

In vitro selection of an archaeal RNase P RNA mimics natural variation

DANIEL WILLIAMS^{1,2} and JAMES W. BROWN^{1,3}

¹ Department of Microbiology, North Carolina State University, Raleigh, NC 27695-7615, USA

² Present address: V.A. Medical Center, Emory University Medical Research, Rm 5A188, 1670 Clarimont Rd., Decatur, GA 30033, USA

³ Author to whom correspondence should be addressed (james_brown@ncsu.edu)

Received September 17, 2003; accepted October 15, 2003; published online November 5, 2003

Summary Archaeal and bacterial RNase P RNAs are similar in sequence and secondary structure, but in the absence of protein, the archaeal RNAs are much less active and require extreme ionic conditions for activity. To assess how readily the activity of the archaeal RNA alone could be improved by small changes in sequence, in vitro selection was used to generate variants of a *Methanobacterium formicicum* RNase P RNA: *Bacillus subtilis* pre-tRNA^{Asp} self-cleaving conjugate RNA. Functional variants were generated with a spectrum of mutations that were predominately consistent with natural variation in this RNA. Variants generated from the selection had cleavage rates comparable to that of wild type; variants with improved cleavage rates or lower ionic requirements were not obtained. This suggests that the RNase P RNAs of Bacteria and Archaea are globally optimized and the basis for the large biochemical differences between these two types of RNase P RNA is distributed in the molecule.

Keywords: Archaea, archaeobacteria, ribonuclease P, ribozyme, SELEX.

Introduction

RNase P RNA is a ribonuclease responsible for the 5' maturation of pre-tRNAs in all cells and subcellular organelles that carry out tRNA synthesis (for review, see Altman 1990, Pace and Smith 1990, Darr et al. 1992). RNase P RNAs identified in each evolutionary domain of life are "ribozymes"; the catalytic subunit of the ribonucleoprotein enzyme is the RNA, not the protein. The RNA subunit of the bacterial RNase P is capable of efficient catalysis in the absence of protein (Guerrier-Takada et al. 1983), whereas those of Eucarya are absolutely dependent on protein for function (Krupp et al. 1986, Lee and Engelke 1989, Arends and Schön 1997, Han et al. 1998). Archaeal RNase P RNAs were initially thought to be dependent on protein for catalytic activity in vitro (Nieuwlandt et al. 1991, LaGrandeur et al. 1993), but the RNase P RNAs of some archaea have recently been shown to be active at low levels, under extreme ionic conditions (3–4 M ammonium acetate, 300–400 mM MgCl₂) (Pannucci et al. 1999).

Bacterial and archaeal RNase P RNA sequences and secondary structures are similar, and most secondary structures fall into the ancestral "type A" class (Harris et al. 2001). These type A archaeal RNase P RNAs contain all of the sequences and structures necessary for substrate recognition and catalysis; their poor catalytic competence is apparently a consequence of inadequate global stabilization (Pannucci et al. 1999, Harris et al. 2001). The most obvious consistent structural difference between bacterial and archaeal type A RNase P RNAs is the presence of helix P18 in Bacteria (absent only in the green sulfur bacteria and altered in other green non-sulfur bacteria and in type B RNAs; Haas et al. 1994) and its uniform absence in the Archaea. The removal of P18 from the *Escherichia coli* RNA results in an increase in the optimum monovalent salt requirement from 1 M to 3 M (ammonium acetate), consistent with the ionic requirement of the archaeal RNA activity in vitro (Haas et al. 1994). However, the affinity for substrate ($1/K_m$) of the mutant bacterial RNA remains at least 1000-fold higher than the most active archaeal RNA, that of *Methanothermobacter thermoautotrophicus* (formerly *Methanobacterium thermoautotrophicum* strain ΔH). The insertion of P18 into the RNase P RNA of *M. thermoautotrophicus*, with or without appropriate sequence changes in helix P8 to maintain the potential for interaction between helices P8 and P18, does not improve catalytic activity (Harris et al. 2001). The differences in activity between bacterial and archaeal RNase P RNAs would seem, then, not to reside in the differences between their core secondary structures, but in more subtle differences between their sequences or tertiary structures.

Catalytic activities of several group I and group II introns that are similarly dependent on protein for efficient cleavage have been dramatically improved by small sequence changes generated during in vitro selection experiments (Chapman and Szostak 1994, Joyce 1994, Kumar and Ellington 1995, Tuschel et al. 2000). To assess how readily the activity of the archaeal RNase P (in the absence of protein) could be improved, we subjected a self-cleaving archaeal RNase P RNA: pre-tRNA conjugate to in vitro selection using sequential rounds of mutagenic PCR and selection for the most rapidly self-cleaving RNAs.

Materials and methods

Selection

The *Methanobacterium formicicum* RNase P RNA:*Bacillus subtilis* pre-tRNA^{Asp} RNA (cpTP RNA) is a circularly permuted enzyme:substrate RNA (Figure 1); this self-cleaving conjugate has been described previously (Pannucci et al. 1999). The selection scheme used to generate cpTP RNA variants is based on the method of Frank et al. (1996). Briefly, the cpTP RNA was transcribed in vitro in the presence of guanosine monophosphorothioate (GMPS) (see below) and covalently linked to iodoacetyl-agarose beads (SulfoLink gel, Pierce). The beads were then washed four times with 16 bed volumes of buffer (3 M ammonium acetate, 50 mM Tris, pH 8, 5 mM EDTA) at room temperature, then once in the same buffer containing 1–3 M ammonium acetate, as specified, at 50 °C. The beads were resuspended in reaction buffer (1–3 M ammonium acetate, 100–300 mM MgCl₂, 50 mM Tris, pH 8) at 50 °C, as indicated in Table 1, to initiate self-cleavage. The RNAs released from the column by self-cleavage were removed by washing with a reaction buffer at 1 h intervals. Fractions containing the first 5–10% of the eluted RNAs were pooled and used in the next round of the selection.

RNA synthesis and purification

Uniformly labeled cpTP RNAs were synthesized in vitro using T7 RNA polymerase (Promega) in the presence of [α -³²P]GTP.

In vitro transcription reactions were modified from Frank et al. (1996) as follows: reactions were incubated at 37 °C overnight in the absence of glycerol and in the presence of 87 mM GMPS (4:1 ratio of GMPS/GTP). Guanosine monophosphorothioate was a gift from Daniel Frank and Norman Pace (University of Colorado). Transcripts were purified by gel electrophoresis in 7.5 M urea/6% polyacrylamide gel with TBE buffer (Sequa-Gel, National Diagnostics, Atlanta, GA; 90 mM Tris, 64.6 mM boric acid, 2.5 mM EDTA, pH 8.3). Full-length transcripts were excised and eluted by diffusion in 10 volumes of TE buffer containing 0.5% SDS at 4 °C overnight. Eluted RNAs were ethanol precipitated, resuspended in 10 mM Tris-HCl, pH 7.5, and quantitated by scintillation spectrophotometry.

Error-prone PCR and reverse transcription

Variant sequences of cpTP RNA were generated by error-prone PCR according to Dieffenbach and Dveksler (1999); this procedure has an overall error rate of approximately 7×10^{-3} errors per nucleotide without apparent sequence bias. Template DNA for cpTP RNA was amplified from plasmid DNA using Taq polymerase (Gibco BRL) and the primers T75'pre (TAATACGACTCACTATAGGGCGAATTGGAGC TCCACCGGTCCGGTAGTTCAG) and Mfo260R (CCCAA GCTTCTGCCTCATA CAGGATTC) under error-prone PCR conditions (Dieffenbach and Dveksler 1999), and gel purified. After the first round of selection, the RNA eluted from beads was subjected to reverse transcription as described by Frank et

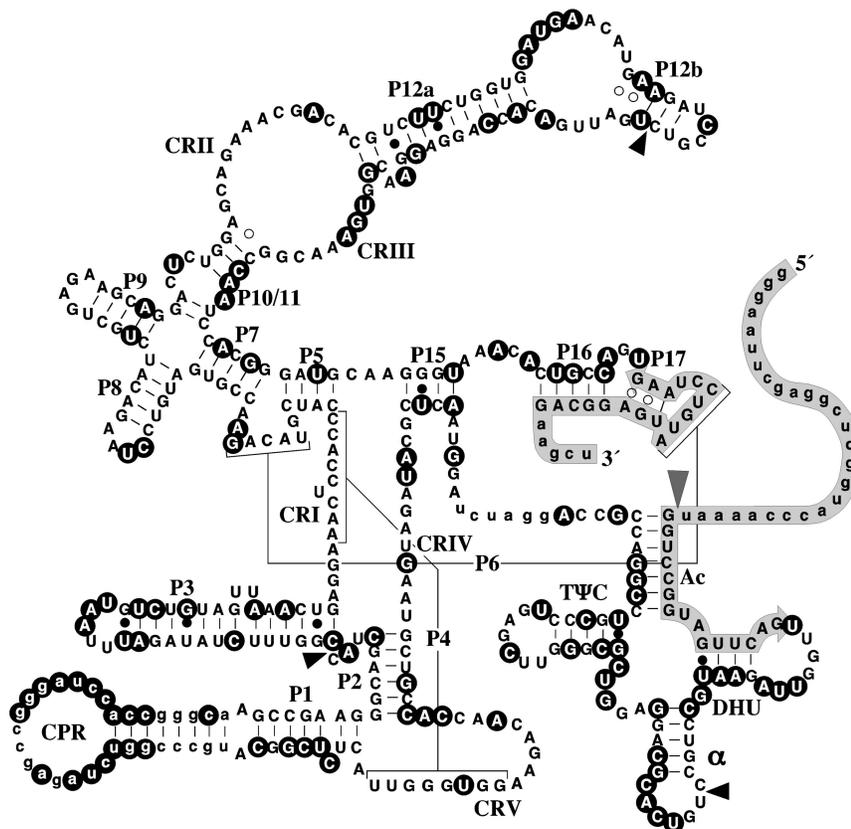


Figure 1. Secondary structure of the circularly permuted *Methanobacterium formicicum* RNase P RNA:*Bacillus subtilis* pre-tRNA^{Asp} RNA. Sequences at the 5' and 3' ends of the RNA enclosed by a gray box are the locations of the oligonucleotide primers used in each round of selection. Lowercase nucleotides represent non-native sequences: the sequence joining the tRNA and RNase P RNA; the sequence joining the native 5' and 3' ends of the RNase P RNA (CPR, distal to P1); and the 5' leader sequence. Helices in the RNase P RNA are designated P1 to P17 according to Haas et al. (1996). Conserved regions in the RNase P RNA are designated CR I to CR V according to Chen and Pace (1997). Transfer RNA helices are labeled Ac (acceptor stem), DHU (dihydrouracil stem), α (anticodon stem) and T Ψ C (T-pseudouracil-C stem). The RNase P cleavage site is indicated by the gray arrow. Nucleotides highlighted in black are in positions at which mutations (transition, transversion, or deletion) were observed among the 50 sequenced active RNAs. Sites of insertions among these sequences are indicated by the black triangles.

Table 1. Selection conditions.

Round	Ammonium acetate concentration (M)	MgCl ₂ concentration (mM)	Reaction time (h)	Active clones sequenced
1	3	300	2	0
2	3	300	4	3
3	3	300	2	3
4	2	200	2	1
5	2	200	2	5
6	2/2*	200/200	2	35
7	1/2	100/200	2	3
8	1/2	100/200	2	0
9	1/2	100/200	2	0
10	1/2	100/200	2	0
11	1/2	100/200	2	0

* Rounds 6–11 were performed twice.

al. (1996) using the reverse primer Mfo260R, and used directly as template in error-prone PCR reactions.

Cloning and sequencing

Reverse transcription-PCR products from each round were cloned into pGemT (Promega) following the manufacturer's protocol. Clones were screened by amplification of an approximately 446 bp product using primers T75'pre and Mfo260R under non-mutagenic PCR conditions (94 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min, 30 amplification cycles) from plasmid DNA minipreps. The RNA transcribed from each clone using T7 RNA polymerase was assayed for self-cleaving activity (see below). A total of 50 RNA variants with activity within twofold of that of the wild-type RNA were sequenced (Genbank accession numbers AF443948 to AF443998). Sequences were aligned manually using Bioedit (Version 5.0.7; <http://www.mbio.ncsu.edu/Bioedit/>); this alignment is available at the RNase P Database (<http://www.mbio.ncsu.edu/RNaseP/>).

Activity assay

We synthesized RNA in the absence of GMPS by in vitro transcription from individual clones and eluted in 50–100 µl distilled water at 4 °C overnight. Eluted RNA (1–2 µl) was assayed for RNase P activity in 1–3 M ammonium acetate, 100–300 mM MgCl₂, 50 mM Tris (pH 8) at 50 °C for 6–8 h. Reaction products were separated by electrophoresis in a 7.5 M urea/12% polyacrylamide gel with TBE buffer and visualized by phosphorimager (Molecular Dynamics, Sunnyvale, CA).

Results

The in vitro selection scheme was designed to assess how readily the archaeal RNase P RNA would be "improved," either by cleaving faster or at lower ionic strength, by small changes in sequence. Random mutations were introduced into an initially completely specified sequence by mutagenic PCR in each round of the selection, and those RNAs that were released most rapidly from the column matrix by self-cleavage

after the addition of MgCl₂ were collected. Increases in the stringency of each round were based on the rate of RNA release from the column (see Table 1). The RNAs from individual clones from each round were assayed for catalytic activity and active clones were sequenced (Figure 2). All active RNAs had activities comparable to that of wild-type RNA: none self-cleaved as much as twofold faster than wild-type RNA at high (3 M) or low (1 M) salt concentration (data not shown). The RNA from all sequenced clones cleaved at least half as fast as wild-type RNA in 3 M ammonium acetate/300 mM MgCl₂. Although the majority of the active clones sequenced were from Round 6 of the mutagenic PCR (in order to maximize the extent of sequence variation observed), at least one representative was sequenced from each of the earlier rounds except for Round 1 (Table 1). Rounds 8–11 generated no active clones. Sequenced DNA from these pools contained no recognizable remnant of the starting sequence, whereas earlier pool sequences were clearly related to the initial sequence (data not shown). The nucleotide sequences of several of these non-active clones were determined; all contained variations of a sequence that was unrelated to the starting sequence or to any sequence available in GenBank. The origin of this sequence is unknown, nor is it clear how the selection process enriches for this sequence over the cpTP RNA variants. To determine if the unrelated RNA sequences were derived from the starting sequence or are the result of an artifact generated during the selection, we repeated the selection again starting from Round 6 without increasing the stringency of the elution, and once again we were unable to recover the starting sequence or any self-cleaving RNAs after the eighth round.

Discussion

A spectrum of sequence variation consistent with wild-type levels of catalytic activity was identified in the 50 clones analyzed. Base transitions (115 mutations) and transversions (73 mutations) were the most common mutations observed, but deletions (30 mutations) and insertions (four mutations) were also represented, averaging 4.44 mutations per molecule. Although the mutations were widely scattered, most occurred pe-

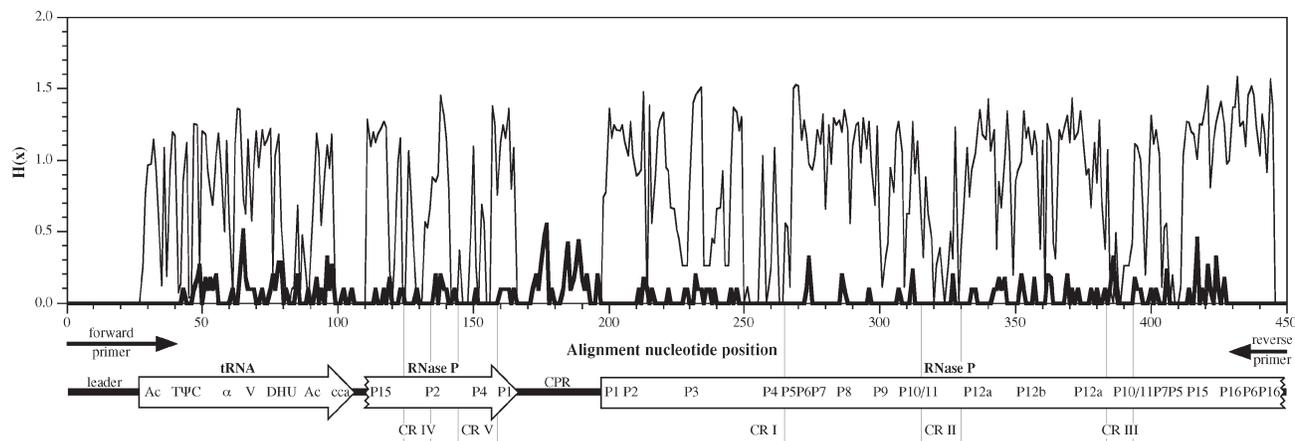


Figure 2. Sequence variation in the selected RNAs compared with naturally occurring variation. The linear sequence of the cpTP RNA, shown here diagrammatically and labeled as in Figure 1, is plotted against sequence variation expressed as $H(x)$; H is an entropy coefficient for each position x in which $H = -\sum_b f_b \ln b$ where b is in set (A, G, U, C, -) expressed in bits of information (Chiu and Kolodziejczak 1991). The coefficient H varies from zero (absolutely conserved) to 1.61 (completely randomized). Observed variation in the 50 sequenced active selection products is shown as a heavy line; naturally occurring variation among the 40 available archaeal RNase P RNAs in the RNase P Database (<http://jwbrown.mbio.ncsu.edu/RNaseP/>) and 155 archaeal tRNAs in the tRNA Database (<http://www.bayreuth.de/departments/biochemie/sprinzl/index.html>) is shown as a light line.

ripheral to the catalytic core of the *M. formicicum* RNase P RNA secondary structure, and this sequence variation is consistent with naturally occurring sequence variation in this RNA (Figure 1). Conserved regions (CR) I, II, IV and V, as well as the cruciform composed of helices P8 to P10, were especially conserved. Because CRI, CRIV and CRV are central to the catalytic site (Chen and Pace 1997), and the cruciform is apparently involved in recognition of the substrate T loop (20), it is not surprising that these regions were resistant to variation in this selection. The mutations that occurred in CRI, CRIV and CRV were in the most variable nucleotides within these conserved sequences, and without exception are substitutions known to occur in nature. Conserved region III is a highly conserved region in RNase P RNA among all three domains of life (Chen and Pace 1997), yet it was a region of relatively high sequence variation among the selected RNAs (see Figure 2). However, CRIII is probably involved in protein binding in bacteria (Loria et al. 1998), and so probably has a similar function in the archaeal RNA. Because the selection was carried out in the absence of protein, the sequence of CRIII was apparently less constrained. Although it seems likely that CRII and CRIII are related in structure and function, CRII (unlike CRIII) was as highly conserved as CRI, CRIV and CRV.

The RNA sequence region with the highest mutation density was the artificial sequence connecting the natural 5' and 3' ends of the RNase P RNA (circularly permuted region; CPR). Given that the CPR is not a native part of the RNA, it is not surprising that sequence variation here is readily tolerated.

The sequences adjacent to the highly conserved J15/16 internal loop of RNase P RNA (i.e., between helices 15 and 16 and the 3' primer sequence) were relatively more variable than in nature. It might be expected that the covalent linkage of the 3' CCA trinucleotide of the pre-tRNA to its binding site in J15/16 would relax the constraints on these sequences because

the interaction that requires these sequences is pre-enforced. On the other hand, the CC of the NCCA tail of the tRNA was unchanged in all of the RNAs, as was the 6-nucleotide artificial sequence that links the substrate and the RNase P RNA.

Mutations in the tRNA sequences were more evenly distributed than mutations in the RNase P RNA sequence, but were maximal in the variable loop, consistent with naturally occurring sequence variation (refer to Figure 2).

Although we have generated a large number of functional sequence variants by *in vitro* selection, none of them has a noticeably improved catalytic rate or ionic strength requirement. Instead, the observed sequence variation is consistent with naturally observed sequence variation except where the idiosyncrasies of the selection might be expected to relax the constraints on variation. This suggests that the RNA as a whole is optimized for its function, and that small changes are unlikely to lead to substantial improvement. This is consistent with the results of a related experiment by Frank et al. (1996) in which the two most highly conserved regions of the *E. coli* RNase P RNA—CRI and CRV—were entirely randomized. After 10 rounds of selection, the only remaining sequence detected was the original wild-type *E. coli* sequence (Frank et al. 1996); even naturally occurring sequence variations were not observed in these regions. We conclude that the RNase P RNAs of Bacteria and Archaea are globally optimized and that the basis for the large biochemical differences between these two types of RNase P RNA is distributed in the molecule.

Acknowledgments

We thank Dan Frank and Jim Nolan for materials and advice on optimizing the selection procedures. This work was supported by the NIH grant GM52894 to JWB.

References

- Altman, S. 1990. Ribonuclease P. Postscript. *J. Biol. Chem.* 265: 20,053–20,056.
- Arends, S. and A. Schön. 1997. Partial purification and characterization of the nuclear ribonuclease P from wheat. *Eur. J. Biochem.* 244:635–645.
- Chapman, K.B. and J.W. Szostak. 1994. In vitro selection of catalytic RNAs. *Curr. Opin. Struct. Biol.* 4:618–622.
- Chen, J. and N.R. Pace. 1997. Identification of the universally conserved core of ribonuclease P RNA. *RNA* 3:557–560.
- Chiu, D.K. and T. Kolodziejczak. 1991. Inferring consensus structure from nucleic acid sequences. *Comput. Appl. Biosci.* 7:347–352.
- Darr, S.C., J.W. Brown and N.R. Pace. 1992. The varieties of ribonuclease P. *Trends Biochem. Sci.* 17:178–182.
- Dieffenbach, C.W. and G.S. Dveksler. 1999. Mutagenesis by PCR. PCR primer laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 581–589.
- Frank, D.N., A.E. Ellington and N.R. Pace. 1996. In vitro selection of RNase P RNA reveals optimized catalytic activity in a highly conserved structural domain. *RNA* 2:1179–1188.
- Guerrier-Takada, C., K. Gardiner, T. Marsh., N.R. Pace and S. Altman. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35:849–857.
- Haas, E.S., D.W. Armbruster, B.M. Vucson, C.J. Daniels and J.W. Brown. 1996. Comparative analysis of ribonuclease P RNA structure in Archaea. *Nucleic Acids Res.* 24:1252–1259.
- Haas, E.S., J.W. Brown, C. Pitulle and N.R. Pace. 1994. Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc. Natl. Acad. Sci. USA* 91:2527–2531.
- Han, S.J., B.J. Lee and H.S. Kang. 1998. Purification and characterization of nuclear ribonuclease P of *Aspergillus nidulans*. *Eur. J. Biochem.* 251:244–251.
- Harris, J.K., E.S. Haas, D. Williams, D.N. Frank and J.W. Brown. 2001. New insights into RNase P RNA structure from comparative analysis of the archaeal RNA. *RNA* 7:220–232.
- Joyce, G.F. 1994. In vitro evolution of nucleic acids. *Curr. Opin. Struct. Biol.* 4:331–336.
- Krupp, G., B. Cherayil, D. Frendewey, S. Nishikawa and D. Söll. 1986. Two RNA species co-purify with RNase P from the fission yeast *Schizosaccharomyces pombe*. *EMBO J.* 5:1697–1703.
- Kumar, P.K.R. and A.E. Ellington. 1995. Artificial evolution and natural ribozymes. *FASEB J.* 9:1183–1195.
- LaGrandeur, T.E., S.C. Darr, E.S. Haas and N.R. Pace. 1993. Characterization of the RNase P RNA of *Sulfolobus acidocaldarius*. *J. Bacteriol.* 175:5043–5048.
- Lee, J.Y. and D.R. Engelke. 1989. Partial characterization of an RNA component that copurifies with *Schizosaccharomyces cerevisiae* RNase P. *Mol. Cell. Biol.* 9:2536–2543.
- Loria, A., S. Niranjankumari, C.A. Fierke and T. Pan. 1998. Recognition of a pre-tRNA substrate by the *Bacillus subtilis* RNase P holoenzyme. *Biochemistry* 37:15,466–15,473.
- Nieuwlandt, D.T., E.S. Haas and C.J. Daniels. 1991. The RNA component of RNase P from the archaeobacterium *Haloferax volcanii*. *J. Biol. Chem.* 266:5689–5695.
- Pace, N.R. and D. Smith. 1990. Ribonuclease P: function and variation. *J. Biol. Chem.* 265:3587–3590.
- Pannucci, J.A., E.S. Haas, T.A. Hall, J.K. Harris and J.W. Brown. 1999. RNase P RNAs from some Archaea are catalytically active. *Proc. Natl. Acad. Sci. USA* 96:7803–7808.
- Tuschl, T., P.A. Sharp and D.P. Bartel. 2000. A ribozyme selected from variants of U6 snRNA promotes 2',5'-branch formation. *RNA* 7:29–43.