*Proc. Natl. Acad. Sci. USA*
Vol. 92, pp. 11461–11464, December 1995
Biochemistry

# *In vitro* trans-splicing in *Saccharomyces cerevisiae*

ANDREA GHETTI AND JOHN N. ABELSON

Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125

*Contributed by John N. Abelson, August 30, 1995*

**ABSTRACT The interactions established at the 5′-splice site during spliceosome assembly are likely to be important for both precise recognition of the upstream intron boundary and for positioning this site in the active center of the spliceosome. Definition of the RNA–RNA and the RNA–protein interactions at the 5′ splice site would be facilitated by the use of a small substrate amenable to modification during chemical synthesis. We describe a trans-splicing reaction performed in *Saccharomyces cerevisiae* extracts in which the 5′ splice site and the 3′ splice site are on separate molecules. The RNA contributing the 5′ splice site is only 20 nucleotides long and was synthesized chemically. The trans-splicing reaction is accurate and has the same sequence, ATP, and $Mg^{2+}$ requirements as cis-splicing. We also report how deoxy substitutions around the 5′-splice site affect trans-splicing efficiency.**

The recognition of the 5′-splice site (5′SS) and its positioning in the spliceosome active site appears to be a multistep reaction (for review, see refs. 1–4). At least three small nuclear RNAs (snRNAs) have been implicated in this process. U1 snRNA is required for the initial intron 5′-end recognition, which is accomplished primarily by Watson–Crick base pairing (5–7). Subsequently, U5 snRNA contributes to the recognition of the 3′ end of the upstream exon, and this interaction also involves base-pairing interactions (8, 9). Before the first step of the splicing reaction, the U6 snRNA displaces U1 at the 5′SS to establish an interaction important for 5′SS definition (10–15). A number of proteins have also been reported to interact with the 5′SS (for review, see ref. 4). Many details of these RNA–RNA and RNA–protein interactions are not known. For example, although the first three bases of introns are completely conserved among the known intron sequences and have been placed in close proximity to $A^{51}$ in U6 snRNA by crosslinking experiments (ref. 15; C.-H. Kim and J.N.A., unpublished results), it is not clear whether U6 specifically recognize these nucleotides. Knowledge of the precise geometry of the U5 and U6 snRNA interactions across the 5′SS could contribute substantially to our understanding of the first step of splicing by constraining the conformation of the substrate.

A small chemically synthesized 5′SS substrate would allow for easy RNA modifications. Introducing crosslinking derivatives at specific positions or changing the hydrogen-bonding potential or steric hindrance at particular sites would allow the probing of many features of the interactions formed at the 5′SS. Hall and Konarska (16) found that a small RNA oligonucleotide bearing the 5′SS consensus sequence can induce assembly of ribonucleoprotein complexes containing the U2, U4, U5, and U6 small nuclear ribonucleoproteins. Interestingly, the interaction between this short 5′SS and the spliceosomal RNAs appears to have many features in common with the interactions formed at the 5′SS during the splicing reaction (17, 18). HeLa cell extracts can splice substrates in which the 5′ and 3′ splice sites are on separate RNA molecules (19–23). In the trans-splicing reactions studied by Chiara and Reed (23), the donor site molecules contained between 45 and 129 nucleotides of upstream exon sequence but product formation has been observed by these authors with as few as 4 nucleotides of intronic sequence at the 5′SS.

In the present report, we describe trans-splicing reactions performed in *Saccharomyces cerevisiae* crude extracts by using a short donor site molecule synthesized chemically. The splicing reaction is accurate, and this experimental system has allowed us to probe the role of all of the 2′-OH groups in consensus nucleotides at the 5′SS.

## METHODS

**RNA Preparation.** The following RNA oligonucleotides were synthesized chemically by using an Applied Biosystems synthesizer: RNA1, 5′-AUGGAUUCUGGUAUGUUCUA-3′; RNA1/int, 5′-AUGGAUUCUGAUAUGUUCUA-3′; RNA1/ex, 5′-AUGGAUUGAGGUAUGUUCUA-3′. All 2′-deoxy-modified RNA oligonucleotides were synthesized chemically. The plasmid construct for transcribing RNA2 was generated as follows. The cloned yeast actin gene was used as a template for a polymerase chain reaction (PCR) with the oligonucleotides Act2 (5′-CGCCGGATCCC*TAATACGACTCACTATAG*GGTGATAAGTGATAG-3′) containing a *Bam*HI restriction site (underlined) and the sequence of the T7 RNA polymerase promoter (in italic type) and ActR (5′-CGCCGGAATTCCTTCATCACCAACGTAGGA-3′) containing an *Eco*RI restriction site (underlined). The PCR product obtained with these two oligonucleotides contains the 104 nucleotides at the 3′ end of the actin gene intron as well as 163 nucleotides of the downstream exon. The PCR products were phenol-extracted; the DNA was precipitated with ethanol and digested with *Eco*RI and *Bam*HI. After purification by agarose gel electrophoresis, the PCR product was cloned in a pUC19 plasmid vector containing *Eco*RI and *Bam*HI ends. The sequence of the cloned actin gene fragment was confirmed by direct plasmid sequencing. For RNA2 synthesis, the plasmid was linearized with *Eco*RI and used as template for *in vitro* transcription using T7 RNA polymerase.

**Trans-Splicing Reaction.** $^{32}P$ internally labeled RNA2 (80 fmol) and the indicated amount of RNA1 were incubated in 10 μl containing 55 mM potassium phosphate (pH 7.20), 2.5 mM $MgCl_2$, 3 mM ATP, 2.5% (wt/vol) PEG ($M_r$ 8000), 1 mM spermidine, 1 mM dithiothreitol, and 60% (vol/vol) yeast extract prepared as described (24). Incubation was at 16°C for the indicated time. Reactions were stopped by addition of 1 μl of a mixture containing 5% (wt/vol) SDS, 100 mM EDTA, proteinase K (100 μg/ml), and yeast tRNA (100 μg/ml). After a 10-min incubation at 37°C, the reaction products were extracted with phenol equilibrated with 0.1 M sodium acetate (pH 5.2), and the RNA was precipitated with ethanol. Samples were dissolved in denaturing buffer (8 M urea/1 mM EDTA), heated 3 min at 80°C, and loaded on denaturing 6% polyacrylamide/7 M urea gels.

**Products Analysis.** Reaction products were isolated from slices of the denaturing gel by elution in 0.3 M sodium acetate/0.1% SDS/1 mM EDTA. The eluted RNA was phe-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: 5′SS, 5′-splice site; snRNA, small nuclear RNA.

nol-extracted and ethanol-precipitated. The isolated products were used as templates in primer-extension reactions with either one of the following DNA oligonucleotides: ActR2 (5′-ACATACCAGAACCGTTATCA-3′) or ActR3 (5′-CT-AAACATATAATATAGCAAC-3′). The Moloney murine leukemia virus reverse transcriptase (Superscript, from BRL) was used according to manufacturer's instructions. The spliced exon product was amplified by PCR after the reverse transcription, using the oligonucleotides ActR and ActF (5′-CGCGCGGGATCCGGGATT-3′) containing a *Bam*HI site (underlined). Thirty-five amplification cycles were performed: denaturation for 1 min at 90°C, annealing for 1 min at 45°C, and elongation for 1 min at 72°C. Annealing conditions were changed during the course of the amplification: in the first 3 cycles, annealing was at 30°C for 1 min and for cycles 4–10, annealing was at 35°C for 1 min. The PCR product was digested with the restriction enzymes *Eco*RI and *Bam*HI and ligated to a pUC19 plasmid vector. Five isolates were sequenced by using the ActR2 primer.

## RESULTS

For the trans-splicing reaction, we used two RNA substrates. RNA1, the first molecule, was synthesized chemically and consists of 10 nucleotides of exon sequence and 10 nucleotides of intron sequence containing the consensus sequence for the 5′SS. RNA2, the second molecule, is a fragment of the actin pre-mRNA containing the branch site, the 3′SS, and the downstream exon (Fig. 1*A*). In the experiment shown in Fig. 1*B*, internally labeled RNA2 was incubated in yeast crude extract, under different conditions, with an excess of unlabeled RNA1. Several products appeared in a time-dependent manner. Product formation required ATP and $Mg^{2+}$ (Fig. 1*B*, lanes 5 and 6). The identity of the first nucleotide of the intron is also an important requirement: when, as 5′SS, RNA1/int was used, which contains a G → A mutation at the first intron nucleotide, product formation was abolished (Fig. 1*B*, lane 3). In contrast a double mutation in the exon sequence just upstream of the 5′SS did not interfere with the reaction (Fig. 1*B*, lane 4). These requirements are identical to those found in normal cis-splicing reactions. Extending the 5′ exon up to 30 nucleotides did not improve the reaction efficiency, but shortening the exon to 3 nucleotides failed to support splicing (data not shown).

To assess their identity, the individual products were purified from the denaturing gels and analyzed by reverse transcription using different 5′-end-labeled primers. The product labeled a in Fig. 1*B* has the size expected for the ligated exons and primer extension with an oligonucleotide complementary to exon 2 supported this assignment (Fig. 2). To confirm the identity of product a and examine the accuracy with which the two exons are spliced, product a was purified from the denaturing gel, reverse-transcribed with a primer complementary to exon 2, and amplified in a polymerase chain reaction (PCR) by adding a primer complementary to the first 5 bases of exon 1 present on RNA1. The PCR product was cloned and five clones were sequenced and found to be identical. Sequencing across the splice junction shows that the two exons are spliced accurately (Fig. 3).

Based on electrophoretic mobility, product b is slightly larger than the linear intron sequence of RNA2, while products c and d are slightly smaller. In primer-extension experiments, RNA b gives the product predicted for the use of the normal branch site (Fig. 2). Although less intense, a second band, 1 nucleotide longer, is also visible. These results indicate that product b is the branched intron generated in the splicing reaction and that branch site selection on RNA2 appears to be accurate. Also consistent with this conclusion is the lack of detectable extension products from a primer complementary to exon 2 (Fig. 2); furthermore, product b was not detected

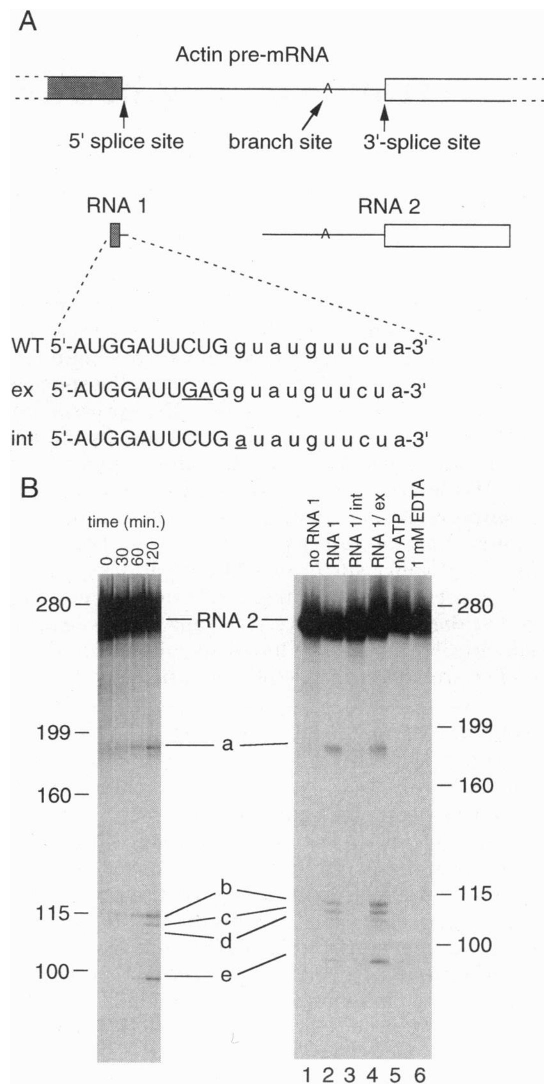


FIG. 1. *In vitro* trans-splicing. (*A*) The trans-splicing substrates are schematically shown in comparison with the actin pre-mRNA. Also shown are the sequences of RNA1, containing the wild-type 5′SS (WT) and two mutant versions (ex and int); the exon sequence is shown in uppercase type. (*B Left*) Internally labeled RNA2 was incubated for the indicated time in the presence of 5 μM RNA1; reaction products are indicated a–e. (*B Right*) Reactions were incubated for 2 hours as above (lane 2). All other reactions were as described in *Methods* but with the following variations. Lanes: 1, RNA1 was omitted from the reaction; 3, mutant 5′SS containing the G1A mutation (oligoribonucleotide RNA1/int, 5 μM); 4, mutant RNA1 containing two mutations in the 5′ exon sequence (oligoribonucleotide RNA1/ex, 5 μM); 5, no ATP; 6, $MgCl_2$ was omitted and 1 mM EDTA was added to the reaction mixture. Numbers on the side of the gels correspond to the mobility of size markers in bases.

when RNA2 was 3′-end-labeled but was detected when it was 5′-end-labeled (data not shown). Primer extension of RNA c gives the same products observed in the reaction with RNA b. Based both on the kinetics of formation of products c, d, and e and the results obtained with end-labeled substrates (data not shown), we believe RNAs c, d, and e represent degradation products of the branched intron.

We have investigated the dependence of the reaction rate on the RNA1 concentration. At the highest RNA1 concentrations (3–7 μM), formation of the spliced exon products reaches a plateau after approximately 2 hours (Fig. 4). The maximum amount of spliced product corresponds to about 6.5% of the labeled RNA2 substrate. The rate vs. time relation appears to be linear up to about 90 min of incubation. Therefore, product

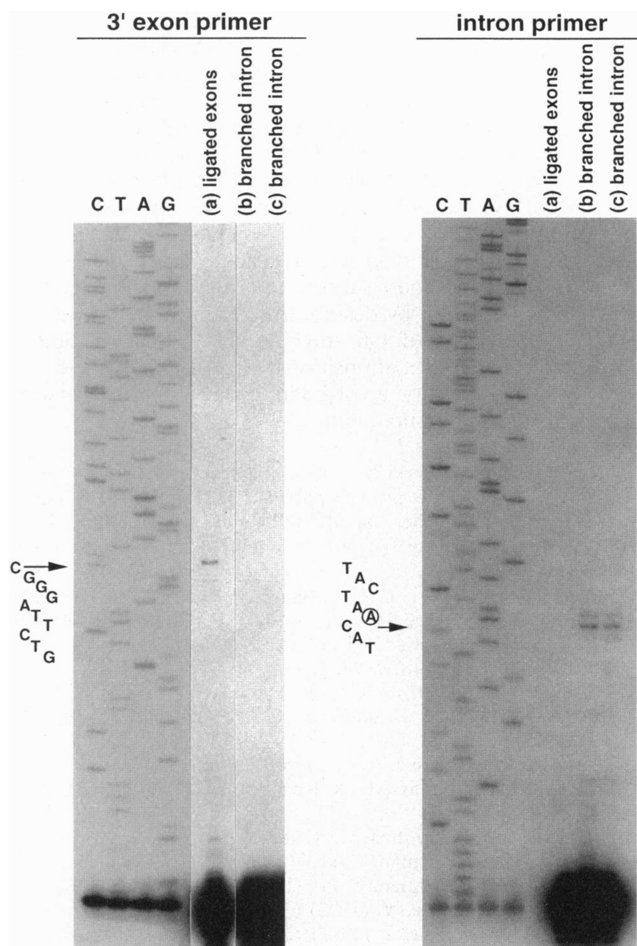


FIG. 2. Identification of the trans-splicing products by primer extension. Two primers were used: 3′ exon primer (oligonucleotide ActR2) and intron primer (oligonucleotide ActR3). The trans-splicing products used as template for the primer-extension reactions are indicated. The same primers were used to generate the sequencing ladders. (*Left*) Sequence of cloned spliced exons is shown. (*Right*) Sequence of cloned RNA2 cDNA is shown. The main primer-extension products are indicated by arrows. Circled base (*Right*) is the canonical branch site used during RNA splicing of the actin pre-mRNA.

accumulation over the first 90 min was used to estimate the reaction rates at different RNA1 concentrations. By plotting

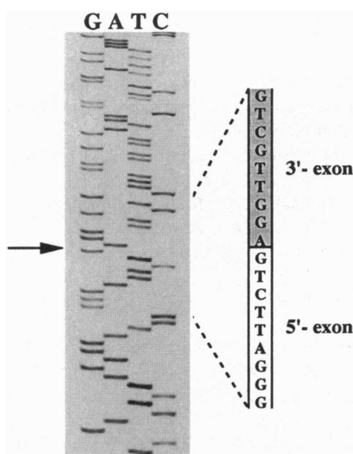


FIG. 3. Splice junction sequence. The spliced exons were reverse-transcribed and the resulting cDNA was amplified by PCR, cloned, and sequenced. The splice junction is indicated by the arrow.

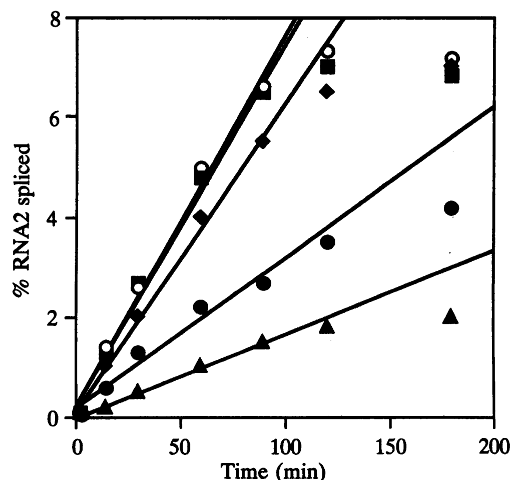


FIG. 4. Trans-splicing reaction rate. Trans-splicing reaction efficiency at various RNA1 concentrations as function of time was estimated by resolving the reaction products on denaturing gels and analyzing the gel by PhosphorImager (Molecular Dynamics). RNA1 concentrations used are the following: ▲, 0.2 μM; ●, 1 μM; ◆, 3 μM; ■, 5 μM; ○, 7 μM. The lines represent linear fitting of the first 90 min of reaction.

the results in an Eadie–Hofstee plot, we estimated the $K_m$ value for RNA1 to be ≈1.6 μM. The purpose of this estimate is not a detailed analysis of the kinetics parameters of the reaction but the definition of a range of RNA1 concentrations that are below saturation (see also *Discussion*).

The effect of multiple and single 2′-deoxy substitutions at the 5′SS was then investigated. In these experiments, the amount of RNA1 used was 1.5 μM, slightly below the estimated $K_m$; under these conditions, the assay is expected to be more sensitive to the effect of the deoxynucleotide substitutions. All of the modified substrates were still converted into products (Fig. 5). 2′-Deoxy substitution at all the nucleotides in exon 1 reduces the efficiency of the reaction by about 70%. Similar results were observed 2′-deoxy substitution at all the nucleotides in the intron sequence at the 5′SS. Single substitutions from positions −3 to +10 had smaller effects. A 2-fold reduction in trans-splicing efficiency was observed for 2′-deoxy

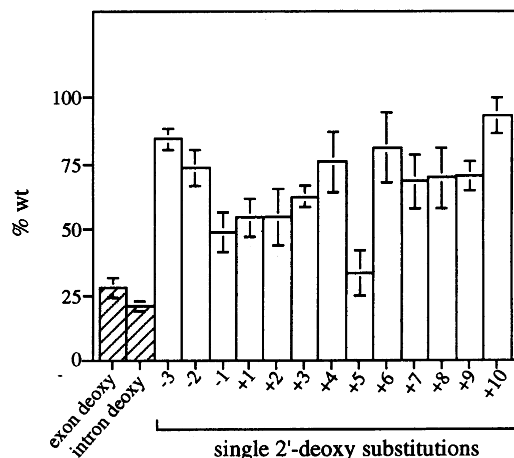


FIG. 5. Effect of 2′-deoxy substitution at the 5′SS. The effect on trans-splicing of 2′-deoxy substitutions in RNA1 was evaluated in reactions containing 1.5 μM RNA1. Reactions were incubated 2 hours; products were resolved on denaturing polyacrylamide gels and quantitated by PhosphorImager (Molecular Dynamics) analysis. Results are shown after normalization to the splicing efficiency of the unmodified RNA1 (wt). Each bar represents the average of the results obtained in three experiments.

substitution at positions −1, +1, and +2. The only position where a single 2′-deoxy substitution has an appreciable effect is position +5, where reaction efficiency was reduced almost 65%.

## DISCUSSION

Trans-splicing has been demonstrated in trypanosomes (for review, see ref. 25), in nematodes (26), and in plant organelles (27). Mammalian extracts have also been reported to catalyze trans-splicing reactions provided a complementary sequence is present on the two substrates forming the 5′ and 3′ splice sites (19, 20), an enhancer sequence is present in the molecule containing the 3′ splice site (22, 23), or the molecule containing the 3′ splice site also has a 5′SS downstream of the exon sequence (23). In addition, it has recently been shown that trans-splicing can be detected in cultured mammalian cells (28). We have investigated the ability of yeast crude extracts to perform trans-splicing by using as a 5′SS substrate, a small RNA oligonucleotide. Short RNA oligonucleotides have been shown to form specific interactions with spliceosomal components in HeLa extract (16–18). Although RNA oligonucleotides of the same size (11 nucleotides) used by Konarska and colleagues (19) did not appear to be substrate for the trans-splicing reaction in our system, slightly longer RNA molecules could be trans-spliced in yeast extracts. The data reported in the present work represent an example of the *in vitro* trans-splicing reaction in *S. cerevisiae*. The small size of the substrates makes this an attractive system for studies in which chemically modified RNAs are used as splicing substrates to probe RNA–RNA and RNA–protein interactions at the splice sites.

Compared to the *in vitro* cis-splicing reaction, which usually reaches a plateau of 50–80% spliced precursor in approximately 40 min at 16°C, trans-splicing appears to be slower and relatively inefficient, reaching a plateau of 6–8% spliced precursor in about 2 hours. The reaction, however, is accurate in terms of branch site selection and exon splice junction. In contrast with studies performed in mammalian cells, our RNA substrates do not contain any obvious complementarity nor do they contain known enhancer sequences.

We have estimated the $K_m$ value for the RNA1 substrate by titrating the RNA1 and measuring the change in reaction rate. The crude estimate obtained ($K_m \approx 1.6$ μM) was a useful reference for the experiments in which the effect of modifications in RNA1 was studied, since saturating amounts of RNA1 may lead to underestimation of the effect of modifications. We tested the effect of 2′-deoxy substitutions around the 5′SS. By analogy with what has been observed in ribozymes, the 2′-OH group could play an important role by activating the neighboring 3′-oxygen leaving group at the 5′SS (29), by coordinating metal cations (30), or by participating in RNA–RNA tertiary interactions (31, 32). From our results, at all of the positions we have tested, the 2′-OH group is not an absolute requirement: all the multiple and single 2′-deoxy substitution were compatible with trans-splicing. However, small but reproducible effects were observed at some positions. Single 2′-deoxy substitution at positions −1 to +2 reduced the reaction efficiency by about 2-fold and a 3-fold reduction is observed by removing the 2′-OH at position +5. The effect of 2′-deoxy substitution at position −1 has previously been investigated in cis-splicing reactions by Moore and Sharp (33), who reported absolutely no effect on *in vitro* splicing for the 2′-deoxy substitution at position −1. The slightly different result we have obtained is likely a consequence of the differences in the experimental system used: our experiments were performed under subsaturating 5′SS concentration that probably makes the assay more sensitive compared to the cis-splicing reaction used by Moore and Sharp (33). It is also possible that the rate limiting step in yeast trans-splicing may be not the same as in mammalian cis-splicing. The small effect we have observed for the 2′-OH group removal at position −1 supports the conclusion by Moore and Sharp (33) that this group is unlikely to play the crucial role in catalysis proposed for the 2′-OH group at −1 in the self splicing Tetrahymena ribozyme (29). The small effects (2- to 3-fold) we have observed for the 2′-deoxy substitutions at positions −1, +1, +2, +3, and +5 have the magnitude expected for duplex destabilization by single deoxy substitution (34); if this is the case these substitutions may interfere with the trans-splicing reaction by destabilizing the duplex formed by the 5′SS with U1 and U6 snRNAs. However, we cannot completely rule out that at positions −1, +1, +2, +3, and +5, the 2′-OH group may be involved in tertiary interactions with other spliceosomal components.

We are grateful to Maria Konarska and Robin Reed for sharing their trans-splicing results prior to publication. We also thank James Bruzik, Christine Guthrie, David McPheeters, Anna Marie Pyle, Christian Siebel, and John Wagner for comments on the manuscript.

1. Moore, M. J., Query, C. C. & Sharp, P. A. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 303–358.
2. Nilsen, T. W. (1994) *Cell* **78**, 1–4.
3. Madhani, H. & Guthrie, C. (1994) *Annu. Rev. Genet.* **28**, 1–26.
4. Horowitz, D. S. & Krainer, A. R. (1994) *Trends Genet.* **10**, 100–106.
5. Zhuang, Y. & Weiner, A. M. (1986) *Cell* **46**, 827–835.
6. Seraphin, B., Kretzner, L. & Rosbash, M. (1988) *EMBO J.* **7**, 2533–2538.
7. Siliciano, P. G. & Guthrie, C. (1988) *Genes Dev.* **2**, 1258–1267.
8. Newman, A. & Norman, C. (1991) *Cell* **65**, 827–835.
9. Newman, A. & Norman, C. (1992) *Cell* **68**, 743–754.
10. Sawa, H. & Shimura, Y. (1992) *Genes Dev.* **6**, 244–254.
11. Sawa, H. & Abelson, J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11269–11273.
12. Wassarman, D. A. & Steitz, J. A. (1992) *Science* **257**, 1918–1925.
13. Kandels-Lewis, S. & Seraphin, B. (1993) *Science* **262**, 2035–2039.
14. Lesser, C. F. & Guthrie, C. (1993) *Science* **262**, 1982–1988.
15. Sontheimer, E. J. & Steitz, J. A. (1993) *Science* **262**, 1989–1996.
16. Hall, K. B. & Konarska, M. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10969–10973.
17. Konforti, B. B., Koziolkiewicz, M. J. & Konarska, M. M. (1993) *Cell* **75**, 863–873.
18. Konforti, B. B. & Konarska, M. M. (1994) *Genes Dev.* **8**, 1962–1973.
19. Konarska, M. M., Padgett, R. A. & Sharp, P. A. (1985) *Cell* **42**, 165–171.
20. Solnick, D. (1985) *Cell* **42**, 157–164.
21. Bruzik, J. P. & Maniatis, T. (1992) *Nature (London)* **360**, 692–695.
22. Bruzik, J. P. & Maniatis, T. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 7056–7059.
23. Chiara, M. D. & Reed, R. (1995) *Nature (London)* **375**, 510–513.
24. Lin, R. J., Newman, A. J., Cheng, S. C. & Abelson, J. (1985) *J. Biol. Chem.* **260**, 4780–4792.
25. Ullu, E., Matthews, K. R. & Tschudi, C. (1993) *Mol. Cell. Biol.* **13**, 720–725.
26. Nilsen, T. W. (1993) *Annu. Rev. Microbiol.* **47**, 413–440.
27. Sharp, P. A. (1991) *Science* **254**, 663.
28. Eul, J., Graessmann, M. & Graessmann, A. (1995) *EMBO J.* **14**, 3226–3235.
29. Herschlag, D. & Cech, T. R. (1990) *Nature (London)* **344**, 405–409.
30. Mei, H.-Y., Kaaret, T. W. & Bruice, T. C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9727–9730.
31. Pyle, A. M. & Cech, T. R. (1991) *Nature (London)* **350**, 628–631.
32. Chastain, M. & Tinoco, I. (1992) *Biochemistry* **31**, 12733–12741.
33. Moore, M. J. & Sharp, P. A. (1992) *Science* **256**, 992–993.
34. Bevilacqua, P. C. & Turner, D. H. (1991) *Biochemistry* **30**, 10632–10640.