854–858 *Nucleic Acids Research, 1996, Vol. 24, No. 5*

The P9.1–P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme

Patrick P. Zarrinkar and James R. Williamson*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received November 28, 1995; Revised and Accepted January 8, 1996

ABSTRACT

We have previously proposed a hierarchical model for the folding mechanism of the *Tetrahymena* ribozyme that may illustrate general features of the folding pathways of large RNAs. While the role of elements in the conserved catalytic core of this ribozyme during the folding process is beginning to emerge, the participation of non-conserved peripheral extensions in the kinetic folding mechanism has not yet been addressed. We now show that the 3′-terminal P9.1–P9.2 extension of the *Tetrahymena* ribozyme plays an important role during the folding process and appears to guide formation of the catalytic core.

INTRODUCTION

Group I introns catalyze their own excision from pre-RNA in a two step transesterification reaction which results in ligation of the flanking exons and release of the free intron (1). They can be found in the genomes of a wide variety of organisms and are characterized by a set of highly conserved base paired regions which together form a core containing all the structural elements required for catalysis (2,3). Based on phylogenetic comparisons of the core sequences from a large number of introns, a model for the three-dimensional architecture of the catalytic core has been proposed (4,5). The model shows the existence of two helical subdomains connected by a triple helical scaffold, which probably orients the subdomains with respect to each other (6–8). In addition to the catalytic core, group I introns contain a series of less well-conserved peripheral extensions, whose variation defines subgroups of related introns (4). While the extensions are not absolutely required for catalytic activity, they are important for stabilizing the catalytic core to allow formation of the active structure of the RNA at physiological Mg^{2+} concentrations (2,9–13). For RNAs lacking some peripheral extensions their function may have been taken over by proteins that specifically recognize the core structure. There is evidence, for example, that loss of the P5abc extension in the *Tetrahymena* ribozyme can be complemented by the CYT-18 protein (14).

The group I intron from *Tetrahymena* pre-rRNA has been extensively characterized and serves as a model system for studying RNA folding and catalysis (Fig. 1) (1). In addition to the self-splicing reaction, a shortened version of the intron missing 21 nt from the 5′-end (the L-21 *Sca*I *Tetrahymena* ribozyme) can catalyze cleavage of short oligoribonucleotides in *trans* with multiple turnover. The L-21 *Sca*I ribozyme is moderately large (388 nt) and, like most ribozymes, requires Mg^{2+} for formation of its active structure and catalysis. In the presence of Mg^{2+} the catalytic core forms a single globular structure (15,16). While the two subunits of the core have previously been referred to as domains (8,10,17), we suggest that the term subdomain more accurately describes their identity as substructures of a globular domain. The P4–P6 subdomain includes the conserved P4 and P6 stems and is independently stable outside of the context of the whole intron (17). The P3–P7 subdomain includes the helices P3 and P7, which are formed by base pairing between regions relatively far apart in the linear sequence of the RNA.

We have previously proposed a model for the kinetic folding pathway of the *Tetrahymena* ribozyme, including both Mg^{2+}-dependent and Mg^{2+}-independent steps, in which the two main structural subdomains form hierarchically (6,18). The P4–P6 subdomain forms first and the overall rate limiting step is a Mg^{2+}-independent rearrangement preceding stable formation of the P3–P7 subdomain. The P3 and P7 helices form in an interdependent manner and the observed structural subdomains appear to correspond to kinetic folding units. While the rate limiting step in our model may represent a number of microscopic folding events, we have shown that one of these events is the formation of the triple helical scaffold (6). This hierarchical model for folding of the group I ribozyme is consistent with studies of Mg^{2+}-induced folding at equilibrium (10,19).

We now explore the role played by a structural element outside the catalytic core during folding. In the *Tetrahymena* intron the 3′-end is formed by several helices that constitute the P9.1–P9.2 peripheral extension (Fig. 1). A ribozyme truncated at a point corresponding to a *Nhe*I site in the plasmid template (L-21 *Nhe*I) lacks the P9.1–P9.2 extension, but, once folded, has catalytic activity equal to that of full-length L-21 *Sca*I RNA (10). At high Mg^{2+} concentrations at equilibrium, deletion of the extension introduces no apparent structural perturbations in the remainder of the molecule (10). However, an increased concentration of Mg^{2+} is required to form the P3–P7 subdomain compared with full-length RNA. The P9.1–P9.2 extension therefore appears to specifically stabilize the P3–P7 subdomain (10). We now show that, in addition to stabilizing the final structure of the RNA, the 3′-terminal extension also plays a role during the folding process. The possible implications of these results for the folding of other large RNAs containing non-conserved peripheral extensions are discussed.

* To whom correspondence should be addressed

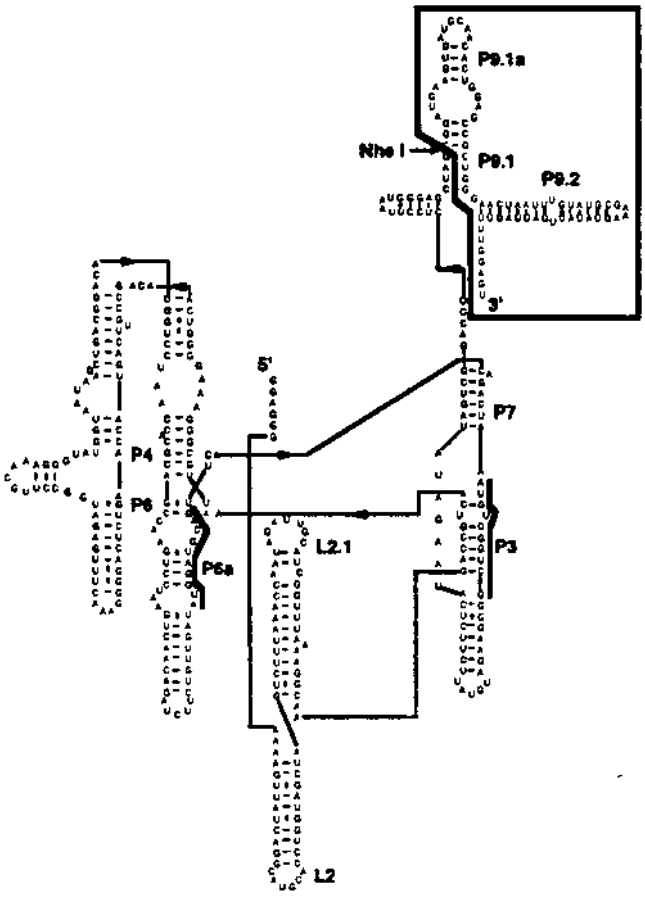

Figure 1. The *Tetrahymena* ribozyme. The 3′-end of the shortened L-21 *Nhe*I RNA is indicated and the P9.1–P9.2 extension deleted in this mutant is outlined by a box. Sequences targeted by the oligonucleotide probes are highlighted by bold lines.

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligodeoxynucleotide probes were synthesized on a 1 µmol scale on an Applied Biosystems DNA synthesizer, deprotected overnight at 65°C in 2 ml concentrated ammonium hydroxide and purified on 20% denaturing polyacrylamide gels. Full-length bands were excised and eluted from the gel overnight into water at 4°C, followed by desalting on C_{18} Sep-Paks® (Waters).

Ribozyme preparation

Full-length L-21 *Sca*I and truncated L-21 *Nhe*I ribozymes were prepared by transcription from plasmid pT7L-21 (20) linearized with *Sca*I or *Nhe*I (New England Biolabs) respectively. Transcription reactions (100 µl) were performed in 40 mM Tris–HCl, pH 7.5, 2 mM spermidine, 10 mM dithiothreitol, 25 mM $MgCl_2$, 1 mM each GTP, CTP and UTP, 0.1 mM ATP and 250 µCi [α-^{32}P]ATP (3000 Ci/mmol; New England Nuclear) for 3.5 h at 37°C using 500 U T7 RNA polymerase (New England Biolabs) and 15 µg linearized plasmid template. Full-length transcription products were purified on 6% denaturing polyacrylamide gels and eluted from the gel overnight at 4°C into buffer containing 10 mM Tris–HCl, pH 7.5, 1 mM EDTA and 0.3 M sodium acetate. The RNA was ethanol precipitated and resuspended in 10 mM Tris–HCl, pH 7.5, 0.1 mM EDTA. Ribozyme concentrations were determined by Cerenkov counting.

Kinetic oligonucleotide hybridization assay

To determine folding rates the oligonucleotide hybridization assay was performed as described (18). RNA (final concentration 1 nM) in 60 µl buffer containing 1 mM Tris–HCl, pH 7.5, 0.01 mM EDTA was annealed by heating to 95°C for 45 s followed by equilibration at 37°C for 3 min. Folding was initiated by addition of an equal volume of 2× folding buffer (1× = 50 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM NaCl, 1 mM dithiothreitol), aliquots (10 µl) were taken at the times indicated and added to 10 µl 1× folding buffer containing oligonucleotide probe and RNase H (final concentration 0.1 U/µl; United States Biochemical). Oligonucleotide binding and RNase H cleavage were allowed to proceed for 30 s before the reaction was quenched with 14 µl stop solution (90 mM EDTA and marker dyes in 82% formamide). The zero time points were obtained by adding oligonucleotide probe and RNase H in 2× folding buffer immediately after annealing in a separate reaction. Final probe concentrations were 20 µM for the P3 probe and 60 µM for the P6 probe. Products were separated on 6% denaturing polyacrylamide gels and quantitated using a Molecular Dynamics PhosphorImager. The data were fitted to a single exponential, $f = f_{equ.} + (f_0 - f_{equ.}) \exp(-k_{obs.}t)$, where f is the fraction cleaved at time t, $f_{equ.}$ is the fraction cleaved at equilibrium, f_0 is the fraction cleaved at time $t = 0$ and $k_{obs.}$ is the observed rate constant. This fitting procedure allows the end points to vary and independent values for the rate constant and equilibrium end point are obtained.

Mg^{2+} concentration dependence of folding

To obtain the Mg^{2+} concentration dependence of folding at equilibrium, the Mg^{2+} concentration in the folding buffer was varied and kinetic experiments as described above were performed at each Mg^{2+} concentration. During the probe binding/RNase H cleavage step the Mg^{2+} concentration was always adjusted to 10 mM. The equilibrium end point from the fit of the data to single exponentials yielded the fraction of RNA folded at equilibrium for each Mg^{2+} concentration. The data were fitted to an expression for two state binding of n Mg^{2+} ions, $f = 1/\{([Mg^{2+}]/[Mg^{2+}]_{1/2})^n + 1\}$, where f is the fraction cleaved at equilibrium, n is the number of Mg^{2+} ions bound per RNA molecule and $[Mg^{2+}]_{1/2}$ is the mid point of the transition.

RESULTS

Measuring the rate of RNA folding

To observe RNA folding kinetically as well as at equilibrium, we have developed a kinetic assay based on hybridization of complementary oligodeoxynucleotide probes and RNase H cleavage (18). Folding is induced by the addition of Mg^{2+}, and the fraction of RNA still accessible to oligonucleotide binding and/or RNase H cleavage is determined at increasing folding times. The rate of the transition from the accessible, unfolded to the inaccessible, folded state can thus be measured. The assay requires that there is a substantial difference in the accessibility of the RNA to the probes at different Mg^{2+} concentrations. This condition is only met for certain regions of the RNA, including the P3 and P7 helices in the P3–P7

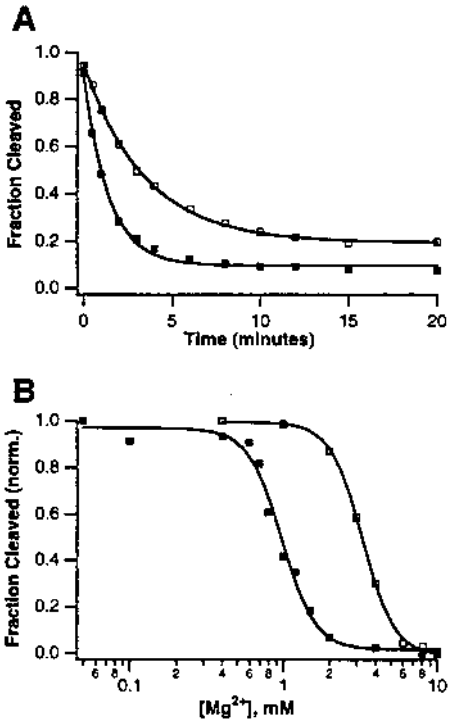

Figure 2. Formation of the P3 helix in L-21 *Sca*I (■) and L-21 *Nhe*I (□) RNA. (**A**) Kinetics of P3 formation at 10 mM Mg^{2+} and 37°C. Observed rates were 0.72 ± 0.14/min for L-21 *Sca*I and 0.32 ± 0.04/min for L-21 *Nhe*I. The error represents the standard deviation from 17 independent experiments for L-21 *Sca*I, as described (18), and the range of values obtained from three independent experiments for L-21 *Nhe*I. (**B**) Mg^{2+} dependence of P3 formation at equilibrium. Transition mid points are 0.97 mM for L-21 *Sca*I, as described (18), and 3.3 mM for L-21 *Nhe*I. The data were normalized to allow a direct comparison.

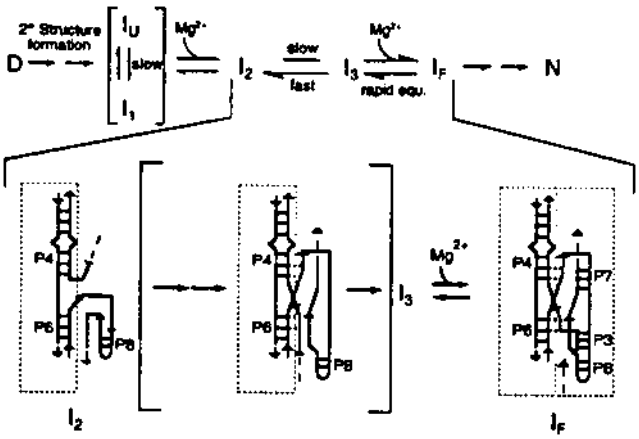

Figure 3. Proposed minimal model for the kinetic folding pathway of the *Tetrahymena* ribozyme (6,18). The two main structural subdomains are outlined.

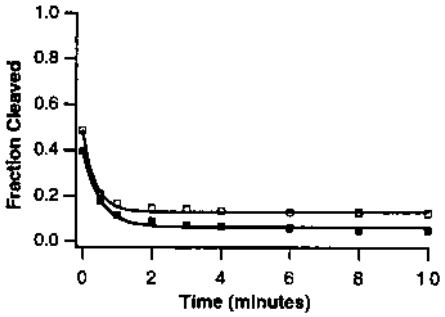

Figure 4. Kinetics of P4–P6 subdomain formation in L-21 *Sca*I (■) and L-21 *Nhe*I (□) RNA.

subdomain and the P4 and P6 helices in the P4–P6 subdomain (18). Using different oligonucleotide probes folding of both the P4–P6 and the P3–P7 subdomains can therefore be followed specifically and independently.

Slow folding of a ribozyme lacking the P9.1–P9.2 extension

To test the effect of the P9.1–P9.2 extension on folding of the P3–P7 subdomain we examined the folding kinetics of the L-21 *Nhe*I ribozyme, in which the extension is deleted, using an oligonucleotide probe targeting the P3 helix (Fig. 1). Since the P3–P7 subdomain forms in an interdependent manner, the probe targeting P3 reports on formation of the entire subdomain (6). The folding rate in L-21 *Nhe*I RNA was decreased compared with full-length L-21 *Sca*I ribozyme (Fig. 2) ($k_{obs.}$ = 0.32 ± 0.04/min for L-21 *Nhe*I versus 0.72 ± 0.14/min for L-21 *Sca*I). Because the rearrangement monitored with this probe is a slow step required for stable formation of the P3–P7 subdomain (formation of intermediate I_3 from I_2; see Fig. 3), the decreased rate suggests that stabilization of the P3–P7 subdomain by the P9.1–P9.2 extension may be involved in this slow step. Alternatively, deleting the 3′-terminal extension may have resulted in a change in the rate limiting step.

To determine whether the Mg^{2+}-dependent formation of I_F from I_3 had become rate limiting (Fig. 3), we measured the folding rate with the P3 probe at different concentrations of Mg^{2+}. No change in the rate was observed, within the error of our experiments, suggesting that Mg^{2+} binding had not become rate limiting (data not shown). To determine whether an earlier step in the folding pathway had become rate limiting, we monitored the kinetics of formation of the P4–P6 subdomain and intermediate I_2 using a probe targeting P6/P6a (Fig. 1). The rate was found to be fast and unchanged from L-21 *Sca*I RNA (Fig. 4), showing that formation of I_2 had also not become rate limiting. Together these results indicate that within our minimal kinetic folding pathway the rate limiting step remained unchanged upon deletion of the P9.1–P9.2 extension (Fig. 3).

Mg^{2+}-dependence of P3–P7 formation

To compare our data with the published equilibrium Mg^{2+}-dependence of P3–P7 formation in the context of the L-21 *Nhe*I ribozyme (10) we determined the fraction of RNA in which P3 is accessible to our oligonucleotide probe at equilibrium at a series of Mg^{2+} concentrations (Fig. 2B). The fraction of ribozyme folded at each Mg^{2+} concentration was obtained from the equilibrium end point of a kinetic experiment as described above. Each point therefore represents an independent experiment. The mid point of the transition ($[Mg^{2+}]_{1/2}$ = 3.3 mM) was shifted to a higher Mg^{2+} concentration compared with L-21 *Sca*I ($[Mg^{2+}]_{1/2}$ = 0.97 mM) (18).

DISCUSSION

We have continued our investigation of the kinetic folding pathway of the *Tetrahymena* group I intron, which serves as a model system for RNA folding, by examining the role of a peripheral extension outside the catalytic core of the ribozyme during folding. Deletion of the P9.1–P9.2 extension at the 3′-end of the ribozyme decreases the rate of the slow step in our proposed minimal kinetic scheme (Fig. 3). This slow rate represents the conversion of intermediate I_2, in which the P4–P6 subdomain but not the P3–P7 subdomain is formed, to the transient intermediate I_3. I_3 is competent to rapidly bind Mg^{2+}, allowing stable formation of the P3–P7 subdomain and thus I_F, an intermediate in which both the P4–P6 and the P3–P7 subdomains are present (18). Folding of the fast forming P4–P6 subdomain, and therefore formation of I_2, is not perturbed in L- 21 *Nhe*I RNA, nor has Mg^{2+} binding become rate limiting. The decreased rate is therefore due to a direct effect on the slow step in the minimal folding mechanism and not to a change in the rate limiting step within our kinetic scheme.

The observed decrease in the folding rate upon deletion of the P9.1–P9.2 extension is only 2-fold and therefore not very dramatic. The folding rate of L-21 *Sca*I RNA was measured numerous times with separate preparations of each reagent over a period of almost 1 year and in none of these experiments was a rate as low as that observed here for L-21 *Nhe*I RNA ever measured. Conversely, the folding rate of the L-21 *Nhe*I ribozyme was independently determined several times using separate preparations of reagents and the rate was never in the range observed for wild-type RNA. The effect is therefore highly reproducible, which leads us to believe that the 2-fold difference, although small, is real.

Chemical modification experiments have provided evidence for a tertiary interaction between the 3′-terminal extension and bases in loops L2 or L2.1, which may help lock the P3–P7 subdomain in place (11) and be responsible for stabilization of P3–P7 (10). In the *sunY* group I intron a similar stabilization of the core by elements of a 3′-terminal domain was observed (9). Our results suggest that the proposed tertiary interaction between the P9.1–P9.2 extension and L2 or L2.1 in the *Tetrahymena* intron is formed during the slow step of the minimal folding mechanism and is therefore present in the transient intermediate I_3. The identity of the nucleotides in P9.1–P9.2 and L2/L2.1 participating in the tertiary interaction has not yet been established and it was therefore not possible to test the involvement of this interaction more specifically. We propose that the P9.1–P9.2 extension is important not only for stabilization of the P3–P7 subdomain, but also helps guide folding of the RNA by limiting the mobility of P3 and P7 and, along with the triple helical scaffold, promoting proper association of the P4–P6 and P3–P7 subdomains.

The slow formation of I_3 from I_2 may consist of several microscopic steps, including formation of the triple helical scaffold connecting the two subdomains and possibly base pair formation in P3 and P7. We have shown that the triple helical scaffold is formed during the slow step (6), but can at present not distinguish whether formation of the triple helix scaffold and stabilization of the P3–P7 subdomain by the P9.1–P9.2 extension are discrete steps or whether they occur in a concerted rearrangement. A double mutant ribozyme in which both the 3′-terminal extension was deleted and the triple helix scaffold was disrupted was too unstable to allow careful analysis of folding kinetics (data not shown).

Determination of the Mg^{2+} dependence of formation of the P3 helix in L-21 *Nhe*I RNA showed a shift of the transition mid point to a higher Mg^{2+} concentration, and confirmed published observations (10,11) that the P3–P7 subdomain is destabilized in this truncated ribozyme. While the exact value of the transition mid point ($[Mg^{2+}]_{1/2}$ = 3.30 mM) differs somewhat from that observed using Fe(II)–EDTA as a footprinting probe ($[Mg^{2+}]_{1/2}$ = 1.83 mM; 10), the difference in the apparent mid points is not surprising given the large difference in size between the two probes used [small hydroxyl radicals generated in the presence of Fe(II)–EDTA versus large oligonucleotides]. The different temperatures at which the experiments were performed (42°C for the hydroxyl radical footprinting, 37°C for the kinetic oligonucleotide hybridizations) may also contribute to this discrepancy. Although formation of I_3 from I_2 itself does not involve binding of Mg^{2+}, the observed increase in the equilibrium Mg^{2+} requirement for P3–P7 subdomain formation can at least in part be accounted for by the decreased rate of formation of I_3. Since rapid Mg^{2+} binding to I_3 drives the otherwise unfavorable conversion of I_2 to I_3 (Fig. 3), a decrease in the rate of this conversion will result in an increase in the Mg^{2+} concentration required for formation of I_F, even if the dissociation constant of Mg^{2+} binding to I_3 is unaltered. The present kinetic results thus complement previous observations made at equilibrium (10,11).

In the cellular environment it may be that production of mature RNA for transcripts containing group I introns is limited by proper folding of the intron, rather than by catalysis. Partially or completely unfolded RNA may be bound rapidly and non-specifically by proteins, which can inhibit formation of the active structure. Even a relatively minor increase in folding rate afforded by a peripheral extension may therefore confer a selective advantage by preventing the core of the RNA from becoming kinetically trapped in unproductive conformations or complexes.

These results may have implications for the folding mechanisms of other large RNAs, many of which also consist of phylogenetically highly conserved core regions surrounded by less conserved peripheral extensions. Such an organization can be found in group I (4,5) and group II introns (21), in the RNA component of RNase P (22) and in rRNAs (23,24). In several cases it has been demonstrated that peripheral extensions can stabilize the core region of a large RNA (2,9,10–12), suggesting that some general functions of the extensions may be conserved between different RNAs. The importance of the P9.1–P9.2 extension during folding of the *Tetrahymena* ribozyme suggests that, in addition to stabilizing the final conformation, non-conserved peripheral extensions may also guide the folding process in large, highly structured RNAs.

ACKNOWLEDGEMENTS

We thank Dan Treiber and Martha Rook for critical reading of the manuscript. This work was supported by grants from the Searle Scholar Program of the Chicago Community Trust, the Rita Allen Foundation and the Camille and Henry Dreyfus Foundation. P.P.Z. is a pre-doctoral fellow of the Howard Hughes Medical Institute.

REFERENCES

1 Cech,T.R. (1993) In Gesteland,R.F. and Atkins,J.F. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 239–269.
2 Beaudry,A.A. and Joyce,G.F. (1990) *Biochemistry*, **29**, 6534–6539.
3 Doudna,J.A. and Szostak,J.W. (1989) *Mol. Cell. Biol.*, **9**, 5480–5483.
4 Michel,F. and Westhof,E. (1990) *J. Mol. Biol.*, **216**, 585–610.
5 Cech,T.R., Damberger,S.H. and Gutell,R.R. (1994) *Nature Struct. Biol.*, **1**, 273–280.

6 Zarrinkar,P.P. and Williamson,J.R. (1996) *Nature Struc. Biol*., submitted for publication.
7 Michel,F., Ellington,A.D., Couture,S. and Szostak,J.W. (1990) *Nature*, **347**, 578–580.
8 Doudna,J.A. and Cech,T.R. (1995) *RNA*, **1**, 36–45.
9 Jaeger,L., Westhof,E. and Michel,F. (1993) *J. Mol. Biol*., **234**, 331–346.
10 Laggerbauer,B., Murphy,F.L. and Cech,T.R. (1994) *EMBO J*., **13**, 2669–2676.
11 Banerjee,A.R., Jaeger,J.A. and Turner,D.H. (1993) *Biochemistry*, **32**, 153–163.
12 Jaeger,L., Westhof,E. and Michel,F. (1991) *J. Mol. Biol*., **221**, 1153–1164.
13 van der Horst,G., Christian,A. and Inoue,T. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 184–188.
14 Mohr,G., Caprara,M.G., Guo,Q. and Lambowitz,A.M. (1994) *Nature*, **370**, 147–150.
15 Wang,Y.-H., Murphy,F.L., Cech,T.R. and Griffith,J.D. (1994) *J. Mol. Biol*., **236**, 64–71.
16 Nakamura,T.M., Wang,Y.-H., Zaug,A.J., Griffith,J.D. and Cech,T.R. (1995) *EMBO J*., **14**, 4849–4859.
17 Murphy,F.L. and Cech,T.R. (1993) *Biochemistry*, **32**, 5291–5300.
18 Zarrinkar,P.P. and Williamson,J.R. (1994) *Science*, **265**, 918–924.
19 Celander,D.W. and Cech,T.R. (1991) *Science*, **251**, 401–407.
20 Zaug,A.J., Grosshans,C.A. and Cech,T.R. (1988) *Biochemistry*, **27**, 8924–8931.
21 Michel,F., Umesono,K. and Ozeki,H. (1989) *Gene*, **82**, 5–30.
22 Darr,S.C., Brown,J.W. and Pace,N.R. (1992) *Trends Biochem. Sci*., **17**, 178–182.
23 Neefs,J.-M., Van de Peer,Y., De Rijk,P., Chapelle,S. and De Wachter,R. (1993) *Nucleic Acids Res*., **21**, 3025–3049.
24 De Rijk,P., Van de Peer,Y., Chapelle,S. and De Wachter,R. (1994) *Nucleic Acids Res*., **22**, 3495–3501.