Article No. jmbi.1998.2447 available online at http://www.idealibrary.com on IDEAL J. Mol. Biol. (1999) 285, 1935–1950

JMB

Trans-splicing Ribozymes for Targeted Gene Delivery

Uwe Köhler[1], Brian G. Ayre[1], Howard M. Goodman[2] and Jim Haseloff[1]*

[1]*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England*

[2]*Department of Genetics Harvard Medical School and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA*

**Corresponding author*

Ribozymes are potential tools for genetic manipulation, and various naturally occurring catalytic RNAs have been dissected and used as the basis for the design of new endoribonuclease activities. While such cleaving ribozymes may work well *in vitro*, they have not proved to be routinely effective in depleting living cells of the chosen target RNA. Recently, *trans*-splicing ribozymes have been employed to repair mutant mRNAs *in vivo*. We have designed modified *trans*-splicing ribozymes with improved biological activity. These allow accurate splicing of a new 3′ exon sequence into a chosen site within a target RNA, and in frame fusion of the exon can result in expression of a new gene product. These *trans*-splicing ribozymes contain catalytic sequences derived from a self-splicing group I intron, which have been adapted to a chosen target mRNA by fusion of a region of extended complementarity to the target RNA and precise alteration of the guide sequences required for substrate recognition. Both modifications are required for improved biological activity of the ribozymes. Whereas cleaving ribozymes must efficiently deplete a chosen mRNA species to be effective *in vivo*, even inefficient *trans*-splicing can allow the useful expression of a new gene activity, dependent on the presence of a chosen RNA. We have targeted *trans*-splicing ribozymes against mRNAs of chloramphenicol acetyltransferase, human immunodeficiency virus, and cucumber mosaic virus, and demonstrated *trans*-splicing and delivery of a marker gene in *Escherichia coli* cells. The improved *trans*-splicing ribozymes may be tailored for virtually any target RNA, and provide a new tool for triggering gene expression in specific cell types.

Keywords: Tetrahymena thermophila; gene therapy; HIV; CMV; bioassay

Introduction

Self-splicing group I introns

The *Tetrahymena thermophila* precursor rRNA contains a group I intron capable of catalysing its own excision (Cech, 1990; Michel & Westhof, 1990; Price & Cech, 1985). Self-splicing of the intron requires presence of a guanosine cofactor and a divalent cation, either Mg^{2+} or Mn^{2+}, and occurs *via* two sequential transesterification reactions (Figure 1(a)). First, a free guanosine is bound to the ribozyme and its 3′ hydroxyl group is positioned to attack the phosphorous atom at the 5′ splice site. The guanosine becomes covalently attached to the 5′ terminus of the intron. Second, the phosphodiester bond located at the 3′ splice site undergoes attack from the newly freed 3′ hydroxyl group of the 5′ exon, giving rise to ligated exon sequences. These successive reactions are chemically similar and appear to occur at a single active site. The reactions of self-splicing are characterized by the formation of alternative RNA helices as the 5′ and 3′ splice sites are each brought to form similar base-paired conformations within the catalytic core of the intron. The intron-exon junctions are aligned along a complementary sequence termed the ''internal guide sequence'' or IGS (Been & Cech, 1986; Michel *et al.*, 1989; Waring *et al.*, 1986). Clea-

Present addresses: U. Köhler, and J. Haseloff, Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, England; B. G. Ayre, Division of Biological Sciences, Section of Plant Biology, Cornell University, Ithaca, NY, USA.

Abbreviations used: CAT, chloramphenicol acetyl transferase; HIV, human immunodeficiency virus; CMV, cucumber mosaic virus; IGS, internal guide sequence; SL-RNA, spliced leader RNA; snRNA, small nuclear RNA; IPTG, isopropyl β-D-thiogalactopyranoside; RT-PCR, reverse transcriptase polymerase chain reaction; GFP, green fluorescent protein.

E-mail address of the corresponding author: jph@mrc-lmb.cam.ac.uk

0022-2836/99/051935–16 $30.00/0

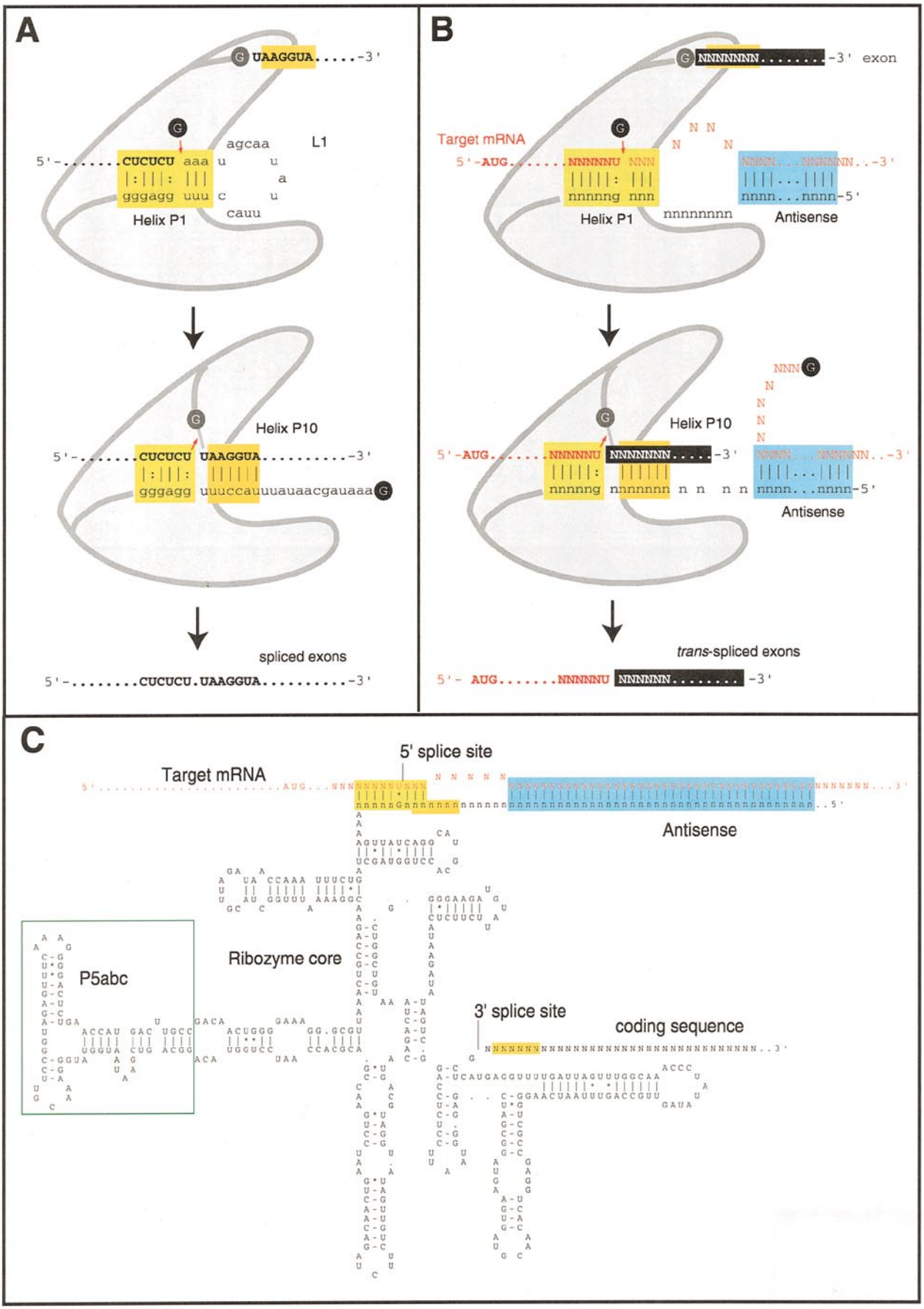

Figure 1 (*legend opposite*)

vage at the 5′ splice site depends on formation of a base-paired helix P1 between the IGS and sequences adjacent the splice site. The presence of a U·G ''wobble'' base-pair positioned four, five or six residues from the base of this helix defines the susceptible phosphodiester bond (Doudna *et al*., 1989; Murphy & Cech, 1989). After cleavage, a portion of the P1 helix is displaced to allow formation of a new helix P10, due to complementarity between the IGS and sequences adjacent the 3′ splice site (Figure 1(a)), and ligation of the exons appears to occur as a reverse of the first cleavage reaction.

Catalytic activities

In order to better define the structural and catalytic properties of the group I introns, exon sequences have been stripped from the ''core'' of the *Tetrahymena thermophila* intron. Such RNAs retain enzyme activity and can interact with substrate sequences *in trans*. For example, when truncated forms of the intron are incubated with sequences corresponding with the 5′ splice junction, the fragments undergo guanosine-dependent cleavage in mimicry of the first step in splicing (Zaug *et al*., 1986). The substrate and endoribonucleolytic intron RNAs base-pair to form helix P1, and cleavage occurs after a U·G base-pair at the expected position. Phylogenetic comparisons and mutational analyses indicate that the nature of the sequences immediately adjacent the conserved uridine residue at the 5′ splice junction are not important for catalysis, provided the base-pairing of helix P1 is maintained (Doudna *et al*., 1989; Murphy & Cech, 1989).

The sequence requirements for 3′ splice site selection appear to lie mainly within the structure of the intron itself, in the form of helix P9.0, and the following guanosine residue which delineates the 3′ intron boundary. However, there is an additional requirement for 3′ exon sequences which are needed for the formation of helix P10 and efficient splicing. The importance of helix P10 has been demonstrated by analysis of intron mutants (Suh & Waring, 1990). In addition, oligonucleotides have been ligated *in trans*, using a truncated form of the intron as a catalyst, an ''external'' guide sequence and oligonucleotides which had been extended by a 5′ guanosine residue. Here, the substrate oligonucleotides corresponding with the 3′ exon sequences were aligned solely by the formation of P10-like helices on an external template prior to ligation (Doudna & Szostak, 1989).

Trans-splicing

Trans-splicing, the joining of exons contained on different RNAs, has been demonstrated in a number of living systems. For example, the 5′ untranslated portions of certain mRNAs in trypanosomes and *Caenorhabditis elegans* are efficiently *trans*-spliced from SL-RNAs (spliced leader RNA; Blumenthal, 1995), which consist of the 5′ exon fused to sequences which can functionally substitute for U1 snRNA (small nuclear RNA) in mammalian snRNP-splicing extracts (Eul *et al*., 1995). Artificial *trans*-splicing has been demonstrated for group I introns that have been split and co-expressed in *Escherichia coli*. In the first case, a bacteriophage T4 thymidylate synthase gene (*td*) containing a group I intron was divided at the loop connecting the intron helix P6a. Transcripts of the *td* gene segments were shown to undergo *trans*-splicing *in vitro*, and to rescue dysfunctional *E. coli* host cells (Galloway-Salvo *et al*., 1990). Known base-pairings (P3, P6 and P6a) and probable tertiary interactions between the segments allowed correct assembly and processing of the gene halves. Second, Sullenger & Cech (1994) have shown that the L21 deleted form of the *T. thermophila* rRNA group I intron is capable of splicing a

Figure 1. (a) Alternate base-pairing of 5′ and 3′ exon sequences within the *T. thermophila* group I intron. Formation of helix P1 (green) is required for recognition of the 5′ splice site, which is positioned immediately 3′ to a conserved G·U wobble base-pair within the helix. Splicing is initiated by binding of a guanosine co-factor (G) to the intron and nucleophilic attack at the 5′ splice site. In the second step of splicing, the distal portion of helix P1 is displaced by sequences from the 3′ exon which form helix P10 (yellow). The 3′ hydroxyl group of the cleaved 5′ exon then attacks the precisely positioned 3′ splice site, resulting in ligation of the exons. Exon sequences are in upper case letters, intron sequences in lower case letters. The sequence of the ligated exons is shown below and an arrow indicates the ligation site. (:) shows wobble base-pairs in the paired region P1. (b) A schematic diagram of base-pairing between a model *trans*-splicing ribozyme and its target mRNA (red). A uridine residue (U) is chosen adjacent to the intended 5′ splice site. Compensating base changes within the ribozyme allow formation of helix P1, and an unpaired region, corresponding with loop 1 (L1) is maintained. Sequences with extended complementarity to the target mRNA are fused to the ribozyme, to allow an extensive base-paired interaction between the molecules. Helices P1, P10, and the antisense region are shaded in green, yellow and blue, respectively. Coding sequences are in upper case letters, intron sequences in lower case letters. After cleavage of the target mRNA, the 3′ exon is ligated to the 5′ portion of the target mRNA, to produce a novel fused messenger RNA. Expression of the 3′ exon depends on *trans*-splicing in frame with the coding sequence of the target mRNA. The sequence of the ligated exons is shown, and an arrow indicates the ligation site. (c) Schematic diagram of a model *trans*-splicing ribozyme base-paired to its target mRNA (red). Required sequences are labelled explicitly, while those which are not constrained are indicated as N. Conserved sequences within the ribozyme are also labelled explicitly, while those which can be altered to maintain conserved base-pairing or to provide a chosen 3′ exon are indicated as n or N. The P5abc region which is deleted in the negative control ribozymes is boxed.

3′ exon into a target RNA containing a sequence corresponding with the normal 5′ splice site (CCCUCU/AA; Jones *et al*., 1996; Jones & Sullenger, 1997). The 5′ exon and 5′ portion of the P1 helix has been deleted from the L21 form of the intron, and this RNA can interact with the target *via* a six nucleotide stretch of sequence complementarity. *Trans*-splicing resulted in the formation of translatable β-galactosidase messenger RNA in *E. coli*. These ribozymes have been shown to be active in mammalian cells (Jones *et al*., 1996; Jones & Sullenger, 1997), and recently have been targeted against human myotonic dystrophy protein kinase and globin mRNAs, with alterations made to the ribozyme sequences that form helix P1 (Lan *et al*., 1998; Phylactou *et al*., 1998). This method allowed the repair of mutant mRNAs *in vivo*, but the specificity and efficiency of the reaction was low: in one case, at least ten different transcripts were recognized as target RNAs and tagged with the 3′ exon (Jones *et al*., 1996). This may be due to the limited extent of the interaction between ribozyme and target RNAs which is comprised solely of a six base-pair region of complementarity (helix P1). Also, the P10 interaction was not allowed for in these ribozymes. While the P10 interaction within the *T. thermophila* intron is not essential for the recognition of the 3′ splice site *in vitro* (Been & Cech, 1985), it appears to increase the efficiency and fidelity of the second step of splicing, and is conserved in most group I introns analysed so far (Jaeger *et al*., 1996; Lehnert *et al*., 1996; Michel & Westhof, 1990; Suh & Waring, 1990).

In addition to mRNA repair, *trans*-splicing ribozymes could provide a route to specifically modifying cells expressing a chosen target mRNA, without the requirement for highly efficient catalysis *in vivo*. The effect of producing a small amount of spliced mRNA product can be substantially amplified by the biological activity of its translation product. If *trans*-splicing ribozymes are to be used as reliable tools for new experiments or therapies, it is their biological effect and specificity that is of importance. Therefore, we have established bioassays to enable rapid testing and improvement of the biological properties of these molecules. Here, we provide the first description of a flexible design for *trans*-splicing ribozymes that have improved biological activity over previous versions. These new ribozymes are: (i) tailored to particular target sites within chosen mRNAs, and, in contrast to previous work; are (ii) mutated to conserve both the P1 and P10 helices; and (iii) contain an additional region of 5′ terminal sequence that is complementary to the intended target RNA. Using a simple bioassay in bacteria, we have shown that it is possible to deliver β-galactosidase activity with a high degree of biological specificity *via trans*-splicing. Single base changes within the target mRNA eliminate ribozyme-mediated gene delivery. We have constructed ribozymes targeted at four different sites in three different RNAs, and shown that they function in a highly precise manner in *E. coli*. These results demonstrate the feasibility of conditional delivery of new gene activities within living cells, depending upon the presence of a chosen target RNA.

Results

Design of improved *trans*-splicing ribozymes

Previous workers have used truncated forms of the group I intron of *T. thermophila* rRNA for *trans*-splicing mediated mRNA repair (Sullenger & Cech, 1994; Jones *et al*., 1996; Jones & Sullenger, 1997; Lan *et al*., 1998; Phylactou *et al*., 1998). The group I intron can also be converted into ribozymes capable of highly specific targeted gene delivery *via trans*-splicing. The ribozymes can base-pair to a chosen target mRNA, catalysing a precise endonucleolytic cleavage, and ligating a new 3′ exon to the target mRNA. For cells bearing the target mRNA, in frame fusion of the 3′ exon sequence will trigger expression of the new spliced gene. In order to construct these *trans*-splicing ribozymes, additional sequences must be fused to a catalytic domain of the *Tetrahymena* intron (nt 28 to 414) to allow recognition of a chosen target RNA, and to provide a new 3′ exon. Several parameters govern the design of the ribozymes.

Firstly, a target RNA and ''5′ splice site'' must be chosen. In the final *trans*-splicing complex, only the 5′ portion of the P1 duplex is contributed by the target RNA. A uridine residue is located here as part of a G·U wobble base-pair (Doudna *et al*., 1989), found adjacent to the 5′ splice sites of all group I introns (Jaeger *et al*., 1996; Lehnert *et al*., 1996; Michel & Westhof, 1990). Therefore, a uridine within the target mRNA is chosen to define the intended splice site. This is the sole sequence requirement in the target RNA (Figure 1(b)).

Secondly, helix P1 must be altered to allow recognition of the chosen 5′ splice site. Having chosen a particular target site, compensating sequence changes must be made to the IGS in the 5′ portion of the ribozyme, in order to allow the formation of a suitable helix P1 between the target and ribozyme RNAs. The helix P1 should contain a G·U base-pair at the intended 5′ splice site located at the sixth position from the base of the helix. P1 extends for an additional three base-pairs past the intended 5′ splice site in the *T. thermophila* intron, and this is maintained in the *trans*-splicing model (Figure 1(a) and (b)).

Thirdly, further adjustments must be made in the IGS to allow the formation of a stable P10 helix. The helices P1 and P10 overlap along the *T. thermophila* intron internal guide sequence (Suh & Waring, 1990), and the second and third residues following both the 5′ and 3′ splice sites are complementary to the same residues in the IGS.

Fourthly, the exon sequences flanking the 3′ splice site must be chosen to allow in frame fusion with the 5′ exon sequences derived from the target mRNA. We have also avoided in frame start

codons positioned near the 3′ splice site, to minimise illegitimate translation of the exon.

Finally, a region of extended complementarity to the target mRNA is placed at the 5′ terminus of the *trans*-splicing ribozyme in order to increase its affinity and specificity for the target RNA. An arbitrary length of around 45 residues was initially chosen. The *T. thermophila* group I intron contains a loop (L1) at the terminus of helix P1, and a short region of unpaired sequence is maintained between helix P1 and the added complementary region (Figure 1(a) and (b)). Sequences similar to the target site, that might form a P1 helix *in cis*, should be avoided.

Ribozymes designed according to these guidelines may be useful for the conditional delivery of new gene activities *in vivo*, and in order to test this we have constructed ribozymes designed to *trans*-splice the *E. coli LacZ* α-peptide coding sequence to sites within the Tn*9* chloramphenicol acetyl transferase (CAT), cucumber mosaic virus (CMV) coat protein and human immunodeficiency virus (HIV) *tat* mRNAs. In each case, the target splice site was located at a uridine residue within the N-terminal coding sequence of the gene, and splicing should result in an in frame fusion of target mRNA and *LacZ* coding sequences. First, ribozyme activity was tested *in vitro*. Incubation of ribozyme and target RNA transcripts resulted in production of the expected *trans*-spliced RNA products (data not shown). The ribozymes were then co-expressed with target sequences in bacteria and the transformed cells were tested for the production of *trans*-spliced mRNAs and protein products *in vivo*.

Bioassay for ribozyme activity

We first constructed a ribozyme targeted against the Tn*9* CAT mRNA, and a uridine residue 18 nt downstream of the AUG translational start codon was arbitrarily chosen as the splice site. A sequence which included the α-peptide of *E. coli* β-galactosidase was selected as the 3′ exon (Figure 2(a)). A *trans*-splicing ribozyme was designed following the rules described above, and constructed using oligonucleotide synthesis and PCR amplification techniques. The altered ribozyme sequences required for splice site recognition are shown in Figure 2(a). The ribozyme and target RNAs were introduced into *E. coli* cells using pET-derived (Novagen) plasmids and pACYC184 (Chang & Cohen, 1978), respectively. The two plasmids bear different antibiotic resistance markers and compatible origins of replication. The ribozyme sequences were placed under control of the *tac* promoter (Amann *et al*., 1983; de Boer *et al*., 1983) and transcription was induced by the addition of IPTG to the medium. Cells expressing ribozyme and target RNAs were analysed for the presence of *trans*-spliced RNA product and β-galactosidase activity. As controls, we: (i) produced a truncated form of the ribozyme which did not contain extended complementarity to the target; (ii) constructed a ribozyme where the P10 interaction could not be formed; (iii) constructed defective ribozymes from which the P5abc region of the intron core had been deleted; and (iv) expressed the target and ribozyme RNAs separately in different cultures. The P5abc domain is required for correct folding of the *T. thermophila* intron, and its deletion results in loss of activity under physiological conditions (Joyce *et al*., 1989; Mohr *et al*., 1994).

Accurate *trans*-splicing was expected to produce a novel fused RNA, and subsequent translation of the new product should result in a chimaeric protein containing the first six amino acid residues of CAT fused to 89 amino acid residues derived from the N terminus of β-galactosidase. This fusion peptide was expected to complement the ω-peptide of β-galactosidase (expressed in XL1-Blue *E. coli* cells carrying the F′ plasmid (*pro*AB+ *lacI*^{q} *lacZ*ΔM15 Tn*10*(tet^{r}))), to produce β-galactosidase activity as a result of *trans*-splicing *in vivo*. The enzyme activity was measured in cell extracts, and by development of blue colonies on X-gal indicator plates (Miller, 1992). Only cells that expressed the target RNA and an active anti-CAT ribozyme showed β-galactosidase activity after overnight growth on indicator plates (Figure 2(b), colony a; Table 1). As controls, we tested colonies that co-expressed the target RNA with the ΔP5abc inactive ribozyme (Joyce *et al*., 1989; Mohr *et al*., 1994), and colonies that expressed either the ribozyme or the target RNA alone. β-Galactosidase activity could not be detected in any of the controls (Figure 2(b), colonies b, c and d; Table 1). This indicated that the β-galactosidase activity detected after co-expression of the ribozyme with the target RNA was due to *trans*-splicing *in vivo*.

Importance of the antisense region

In order to gauge the importance of sequence complementarity between the ribozyme and target mRNA for efficiency of the *trans*-splicing reaction, anti-CAT ribozymes with various length antisense domains were tested for activity (Figure 3). The extent of complementarity between ribozyme and target varied from only nine nucleotides (helix P1) for a truncated ribozyme to a total of 55, 110, and 210 nt for ribozymes with additional antisense regions. Co-expression of the target RNA with the ribozyme without an extended antisense region gave rise to ten units of β-galactosidase activity, while the addition of 46 nt of complementary sequence resulted in a 39-fold increase of activity (412 units). Increasing the length of sequence complementarity to a total of 110 and 210 nucleotides resulted in 523 and 426 units of β-galactosidase activity, respectively (Figure 3; Table 1). The antisense region of the ribozyme is clearly important for increased biological efficiency of the ribozyme. We also exchanged the antisense region for an unrelated sequence that is complementary to the HIV *TAT* gene (see below). This anti-CAT ribozyme with an HIV-specific antisense region pro-

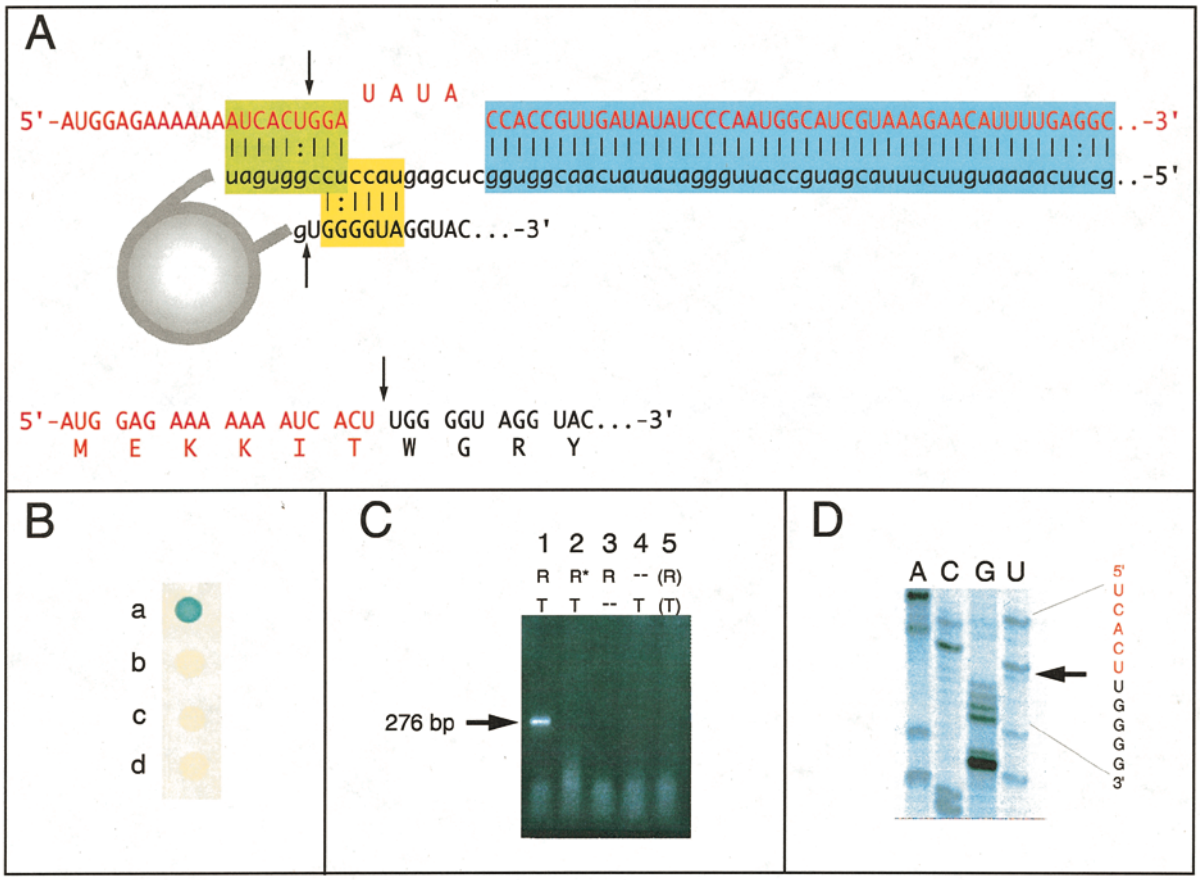

Figure 2. (a) Schematic diagram of the base-pairing between the CAT mRNA (red) and the *trans*-splicing ribozyme. Target and ribozyme RNA sequences are shown, helices P1 (green) and P10 (yellow) are indicated, together with the 46 nucleotide long complementary region (blue). (:) shows wobble base-pairs in the paired regions, arrows show 5′ and 3′ splice sites. A wobble base-pair was introduced by mutation of P10 within the anti-CAT ribozyme, to avoid formation of an in frame stop codon after *trans*-splicing. Coding sequences are in upper case letters, intron sequences in lower case letters. The sequence of the ligated exons is given (an arrow indicates the site of ligation) with the N-terminal portion of the predicted translation product. (b) The β-galactosidase assay of *E. coli* colonies containing the CAT mRNA and/or various targeted ribozyme RNAs. Cells co-expressing the ribozyme harbouring the antisense region with the target RNA (lane a); the inactive ΔP5abc control ribozyme with the target RNA (lane b); expressing the ribozyme alone (lane c), and the target RNA alone (lane d). (c) RNAs were extracted from cells expressing ribozyme and target mRNAs and assayed by RT-PCR for the RNA products of *trans*-splicing. Lane 1, cells expressing the ribozyme (R) and the target (T) RNAs; lane 2, cells co-expressing the ΔP5abc ribozyme (R*) with the target RNA; lane 3, cells expressing the ribozyme alone; lane 4, cells expressing the target RNA alone; lane 5, control, where cells expressing the ribozyme were mixed with others expressing the target RNA prior to analysis. The arrow indicates amplified spliced product. (d) The sequence of the amplified RT-PCR band was determined to confirm the identity of the spliced product ((c), lane 1). A portion of the gel corresponding with the splice site (arrow) is shown.

duced less β-galactosidase activity (one unit) than the ribozyme lacking an antisense region (ten units). This demonstrates that the CAT-antisense region increases the efficiency in a sequence-specific manner, and suggests that the presence of non-specific sequences may impair *trans*-splicing *in vivo*. No significant levels of β-galactosidase activity were seen in the absence of target RNA, for all of the ribozymes (Table 1).

Importance of helix P10

The design of these *trans*-splicing ribozymes also differs from those previously described (Jones *et al.*, 1996; Lan *et al.*, 1998; Phylactou *et al.*, 1998; Sullenger & Cech, 1994) in the addition of a P10 helix. In the wild-type *Tetrahymena* intron, the IGS base-pairs with the 5′ and 3′ exons and aligns them for the second step of splicing. The formation of a P10 helix has been shown to increase 3′ splice site specificity (Suh & Waring, 1990). Mutation of sequences within the 3′ exon that are involved in formation of helix P10 (from 5′-GGGGUA-3′ to 5′-UUUCCC-3′; Figure 4(a) and (d)), caused a 50-fold drop in the production of β-galactosidase activity in our bioassay (Table 1). Because of the two nucleotide overlap of helices P1 and P10, the first two nucleotides of helix P10 must be adapted to the target sequence. Therefore, we tested a ribozyme where the first two base-pairs of P10 were not allowed for and helix P10 was comprised of

Table 1. Biological activity of anti-CAT ribozyme variants

	α-Sense (nt)	P1 (nt)	P10 (nt)	β-Gal. units	SD
A. *Ribozymes and target co-expressed*					
Anti-CAT rz	46	9	6	411.9	132.7
Anti-CAT ΔP5 rz	46	9	6	0.3	0.3
Anti-CAT rz without antisense	0	9	6	10.4	9.1
Anti-CAT rz 101 nt antisense	101	9	6	523.4	57.6
Anti-CAT rz 201 nt antisense	201	9	6	425.7	59.0
Anti-CAT rz 45 nt non-specific	(45)	9	6	1.0	0.8
Anti-CAT rz 4 nt P10	46	9	4	419.8	60.4
Anti-CAT rz 2 nt P10	46	9	2	8.4	3.7
Anti-CAT rz no P10	46	9	0	8.2	4.9
Anti-CAT rz P1 6 nt	46	6	4	101.0	36.2
B. *Ribozymes and target expressed alone*					
Anti-CAT wt rz				0.3	0.4
CAT target				0.0	-
Anti-CAT rz without antisense				0.1	0.2
Anti-CAT rz 101 nt antisense				0.5	0.6
Anti-CAT rz 201 nt antisense				0.1	0.2
Anti-CAT rz 4 nt P10				0.2	0.2
Anti-CAT rz no P10				0.0	-

Units of β-galactosidase measured in assays of bacterial cultures expressing the constructs indicated. The values are mean values of three to seven experiments. rz, ribozyme; wt, wild-type; α-Sense, antisense; nt, nucleotides; SD, standard deviation; β-Gal., β-galactosidase.

four nucleotides only (Figure 4(b)). This ribozyme, when tested in cells co-expressing the target, resulted in 420 units of β-galactosidase activity, similar to the construct with a six nucleotide-long helix P10 (Figure 4(a); Table 1). In addition, P10 was reduced to two and zero base-pairs, and this resulted in a severe drop of β-galactosidase activity, to 8.4 and 8.2 units, respectively. While the sequences involved in P10 formation are required for efficient *in vivo* activity, it should be feasible to use ribozymes with shorter P10 helices. This would allow the re-use of 3′ intron sequences for different *trans*-splicing ribozymes, as helix P10 would not need to be adapted for different target RNAs.

Importance of an extended helix P1

The previously described *trans*-splicing ribozymes (Jones *et al*., 1996; Lan *et al*., 1998; Phylactou *et al*., 1998; Sullenger & Cech, 1994) utilized a helix P1 that consisted of six nucleotides only. The helix P1 in the original *cis*-splicing *Tetrahymena* intron extends for another three base-pairs downstream of the 5′ splice site to a total of nine base-pairs (Figure 1(a)). This was maintained in our design (see above). If extension of helix P1 was not necessary, designing new *trans*-splicing ribozymes would be easier as only six nucleotides rather then nine would have to be adapted to a new target. To test this, we constructed a ribozyme where helix P1 was allowed for six nucleotides only (Figure 4(e)), with a P10 helix of four base-pairs. Co-expression of this ribozyme with the target resulted in a drop of β-galactosidase units from 412 to 101 units (25 %). This indicates that a ''full-length helix'' P1 of nine nucleotides is important for higher biological efficiency of the reaction.

Trans-spliced RNA products

Cells expressing the various ribozyme and target RNAs were grown overnight with 2 mM IPTG to induce transcription of the ribozymes, harvested, and RNA was extracted. Spliced RNA products were amplified using reverse transcription-polymerase chain reaction (RT-PCR) analysis. Putative spliced RNAs were reverse transcribed using a primer specific to the 3′ region of the ribozyme transcription unit. The resulting cDNA was PCR amplified, using a 3′ nested primer and a target-specific 5′ primer. Amplification of the correctly spliced product would result in a 276 bp DNA product, containing the two ligated exon sequences. Figure 2(c) shows the result of a representative experiment. Products of expected size could only be detected in cells that expressed both the ribozyme (with 46 nt complementarity) and the target RNA (Figure 2(c), lane 1). No bands corresponding with spliced product were amplified from controls which contained the ΔP5abc inactive form of the ribozyme (lane 2), or where either the ribozyme or the target RNA were expressed alone (lanes 3 and 4, respectively). In order to exclude the possibility that the spliced product in lane 1 resulted from *in vitro* splicing during RNA extraction and/or reverse transcription, a culture expressing the ribozyme was mixed with a culture expressing the target RNA before analysis (Figure 2(c), lane 5). Any product would be the result of *in vitro* splicing. No spliced product was detected, indicating that the band in lane 1 resulted from *trans*-splicing *in vivo*. Small amounts of correctly spliced products could also be amplified from cells expressing the target RNA and ribozymes either with no antisense region, with the HIV-specific antisense region, or with a mutated helix P10 region. Quantitative RT-

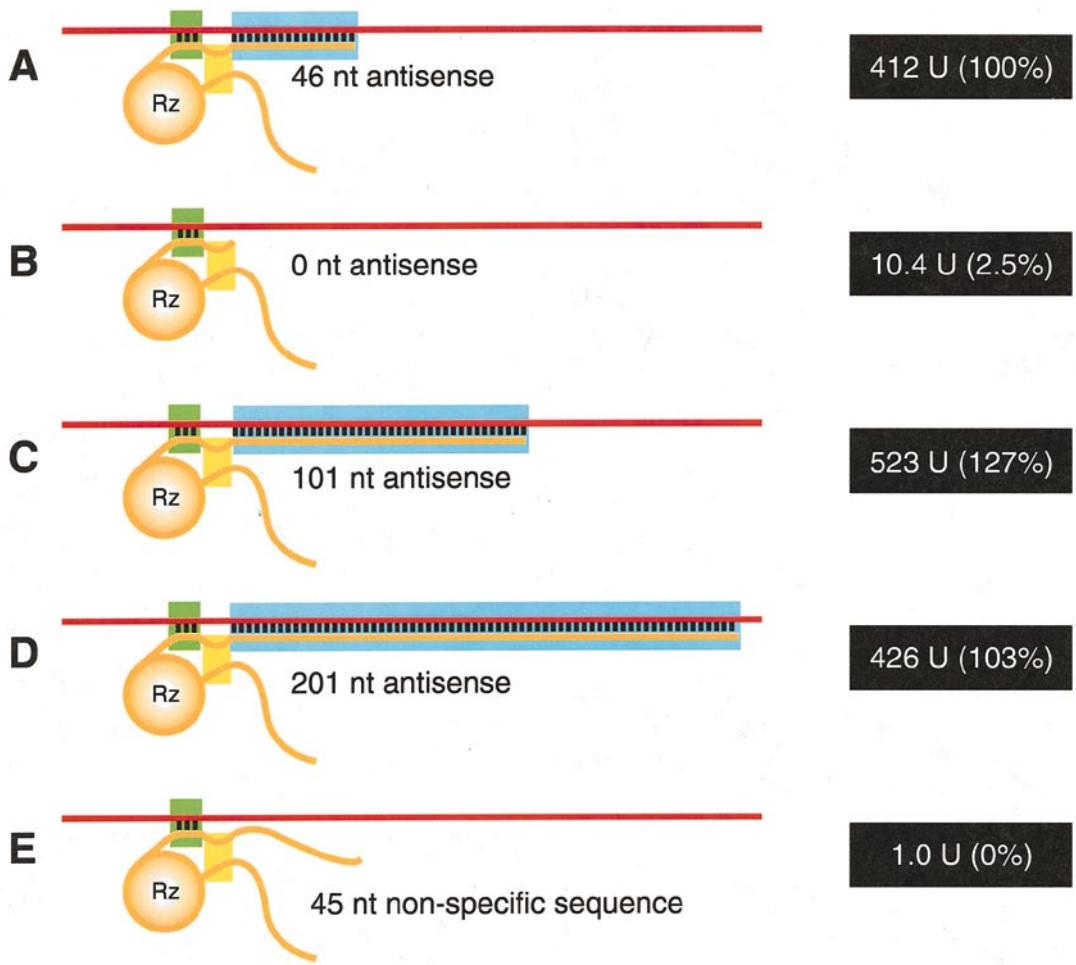

Figure 3. Schematic presentation of different anti-CAT ribozymes tested in respect of their antisense region. All ribozymes shown have the ability to form helices P1 and P10 for nine and six base-pairs, respectively (boxed). The constructs differ in the length and sequence of the antisense region. (a) A ribozyme with 46 nt long antisense region of complementary sequence to the target (CAT mRNA). This ribozyme was constructed first and arbitrarily set to 100 %. (b) A ribozyme lacking the antisense region. (c) A ribozyme with 101 nt of CAT-specific antisense region complementary to CAT mRNA. (d) A ribozyme with a 201 nt antisense region (e) A ribozyme with 45 nt of non-specific sequence. The measured units of β-galactosidase activity and the % values relative to the consruct shown in (a) are given.

PCR confirmed that the products were obtained at lower yield than for the intact ribozyme, and lower β-galactosidase activities were found (Table 1, and data not shown). RT-PCR was also performed with the other described ribozymes when co-expressed with the target. In all cases, an amplified band corresponding with the spliced product could be detected, as expected. In order to check the fidelity of *trans*-splicing, RT-PCR products were purified, amplified, and directly sequenced using a 3′ *LacZ* internal primer. The sequence of the RT-PCR amplified fragments always corresponded with that of the predicted spliced product (Figure 2(a) and (d); data not shown).

The two anti-CAT ribozymes with 101 and 201 nt antisense regions were expressed without the target. No β-galactosidase activity could be detected, indicating that little productive *cis*-splicing occured, where the longer antisense region would fold back to form an illegitimate P1. To investigate the possibility of *cis*-splicing that would result in out-of-frame or non-translatable RNAs, RT-PCR analysis was performed with primers that would amplify any *cis*-spliced RNA product. The 5′ primer was homolgous to the 5′ end of the extended antisense region. No product could be amplified (data not shown), indicating that the increased efficiency measured in the assays when the ribozymes were co-expressed with the target is due to more efficient *trans*-splicing.

Targeting different RNAs

In order to test whether our model for the design of *trans*-splicing ribozymes is a general one, we constructed ribozymes that were directed against other target RNAs. The design with a nine base-pair helix P1, a six base-pair helix P10, and antisense regions of 42 and 45 nt were used. *Trans*-splicing ribozymes could be useful for triggering gene expression in response to virus infection of a cell,

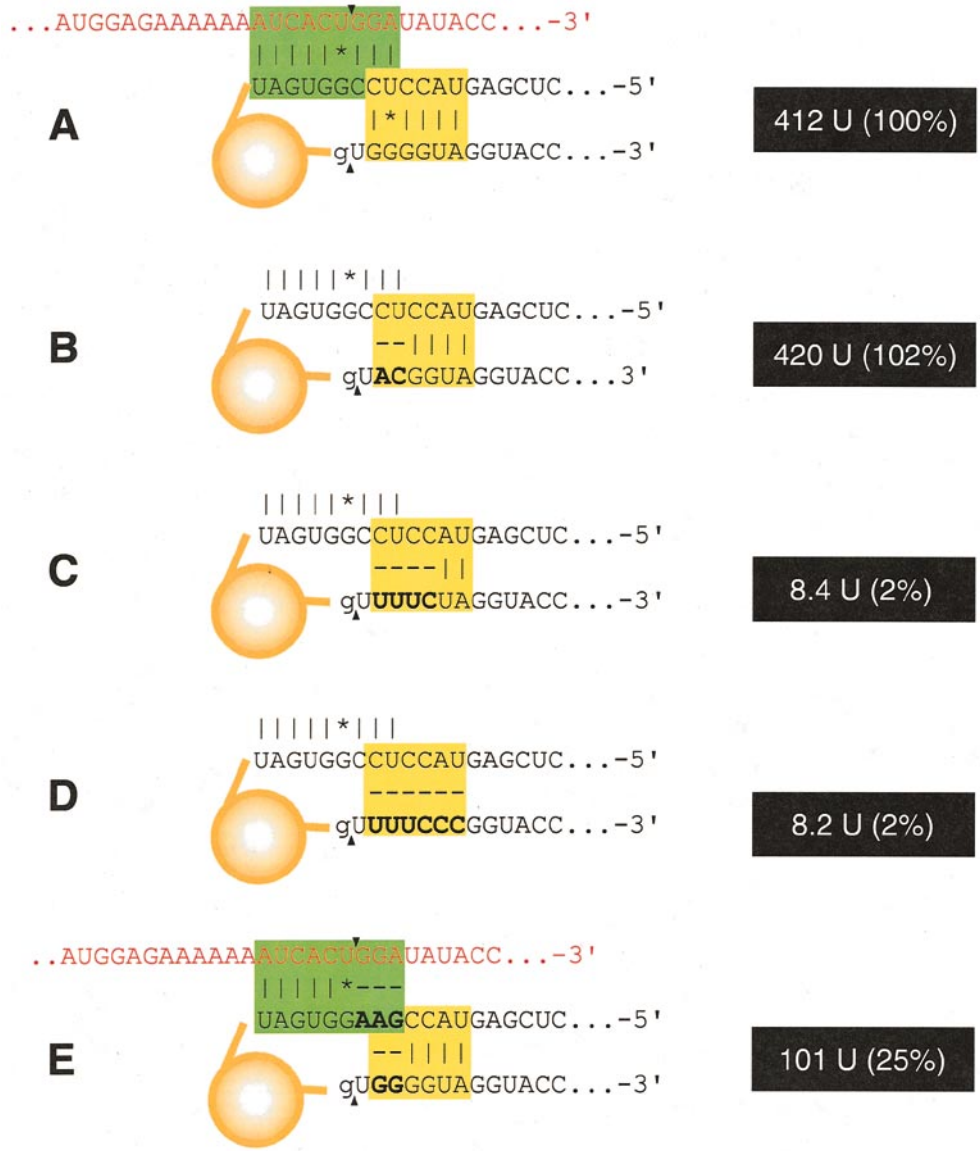

Figure 4. Schematic presentation of different anti-CAT ribozymes with different sequences of helices P10 and P1. All constructs shown harbour the 46 nucleotides long antisense region. (a)-(d) All ribozyme have the ability to form a nine nucleotides long helix P1 but differ in the length of helix P10: (a) six nucleotides; (b) four nucleotides; (c) two nucleotides; (d) no P10 can be formed. (e) A ribozyme where P1 is allowed for six nucleotides and P10 for four nucleotides only. Measured β-galactosidase units and % values relative to the construct shown in A (arbitrarily set to 100 %) are shown. Mutated residues in (b)-(e) are shown in bold type.

and we have targeted *trans*-splicing ribozymes at sites within a plant and a human virus, cucumber mosaic virus and human immunodeficiency virus type I, respectively. We chose two splice sites within the CMV coat protein mRNA (Schwinghammer & Symons, 1977), and one splice site in the mRNA encoding the transactivator protein (TAT) of HIV-I (Zhou *et al*., 1994).

The two chosen splice sites on the CMV coat protein mRNA were located after uridine residues positioned 10 and 12 residues downstream of the start codon for translation. Two *trans*-splicing ribozymes were constructed, and termed anti-CMV(10) and anti-CMV(12), respectively. The nucleotide sequence of the 5′ region of the coat protein mRNA is highly conserved between sequenced CMV isolates, and we expect the ribozymes to exhibit a broad specificity for different viral strains. A splice site was also chosen within the HIV transactivator (TAT) protein mRNA, located next to a uridine positioned 18 residues after the start codon. The target region chosen for the helix P1 interaction is entirely conserved between sequenced HIV-I isolates, and the sequences recognized by the antisense portion of the ribozyme are also highly conserved. We chose to keep β-galactosidase α-peptide coding sequences within the 3′ exon sequence, and the *trans*-splicing ribozymes were designed using the techniques described above. The anti-CMV(10) and anti-CMV(12) ribozymes were constructed with antisense domains of 42 nt, while that of the HIV ribozyme was 45 nt

long. Schematic representations of the ribozyme-target complexes are shown (Figure 5(a), (e) and (i)). In addition, defective ΔP5abc ribozymes were produced as negative controls.

Two artificial target mRNAs were also constructed by fusing a 60 nt CMV sequence or a 67 nt of HIV sequence to the N terminus of the green fluorescent protein (GFP) gene. Thus, the ribo-

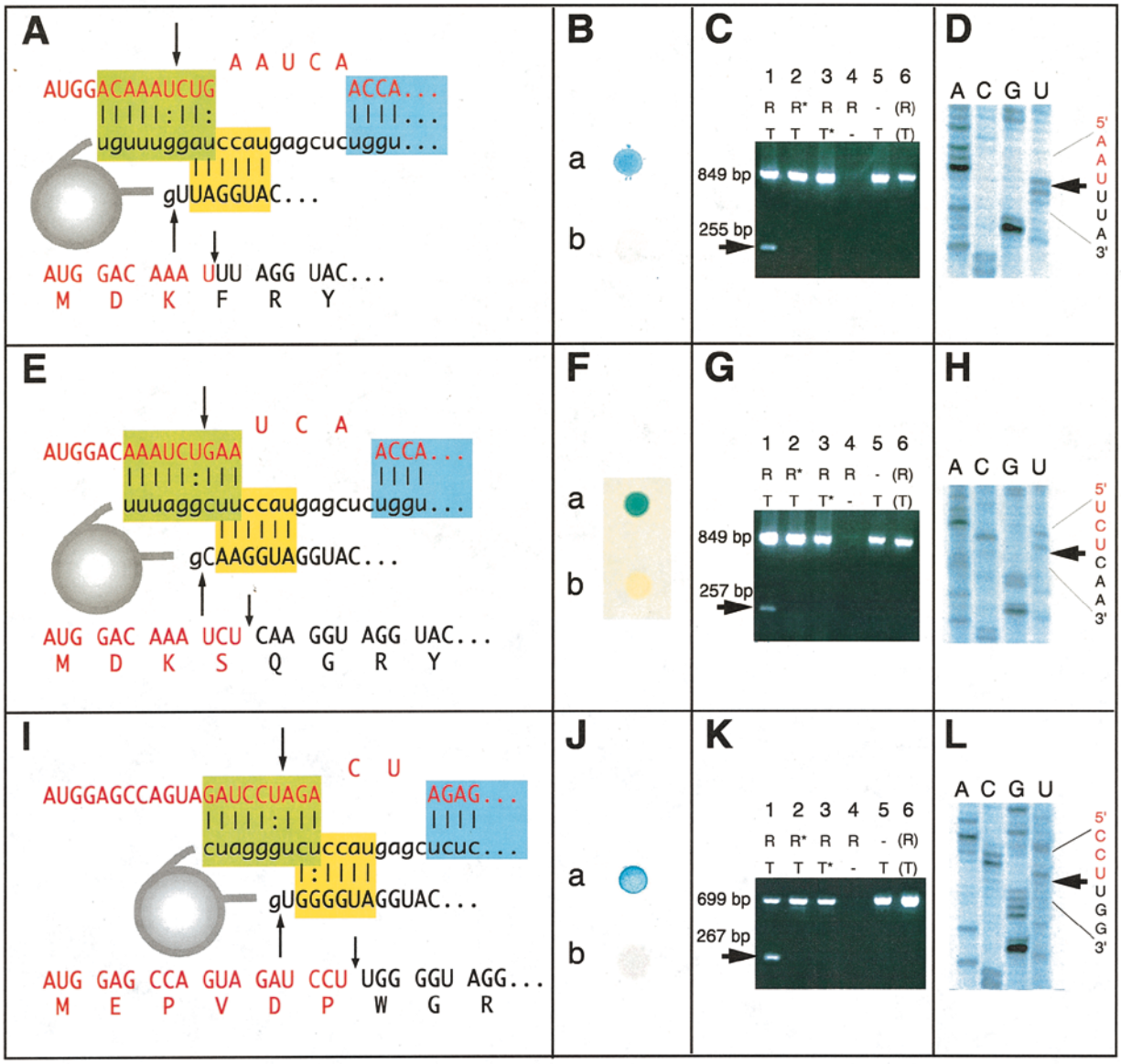

Figure 5. Sequences of base-paired ribozyme and target mRNAs are shown with the results of assays for *trans*-splicing of the anti-CMV(10) (a)-(d), anti-CMV(12) (e)-(h), and anti-HIV(i)-(l) ribozymes. (a), (e) and (i) Sequences forming the P1 (green) and P10 (yellow) interactions and the antisense region (blue) are shown for the *trans*-splicing ribozymes with their corresponding targets. (:) shows wobble base-pairs in the paired regions P1 and/or P10, and arrows show 5′ and 3′ splice sites. The wobble base-pair in P10 of the anti-HIV ribozyme was introduced to avoid an in frame stop codon after the *trans*-splicing reaction. Coding sequences are in upper case letters, intron sequences in lower case letters. The sequence of the ligated exons is given and an arrow indicates the ligation site. (b), (f) and (j) β-Galactosidase assays. The upper colony co-expressed the ribozyme with the target, the lower one expressed the inactive ΔP5abc control ribozyme with the target. Colonies containing the active ribozymes were incubated for different time periods to produce intense histochemical staining (overnight for anti-HIV, and two days for anti-CMV). (c), (g) and (t) RT-PCR assays. Results for RNA analysis of the following cultures are shown: lane 1, culture co-expressing the ribozyme (R) with the correct target RNA (T); lane 2, culture co-expressing the ΔP5abc ribozyme (R*) with the target RNA; lane 3, culture co-expressing the ribozyme with the mutated target RNA (T*); lane 4, culture expressing the ribozyme alone; lane 5, culture expressing the target RNA alone; lane 6, control, where a culture expressing the ribozyme was mixed with a culture expressing the target RNA prior to analysis. Arrows indicate amplified spliced product, larger products derived from the target mRNA are also evident. (d), (h) and (l) The products derived from spliced products were directly sequenced. Sequences adjacent to splice sites are shown and arrows indicate the ligation site.

zymes could be challenged with long mRNA transcripts and cells expressing the chimaeric target RNA be easily visualized. We also produced mutated forms of the target mRNAs in order to provide a more stringent test of ribozyme specificity. A uridine residue near each of the target splice sites was changed to guanosine, which prevented formation of an essential G·U wobble base-pair preceding the splice sites for the anti-CMV(10) and anti-HIV ribozymes, and which introduced a base-pair mismatch within helix P1 in the case of the anti-CMV(12) ribozyme.

Two approaches were taken for the initial testing of these ribozymes. In one set of experiments the anti-CMV ribozymes and the target mRNAs were maintained on separate compatible plasmids in *E. coli*, similar to the experiments described above for the anti-CAT ribozymes. In a second approach, the anti-CMV(10) and anti-HIV ribozymes were cloned with their corresponding target genes on single plasmids, with the genes orientated for divergent transcription. Similar results were seen for both approaches. All of the ribozyme and target genes had been placed under control of the *tac* promoter (Amann *et al*., 1983; de Boer *et al*., 1983), and expression was induced in *E. coli* using IPTG. RT-PCR techniques were used to detect the expected spliced RNA products, and histochemical staining and enzyme assays were used to test for β-galactosidase activity in the transformed cells. RNA products and β-galactosidase activity were seen after coexpression of each of the three active ribozymes with their corresponding target RNAs (Figure 5), whereas no detectable protein activity or RNA products were produced with the negative controls. For example, mutation of the CMV-target mRNA to alter a single nucleotide, positioned two residues upstream of the chosen splice site, eliminated *trans*-splicing with the anti-CMV(12) ribozyme (Figure 5(g)).

In order to quantitate the levels of enzyme activity produced after co-expression of the active ribozymes and target mRNAs, we assayed bacterial extracts. Expression of the anti-HIV ribozyme with its target gave rise to nine units of β-galactosidase activity, while the anti-CMV ribozymes produced little activity above background with this assay, which was less sensitive than histochemical staining (data not shown, but see Figure 5(b) and (f)). In contrast, the anti-CAT ribozyme produced 412 units of β-galactosidase activity after expression with its target. The reason for the observed differences in activity is not clear; however, it is possible that ribozymes have different affinities for their target RNAs, or that different sites in the target RNAs are more easily accessible to the ribozyme. Also, sequence alterations within helices P1 or P10 might affect catalytic efficiencies. Alternatively, N-terminal fusion of amino acid residues (encoded by the target mRNA) may modulate the α-peptide activity of the fused polypeptide product, and thereby alter β-galactosidase activity. Co-expression of the anti-CMV(10), anti-CMV(12), and anti-HIV ribozymes with their respective targets, resulted in detection of RT-PCR products of the size expected to result from correct splicing of the target and ribozyme RNAs (255, 257, and 267 base-pairs, respectively; Figure 5(c), (g) and (k)). Sequence analysis confirmed that the products all contained the expected splice junction (Figure 5(d), (h) and (l)). RT-PCR amplification of cell extracts also resulted in longer products corresponding with the target RNAs (849 bp for the CMV target and 699 bp for the HIV target RNA), since these also contain binding sites for the primers (Figure 5(c), (g) and (k)).

All of the *trans*-splicing ribozymes produced detectable *trans*-spliced RNA products and β-galactosidase activity when expressed in cells containing the intended target mRNA. Thus, the model for the design of *trans*-splicing ribozymes appears to be robust. It is of equal importance that no detectable RNA or β-galactosidase products were formed when the ribozymes (or target mRNAs) were expressed alone, or when the defective ΔP5abc forms of the ribozymes or mutated target mRNAs were used (Figures 2 and 5). There is no evidence for expression of the 3′ exon sequences due to illegitimate splicing to other mRNAs and subsequent translation. However, we cannot exclude the possibility that inactive products might be produced by *trans*-splicing to other RNAs.

Mutation of a single base at or near the chosen splice site in the target mRNA eliminates the production of *trans*-spliced RNA product (see Figure 5(c), (g) and (k), lanes 3) and β-galactosidase activity (data not shown). In the case of the anti-CMV(10) and anti-HIV ribozymes and their mutated target mRNAs, an essential G·U wobble base-pair (Michel & Westhof, 1990) preceding each chosen splice site is disrupted, and the loss of catalytic activity is expected. In the case of the anti-CMV(12) ribozyme and its mutant target, a single base-pair, located two nucleotides upstream from the splice site in helix P1, is disrupted, and this also abolishes activity (Figure 5(g), lane 3; the U located two nucleotides upstream of the splice site was mutated to a G). This demonstrates a high degree of specificity for the *trans*-splicing reaction and should allow the use of *trans*-splicing ribozymes to discriminate between closely related mRNA targets in other cells. Recently, we have established a *trans*-splicing ribozyme-based system for targeted gene delivery in yeast, using the diphtheria toxin A chain. It has been possible to selectively kill cells differing by a single nucleotide within a chosen target mRNA. Delivery of a toxin provides a considerably more sensitive indicator for illegitimate expression, and we have confirmed that addition of an extended antisense domain to the ribozyme is a requirement for efficient and selective cytotoxicity in this system (B.G.A., U.K., H.M.G. & J.H., unpublished results).

Discussion

Design of improved *trans*-splicing ribozymes

We have established an improved method for the design of *trans*-splicing ribozymes, based on the *T. thermophila* large rRNA group I intron. The ribozymes are built using the catalytic domain of the intron, and sequences are fused to the 5′ terminus to confer specificity towards the chosen target RNA, and fused to the 3′ terminus to provide the new 3′ exon. The ribozymes can be precisely targeted at a chosen site within any given mRNA, provided that several factors are taken into account during the design of a new ribozyme. First, a target RNA and 5′ splice site must be chosen. In the final *trans*-splicing complex, only the 5′ portion of the P1 duplex is contributed by the target RNA. A single conserved uridine residue is required immediately 5′ of the intended splice site. This is the sole sequence requirement in the target RNA. Second, having chosen a particular target sequence, compensating sequence changes must be added to the 5′ section of the ribozyme, in order to allow the formation of a suitable helix P1 between the target and ribozyme RNAs. The helix P1 should contain a U·G base-pair at the intended 5′ splice site, ideally positioned at the sixth position from the base of the helix. Third, the exon sequences flanking the 3′ splice site must be chosen, and adjustments made in the 5′ section of the ribozyme to allow the formation of a stable P10 helix. Finally, a region of complementarity to the target RNA is placed at the 5′ terminus of the *trans*-splicing ribozyme in order to increase the potential for interaction between the ribozyme and target.

Trans-splicing ribozymes may be useful for the conditional delivery of new gene activities *in vivo*, as *trans*-spliced product can only be produced in cells containing the chosen target mRNA. A major practical advantage of *trans*-splicing ribozymes is that even inefficient catalysis *in vivo* can give rise to a RNA product that is unique and easily detected, and may produce an easily assayed translation product. In order to test these new ribozymes, we have developed a simple biological assay in *E. coli*. This assay is based upon ribozymes that *trans*-splice the *LacZ* α-peptide coding sequence into various chosen target mRNAs. *Trans*-splicing gives rise to β-galactosidase activity, which can be readily measured. We have constructed four ribozymes targeted at different sites in the CAT, CMV coat protein and HIV TAT mRNAs. In each case, the ribozyme recognized the intended target and accurately spliced the 3′ exon sequence to it, producing the correct spliced product and β-galactosidase activity. In control experiments, the ribozymes were expressed in the absence of the correct target RNA and with defective catalytic domains. Cryptic splicing and illegitimate expression of β-galactosidase were not evident, and the rules for the design of these new ribozymes appear to be general and robust. We have also modified ribozymes targeted at the CAT mRNA, and assayed these to gauge whether the salient features of the ribozyme design can be further improved.

Antisense region

Previously described *trans*-splicing ribozymes for mRNA repair (Sullenger & Cech, 1994) contain a region of six base-pairs (helix P1) of sequence complementary to the target mRNA. It is not suprising that these ribozymes have been shown to produce illegitimate splice products *in vivo* (Jones *et al*., 1996), as a homologous target sequence will occur on average once in every 4096 nucleotides. We have shown that helix P1 can be extended to nine base-pairs, similar to the wild-type intron, with a beneficial effect on the biological activity of the ribozymes (Table 1). In order to confer further improved specificity for the target mRNA, we have fused additional complementary sequences to the 5′ terminus of the ribozymes. We first chose an arbitrary length of about 45 nt, and then compared the biological activity of ribozymes with various lengths of additional complementary sequence (0, 101 and 201 nt). The addition of 45 to 201 nucleotides of complementary sequence to the ribozyme caused a dramatic 40 to 50-fold increase in β-galactosidase activity in *E. coli* bioassays. Replacement of this antisense domain with non-specific sequences caused loss of activity. Increased complementarity between the ribozyme and the target RNAs results in improved biological activity. The antisense sequences are not expected to contribute directly to the ribozyme catalysed reaction, and it is likely that they aid the association of the ribozyme and target RNAs *in vivo*.

On the face of it, this observation contradicts theoretical studies, which predict that an increased length of complementarity would promote illegitimate interactions, rather than enhance ribozyme specificity (Herschlag, 1991). However, a plausible explanation for our results is that fortuitous secondary and tertiary interactions between the antisense domain and the internal guide sequence might exclude non-specific mRNAs from entering the active site, even those which might contain a suitable target site. However, the intended target could still interact with any accessible portion of the antisense domain, to initiate branch migration and formation of an extended duplex, which would then unmask the IGS and allow formation of helix P1. If this explanation is correct, it should be feasible to engineer or to select for shorter antisense domains which act in a similar way to confer improved ribozyme specificity, and we are currently testing this. Similar shielding of complementary sequences has been noted in natural antisense RNAs (Simons & Kleckner, 1988).

Helices P1 and P10

A novel feature of the ribozymes described here is the maintenance of an artificial helix P10,

through compensating base changes. The P10 sequence element is conserved among group I introns related to that of *Tetrahymena*, and has been shown to be important for 3′ splice site selection during splicing (Suh & Waring, 1990). In our initial model, we maintained the wild-type arrangement, with the potential for the formation of six base-pairs. We subsequently shortened this to four, two and zero base-pairs, and tested the biological activity of the variants. Reduction of helix P10 to zero or two base-pairs caused an almost 50-fold drop in β-galactosidase activity, showing that P10 is clearly important for ribozyme activity *in vivo*. Two of the nucleotides required for formation of helix P10 also form part of helix P1, and we have shown that this portion of helix P10 is dispensible, provided that the remaining four base-pairs are maintained. The P1 sequence is determined by the choice of target site. Therefore, in future it should not be necessary to adjust the P10 sequences within the 3′ exon when targeting the ribozymes at different mRNAs. This will allow re-use of the 3′ exon sequences.

Biological specificity

The ribozymes appear to be biologically highly specific. Alteration of a single nucleotide near the chosen mRNA target site eliminates the production of β-galactosidase activity. We have mutated the conserved uridine adjacent to the intended splice site within the HIV and CMV(10) target mRNAs, and the residue positioned two nucleotides upstream of the CMV(12) splice site. The latter might be regarded as a structural rather than catalytic mutant. The use of mutant target RNAs resulted in the loss of detectable *trans*-spliced product and β-galactosidase activity. This is a strong indication of the biological specificity of these ribozymes, and suggests that the technique should allow distinction between closely related transcripts. We have seen the same high degree of specificity in yeast experiments, where single base changes in the target mRNA can produce immunity to ribozyme-mediated cell ablation (Ayre *et al*., unpublished results).

Ribozyme-mediated gene delivery

These ribozymes catalyse *in vivo trans*-splicing in an effective and reproducible fashion. It may be possible to target *trans*-splicing ribozymes against virtually any RNA in a cell and to introduce new gene activities depending upon existence of the specific target RNA. This modified design may also be useful for the construction of improved *trans*-splicing ribozymes for the repair of mutant mRNAs (Lan *et al*., 1998; Phylactou *et al*., 1998), but this requires efficient catalysis. It seems likely that inclusion of a helix P10 and extension of P1 will be of benefit; however, addition of an antisense region may be problematic due to potential base-pairing with the 3′ exon sequences which would be identical to the target RNA sequences. This might be circumvented by alteration of the codon usage of the 3′ exon or by use of a related gene, to reduce the amount of intra-ribozyme base-pairing between the antisense region and the 3′ exon.

The ''hammerhead'' class of self-cleaving RNAs was used as a basis for the design of highly specific endoribonucleases, which could be targeted at a chosen mRNA (Haseloff & Gerlach, 1988). It was hoped that these might find application for inhibiting gene expression *in vivo*. However, cleaving ribozymes must effectively deplete the target mRNA pool within the cell, and efficient access to the substrate and subsequent catalysis is essential for their use *in vivo*. In contrast, even inefficient delivery of a *trans*-splicing ribozyme might allow the useful expression of new gene activities, dependent on the presence of a chosen RNA. The effects of producing a small amount of spliced mRNA product can be substantially amplified by the biological activity of its translation product. This has two benefits. First, *trans*-splicing ribozymes are easier to assay. Even inefficient catalysis can give rise to a fused RNA product that itself is unique and easily detected, and which may give rise to an assayable translation product. This is a considerable improvement compared with the problems of assaying the activity of endoribonucleolytic ribozymes, where *in vivo* catalysis results in merely depletion of the pool of target RNAs and generation of unstable and inactive products. We have already established simple bioassays for ribozyme delivery in *E. coli* and yeast (B.G.A., U.K., H.M.G. & J.H., unpublished results), and this will allow more rapid testing of methods for improved ribozyme expression *in vivo*. Second, the *trans*-splicing ribozymes might be useful as ''RNA sensors''. *Trans*-splicing ribozymes could be used for the delivery of new gene activities in living cells, conditional upon the presence of a chosen mRNA species. For example, we have constructed ribozymes capable of splicing the diphtheria toxin A chain coding sequence into the CMV coat protein mRNA, and shown that yeast cells expressing the target RNA can be specifically ablated (Ayre *et al*., unpublished results). Ribozyme-targeted delivery of new gene activities could be used to specifically mark or kill cells that are virus-infected or of a particular differentiated or malignant type.

Materials and Methods

Ribozyme construction

All plasmids were constructed using standard techniques (Sambrook *et al*., 1989). The ribozymes were cloned into a modified pET17b vector (Novagen). The *Xho*I site in the multiple cloning site was destroyed by end-filling, followed by religation. A *tac* promoter (Amann *et al*., 1983; de Boer *et al*., 1983) sequence TCT AGA TCT GTT GAC AAT TAA TCA TCG GCT CGT ATA ATG TGT GGA ATT GTG AGC GGA TAA CAA TTT CAC ACA AGC TT was synthesised and inserted

downstream of the T7 promoter between the *Xba*I and *Hin*dIII sites of the polylinker. This insertion replaced ATG sequences which could potentially be utilised as translational initiation sites. The ribozyme genes were constructed in a modular fashion to allow easy replacement or re-use of functional domains. PCR techniques were used to amplify and to insert flanking restriction endonuclease sites of *Hin*dIII and *Xho*I for the antisense region, *Xho*I and *Kpn*I for the intron, and *Kpn*I and *Sac*I for the 3′ exon. For construction of the different ribozymes, the intron was PCR amplified using a 5′ primer that contains the recognition sequence for *Hin*dIII, the target-RNA-specific antisense region, the recognition sequence for *Xho*I, and the new IGS with the capability of forming the P1 and P10 interactions; the 3′ primer contains the first nucleotides of the 3′ exon, complementary to the IGS for forming the P10 interaction and the recognition sequence for *Kpn*I. A cloned sequence of *T. thermophila* large subunit rRNA intervening sequence (Zaug & Cech, 1986) served as template for PCR, and the active ribozymes contained nt 28 to 414 of the intron. A deleted form of the intron, corresponding with the ΔP5abc mutation (Joyce *et al*., 1989; Mohr *et al*., 1994) was used to provide an inactive control. The sequences of the primers used are as follows: CAT 5′ primer: TCA GAA GCT TCA AAA TGT TCT TTA CGA TGC CAT TGG GAT ATA TCA ACG GTG GCT CGA GTA CCT CCG GTG ATA AAA GTT ATC AGG CAT GCA CCT GG; CAT 3′ primer: ATC AGG TAC CTA CCC CAC GAG TAC TCC AAA ACT AAT CA. Anti-CMV(10): 5′ primer: CAG AAG CTT GGA TCC ACC ACG ACG CAA ACG ACG TCG ACG GTT ACG ACC AGC ACT GGT CTC GAG TAC CTA GGT TTG TAA AGT TAT CAG GCA TGC ACC TGG; 3′ primer: CTA AGG TAC CTA ACG AGT ACT CCA AAA CTA ATC. Anti-CMV(12): 5′ primer CAG AAG CTT GGA TCC ACC ACG ACG CAA ACG ACG TCG ACG GTT ACG ACC AGC ACT GGT ACT CGA GTA CCT TCG GAT TTA AAA GTT ATC AGG CAT GCA CCT GG and 3′ primer CTA AGG TAC CTA CCT TGC GAG TAC TCC AAA ACT AAT C. HIV 5′ primer: TCA GAA GCT TAC AAG CAG TCC TAG GTT GAC TTC CTG GAT GCT TCC AGG GCT CTC GAG TAC CTC TGG GAT CAA AAG TTA TCA GGC ATG CAC CTG G; HIV 3′ primer: same as for CAT 3′.

The PCR products were purified, digested with *Hin*dIII and *Kpn*I, and cloned into the modified pET derivative cut with the same restriction enzymes. The α-peptide of β-galactosidase was used as 3′ exon, giving the possibility to detect successful *trans*-splicing as β-galactosidase activity in the bacteria. PCR was used to create a fragment with pGEM3Zf (Promega) as a template. The 5′ oligo was GGC GGT ACC TGA TTA CGG ATT CAC TGG CCG TCG TTT TAC AAC G and the 3′ oligo was GGC GAA TTC GAG CTC TTA TTC GCC ATT CAG GCT GCG CAA CTG TTG GGA AGG. This gene fragment was cloned into the *Kpn*I and *Sac*I sites downstream of the ribozyme sequences in plasmid pET.

A truncated form of the anti-CAT ribozyme without the antisense domain was constructed by digestion of the ribozyme-containing plasmid with *Hin*dIII and *Xho*I followed by a fill-in reaction using T4 polymerase (New England Biolabs) and religation. The anti-CAT ribozymes with 101 and 201 nt antisense regions were constructed by cloning PCR fragments of the CAT gene into the *Hin*dIII and *Xho*I sites of the ribozyme, thereby replacing the 46 nt antisense region. The CAT fragments were generated using the following primers: 5′ primer: ATC GCT CGA GAC CAC CGT TGA TAT ATC CCA ATG G and 3′ primers for the 101 and 201 nt fragments: ATC GAA GCT TCG TAA TAT CCA GCT GAA CGG TC and ATC GAA GCT TCC ATA CGG AAT TCC GGA TGA GC, respectively.

For construction of a ribozyme with a mutated P10 helix, the anti-CAT ribozyme was used as a template for PCR with the above 5′ primer and a new 3′ primer where the six nucleotides involved in the P10 base-pairing were mutated. The oligo was ATC AGG TAC C*GG GAA A*AC GAG TAC TCC AAA ACT AAT CA. Construction of the other P10 mutated ribozymes was done in the same way with primers introducing the base changes shown in Figure 4.

A ribozyme with shortened helices P1 and P10 with six and four nucleotides only, respectively, was constructed using the anti-CAT ribozyme as a template for PCR with the following primers: GTG GCT CGA GTA CCG AAG GTG ATA AAA GTT ATC AGG CAT GCA CCT GG and ATC AGG TAC CTA CCG TAC GAG TAC TCC AAA ACT AAT CA.

For further analysis, derivatives of vector pBBR122 (MoBiTec) were constructed for the anti-CMV ribozymes. This broad host range vector is compatible with the pET plasmids in *E. coli*. For cloning of the ribozyme transcription units into pBBR122, they were PCR amplified with the pET vectors as template using primers that generate recognition sites for *Nco*I and *Ssp*I at both the 5′ and 3′ ends of the amplification products (5′ primer: TCC ACC ATG GAA TAT TGC GCA CCC GTG GCC AGG ACC CAA CGC TGC CCG; 3′ primer: TTG ACC ATG GAA TAT TAT CCG GAT ATA GTT CCT CCT TTC AGC AAA AAA CC). The PCR products were purified, digested with *Nco*I, and subcloned into *Nco*I-digested pBBR122.

Target mRNAs

The CAT mRNA was expressed from plasmid pACYC184 (Chang & Cohen, 1978) . The vectors for expression of the CMV and HIV targets are based on the modified pET17b vector harbouring the *tac* promoter (see above). The CMV target contains the first 60 coding nucleotides of the coat protein mRNA of CMV. A target site was chosen within a region that is highly conserved between the coat protein mRNAs of CMV strains Y, WD, Fny, M, I, O, D, and C. A Shine-Dalgarno sequence was added upstream of the AUG start codon for efficient translation in *E. coli*. The coding region of green fluorescent protein (mGFP5; Siemering *et al*., 1996) was fused in frame to the 3′ end of the CMV sequences (*Bgl*II-*Bam*HI fusion), allowing the expression of the target gene to be monitored in the bacteria due to the fluorescent protein product. This DNA fragment was cloned into the modified pET17b as a *Hin*dIII/*Sac*I fragment. A mutated version of the target, serving as an inactive control, was constructed, where the tenth nucleotide (counted from the A of the AUG start codon) U is changed to G, to interfere with *trans*-splicing.

The HIV target was cloned into the CMV target vector, replacing the CMV region with the first 67 nt encoding the TAT protein of HIV-I. A region of exact homology between known isolates of HIV-I was chosen as the portion of the target that forms helix P1. The region 3′ of the splice site that is recognized by the antisense portion of the ribozyme also shows a high degree of homology between the known HIV-I sequences. The 5′ and 3′ restriction enzyme sites for *Hin*dIII and *Nco*I, respectively, were added to the HIV sequences and the

digested DNA fragment was cloned into the CMV target vector cut with the same enzymes, thereby replacing the CMV sequences. Translation of the HIV target construct does not result in fluorescent protein as a part of the 5′ portion of GFP is removed. A mutated version was constructed, changing the 18th nucleotide U to G which, again, interrupts the essential G·U wobble base-pair preceding the 5′ splice site and should inhibit splicing.

For analysis of the anti-CMV(10) and anti-HIV ribozymes, vectors that harbour the ribozyme and target transcription units in opposite direction were constructed: employing the same primers used for amplifying the ribozyme transcription units for cloning into pBBR122 (see above), the target transcription units were PCR amplified. The PCR fragments were purified, cut with *Ssp*I and cloned into the *Bsa*AI site of the pET vectors containing the ribozymes.

Detailed restriction maps of the vectors are available upon request.

Bioassay

Bacterial cultures were grown overnight at 37 °C containing the appropriate antibiotics and 2 mM IPTG to induce transcription from the *tac* promoter. Cells were harvested by centrifugation and freezing on solid CO_2. The pellet was resuspended in 300 µl of lysis buffer (5 % (w/v) SDS, 10 mM Tris-HCl (pH 7.5), 20 mM EDTA) and 300 µl of phenol was added. The mixture was vortexed and incubated at 65 °C for ten minutes with repeated vortexing. Two additional phenol/chloroform extractions and one chloroform extraction were performed. The RNA was ethanol-preciptitated and treated with RNase-free DNase (Promega) for 15 minutes at 37 °C, followed by phenol/chloroform and chloroform extractions, and ethanol precipitation. Reverse transcription was performed with a 3′ specific primer (GCC GGA TCT GCT CGA TCG AGC GG) priming in the 3′ transcribed region of pET17b and Super RT (H T Biotechnology) following the supplier's protocol in a volume of 50 µl; 25 µl of the first strand cDNA were heated at 93 °C for five minutes and then used directly in the following PCR using a 3′ nested primer (TGC AGA ATT CCA GCA CAC TGG CG), complementary to the 3′ transcribed region of pET17b and a 5′ target-specific primer (CAT: GAG ATT TTC AGG AGC TAA GGA; CMV and HIV (the first three nucleotides 5′ are not homologous to the HIV target): TCC AAG GAG ATA TAT AAC AAT GG). Amplifications were carried out in a volume of 50 µl for 35 cycles using *Taq* polymerase (Promega) with the following parameters: 30 seconds at 93 °C, 90 seconds at 55 °C, 90 seconds at 72 °C. Aliquots of the reaction mixture were analysed on 2 % agarose gels stained with ethidium bromide. For quantitative RT-PCR analysis, the β-lactamase gene present on the pET plasmids was used as an internal control (employing β-lactamase specific primers) and products were analysed after 20, 25, and 30 cycles of amplification.

Sequence analysis: the RT-PCR bands were gel eluted and reamplified with the same primers used in the RT-PCR. Aliquots of the reamplification mixture were directly sequenced using the Thermo Vent Cycling Kit (New England Biolabs) following the supplier's protocol employing a primer that is homologous to a region in the *LacZ* gene (ACG CCA GCT GGC GAA AGG GG).

β-Galactosidase assays were performed as described (Miller, 1992).

Acknowledgements

We thank Andrew Newman for critical comments on the manuscript, and Andrea Brand, Norbert Perrimon and Andrew Ellington for materials and helpful discussions. The work was supported by grants from AgrEvo (J.P.H.) and Hoechst AG (H.M.G.). We thank Ed Ullman and Sheba Mahmoodian for technical assistance.

References

Amann, E., Brosius, J. & Ptashne, M. (1983). Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene,* **25**, 167-178.

Been, M. D. & Cech, T. R. (1985). Sites of circularization of the *Tetrahymena* rRNA IVS are determined by sequence and influenced by position and secondary structure. *Nucl. Acids Res.* **13**, 8389-8407.

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

Blumenthal, T. (1995). *Trans*-splicing and polycistronic transcription in *Caenorhabditis elegans*. *Trends Genet.* **11(4)**, 132-136.

Cech, T. R. (1990). Self-splicing of group I introns. *Annu. Rev. Biochem.* **59**, 543-568.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* **134(3)**, 1141-1156.

de Boer, H. A., Comstock, L. J. & Vasser, M. (1983). The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. *Proc. Natl Acad. Sci. USA,* **80**, 21-25.

Doudna, J. A. & Szostak, J. W. (1989). RNA-catalysed synthesis of complementary-strand RNA. *Nature,* **339**, 519-522.

Doudna, J. A., Cormack, B. P. & Szostak, J. W. (1989). RNA structure, not sequence, determines the 5′ splice-site specificity of a Group I intron. *Proc. Natl Acad. Sci. USA,* **86**, 7402-7406.

Eul, J., Graessmann, M. & Graessman, A. (1995). Experimental evidence for RNA *trans*-splicing in mammalian cells. *EMBO J.* **14(13)**, 3226-3235.

Galloway-Salvo, J. L., Coetzee, T. & Belfort, M. (1990). Deletion tolerance and *trans*-splicing of the bacteriophage T4 *td* intron. Analysis of the P6-L6a region. *J. Mol. Biol.* **211**, 537-549.

Haseloff, J. & Gerlach, W. L. (1988). Simple RNA enzymes with new and highly specific endoribonuclease activities. *Nature,* **334**, 585-591.

Herschlag, D. (1991). Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules *in vivo*: more isn't always better. *Proc. Natl Acad. Sci. USA,* **88**, 6921-6925.

Jaeger, L., Michel, F. & Westhof, E. (1996). The structure of group I ribozymes. In *Catalytic RNA. Nucleic Acids and Molecular Biology* (Eckstein, F. & Lilley, D., eds), vol. 10, pp. 33-51, Springer Verlag, Berlin, Heidelberg, New York.

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

Jones, J. T., Lee, S. W. & Sullenger, B. A. (1996). Tagging ribozyme reaction sites to follow *trans*-splicing in mammalian cells. *Nature Med.* **2(6)**, 643-648.

Joyce, G. F., van der Horst, G. & Inoue, T. (1989). Catalytic activity is retained in the *Tetrahymena* group I intron despite removal of the large extension of element P5. *Nucl. Acids Res.* **17(19)**, 7879-7889.

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

Lehnert, V., Jaeger, L., Michel, F. & Westhof, E. (1996). New loop-loop tertiary interactions in self-splicing introns subgroup IC and ID: A complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.* **3**, 993-1009.

Michel, F. & Westhof, E. (1990). Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216(3)**, 585-610.

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

Miller, J. H. (1992). *A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for* Escherichia coli *and Related Bacteria*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mohr, G., Caprara, M. G., Guo, Q. B. & Lambowitz, A. M. (1994). A tyrosyl-tRNA synthetase can function similarly to an RNA structure in the *Tetrahymena* ribozyme. *Nature,* **370(6485)**, 147-150.

Murphy, F. L. & Cech, T. R. (1989). Alteration of the substrate specificity for the endoribonucleolytic cleavage of RNA by the *Tetrahymena* ribozyme. *Proc. Natl Acad. Sci. USA,* **86(23)**, 9218-9222.

Phylactou, L., Darrah, C. & Wood, M. (1998). Ribozyme-mediated *trans*-splicing of a trinucleotide repeat. *Nature Genet.* **18**, 378-381.

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

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schwinghammer, M. W. & Symons, R. H. (1977). Translation of the four major RNA species of cucumber mosaic virus in plant and animal cell-free systems and in toad oocytes. *Virology,* **97**, 88-108.

Siemering, K. R., Golbik, R., Sever, R. & Haseloff, J. (1996). Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr. Biol.* **6(12)**, 1653-1663.

Simons, R. W. & Kleckner, N. (1988). Biological regulation by antisense RNA in prokaryotes. *Annu. Rev. Genet.* **22**, 567-600.

Suh, E. R. & 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. Cel. Biol.* **10(6)**, 2960-2965.

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

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

Zaug, A. J. & Cech, T. R. (1986). The intervening sequence RNA of *Tetrahymena* is an enzyme. *Science,* **232**, 470-475.

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

Zhou, C., Bahner, I. C., Larson, G. P., Zaia, J. A., Rossi, J. J. & Kohn, D. B. (1994). Inhibition of HIV-1 in human T-lymphocytes by retrovirally transduced anti-*tat* and *rev* hammerhead ribozymes. *Gene,* **149**, 33-39.

Edited by J. Karn

(Received 22 September 1998; received in revised form 30 November 1998; accepted 30 November 1998)