quantitation not shown).

Each RNA sample also contained unspliced precursor RNA (Figure 2A); because the intron contains multiple termination codons in all reading frames, these unspliced RNAs would be expected to give rise to truncated polypeptides (Figure 2D). Truncated products were observed for Luc(553), Luc(958) and Luc(1578) and corresponded to the predicted sizes of 23, 38 and 58 kDa, respectively (Figure 2C). Any truncated polypeptides produced by unspliced Luc(57) or Luc(257) precursor RNA (Figure 2C, lanes 1 and 2) would be too short to be visualized on the gel. Competition for translation of precursor RNA (leading to the truncated polypeptides) and spliced exons (leading to full-length luciferase) may have contributed to the 3-fold underrepresentation of luciferase compared with Luc(WT), which would not have such competition.

The translatability of the spliced exons was also confirmed by luciferase enzyme activity (data not shown). Seven serial dilutions of each spliced RNA sample were translated *in vitro*, and luciferase activity was measured and normalized to the amount of ligated exons added. The assay was linear for the more dilute samples and then leveled off, suggesting that the ribosomes were becoming saturated. Based on the linear portion of the assay, the luciferase activity from translation of each ligated exon was determined to be within 2- to 5-fold of the activity from translation of Luc(WT) RNA. Because all the spliced exons programmed the translation of correct size, active luciferase protein with efficiencies approaching that of the intronless mRNA, we proceeded to investigate self-splicing and translation in mammalian cells.

Exon context effects on self-splicing in cells: discontinuity between spliced RNA and protein levels

Having shown that all RNA constructs were able to self-splice *in vitro*, albeit with different rates, and to produce translatable mRNA, we compared their splicing in human cells. Each intron-containing DNA was transfected into 293 cells along with DNA that encodes β-galactosidase, similarly driven by the CMV promoter. The β-galactosidase mRNA could then be used to correct for variations in transfection efficiency and RNA recovery. The steady-state level of RNA transcribed from each gene was measured by an RNase protection assay. The first antisense

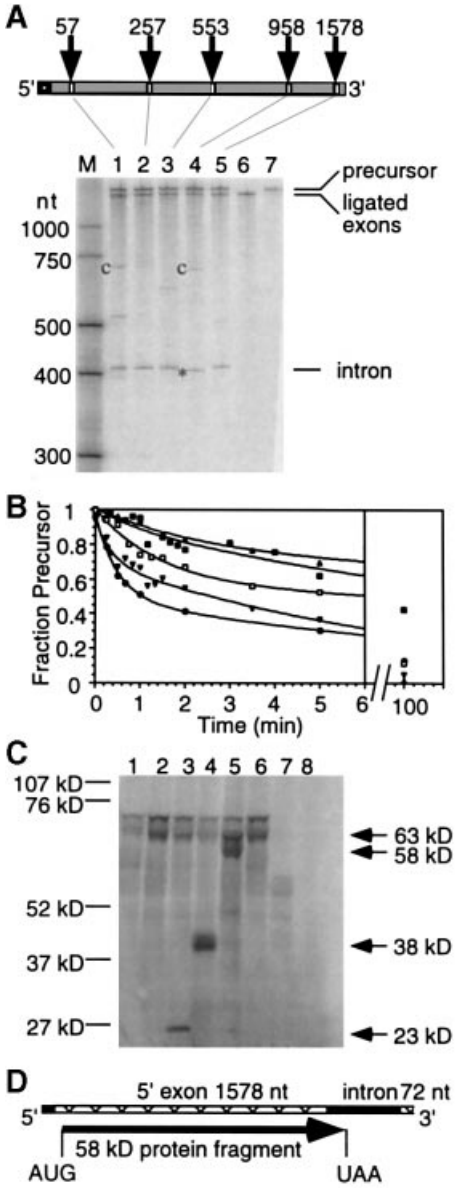

Fig. 2. Effect of exon context on RNA self-splicing and translation *in vitro*. (**A**) Top: representation of the luciferase mRNA (5′ UTR in black, ORF in gray) indicating the positions of insertion of the *Tetrahymena* intron. Bottom: *in vitro* self-splicing. Lanes 1–5, each radiolabeled precursor RNA was self-spliced with 0.2 mM GTP at 37°C for 2 h and products analyzed by 4% polyacrylamide–7 M urea gel electrophoresis; lane 6, intronless Luc(WT) RNA; lane 7, Luc(257)-90 nucleotide intron RNA, a mutant which is incapable of self-splicing, was incubated identically. c, circularized intron reaction products. *, the intron excised from Luc(958) underwent efficient circularization and circle reopening, so the form that has accumulated after 2 h incubation is the reopened circle (15 nucleotides shorter than the intact intron). Other faint bands between the ligated exons and the excised intron correspond in size to splicing intermediates or products of hydrolysis at splice sites. (**B**) Self-splicing kinetics. Self-splicing (as in A) was analyzed by gel electrophoresis. After quantitation of products, unreactive precursor was subtracted from each time point and the reaction was fit to a double exponential. From slowest to fastest, the curves represent (■) Luc(57); (▲) Luc(553); (□) Luc(1578); (▼) Luc(958); and (●) Luc(257). (**C**) *In vitro* translation in rabbit reticulocyte lysates of self-spliced RNAs. Lanes 1–6, *in vitro* translation of the corresponding RNA samples shown in (B), adjusting the quantities to equalize the amount of ligated exons; lane 7, *in vitro* translation of the splicing-defective Luc(257)-90 nucleotide intron RNA; lane 8, *in vitro* translation sample with no added mRNA. Proteins were labeled with [^{35}S]methionine. The 23, 38 and 58 kDa arrows indicate the calculated sizes of truncated polypeptides expected for translation of unspliced precursor RNAs from Luc(553), Luc(958) and Luc(1578), respectively. (**D**) Schematic diagram showing synthesis of the 58 kDa translation product expected from unspliced Luc(1578) RNA.

Table II. Self-splicing of a pol II transcript *in vitro* and in human cells

Intron position	Self-splicing *in vitro*			Splicing and translation in cells		
	k_1 (per min)[a]	f_1[a]	% P at 2 h[b]	% Ligated exons[c]	Luciferase activity[d] [% of Luc(WT)]	Transl. effic.[e] [% of Luc(WT)]
Luc(57)	0.40 ± 0.03	0.37	26 ± 2	50 ± 1	0.6 ± 0.1	1.2
Luc(257)	0.98 ± 0.13	0.63	39 ± 6	40 ± 1	2.1 ± 0.3	5.3
Luc(553)	0.22 ± 0.07	0.28	25 ± 8	60 ± 1	0.9 ± 0.2	1.5
Luc(958)	0.77	0.50	38	11 ± 2	1.7 ± 0.03	16
Luc(1578)	0.47 ± 0.09	0.49	44 ± 4	5 ± 3	0.08 ± 0.01	2

[a]Self-splicing time course was performed in 5 mM $MgCl_2$ at 37°C after refolding for 10 min at 50°C. Unreactive precursor (based on a 2 h reaction) was subtracted from each point, after which data were fit to a double exponential: $F_s = f_1e^{(-k_1t)} + f_2e^{(-k_2t)}$ where F_s is the fraction precursor at each time, and f_1 and f_2 are the fractions of the precursor reacting with rate constants k_1 and k_2, respectively. The initial rate constant k_1 is given here; k_2 ranged from 0.002 to 0.02/min. In those cases where multiple determinations were made, the data are presented as the mean ± SE.
[b]The percentage product formed after a 2 h reaction.
[c]RNA was extracted from transfected cells and subjected to RNase protection analysis. The percentage of ligated exons was calculated by quantitating the ligated exons band, the 5′ precursor band and the 3′ precursor band in the linear range of the titration (Figure 4C) and adding them together. The ligated exons band was divided by the total (×100). The values reported are the average (± range) of two separate determinations using total cellular RNA from two separate transfections.
[d]At 40 h after transfection, cells were harvested and resuspended in lysis buffer, and luciferase assays were performed as described in Materials and methods. Extract amounts were normalized to β-galactosidase activity. Mean ± SD of four replicates.
[e]Translational efficiency, calculated as 100 × % luciferase activity/% ligated exons.

RNA probe was designed to hybridize equivalently to each spliced and unspliced RNA, allowing total RNA levels to be assessed (Figure 3A). A titration of total cellular RNA in the presence of excess amounts of the two probes (luciferase and β-galactosidase) was first performed to determine the linear range of the assay, which extended to 0.25–0.5 μg of total RNA for each construct (Hagen, 1999). Assays with a single amount of total cellular RNA (0.25 μg) from each sample of transfected cells are compared in Figure 3B. Comparison of the steady-state level of luciferase RNA with the β-galactosidase RNA revealed no differences between constructs or between intron-containing luciferase genes and Luc(WT). These results suggest that neither transcription nor RNA stability is affected by the presence of the self-splicing intron.

The fraction of spliced luciferase RNA was also determined by RNase protection assays. A different probe was designed for each precursor RNA such that it spanned the site of intron insertion. When such a probe hybridizes to ligated exons, it protects a continuous stretch of RNA, but when it hybridizes to the unspliced precursor, it is cleaved by RNase at the splice site to yield two smaller fragments (Figure 4A). Total RNA from each transfected sample of cells was titrated with a constant, excess amount of labeled probe. In the sample data shown in Figure 4B, it is apparent that the signals for the precursor and for the ligated exons increased with increasing amount of cellular RNA. The data were quantitated, and a linear response between amount of cellular RNA and amount of protected RNA was obtained (Figure 4C). The percentage of spliced RNA was then calculated by dividing the signal for the ligated exons by the total detected luciferase RNA (ligated exons + 5′ exon + 3′ exon) from data taken within the linear range (Table II). The intron splices in mammalian cells, but the efficiency varies with the position of the intron. Splicing is most efficient (40–60%) from Luc(57), Luc(257) and Luc(553), less efficient (11%) from Luc(958) and least efficient (5%) with the intron near the 3′ end (Table II). A control experiment described in Materials

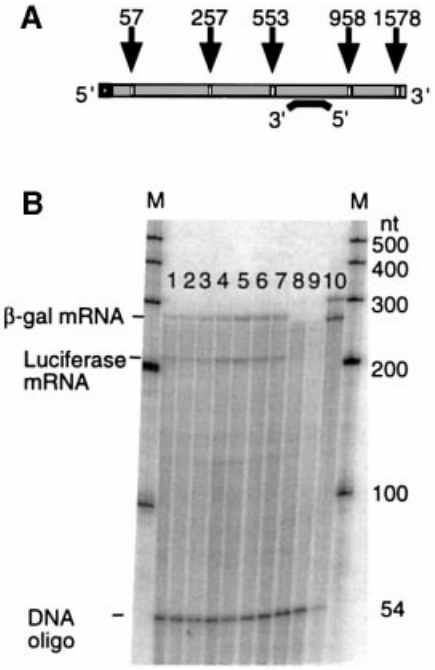

Fig. 3. Equivalent steady-state levels of luciferase RNAs produced after transfection of genes into 293 cells. (**A**) Schematic of the mRNA showing five positions of insertion of the *Tetrahymena* intron and (below) indicating the position of the RNase protection probe, designed to hybridize equally to nucleotides 664–889 of all spliced and unspliced RNAs. (**B**) RNase protection assays. RNA extracted from cells transfected with various luciferase genes and a β-galactosidase gene (to control for transfection efficiency) was hybridized with excess amounts of radiolabeled luciferase and β-galactosidase probes, digested with RNases A and T1, mixed with a 54 nucleotide DNA oligo as a precipitation control, and the products analyzed on a 4% polyacrylamide–7 M urea gel. The assay was determined to be linear to 0.5 μg of total cellular RNA/assay (data not shown); the data shown here used a single concentration of 0.25 μg. Lanes 1–5, Luc(57)–Luc(1578); lane 6, Luc(WT); lane 7, Luc(257)-90 nucleotide intron; lane 8, untransfected 293 cell RNA; lane 9, RNase digestion of probes in the absence of cellular RNA; lane 10, undigested mixture of 275 nucleotide luciferase probe and 300 nucleotide β-galactosidase probe.

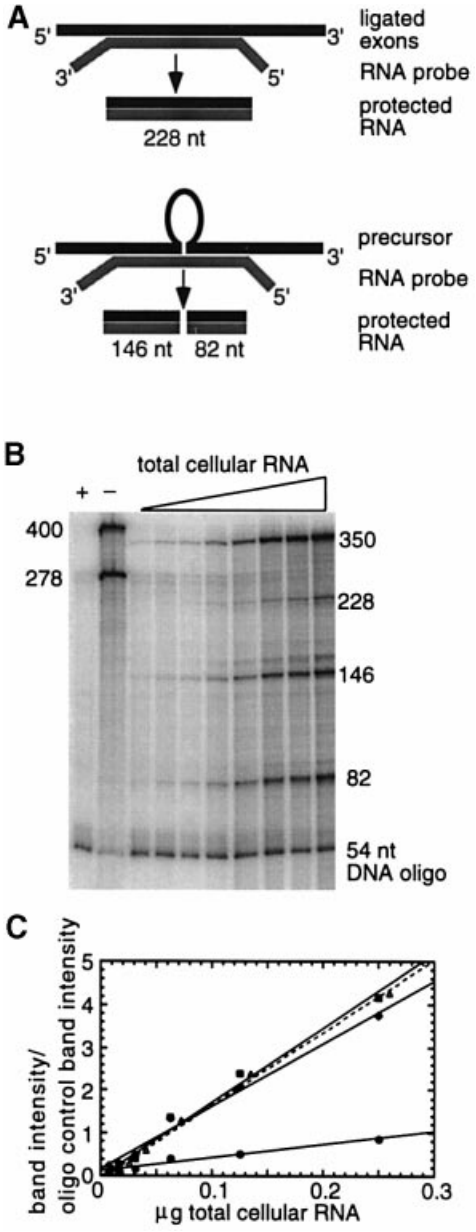

Fig. 4. Detection of self-spliced and unspliced luciferase RNAs in 293 cells. (**A**) RNase protection assay to distinguish spliced RNA (228 nucleotide protection product) from unspliced precursor (146 nucleotide + 82 nucleotide protection products). Thick black bars, luciferase RNA exons; thin black loop, intron; gray bars, RNase protection probe. A separate probe was synthesized for each intron position. (**B**) Titration of increasing amounts (0.0078–2.0 µg) of RNA from Luc(958)-transfected cells. Following RNase protection analysis, the products were analyzed by 4% polyacrylamide–7 M urea gel electrophoresis. Lane +, unhybridized probes digested with RNases A and T1; lane –, 278 nucleotide luciferase probe plus 400 nucleotide β-galactosidase probe, undigested. A 54 nucleotide DNA oligo was added to each of the reactions as a control for recovery.
(**C**) Quantitation of the data from (B), showing the linear portion only. (△) β-galactosidase RNA; (■) and (◆), 146 and 82 nucleotide fragments resulting from unspliced luciferase precursor RNA; (●), 228 nucleotide fragment resulting from luciferase-ligated exons.

and methods confirmed that the observed RNA splicing occurred in the cells, not during RNA extraction.

Protein production, which requires both splicing and translation, was measured by luciferase activity of cell lysates. Luciferase activity was normalized to

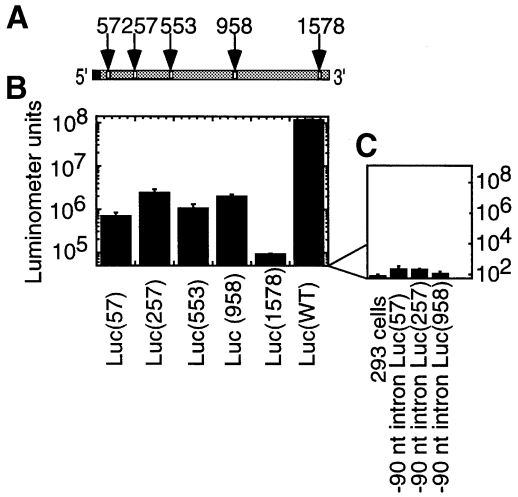

Fig. 5. Luciferase activity due to self-splicing and translation in 293 cells. Luciferase genes containing the *Tetrahymena* intron inserted into five positions (**A**) were transfected into 293 cells, along with a β-galactosidase gene. At 40 h, cells were harvested, lysed and assayed for luciferase. (**B**) Luciferase activity (mean ± SD of four assays) normalized to β-galactosidase activity. (**C**) Untransfected 293 cells and cells transfected with genes carrying splicing-defective mutant introns (the –90 nucleotide versions) gave only background levels of luciferase activity.

β-galactosidase activity resulting from the co-transfected β-galactosidase gene. As shown in Figure 5, all intron-containing DNA constructs produced levels of luciferase that were far above the background levels found in untransfected 293 cells or cells transfected with luciferase genes interrupted by splicing-defective mutant introns. Four of the five intron positions gave approximately equivalent amounts of luciferase activity, while the 3′-most position [Luc(1578)] gave a 10-fold lower level. Most surprising was the observation that none of the constructs gave more than ~2% of the luciferase activity of the uninterrupted Luc(WT) construct.

One simple explanation for the unexpectedly inefficient translation of self-spliced mRNA would be if missplicing of the RNA occurred in cells, resulting in a frameshift or insertion/deletion mutation. Such missplicing was undetectable *in vitro*, as evidenced by the good correspondence between levels of *in vitro* translated protein and luciferase activity. Nevertheless, we could not rule out the possibility of missplicing of RNA in cells transfected with Luc(57), Luc(257), Luc(553) and Luc(1578), which showed low translational efficiency (Table II). Therefore, RNA from these cells was amplified by RT–PCR. Sequencing across the splice junction of the ligated exons revealed only correctly spliced mRNA (data not shown). Thus, missplicing cannot explain the inefficient translation in cells.

Low splicing efficiency of the intron near the 3′ end

Luc(1578) had the least activity in cells and also had the lowest percentage of ligated exons as determined by RNase protection. This contrasts with its high *in vitro* splicing activity (Figure 2B, Table II) and good *in vitro* translatability (Figure 2C). Because the 3′ exon is only 72 nucleotides and the intron is closest to the polyadenylation site, it seemed possible that the polyadenylation machinery could interfere with correct folding of the

intron in cells. Alternatively, the low efficiency could be a problem particular to these exons.

To determine if proximity of an intron to the 3′ end of a transcript affects splicing in cells, the 3′ exon of a more efficiently spliced RNA, i.e. Luc(553), was reduced to 75, 143, 202 or 409 nucleotides. If proximity to the 3′ end reduces splicing efficiency, then shortening of the 3′ exon of Luc(553) should inhibit splicing. The four DNA constructs were each transfected into 293 cells and RNase protection was performed to determine the percentage of ligated exons. The linear range of the ligated exons was the same for each, and the introns spliced with the same efficiency, 50–60% (data not shown). This splicing efficiency was similar to the 60% value obtained with the parental Luc(553) RNA. Therefore, a short 3′ exon or proximity to the polyadenylation site does not always inhibit splicing. This suggests that the particular 5′ and/or 3′ exon sequence in the neighborhood of position 1578 reduces splicing efficiency in cells.

Similar intron positional effects in three cell lines

To determine if self-splicing and subsequent translation were cell line specific, we examined the same set of intron-containing luciferase genes in two other cell lines, Cos-7 (monkey kidney cells) and NIH 3T3 (mouse fibroblasts). One of these cell lines was of the same tissue origin as the human 293 cells, and the other of a different origin. The luciferase activity in both cell lines showed a remarkably similar trend to the results obtained in 293 cells (data not shown; see Hagen, 1999), suggesting that the cause of intron positional dependence of gene expression is the same. The most luciferase activity was obtained by splicing of Luc(257) in Cos-7 cells, where the activity was still only 8% of the activity from Luc(WT).

With the intron in the 5′ UTR, translation is independent of self-splicing

Luc(–28*) had the self-splicing intron inserted in the 5′-untranslated region (5′-UTR), 28 nucleotides preceding the translational initiation codon. The luciferase activity of Luc(–28*) was greater than that of any of the other constructs (Figure 6), although it was still 10-fold lower than that produced by Luc(WT) transcript. *In vitro* self-splicing of Luc(–28*) was much slower than that of Luc(958*); it was so unreactive that it was difficult to quantitate. (Both RNAs contained the A187G mutation, which decreases splicing.) We suspected that the high luciferase activity of Luc(–28*) was not due to the intron being spliced to a greatly increased extent in cells, but instead that the unspliced precursor was translated. To test this hypothesis, two separate mutations were made in the intron to eliminate self-splicing activity, G250C and G264A. G250 is part of the tetraloop receptor of L5b (Figure 1B) and, from the crystal structure, it interacts with the tetraloop. G264 is part of the G-binding site required for catalysis. These mutations did not have an effect on the luciferase activity (Figure 6), indicating that splicing was not required for translation of the gene in this construct. For comparison, the introduction of either mutation G250C or G264A reduced the luciferase activity of Luc(958*) by 1000-fold (Figure 6), showing that these mutations are capable of disrupting self-splicing.

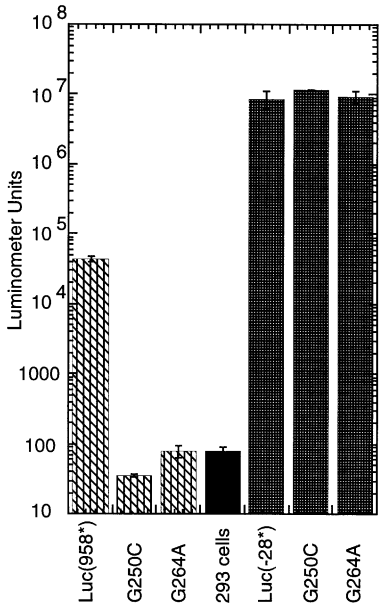

Fig. 6. With the intron in the 5′ UTR, splicing is not required for translation of the adjacent ORF. Human 293 cells transfected with the indicated genes were assayed for luciferase activity as in Figure 5. Cross-hatched bars, Luc(958*) and two splicing-defective mutant versions thereof; solid bar, untransfected 293 cells; shaded bars, Luc(–28*) and two splicing-defective mutant versions. Each bar shows the mean ± SD of three assays, normalized to the activity of a co-transfected β-galactosidase gene.

Discussion

Self-splicing of mRNA in mammalian cells

Group I introns inhabit nuclear rRNA genes in eukaryotes including *Tetrahymena* and *Physarum*. They inhabit mRNA genes in bacteriophage and fungal mitochondria (reviewed in Lambowitz and Belfort, 1993). However, there is no known case of a nuclear protein-coding gene interrupted by a group I intron, and no group I introns have been reported in any mammal. Group I ribozymes are active when produced by phage T7 transcription in the cytoplasm of mouse cells (Jones *et al.*, 1996) or when the RNA is transfected into human blood cells (Lan *et al.*, 1998). However, self-splicing had not been shown for a group I intron transcribed as part of an mRNA in the nucleus of a mammalian cell, where RNP formation, RNA trafficking and the intracellular milieu might or might not be conducive to self-splicing. We now report such activity in human, monkey and mouse cell lines. The activity is sensitive to mutations that destroy the essential guanosine-binding site of the intron or that destabilize folding of the catalytically competent RNA structure. Thus, the splicing that occurs in mammalian cells requires the self-splicing activity inherent to the intron.

Recently, Long and Sullenger (1999) have also reported splicing of the *Tetrahymena* intron in human cell lines, in their case from retroviral genome transcripts. They found that transcript context had a much larger effect on splicing efficiency in cells than on splicing *in vitro*, as in our study.

Self-splicing is inefficient in cells

Although self-splicing of luciferase mRNA in mammalian cells was readily detected, it was not efficient. The steady-state amount of the RNA present as ligated exons ranged

from 5 to 60%, depending on the position of intron insertion. Thus, 40–95% of the luciferase transcripts present in the steady-state were unspliced precursor RNA. The total RNA level (spliced + unspliced) was similar to that of an intronless Luc(WT) control, suggesting that the low steady-state levels of ligated exons were due to inefficient splicing, not to changes in transcription or RNA degradation. The data do not distinguish between two types of inefficient splicing: splicing could be very slow (of the order of the half-life of luciferase mRNA) or it could be heterogeneous, one population being spliced quickly and another population not spliced at all.

The inefficient splicing of the *Tetrahymena* intron inserted in luciferase mRNA in mammalian cells is in contrast to the efficient splicing of the same intron in its natural pre-rRNA context in *Tetrahymena*. In *Tetrahymena*, splicing is so efficient that the steady-state population of unspliced precursor RNA is $<<1\%$ of that of the mature, spliced 26S rRNA (Cech and Rio, 1979; Din *et al.*, 1979). The half-life for pre-rRNA splicing *in vivo* in *Tetrahymena* is extremely short, estimated as ~2 s (Brehm and Cech, 1983).

Under largely optimized conditions *in vitro*, self-splicing is not highly efficient for either the artificial luciferase mRNAs or the *Tetrahymena* pre-rRNA. The percentage of the RNA that reacts within ~2 h is 24–44% for the luciferase RNAs (Table II) and of the order of 50–80% for the *Tetrahymena* pre-rRNA (Been and Cech, 1986; Barfod and Cech, 1989; Woodson and Cech, 1991). The first-order rate constant for the most reactive fraction is $\leqslant$1/min for the luciferase RNAs (Table II) and 0.9/min for the *Tetrahymena* pre-rRNA (Cech and Bass, 1986). The difference is that the intron has a much increased splicing efficiency in pre-rRNA in *Tetrahymena* nucleoli, whereas the same intron present in pre-mRNA transcripts does not derive this large increase in splicing efficiency in mammalian cells. The same high rate and efficiency of splicing observed in *Tetrahymena* are obtained when the intron is inserted into the corresponding site in the 23S rRNA in *E.coli*; thus, *in vivo* facilitation of pre-rRNA self-splicing does not require a species-specific protein (Zhang *et al.*, 1995).

The much greater efficiency of *Tetrahymena* intron splicing in its natural pre-rRNA context than in an mRNA context implicates facilitation of folding of the catalytic RNA structure in the former case. RNA folding may be assisted by the local structure of the rRNA itself (Woodson and Emerick, 1993), by co-transcriptional binding of ribosomal proteins to the exons or by non-specific RNA-binding proteins or ribosomes acting as chaperones (Coetzee *et al.*, 1994; Semrad and Schroeder, 1998).

Exon context effects and predictive ability of *in vitro* tests

Although all five intron positions underwent splicing in mammalian cells, the efficiency of splicing was affected by intron location. Luc(958) was ~5 times less efficient at accumulation of ligated exons as the other intron-containing RNAs, and Luc(1578) was ~10 times less efficient (Table II). Truncation of the 3′ exon of the more efficiently spliced Luc(553) RNA showed that proximity to the 3′ end *per se* was not inhibitory to splicing. Thus, the inefficient splicing of the Luc(958) and Luc(1578) introns appears to be due to their local exonic environment, not their position within the pre-mRNA. It seems likely that these exon sequences interfere with intron folding, as has been observed in other group I intron transcripts (Woodson and Cech, 1991; Semrad and Schroeder, 1998; Clodi *et al.*, 1999).

Strikingly, the efficiency of self-splicing of the same transcripts *in vitro* was not at all predictive of their activity in mammalian cells. Luc(553) and Luc(57) were the worst *in vitro* self-splicing RNAs by three criteria: slowest rate constant k_1, smallest fraction (f_1) of the RNA reacting in the rapid kinetic phase and lowest amount of splicing after 2 h (Table II). Yet these two RNAs gave the highest percentage of ligated exons in mammalian cells. Thus, *in vitro* self-splicing was not predictive of exon context effects in cells. It is, however, a good qualitative test of the effect of intron mutations on splicing activity in cells: catalytic active site mutations that destroy *in vitro* catalytic activity prevent splicing in cells, structural mutations that reduce *in vitro* activity reduce splicing in cells, and mutations of peripheral regions that are neutral *in vitro* have the smallest effects in cells (Table I).

Self-spliced mRNA is not translated efficiently in cells

We expected that those luciferase transcripts that underwent self-splicing in mammalian cells would then be translated efficiently. Indeed, the same RNAs, when self-spliced *in vitro*, were translated in reticulocyte lysates with efficiencies approaching that of the intronless Luc(WT) RNA. However, the translation efficiency of the spliced transcripts in 293 cells was low, varying from 1.2 to 16% of Luc(WT) RNA (Table II). Similarly, low luciferase levels were obtained in two other cell lines.

It was quite unexpected that the prior presence of an intron affects the translatability of mRNA even after splicing has occurred. A number of possibilities could account for this result, two of which are as follows. Because self-splicing appears to be quite slow, perhaps much of the RNA is transported to the cytoplasm before it can be spliced. The RNA, even if it later underwent splicing, might miss a limited window of time to engage in a translation cycle, or might undergo poly(A) tail shortening such that translation was inefficient. A second possibility involves the mRNA surveillance pathway, which guards against mRNAs with premature termination codons (Hilleren and Parker, 1999). The intron contains stop codons in the luciferase ORF, and in fact in all three reading frames. Although a premature stop codon normally results in mRNA degradation, perhaps in this case (e.g. because of the structured intron) the RNA is instead sequestered from the translational machinery. Possibly related are the findings of Zhang *et al.* (1998a,b) that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. Cytological localization of the spliced and unspliced RNA may be informative in illuminating the cause of translational inefficiency.

Implications for trans-splicing ribozymes

One of the motivations for this study was to understand better the robust features and the limitations of *trans*-splicing ribozymes, which are promising reagents for gene therapy. By circumventing the need for ribozyme delivery

and ribozyme binding to its target mRNA, the *cis*-splicing reaction studied here reduces the required steps to the following: RNA splicing, transport, stability and translation. We found that group I RNA self-splicing was not highly efficient in mammalian cells, but nevertheless gave steady-state levels of ~50% spliced RNA in three out of the five intron positions tested. The two other positions gave only 5 and 11% spliced mRNA, and the data suggested that the exon context was inhibitory for splicing. Thus, in *trans*-splicing experiments, testing more than one target site would appear to be wise. We also found that *in vitro* self-splicing tests were not informative in guiding the choice of which target site might be most reactive in cells, suggesting that RNA–protein interactions alter exon structure or exon–intron interactions in cells.

Several of our results provide optimism for the *trans*-splicing approach. Self-splicing was versatile with respect to target site: all of our engineered IGS–exon interactions were active. Splicing was accurate in cells and the presence of a self-splicing intron in the mRNA did not affect its steady-state level, so it presumably did not affect RNA stability.

One potential challenge for the *trans*-splicing approach was also revealed. The spliced mRNAs were translated with only 1–16% the efficiency of the intronless Luc(WT) mRNA, presumably due to perturbation of mRNA transport. It is unknown whether this same inefficient translation would pertain to *trans*-spliced 'repaired' mRNAs. Thus, in *trans*-splicing experiments, it will be important to assess not just *trans*-spliced mRNA levels, but protein production as well.

Materials and methods

Construction of DNA plasmids

Luc(WT), which has the firefly luciferase gene cloned into the polylinker of pcDNA (Invitrogen), was a gift from Thomas Campbell. The *Tetrahymena* intron was cloned into various positions of the luciferase gene using PCR. The regions corresponding to the desired 5′ exon and 3′ exon were amplified from Luc(WT) as separate DNA fragments, and the intron was amplified from plasmid pT/S (Sullenger and Cech, 1994). Amplifications were performed using Vent polymerase (NEB) or Pwo polymerase (Boehringer), both high fidelity polymerases. The 5′ exon was re-amplified with the intron, and this fragment was purified from an agarose gel using the Qiagen gel extraction kit. The 5′ exon–intron fragment was then amplified with the 3′ exon DNA fragment to form the intron-containing luciferase gene. Each gene was cloned into the *Kpn*I–*Eco*RI restriction sites of the vector pcDNA and confirmed by automated sequencing (Applied Biosystems). Intron mutagenesis was performed by the method of Kunkel *et al.* (1987) and confirmed by sequencing using thermosequenase (Amersham).

In vitro kinetics

Each DNA construct was cut with *Eco*RI and transcribed using T7 RNA polymerase in a reaction that contained 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 4 mM ATP, 4 mM GTP, 4 mM CTP, 4 mM UTP, [α-^{35}S]UTP (800 Ci/mmol, NEN) and 30 mM Tris–HCl, pH 8, at 37°C for 2 h. The labeled transcript was pre-incubated in buffer consisting of 5 mM $MgCl_2$, 0.15 M KCl and 10 mM MES pH 5.5, for 10 min at 50°C and 2 min at 37°C. The reactions were initiated by adjusting the solution to 0.2 mM GTP, 5 mM $MgCl_2$, 0.10 M KCl and 50 mM HEPES pH 7.4. Reaction products were analyzed on a 4% polyacrylamide–7 M urea gel and quantitated using a PhosphorImager (Molecular Dynamics).

Transient transfections

Dishes (60 mm) were seeded with 3×10^5 293 cells and incubated overnight at 37°C/5% CO_2. Five micrograms of each luciferase gene plus 1 μg of a β-galactosidase gene, also with a CMV promoter (gift of N.Ahn laboratory), were transfected onto each plate of cells using Superfect (Qiagen). After 40 h incubation, cells were washed with 4 ml of phosphate-buffered saline pH 7.2. Cells were lysed into 1× reporter lysis buffer (Promega); cell extracts were freeze–thawed once and cell debris was removed by centrifugation. Luciferase assays were conducted using luciferin as the substrate (Promega) in the presence of ATP and analyzed using a luminometer (Berthhold). β-galactosidase assays used the Promega β-galactosidase assay system.

RNase protection assays

For RNA extractions, the medium was removed from each plate, and cells were resuspended in 1 ml of Trizol (Gibco-BRL). The RNA was extracted from the cellular debris and precipitated. RNase protection assays were performed using an RPA II kit (Ambion). Protected RNA fragments were separated on a 4% polyacrylamide–7 M urea gel and quantitated using a PhosphorImager. Templates for *in vitro* transcription of antisense probes were amplified by PCR using 100 ng of luciferase DNA template. The 3′ primer contained a T7 promoter. PCR products were analyzed by agarose gel electrophoresis and purified using a Qiagen gel extraction column. PCR products were transcribed by T7 polymerase in a reaction that contained 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, 1 mM UTP, 1 mM GTP, 50 μCi of [α-^{32}P]CTP (800 Ci/mmol) and 30 mM Tris–HCl pH 8. RNA was transcribed for 2 h and then subjected to 2 U of DNase I (Ambion) for 20 min at 37°C. RNA probes were gel purified on a 4% polyacrylamide–7 M urea gel and eluted overnight at 37°C in probe elution buffer (Ambion).

To assess whether self-splicing might have occurred during RNA extraction, [α-^{35}S]UTP-labeled Luc(257) precursor RNA was added directly to a plate of cells after the medium was removed, and RNA was extracted with Trizol. Self-splicing did not occur during the extraction process, as the extracted precursor remained unspliced (Hagen, 1999). Labeled, *in vitro* spliced RNA was similarly added to a plate of cells. The 414 nucleotide intron was recovered, indicating that if self-splicing were to occur during extraction, it would be detected.

Acknowledgements

We thank Bruce Sullenger, Tom Campbell, Natalie Ahn and Leslie Leinwand for advice and materials, and Anne Gooding for T7 RNA polymerase. M.H.was supported by NIH training grant HL07851. This work was supported by NIH grant GM28039. T.R.C. is an investigator of the Howard Hughes Medical Institute.

References

Barfod,E.T. and Cech,T.R. (1988) Deletion of nonconserved helices near the 3′ end of the rRNA intron of *Tetrahymena thermophila* alters self-splicing but not core catalytic activity. *Genes Dev.*, **2**, 652–663.

Barfod,E.T. and Cech,T.R. (1989) The conserved U·G pair in the 5′ splice site duplex of a group I intron is required in the first but not the second step of self-splicing. *Mol. Cell. Biol.*, **9**, 3657–3666.

Been,M.D. and Cech,T.R. (1986) One binding site determines sequence specificity of *Tetrahymena* pre-rRNA self-splicing, *trans*-splicing and RNA enzyme activity. *Cell*, **47**, 207–216.

Brehm,S.L. and Cech,T.R. (1983) Fate of an intervening sequence ribonucleic acid: excision and cyclization of the *Tetrahymena* ribosomal RNA intervening sequence *in vivo*. *Biochemistry*, **22**, 2390–2397.

Campbell,T.B. and Cech,T.R. (1995) Identification of ribozymes within a ribozyme library that efficiently cleave a long substrate RNA. *RNA*, **1**, 598–609.

Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.

Cech,T.R. and Bass,B.L. (1986) Biological catalysis by RNA. *Annu. Rev. Biochem.*, **55**, 599–629.

Cech,T.R. and Golden,B.L. (1999) Building a catalytic active site using only RNA. In Gesteland,R., Cech,T.R. and Atkins,J. (eds), *The RNA World*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 321–349.

Cech,T. and Rio,D.C. (1979) Localization of transcribed regions on the extrachromosomal ribosomal RNA genes of *Tetrahymena thermophila* by R-loop mapping. *Proc. Natl Acad. Sci. USA*, **76**, 5051–5055.

Clodi,E., Semrad,K. and Schroeder,R. (1999) Assaying RNA chaperone activity *in vivo* using a novel RNA folding trap. *EMBO J.*, **18**, 3776–3782.

Coetzee,T., Herschlag,D. and Belfort,M. (1994) *Escherichia coli* proteins, including ribosomal protein S12, facilitate *in vitro* splicing of phage T4 introns by acting as RNA chaperones. *Genes Dev.*, **8**, 1575–1588.

Costa,M. and Michel,F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.*, **14**, 1276–1285.

Couture,S. *et al*. (1990) Mutational analysis of conserved nucleotides in a self-splicing group I intron. *J. Mol. Biol.*, **215**, 345–358.

Din,N., Engberg,J., Kaffenberger,W. and Eckert,W. (1979) The intervening sequence in the 26S rRNA coding region of *T.thermophila* is transcribed within the largest stable precursor for rRNA. *Cell*, **18**, 525–532.

Fedor,M.J. and Uhlenbeck,O.C. (1990) Substrate sequence effects on 'hammerhead' RNA catalytic efficiency. *Proc. Natl Acad. Sci. USA*, **87**,1668–1672.

Ford,E. and Ares,M.,Jr (1994) Synthesis of circular RNA in bacteria and yeast using RNA cyclase ribozymes derived from a group I intron of phage T4. *Proc. Natl Acad. Sci. USA*, **91**, 3117–3121.

Hagen,M. (1999) Self-splicing of the *Tetrahymena thermophila* intron in mammalian cells. PhD thesis, University of Colorado, Boulder, CO.

Heidenreich,O. and Eckstein,F. (1992) Hammerhead ribozyme-mediated cleavage of the long terminal repeat RNA of human immunodeficiency virus type l. *J. Biol. Chem.*, **267**, 1904–1909.

Herschlag,D. (1995) RNA chaperones and the RNA folding problem. *J. Biol. Chem.*, **270**, 20871–20874.

Herschlag,D., Khosla,M., Tsuchihashi,Z. and Karpel,R.L. (1994) An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis. *EMBO J.*, **13**, 2913–2924.

Hilleren,P. and Parker,R. (1999) mRNA surveillance in eukaryotes: kinetic proofreading of proper translation termination as assessed by mRNP domain organization? *RNA*, **5**, 711–719.

Jacquier,A. and Michel,F. (1987) Multiple exon-binding sites in class II self-splicing introns. *Cell*, **50**, 17–29.

Jones,J.T. and Sullenger,B.A. (1997) Evaluating and enhancing ribozyme reaction efficiency in mammalian cells. *Nature Biotechnol.*, **15**, 902–905.

Jones,J.T., Lee,S.-W. and Sullenger,B.A. (1996) Tagging ribozyme reaction sites to follow *trans*-splicing in mammalian cells. *Nature Med.*, **2**, 643–648.

Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.*, **154**, 367–382.

Lambowitz,A.M. and Belfort,M. (1993) Introns as mobile genetic elements. *Annu. Rev. Biochem.*, **62**, 587–622.

Lan,N., Howrey,R.P., Lee,S.-W., Smith,C.A. and Sullenger,B.A. (1998) Ribozyme-mediated repair of sickle β-globin mRNAs in erythrocyte precursors. *Science*, **280**, 1593–1596.

Latham,J.A. and Cech,T.R. (1989) Defining the inside and outside of a catalytic RNA molecule. *Science*, **245**, 276–282.

Legault,P., Herschlag,D., Celander,D.W. and Cech,T.R. (1992) Mutations at the guanosine-binding site of the *Tetrahymena* ribozyme also affect site-specific hydrolysis. *Nucleic Acids Res.*, **20**, 6613–6619.

Long,M.B and Sullenger,B.A. (1999) Evaluating group I intron catalytic efficiency in mammalian cells. *Mol. Cell. Biol.*, **19**, 6479–6487.

Michel,F., Hanna,M., Green,R., Bartel,D.P. and Szostak,J. (1989) The guanosine binding site of the *Tetrahymena* ribozyme. *Nature*, **342**, 391–395.

Murphy,F.L. and Cech,T.R. (1994) GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J. Mol. Biol.*, **236**, 49–63.

Nakamura,T.M., Wang,Y.-H., Zaug,A.J., Griffith,J.D. and Cech,T.R. (1995) Relative orientation of RNA helices in a group I ribozyme determined by helix extension electron microscopy. *EMBO J.*, **14**, 4849–4859.

Price,J.V. and Cech,T.R. (1985) Coupling of *Tetrahymena* ribosomal RNA splicing to β-galactosidase expression in *Escherichia coli*. *Science*, **228**, 719–722.

Semrad,K. and Schroeder,R. (1998) A ribosomal function is necessary for efficient splicing of the T4 phage thymidylate synthase intron *in vivo*. *Genes Dev.*, **12**, 1327–1337.

Strobel,S.A. and Cech,T.R. (1995) Minor groove recognition of the conserved G·U pair at the *Tetrahymena* ribozyme reaction site. *Science*, **267**, 675–679.

Suh,E.R. and Waring,R.B. (1990) Base pairing between the 3′ exon and an internal guide sequence increases 3′ splice site specificity in the *Tetrahymena* self-splicing rRNA intron. *Mol. Cell. Biol.*, **10**, 2960–2965.

Sullenger,B.A. (1997) RNA repair: the next wave of gene therapy. *J. NIH Res.*, **9**, 37–40.

Sullenger,B.A. and Cech,T.R. (1994) Ribozyme-mediated repair of defective mRNA by targeted *trans*-splicing. *Nature*, **371**, 619–622.

Waring,R.B., Scazzocchio,C., Brown,T.A. and Davies,R.W. (1983) Close relationship between certain nuclear and mitochondrial introns. *J. Mol. Biol.*, **167**, 595–605.

Waring,R.B., Ray,J.A., Edwards,S.W., Scazzocchio,C. and Davies,R.W. (1985) The *Tetrahymena* rRNA intron self-splices in *E.coli*: *in vivo* evidence for the importance of key base-paired regions of RNA for RNA enzyme function. *Cell*, **40**, 371–380.

Waring,R.B., Towner,P., Minter,S.J. and Davies,R.W. (1986) Splice-site selection by a self-splicing RNA of *Tetrahymena*. *Nature*, **321**, 133–139.

Woodson,S.A. (1992) Exon sequences distant from the splice junction are required for efficient self-splicing of the *Tetrahymena* IVS. *Nucleic Acids Res.*, **20**, 4027–4032.

Woodson,S.A. and Cech,T.R. (1991) Alternative secondary structures in the 5′ exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry*, **30**, 2042–2050.

Woodson,S.A. and Emerick,V.L. (1993) An alternative helix in the 26S rRNA promotes excision and integration of the *Tetrahymena* intervening sequence. *Mol. Cell. Biol.*, **13**, 1137–1145.

Zaug,A.J., Grabowski,P.J. and Cech,T.R. (1983) Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage–ligation reaction. *Nature*, **301**, 578–583.

Zaug,A.J., Kent,J.R. and Cech,T.R. (1984) A labile phosphodiester bond at the ligation junction in a circular intervening sequence RNA. *Science*, **224**, 574–578.

Zaug,A.J., Been,M.D. and Cech,T.R. (1986) The *Tetrahymena* ribozyme acts like a RNA restriction endonuclease. *Nature*, **324**, 429–433.

Zhang,J., Ramsay,E.S. and Woodson,S.A. (1995) *In vivo* facilitation of *Tetrahymena* group I intron splicing in *Escherichia coli* pre-ribosomal RNA. *RNA*, **1**, 284–292.

Zhang,J., Sun,X., Qian,Y., LaDuca,J.P. and Maquat,L.E. (1998a) At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. *Mol. Cell. Biol.*, **18**, 5272–5283.

Zhang,J., Sun,X., Qian,Y. and Maquat,L.E. (1998b) Intron function in the nonsense-mediated decay of β-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. *RNA*, **4**, 801–815.

Received July 19, 1999; revised September 10, 1999; accepted September 22, 1999