RNA (1997), *3*:613–623. Cambridge University Press. Printed in the USA.

The protein cofactor allows the sequence of an RNase P ribozyme to diversify by maintaining the catalytically active structure of the enzyme

JOSEPH J. KIM,[1] AHMED F. KILANI,[1] XIAOYAN ZHAN,[1] SIDNEY ALTMAN,[2] and FENYONG LIU[1]

[1]School of Public Health, University of California, 140 Earl Warren Hall, Berkeley, California 94720, USA
[2]Department of Biology, Yale University, 266 Whitney Avenue, New Haven, Connecticut 06511, USA

ABSTRACT

To study the effect proteins have on the catalysis and evolution of RNA enzymes, we simulated evolution of RNase P catalytic M1 RNA in vitro, in the presence and absence of its C5 protein cofactor. In the presence of C5, functional M1 sequence variants (not catalytically active in the absence of C5) were selected in addition to those identical to M1. C5 maintains the catalytically active structure of the variants and allows for an enhanced spectrum of M1 molecules to function in the context of a ribonucleoprotein (RNP) complex. The generation of an RNP enzyme, requiring both RNA and protein components, from a catalytically active RNA molecule has implications for how modern RNP complexes evolved from ancestral RNAs.

Keywords: catalytic RNA; C5 protein; in vitro evolution; RNA catalysis; RNase P; tRNA

INTRODUCTION

RNA is postulated to have been the first biological macromolecule in a primeval environment referred to as the RNA world (Gilbert, 1986; Joyce, 1989). Following the advent of protein biosynthesis, RNA molecules evolved in concert with specific protein cofactors into modern ribonucleoprotein complexes, such as ribonuclease P (Altman et al., 1993), spliceosomes (Moore et al., 1993), and ribosomes (Noller et al., 1992). Although many of the RNA components of these modern RNP enzymatic complexes have lost their ability to function as catalytically active ribozymes, the RNA world hypothesis proposes that all of these RNAs once possessed intrinsic enzymatic activity (Gilbert, 1986; Joyce, 1989). How these ancestral RNA enzymes evolved into RNPs that require both protein and RNA components for catalysis remains a mystery. Equally unclear is how the first proteins interacted with the RNA enzymes, already present in the earliest biochemical systems, and what impact these interactions had on the catalysis and evolution of the RNA enzymes.

Although it is impossible at present to identify with any certainty the components of the RNA world, in vitro evolution procedures provide a useful means to simulate the generation of catalytically active RNA molecules (Bartel & Szostak, 1993; Lehman & Joyce, 1993; Illangasekare et al., 1995; Wilson & Szostak, 1995; Ekland & Bartel, 1996). In this procedure, a large, heterogeneous pool of RNAs is subjected to multiple rounds of selection, amplification, and mutation. Variant molecules that display the desired phenotype, such as enzymatic activity or the capability to serve as substrates for an enzyme, are selected, isolated, and analyzed. Most of these studies, focused on evolving RNA catalysts in the absence of protein cofactors, have led to the generation of RNA enzymes displaying novel catalytic activities and allowed us to study the evolution of these molecules (Bartel & Szostak, 1993; Lehman & Joyce, 1993; Illangasekare et al., 1995; Wilson & Szostak, 1995; Ekland & Bartel, 1996; Frank et al., 1996). Moreover, by studying evolution in vitro of RNA molecules in the presence and absence of protein cofactors, we can investigate the effect proteins have on the evolu-

Reprint requests to: Fenyong Liu, School of Public Health, University of California, 140 Earl Warren Hall, Berkeley, California 94720, USA; e-mail: liu_fy@uclink4.berkeley.edu.

Abbreviations: ptRNA, precursor tRNA; p4.5S, RNA precursor to 4.5S RNA; RNP, ribonucleoprotein; RNase P, ribonuclease P.

tion of RNA catalysis and understand the function of the proteins in the catalysis and evolution of the RNA molecules (Liu & Altman, 1994).

RNase P offers a unique system to study the interactions between protein cofactors and catalytic RNA molecules. This enzyme, found in all living organisms examined to date, is a ribonucleoprotein complex responsible for the maturation of the 5′ termini of tRNAs (Gopalan et al., 1995; Pace & Brown, 1995). It is essential for cell viability and is proposed to be one of the ancient enzymatic activities (Weiner & Maizels, 1987; Liu & Altman, 1994; Maizels & Weiner, 1994). In *Escherichia coli*, the holoenzyme consists of M1 RNA and C5 protein, both of which are required for activity in vivo, although M1 RNA can carry out catalysis by itself in vitro (Guerrier-Takada et al., 1983).

Phylogenetic analyses, combined with biochemical studies and computer modeling, have established the models for the secondary and three-dimensional structure of M1 RNA (Haas et al., 1994; Harris et al., 1994; Westhof & Altman, 1994). The P_{10} helix region (positions 61–75 and 347–361) of M1 RNA (Fig. 1) has been shown to be a part of the active site and has been proposed to exist in the common ancestor of all RNase P RNAs (Guerrier-Takada et al., 1989; Haas et al., 1994; Harris et al., 1994; Westhof & Altman, 1994; Frank et al., 1996; Loria & Pan, 1996; Westhof et al., 1996). Both the structure and primary nucleotide sequence of this region have been found to be highly conserved among all known RNase P RNAs (Haas et al., 1994, 1996). Although disruption of the P_{10} helix structure has been shown to abolish the ribozyme activity of M1 RNA (Haas et al., 1991), little is known about whether the sequence conservation of this region is required for RNase P activity and what function these conserved nucleotides play in RNase P catalysis (Harris & Pace, 1995; Hardt et al., 1996).

Because the P_{10} helix region is highly conserved, functionally important, and probably existed in the common ancestor of RNase P RNAs, studies focused on this region should allow us to investigate the catalytic mechanism and evolution of RNase P RNAs. By performing evolution in vitro to select functional ribozyme variants from M1 RNA molecules that contain randomized sequences in the P_{10} region, we can examine what structure and sequence in the region is required for RNase P catalytic activity and study the evolution of RNase P catalysis. Furthermore, these evolution in vitro studies, both in the absence and presence of C5 protein, should reveal the effect the protein has on catalysis and evolution of the catalytic M1 RNA.

In this report, we simulated the evolution of M1 RNA in vitro, in the absence and presence of C5 protein. A part of the P_{10} region of M1 RNA was randomized to create a pool of RNA molecules and catalytically active variants were selected in the presence and absence of C5 protein. Ribozymes that evolved in the absence of C5 were identical in sequence to M1 RNA. In the presence of C5 protein, however, M1 RNA sequence variants were selected in addition to those identical to M1 RNA. Although these ribozyme variants displayed RNase P activity in an RNP complex with C5, they were not catalytic RNA molecules in the absence of C5. Our results show that the conservation of the M1 RNA sequence is required for maintenance of the catalytically active conformation of the ribozyme and that C5 protein allows the sequence of the ribozyme to diversify by maintaining the catalytically active structure of the variants. These results illustrate how an RNase P RNP enzyme, requiring both RNA and protein components, can be generated from a catalytic, active M1 RNA molecule. They may have general implications for how modern RNP enzymatic complexes, including ribosomes and spliceosomes, evolved from ancestral catalytic RNAs.

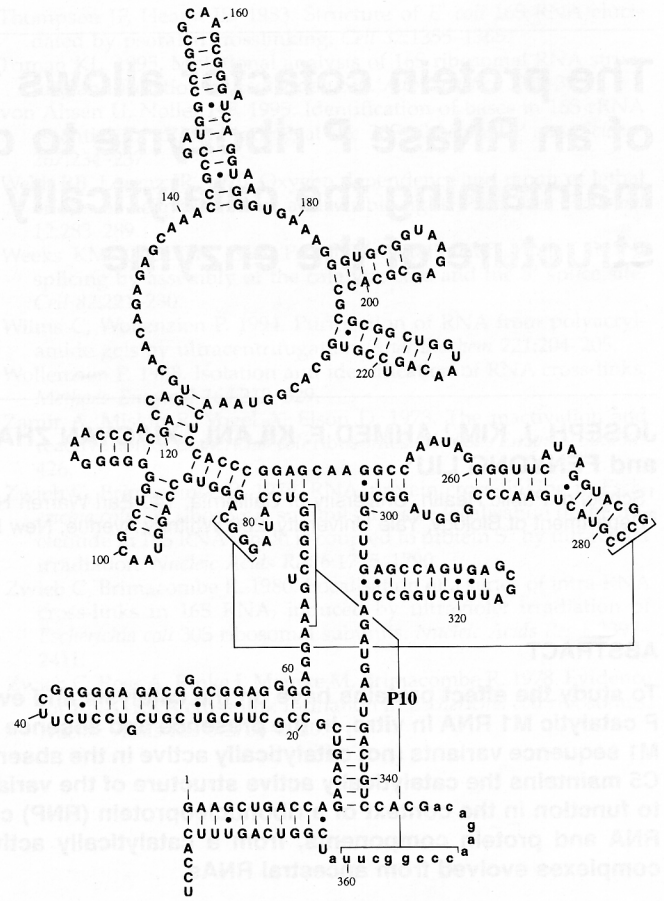

FIGURE 1. Proposed secondary structure of the RNA subunit (M1 RNA) of RNase P from *E. coli* (Haas et al., 1994). The sequence of the region (positions 347–361) in M1 RNA, which was randomized in our studies, is shown as lower-case letters. The sequence and structure of this region is highly conserved in all known RNase P RNA subunits from eubacteria and archaea (Haas et al., 1994, 1996). In particular, positions (A_{347}, C_{348}, A_{349}, A_{351}, A_{352}, G_{356}, G_{357}, C_{358}, U_{360}, and A_{361}) in M1 RNA are universally conserved among all known RNase P RNA subunits from eubacteria (Haas et al., 1994).

RESULTS

Randomized nucleotides were introduced into 15 positions of M1 RNA (positions 347–361), which are located at the 3′ half of the P_{10} helix region (see Fig. 1), to create an RNA pool of 4^{15} (2×10^9) different sequences. The ribozyme pool was subjected to evolution in vitro to cleave a pre-tRNA substrate (i.e., ptRNA^{Tyr} from *E. coli*) by a selection procedure (Fig. 2) (Joyce, 1992; Gold et al., 1993; Szostak & Ellington, 1993). The selection procedure (Fig. 2) involved synthesis of M1 RNA molecules that contained randomized sequences, parallel selection experiments in vitro for functional ribozymes to cleave ptRNA^{Tyr} in the absence and presence of C5 protein, followed by purification and amplification of selected ribozymes. To facilitate the selection of active M1 RNA variants, a leader sequence (LS) was covalently linked to the 3′ end of M1 RNA (Fig. 2). The constructed M1-LS ribozymes were annealed to an RNA molecule (external guide sequence or EGS^{Tyr}) that resembled three-quarters of pre-tRNA^{Tyr} (Yuan & Altman, 1994). The complex formed between the LS and EGS^{Tyr} resembled the structure of pre-tRNA^{Tyr}. Only active ribozymes recognized the tRNA^{Tyr}-like molecule and cleaved at the LS (Fig. 2). Therefore, the selected ribozymes were expected to exhibit efficient catalytic activity to cleave their own LS when it complexed with a EGS^{Tyr} (in *cis*) or the LS of another ribozyme complex (in *trans*). This procedure was repeated nine times, until no apparent enhancement of cleavage rate of the ribozyme population was observed. The enhancement of the cleavage rate between cycle 0 and cycle 9 was about 1 million-fold (data not shown). Sixteen sequences coding for ribozymes selected in the absence of C5 protein and 28 sequences coding for those selected in the presence of the protein cofactor were cloned and characterized (Table 1). The salient features of the results are as follows.

1. Both the wild-type and variant sequences were found in the randomized region of ribozymes selected in the presence of C5 protein, whereas only the wild-type sequence was found in those selected in the absence of C5 protein. This observation raises the possibility that the sequence (from 347–361) of M1 RNA is a remnant of the sequence of the primitive M1 RNA molecules in the RNA world and existed before the appearance of the protein subunit of RNase P (e.g., C5 protein).

2. The ribozymes selected in the presence of C5 protein were divided into three sets. The first set consisted of 10 clones and had the wild-type sequence. The second set consisted of 16 clones and had base substitutions, insertions, and deletions in comparison with the wild-type sequence. More than half of these mutations were at nt A_{351} and A_{352}, both of which are absolutely conserved among all known bacterial and archaeal RNase P RNA subunits (Haas et al., 1994, 1996). These results suggest that the conservation of these two nucleotides is not required for the functional activity of RNase P holoenzyme. The third set consisted of two clones with sequences significantly different from the wild-type M1 RNA sequence.

To analyze the effect of these mutations on the structure of these ribozymes, the secondary and tertiary structures of ribozymes R7, C83, and C102, representatives of each set of the selected ribozymes, were mapped by RNase and Fe(II)-EDTA cleavage approaches, which detect the accessibility of the base and phosphoribose backbone, respectively (Latham & Cech, 1989; Knap et al., 1990; Liu & Altman, 1994; Loria & Pan, 1996; Westhof et al., 1996) (Fig. 3; Table 1). Furthermore, Fe(II)-EDTA cleavage mapping was performed both in the absence and presence of Mg^{2+} ions

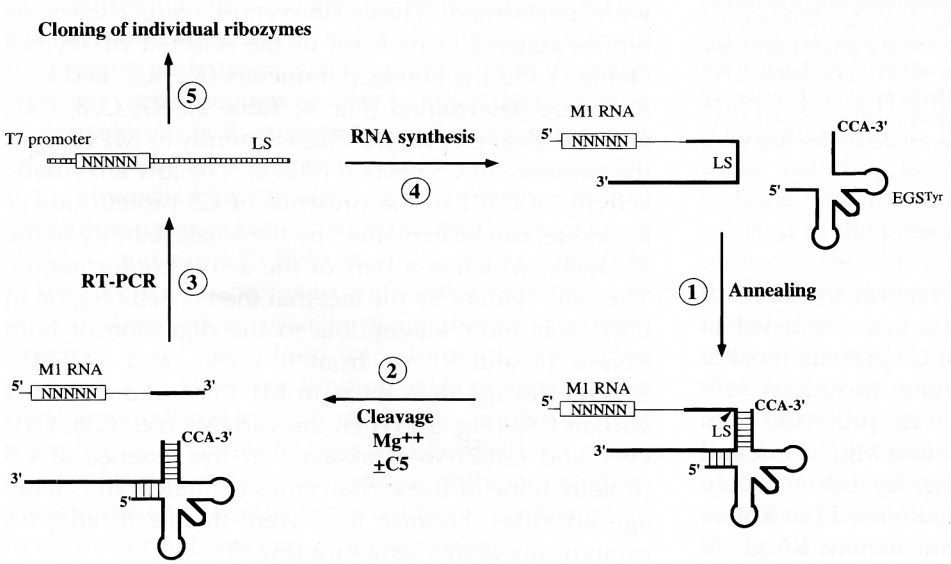

FIGURE 2. Schematic representation of the evolution in vitro procedures to select M1 RNA variants in the absence and presence of C5 protein. The selection procedure involved synthesis of ribozyme molecules, parallel selection experiments in vitro for functional ribozymes to cleave ptRNA^{Tyr} in the absence and presence of C5 protein, followed by purification and amplification of selected ribozymes. The site of cleavage at the leader sequence (LS) by the selected enzymes is marked with a filled arrow. The complex formed between the EGS^{Tyr} sequence and the leader sequence, which resembles the structure of ptRNA^{Tyr}, is highlighted. Experimental details are described in Materials and Methods.

TABLE 1. Sequence analyses of the M1 RNA variants selected in the absence and presence of C5 protein.[a]

Clone #	Randomized region (347–361)		
(1) RNA selected in the absence of C5 protein			
R7	ACAGA	ACCCG	GCTTA
(2) RNA selected in the presence of C5 protein			
Set I			
C33	ACAGA	ACCCG	GCTTA
C40	ACAGA	ACCCG	GCTTA
C43	ACAGA	ACCCG	GCTTA
C49	ACAGA	ACCCG	GCTTA
C55	ACAGA	ACCCG	GCTTA
C70	ACAGA	ACCCG	GCTTA
C77	ACAGA	ACCCG	GCTTA
C108	ACAGA	ACCCG	GCTTA
C109	ACAGA	ACCCG	GCTTA
C124	ACAGA	ACCCG	GCTTA
Set II			
C17	ACCG**C**	ACCC**T**	**A**CTTA
C23	ACCG**A**	ACCCG	GCTTA
C35	ACAGA	G**A**CCCG	GCTTA
C36	ACA**A**A	C**G**CCG	**A**CTTA
C38	ACAG	ACCCG	GCTTA
C41	ACAGA	G**A**CCCG	GCTTA
C44	ACAGA	G**A**CCCG	GCTTA
C45	AGAGA	ACCCG	CCCCA
C78	ACAG**T**	ACCCG	GCTTA
C83	ACAGA	G**C**CCG	GCTTA
C89	ACAGA	ACC**T**G	GCTTA
C111	ACAG**G**	ACCCG	GCTTA
C112	ACAGA	ACCCG	GCTTA**G**
C126	ACAGA	T**C**CCG	GCTTA
C129	ACAGA	G**C**CCG	GCTTA
C137	ACAGA	G**C**CCG	GCTTA
Set III			
C102	**CTAGC**	**AGGAT**	**GATTC**
C128	**GCATC**	ACCC**C**	**ACCCA**

[a]DNA sequences of the ribozymes at the randomized region (position 347–361) after selection in the absence and presence of C5 protein are shown. Only 1 (R7) of the 16 ribozymes selected in the absence of C5 protein, which are all identical to the wild-type M1 RNA sequence, is listed. The positions (353–360) underlined are those that form the P_{10} helix with positions 66–74 (Fig. 1A). Mutated nucleotides in the set II and III ribozyme variants are bold. Ribozyme variants C35, C41, C44 contain single-nucleotide insertional mutations between A_{351} and A_{352}. C112 has a single-nucleotide insertional mutation between A_{361} and U_{362}. C38 has a single-nucleotide deletional mutation at A_{351} or A_{352} (exact position of the deletion not conclusive due to identical nucleotides at both positions). Sequence analyses were performed with a PCR Silver sequencing kit (Promega Inc., Madison, Wisconsin). Experimental details are described in Materials and Methods.

and the results were compared to reveal the effect of Mg^{2+} on the folding of these ribozymes that led to catalytic activity. In the absence of C5 protein, most of the nucleotides in the selected region (positions 347–361) of R7 (identical to M1 RNA) were protected from cleavage by Fe(II)-EDTA upon binding Mg^{2+} ions, and were either susceptible to cleavage by cobra venom RNase V, or resistant to cleavage by RNase T1 or RNase from *Bacillus cereus*. Because cobra venom RNase V only recognizes base paired structures, and T1 along with *B. cereus* RNase cleaves at single-stranded regions, this region is most likely base paired and folded (Fig. 3) (Knap et al., 1990; Harris et al., 1994; Westhof & Altman, 1994; Loria & Pan, 1996; Westhof et al., 1996). In contrast, the same region in variants C83 (e.g., C_{348} and A_{349}) and C102 (e.g., G_{346}, U_{348}, and G_{353}) adopted a structure that differed from M1 RNA. This region in the variants was either less susceptible to cleavage by cobra venom RNase V, or more sensitive to attack by RNase T1, RNase from *B. cereus*, and by Fe(II)-EDTA than that of M1 RNA (Fig. 3). Meanwhile, three discrete regions that consist of nt 20–60, 140–190, and 260–300 in R7, and the variants C83 and C102 were all altered in their susceptibility to cleavage by RNases and Fe(II)-EDTA upon binding C5 protein (Fig. 3). This is consistent with previous observations suggesting that these regions either contact the C5 protein directly or undergo a conformational change upon binding C5 (Vioque et al., 1988; Talbot & Altman, 1994; Westhof et al., 1996). Little change in the cleavage susceptibility was observed for the sequence selected at nt 347–361 in R7, suggesting that R7 RNA folds into a catalytically active conformation both in the absence and presence of the C5 protein. In the variants C83 and C102, however, the susceptibility of nt 347–361 to cleavage reagents was altered substantially by the presence of the C5 polypeptide. The variant P_{10} helix region (G_{72}, C_{348}, A_{349}, and G_{350} in variant C83; G_{68}, G_{353}, G_{354}, and U_{356} in variant C102) adopts a structure similar to that of R7, upon addition of C5 protein (Fig. 3). Therefore, the C5 protein functions to maintain the active conformation of the ribozyme and allows the variant RNAs to assume a catalytically active structure.

To analyze further the relationship between the sequences and structures of the selected ribozymes and their enzymatic activities, kinetic analyses of the cleavage activities of five selected ribozymes (R7, C38, C41, C83, C102) in the absence and presence of C5 protein were performed. These ribozymes were chosen as representatives of each set of the selected ribozymes (Table 1) and the kinetic parameters (K_m, k_{cat}, and k_{cat}/K_m) were determined (Fig. 4; Table 2). R7, C38, C41, and C83 cleaved ptRNA^{Tyr} as efficiently as M1 RNA in the presence of C5 protein (Fig. 4). The low enzymatic activity of C102 in the presence of C5 protein (lower k_{cat} value) can be explained by the weak stability of the P_{10} helix, which is a part of the active conformation. This was evident by the fact that the P_{10} helix region in C102 was more susceptible to the digestion of both RNase T1 and RNase from *B. cereus*, and to Fe(II)-EDTA cleavage than those in M1 RNA and the set II variant C83 (Fig. 3). When the variants (i.e., C38, C41, C83, and C102) were assayed in the absence of C5 protein, none of these ribozymes exhibited any cleavage activities, because they were unable to adopt a catalytically active structure (Fig. 3).

To determine whether some of the variants (e.g., those with mutations at the universally conserved A_{351} and A_{352}) can complement the activity of RNase P both in vitro and in vivo, two experiments were performed. First, to determine whether the variants can cleave other natural substrates for *E. coli* RNase P holoenzyme, C38, C41, and C83 were assayed for their activities to cleave p4.5S RNA, another natural substrate for *E. coli* RNase P (Bothwell et al., 1976). In the presence of C5 protein, these ribozymes cleaved as efficiently as M1 RNA and R7, but did not exhibit any cleavage activities in the absence of C5 protein (data not shown). Second, to determine whether C83 (mutation of A_{352} to G) can complement the function of M1 RNA in vivo, this variant was introduced into *E. coli* NHY322(rnpA_{49}), which exhibits a temperature sensitive (*ts*) growth phenotype as a result of a single amino acid change in C5 protein (Kirsebom et al., 1988). This mutation at C5 precludes the assembly of a functional holoenzyme at the nonpermissive temperature. Complementation, however, can be achieved by overexpression of M1 RNA (Kirsebom et al., 1988; Altman et al., 1993). In this experiment, C83 rescued the *ts* phenotype and supported the growth of NHY322(rnpA_{49}) at the nonpermissive temperature (42 °C) (data not shown). Therefore, the conservation of the identity of the base at position 352 is not required for the functional activity of RNase P holoenzyme either in vitro or in vivo, but is crucial for the activity of the RNA component alone in vitro.

DISCUSSION

Function of C5 protein in the catalysis and evolution of M1 RNA

We have extended our understanding of the evolution of the catalytic RNA subunits of RNase P and their interactions with protein cofactors by randomization of nucleotides in a conserved region (position 347–361) of M1 RNA and selection of active ribozymes in vitro in the absence and presence of C5 protein. Both the wild-type and variant sequences were found in the RNA molecules selected in the presence of C5 protein, whereas only the wild-type sequence was found in those selected in the absence of the protein cofactor. In the presence of C5 protein, most of the variants (i.e., set II) cleaved pre-tRNA^{Tyr} and pre-4.5S RNA as efficiently as M1 RNA. In the absence of C5 protein, the variants were not catalytic and the sequences in the region that had been randomized were not folded into the wild-type active conformation. However, binding of C5 protein to these ribozymes induced a structural change, as inferred from mapping by different RNases and Fe(II)-EDTA cleavage, and restored the conformation of the mutated regions sufficiently similar to the wild-type conformation to allow normal function of these variants.

C5 protein, when in a complex with M1 RNA, has been shown to result in a structural change of M1 RNA at three discrete regions (i.e., nt 20–60, 140–190, 260–300) and its binding site to M1 RNA is not in proximity to either the active site or the substrate binding site of the ribozyme (Vioque et al., 1988; Talbot & Altman, 1994). Therefore, C5 protein neither participates directly in catalysis by contributing functional groups to the active site nor has contact with ptRNA^{Tyr} (Gopalan et al., 1995). Our data showed that similar structural changes in these three regions were found in M1 RNA and the variants when they were in a complex with C5 protein (Fig. 3). These results suggest that C5 protein interacts with the variants in a similar way as with M1 RNA. Moreover, in the presence of C5 protein, the overall cleavage patterns of M1 RNA and the variants by different RNases and Fe(II)-EDTA are very similar to each other (we note the subtle difference at the P_{10} helix and other regions) (Fig. 3, data not shown), suggesting that the overall structures of these variants are similar to that of M1 RNA when they are in a complex with C5 protein. These results imply that, as in the case with M1 RNA, C5 protein is not in proximity to either the active site or the substrate binding site of the variants, and does not participate directly in catalysis nor have contact with ptRNA^{Tyr}. Our suggestion that C5 protein does not contact the substrate when in a complex with the variants is further supported by the observations that the values of K_m for the reactions catalyzed by the variants in the presence of C5 are very similar to that of M1 RNA under the same condition (Table 2). Therefore, our results indicate that C5 functions to maintain the active conformation of the ribozyme variants.

Little is known about why the nucleotides are highly conserved at positions 347–361 and what function this sequence conservation plays in the catalysis of RNase P and its catalytic RNA. Our findings suggest that the conserved base moieties function only to maintain the catalytically active conformation of the ribozyme, but do not participate directly in catalysis. For example, A_{351} and A_{352} are absolutely conserved among all known RNase P RNA subunits in bacteria and archaea (Haas et al., 1994, 1996). Mutations at these nucleotides, including base substitutions, deletions, and insertions, were found in 14 of the 28 ribozyme sequences selected in the presence of C5 protein (Table 1). These observations indicated that these two positions and the conservation of their nucleotide identity are not required for catalysis. Of particular interest is variant C38, which contains a single-nucleotide deletion at either A_{351} or A_{352} (exact position of the deletion not conclusive due to identical nucleotides at both positions). The phosphate oxygens and 2′ hydroxyl groups of the phosphoribose backbone at these two positions have been shown to be important for catalysis, possibly by coordinating the catalytically active Mg^{2+} ions

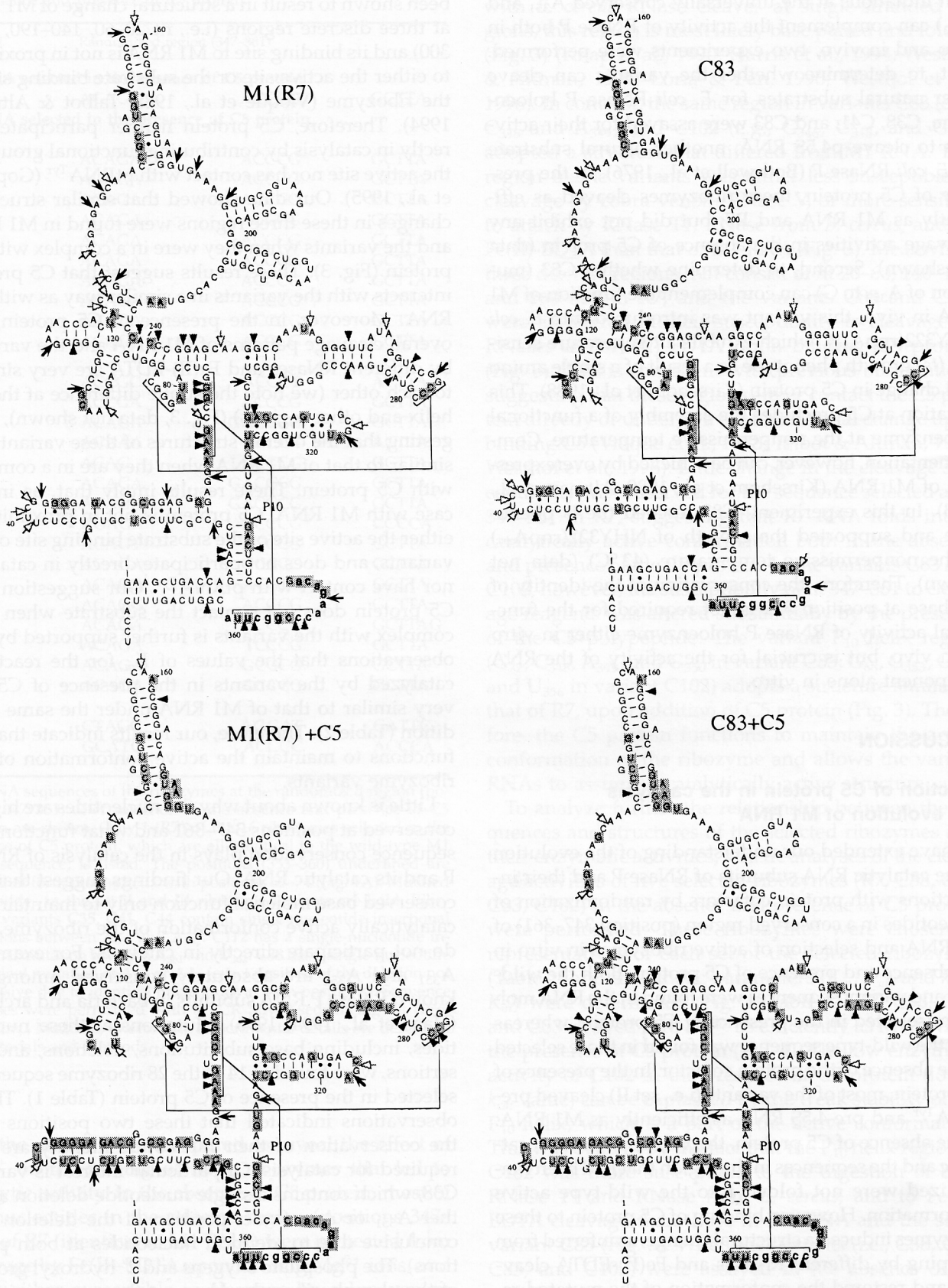

FIGURE 3. (*Figure continues on facing page.*)

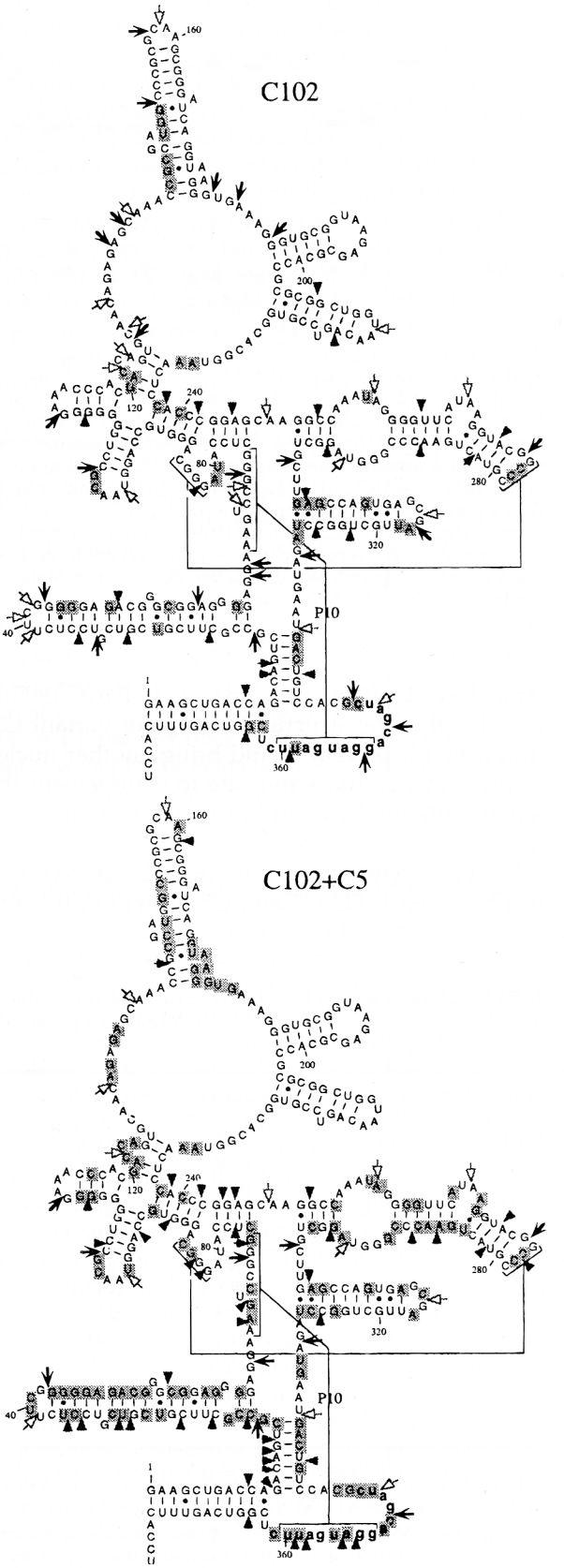

FIGURE 3. *Continued.* Structural analyses of the M1 RNA variants selected in the absence and presence of C5 protein. The secondary and tertiary structures of three selected ribozymes (i.e., R7, C83, and C102) were mapped in the absence (top) and presence (bottom, +C5) of C5 protein by RNase and Fe(II)-EDTA cleavage approaches, respectively. The sequences selected at the randomized positions are shown in bold type and lower-case letters. Sites of cleavage by RNase T1 are indicated by solid arrows. Hollow arrows indicate sites of cleavage by RNase from *B. cereus*. Both RNases T1 and RNase from *B. cereus* were used under conditions that allow them to cleave preferentially at a G or a pyrimidine in single-stranded regions, respectively. Solid triangles denote sites of cleavage by cobra venom ribonuclease V, which is specific for regions that are base paired. Shaded regions are those that were more protected from cleavage by Fe(II)-EDTA in the presence of Mg^{2+} ions than in the absence of the divalent ions. These regions were possibly folded into the active conformation upon binding Mg^{2+} ions. Experimental details are described in Materials and Methods.

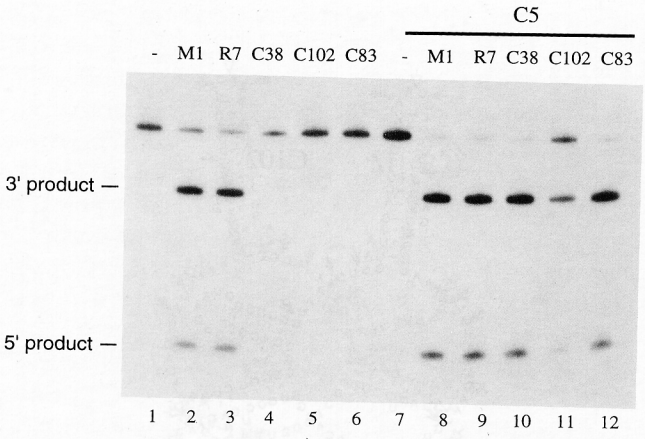

FIGURE 4. Cleavage of pre-tRNA^{Tyr} by the selected ribozymes in the absence and presence of C5 protein. RNA substrate ptRNA^{Tyr} (10 nM) was incubated either alone (lanes 1, 7) or with different ribozymes (5 nM) (lanes 2–6, 8–12). Cleavage reactions without and with C5 protein were performed at 37 °C for 20 min in buffer A (50 mM Tris·HCl, pH 7.5, 100 mM NH_4Cl, 100 mM $MgCl_2$) (lanes 1–6) and buffer B (50 mM Tris·HCl, pH 7.5, 100 mM NH_4Cl, 10 mM $MgCl_2$, 50 nM C5 protein) (lanes 7–12), respectively. Cleavage products were separated in 8% polyacrylamide gels containing 8 M urea. Experimental details are described in Materials and Methods.

(Hardt et al., 1995, 1996; Harris & Pace, 1995). It is possible that the structural change of variant C38 induced by C5 protein would bring another nucleotide in proximity to the active site to complement the deletion mutation by contributing the similar functional groups.

In our selection procedure (Fig. 2), ribozymes were incubated with EGS^{Tyr} and C5 protein in the absence of Mg^{2+} ions before the cleavage reaction proceeded. It has been shown that the wild-type M1 RNA does not fold or bind to C5 protein correctly in the absence of Mg^{2+} ions (Guerrier-Takada et al., 1983; Talbot & Altman, 1994). Therefore, it is possible that the ribozyme variants might be selected for more efficient catalysis and, in addition, better folding or binding to C5 protein and the substrate. It will be interesting to determine how these selected ribozymes fold and interact with the substrate and C5 protein under different buffer conditions (e.g., no $MgCl_2$ or 100 mM $MgCl_2$).

We have also extended our understanding of the function of the protein cofactors in the evolution of RNase P and its RNA subunit. We showed that the variety of the selected M1 RNA sequences, including base substitution, deletion, and insertion, can be attributed to the ability of C5 polypeptide to structurally alter the variant RNAs such that they adopt a catalytically active conformation. The proteins that enhance the rate of the reactions of group I and group II intron ribozymes have been shown to facilitate catalysis of the ribozymes by stabilizing their catalytically active structures (Lambowitz & Perlman, 1990; Mohr et al., 1992; Weeks & Cech, 1995). Therefore, protein cofactors may have had a similar influence on the evolution of the diverse sequences of these ribozymes and, possibly, other functional RNAs in the corresponding ribonucleoprotein complexes, such as ribosomes (Noller et al., 1992) and mammalian spliceosomes (Moore et al., 1993). Our findings that RNase P holoenzyme variants requiring both RNA and protein components can be generated from catalytically active M1 RNA might have implications for how these modern RNP complexes evolved from the ancestral catalytic RNAs in the RNA world.

TABLE 2. Kinetic parameters of the cleavage reactions catalyzed by the selected ribozymes either in the absence (−) or presence (+) of C5 protein.[a]

Ribozyme	C5 protein	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m
M1 RNA	−	0.14	0.06	0.4
	+	0.18	2.4	13
R7	−	0.15	0.07	0.5
	+	0.23	4	17
C38	−	N.D.	N.D.	<0.005
	+	0.06	0.7	12
C41	−	N.D.	N.D.	<0.005
	+	0.15	2.9	19
C83	−	N.D.	N.D.	<0.005
	+	0.09	1.8	20
C102	−	N.D.	N.D.	<0.005
	+	0.16	0.04	0.3

[a]Reactions without and with C5 protein were incubated at 37 °C in buffer A (50 mM Tris, pH 7.5, 100 mM NH_4Cl, 100 mM $MgCl_2$, 4% PEG) and buffer B (50 mM Tris, pH 7.5, 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5–1 μM C5 protein), respectively. Kinetic parameters were measured under both single- and multiple-turnover conditions, as described in Materials and Methods. The values of kinetic parameters determined in different experiments exhibited a variation of less than 20%. N.D. denotes not determined.

Evolution of a M1 RNA sequence in the RNA world

Our results indicate that the complete conservation of the wild-type sequence (nt 347–361) was observed in ribozymes evolved in the absence of C5 protein, whereas a variety of sequences in addition to the wild-type sequence were found in ribozymes selected in the presence of C5 protein. This conservation (e.g., A_{352}) was required for the activity of the ribozyme alone, but not for the ribonucleoprotein complex. If our experiments have any relation to evolutionary events in the natural world, our results suggest that this conserved sequence of M1 RNA evolved independently of C5 protein and existed before the appearance of the protein cofactor. The selective advantage of this sequence in nature is likely only because an RNA world existed and this sequence was the first of many possible sequences to have evolved before the appearance of the protein cofactor, although one cannot exclude completely the possibility that the conservation of this

sequence is required for the optimal function of RNase P in vivo (e.g., intracellular stability and folding). This hypothesis is consistent with the observation that the contemporary sequence is highly conserved among all known RNase P RNAs from bacteria and archaea (Altman et al., 1993; Pace & Brown, 1995) and with the suggestion that M1 RNA-like activity existed and co-evolved with its ancestral tRNA substrates in the RNA world (Weiner & Maizels, 1987; Liu & Altman, 1994). This hypothesis is further in agreement with the recent observations that a part of this sequence plays a central role in catalysis by forming the P_{10} helix and possibly coordinating the catalytically active Mg^{2+} ions, and that only the wild-type sequence of M1 RNA was found when a part of this region was randomized and evolved in vitro in the absence of C5 protein (Hardt et al., 1995, 1996; Harris & Pace, 1995; Frank et al., 1996; Westhof et al., 1996).

We note that selection experiments conducted with present RNase P components may not necessarily reflect the chronology of events that involved the ancestors of these components. Because it is no longer possible to determine the components of the RNA world and chronicle the events during its transition to the RNA–protein world with complete certainty, evolution in vitro studies on M1 RNA and other RNase P RNA subunits and their respective protein cofactors may still provide a useful mean to investigate the nature of these macromolecules and their interactions in the past.

MATERIALS AND METHODS

Evolution in vitro

The double-stranded DNA templates coding for the RNA pool that contained randomized sequences in M1 RNA was constructed by PCR with pFL117 as the template, which contained the sequence coding for M1 RNA (Liu & Altman, 1995). In the PCR reaction, Oligo 101 (5′-GGAATTC**TAAT ACGACTCACTATAG**AAGCTGACCAGACAGTCGCC-3′) was used as the 5′ primer. Two 3′ primers were used: Oligo 118 (5′-CGCAAGGGAACCCCACCGAATACCTATGACCA TGATTACGCCAAGCTTCAGGTGAA ACTGACCGANNNN NNNNNNNNNNNCGTGGACAGTCATTCATCT-3′) and Oligo 119 (5′-GACCCCTGCCATCAACACGCGTCTGCGT TCGAAGGATCCGCAAGGGAACCCCACCGAATACCTA TGACCATGATTACGCC-3′). Oligo 101 contains the 5′ sequence of M1 RNA and, in addition, the promoter sequence for the bacteriophage T7 RNA polymerase, which is in bold type. Oligo 118 contains a randomized sequence indicated as $(N)_{15}$ and underlined. Oligo 119 contains the LS and the 5′ sequence of $tRNA^{Tyr}$. Ribozyme molecules were synthesized in vitro from the DNA template with T7 RNA polymerase and were further purified in denaturing 8% polyacrylamide gels prior to use. Transcription reactions were performed in the reaction mixture (40 mM Tris·HCl, pH 7.9, 6 mM $MgCl_2$, 10 mM DTT, 2 mM spermidine, 2 mM NTPs, 50 μCi α-[^{32}P]-GTP) and in the presence of T7 RNA polymerase for 16 h at 37 °C. C5 protein was prepared as described previously (Vioque et al., 1988).

The selection procedure is as follows (Fig. 2). Step 1: Ribozymes (10–100 pmol) that contained randomized mutations were annealed in the annealing buffer (50 mM Tris, pH 7.5, 100 mM NH_4Cl) to the same amount of EGS^{Tyr}, resulting in a $tRNA^{Tyr}$-like complex between the LS and EGS^{Tyr} (Yuan & Altman, 1994). Mg^{2+} ions, which are required for M1 RNA catalysis (Guerrier-Takada et al., 1983), were omitted in the annealing buffer to prevent cleavage. Step 2: Reaction buffer that contained $MgCl_2$ without or with C5 protein was added to the annealing mixture to allow the cleavage reaction to proceed. The reactions without and with C5 protein were performed in buffer A (50 mM Tris, pH 7.5, 100 mM NH_4Cl, 100 mM $MgCl_2$, 4% PEG) and buffer B (50 mM Tris, pH 7.5, 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5–1 μM C5 protein), respectively. Only active ribozymes recognized the $tRNA^{Tyr}$-like complex and cleaved at the LS region (Fig. 2). Incubation times for the first two cycles, cycles 3–4, cycles 5–6, and cycles 7–9 were 720, 120, 30, and 10 min, respectively. Steps 3–4: The cleaved RNA molecules were separated in denaturing gels and recovered. cDNA was synthesized from these RNA molecules with primer Oligo 119 in the presence of AMV reverse transcriptase and further amplified by PCR in the presence of primers Oligo 101 and Oligo 119. The PCR products contained the promoter sequence for T7 RNA polymerase and the LS region and were used for synthesis of the RNA molecules for the next round of selection (Fig. 2) (Liu & Altman, 1994). The entire procedure was repeated nine times until no further increase was apparent in the cleavage rate of the ribozyme population. cDNA that contained ribozyme sequences was cloned into plasmid vector pUC19. Sequencing analyses were conducted with a PCR Silver Sequencing kit (Promega Inc., Madison, Wisconsin).

Structural analyses of the selected ribozymes

All ribozymes were synthesized by T7 RNA polymerase (Promega Inc., Madison, Wisconsin) and purified in 4% polyacrylamide gels that contained 8 M urea. The 5′ termini of the ribozymes were radiolabeled with γ-[^{32}P]-ATP and polynucleotide kinase, whereas the 3′ ends were labeled in the presence of [^{32}P]pCp and T4 RNA ligase. Prior to structural analyses, ribozymes were incubated at 80 °C for 2 min and then cooled slowly to room temperature. The assembly of the RNase P holoenzyme was performed by mixing the ribozyme and C5 protein in a ratio of 1:20.

RNase mapping was conducted in buffer A (50 mM Tris, pH 7.5, 100 mM NH_4Cl, 100 mM $MgCl_2$, 4% PEG) (without C5 protein) and in buffer B (50 mM Tris, pH 7.5, 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 μM C5 protein) (with C5 protein), as described previously (Liu & Altman, 1994). Briefly, RNases were diluted and incubated with ribozymes (either 5′ or 3′ end-labeled) in the absence and presence of C5 protein. After different periods of time, the digestion reactions were stopped by adding 8 M urea. The alkaline treatment of ribozymes for preparation of a sequence ladder was performed as described elsewhere (Liu & Altman, 1994). The cleavage products were separated in urea–polyacrylamide gels. All RNases used in the study were purchased from Pharmacia Biotech. (Piscataway, New Jersey).

Fe(II)-EDTA mapping of the structures of M1 RNA and the variants were performed according to the procedures de-

scribed previously (Loria & Pan, 1996; Westhof et al., 1996). Briefly, the end-labeled ribozymes were incubated for 120 min in the reaction buffers supplemented with 2.5 mM DTT, 2 mM $(NH_4)_2Fe(SO_4)_2$, and 4 mM EDTA. The cleavage products were separated in denaturing polyacrylamide gels and quantitated by a phosphorimager (Molecular Dynamics, Sunnyvale, California). Cleavage reactions were performed in buffer A and B for reactions without and with C5 protein, respectively.

Furthermore, Fe(II)-EDTA cleavage of ribozymes was conducted in these buffers minus $MgCl_2$ to reveal the folded structure of the ribozymes in the absence of Mg^{2+} ions. Comparison of the mapping results obtained in the presence and absence of $MgCl_2$ revealed the regions that were folded in the presence of Mg^{2+} ions and protected from Fe(II)-EDTA cleavage. The protection factor at a nucleotide position was calculated as the ratio of the percentage of the cleavage product at this position in the absence of Mg^{2+} ions to that in the presence of the divalent ions (Loria & Pan, 1996; Westhof et al., 1996). Those regions in which the values of the protection factor were more than 1.25 were considered to be folded upon binding of Mg^{2+} ions and are shaded in Figure 3 (Loria & Pan, 1996; Westhof et al., 1996).

Kinetic analyses of the reactions catalyzed by the selected ribozymes

The reactions of the cleavage of substrates $ptRNA^{Tyr}$ and p4.5S RNA by different ribozymes in the absence and presence of C5 protein were performed as described previously (Liu & Altman, 1994). Reactions without and with C5 protein were incubated at 37 °C in buffer A and B, respectively. The cleavage products were separated on 8% denaturing gels, which were autoradiographed and/or dried and quantitated with a phosphorimager (Molecular Dynamics, Sunnyvale, California). Assays to determine kinetic parameters were performed as described previously (Liu & Altman, 1994). In brief, the cleavage of substrates was assayed at various concentrations of substrates, both above and below the K_m for each respective enzyme. Aliquots were withdrawn from reaction mixtures at regular intervals and analyzed in polyacrylamide–urea gels. Values of K_m and k_{cat} were obtained from Lineweaver–Burk double-reciprocal plots. The estimates of the values of k_{cat}/K_m for C38, C41, C83, and C102 in the absence of C5 protein were calculated by assaying the observed cleavage rate of these ribozyme variants under single-turnover conditions (Liu & Altman, 1994). The values of kinetic parameters determined in different experiments exhibited a variation of less than 20%.

ACKNOWLEDGMENTS

We are grateful to S. Lu, I. Mohr, C. Guerrier-Takada, V. Gopolan, and Y. Yuan for their invaluable advice, critical reading of the manuscript, and numerous reagents (C5 protein and *E. coli* NHY322(rnpA$_{49}$) strain). A.F.K. is a recipient of Hillel and Rose Levine Graduate Fellowship at the University of California at Berkeley. This work is supported by a grant from University of California Cancer Research Coordinating Committee (F.L.), and NIH grants (S.A. and F.L.).

Manuscript accepted without revision March 24, 1997

REFERENCES

Altman S, Kirsebom L, Talbot SJ. 1993. Recent studies of RNase P. *FASEB J 7*:7–15.

Bartel DP, Szostak JW. 1993. Isolation of new ribozymes from a large pool of random sequences. *Science 261*:1411–1418.

Bothwell AL, Garber RL, Altman S. 1976. Nucleotide sequence and in vitro processing of a precursor molecule to *Escherichia coli* 4.5 S RNA. *J Biol Chem 251*:7709–7716.

Ekland EH, Bartel DP. 1996. RNA-catalyzed RNA polymerization using nucleoside triphosphates. *Nature 382*:373–376.

Frank DN, Ellington AE, Pace NR. 1996. In vitro selection of RNase P RNA reveals optimized catalytic activity in a highly conserved structural domain. *RNA 2*:1179–1188.

Gilbert W. 1986. The RNA world. *Nature 319*:618.

Gold L, Allen P, Binkley JDB, Schneider D, Eddy SR, Tuerk C, Green L, MacDougal S, Tasset D. 1993. RNA: The shape of things to come. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 497–509.

Gopalan V, Talbot SJ, Altman S. 1995. RNA–protein interactions in RNase P. In: Nagai K, Mattaj I, eds. *RNA–protein interactions*. Oxford: Oxford University Press. pp 103–127.

Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell 35*:849–857.

Guerrier-Takada C, Lumelsky N, Altman S. 1989. Specific interactions in RNA enzyme–substrate complexes. *Science 246*:1578–1584.

Haas ES, Armbruster DW, Vucson BM, Daniels CJ, Brown JW. 1996. Comparative analysis of ribonuclease P RNA structure in Archaea. *Nucleic Acids Res 24*:1252–1259.

Haas ES, Brown JW, Pitulle C, Pace NR. 1994. Further perspective on the catalytic core and secondary structure of ribonuclease P RNA. *Proc Natl Acad Sci USA 91*:2527–2531.

Haas ES, Morse DP, Brown JW, Schmidt FJ, Pace NR. 1991. Long-range structure in ribonuclease P RNA. *Science 254*:853–856.

Hardt WD, Erdmann VA, Hartmann RK. 1996. Rp-deoxy-phosphorothioate modification interference experiments identify 2′-OH groups in RNase P RNA that are crucial to tRNA binding. *RNA 2*:1189–1198.

Hardt WD, Warnecke JM, Erdmann VA, Hartmann RK. 1995. Rp-phosphorothioate modifications in RNase P RNA that interfere with tRNA binding. *EMBO J 14*:2935–2944.

Harris ME, Nolan JM, Malhotra A, Brown JW, Harvey SC, Pace NR. 1994. Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of ribonuclease P RNA. *EMBO J 13*:3953–3963.

Harris ME, Pace NR. 1995. Identification of phosphates involved in catalysis by the ribozyme RNase P. *RNA 1*:210–218.

Illangasekare M, Sanchez G, Nickles T, Yarus M. 1995. Aminoacyl-RNA synthesis catalyzed by an RNA. *Science 267*:643–647.

Joyce GF. 1989. RNA evolution and the origins of life. *Nature 338*:217–224.

Joyce GF. 1992. Directed molecular evolution. *Sci Am 267*:90–97.

Kirsebom LA, Baer MF, Altman S. 1988. Differential effects of mutations in the protein and RNA moieties of RNase P on the efficiency of suppression by various tRNA suppressors. *J Mol Biol 204*:879–888.

Knap AK, Wesolowski D, Altman S. 1990. Protection from chemical modification of nucleotides in complexes of M1 RNA, the catalytic subunit of RNase P from *E. coli*, and tRNA precursors. *Biochimie 72*:779–790.

Lambowitz AM, Perlman PS. 1990. Involvement of aminoacyl tRNA synthetases and other proteins in Group I and Group II intron splicing. *Trends Biochem Sci 15*:440–444.

Latham JA, Cech TR. 1989. Defining the inside and outside of a catalytic RNA molecule. *Science 245*:276–282.

Lehman N, Joyce GF. 1993. Evolution in vitro of an RNA enzyme with altered metal dependence. *Nature 361*:182–185.

Liu F, Altman S. 1994. The differential evolution of substrates for an RNA enzyme in the presence and absence of its protein cofactor. *Cell 77*:1083–1100.

Liu F, Altman S. 1995. Inhibition of viral gene expression by the catalytic RNA subunit of RNase P from *Escherichia coli*. *Genes & Dev 9*:471–480.

Loria A, Pan T. 1996. Domain structure of the ribozyme from eubacterial ribonuclease P. *RNA* 2:551–563.

Maizels N, Weiner AM. 1994. Phylogeny from function: Evidence from the molecular fossil record that tRNA originated in replication, not translation. *Proc Natl Acad Sci USA 91*:6729–6734.

Mohr G, Zhang A, Gianelos JA, Belfort M, Lambowitz AM. 1992. The neurospora CYT-18 protein suppresses defects in the phage T4 td intron by stabilizing the catalytically active structure of the intron core. *Cell 69*:483–494.

Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNA by spliceosomes. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–357.

Noller HF, Hoffarth V, Zimniak L. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science 256*:1416–1419.

Pace NR, Brown JW. 1995. Evolutionary perspective on the structure and function of ribonuclease P, a ribozyme. *J Bacteriol 177*:1919–1928.

Szostak JW, Ellington AD. 1993. In vitro selection of functional RNA sequences. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 511–533.

Talbot SJ, Altman S. 1994. Gel retardation analysis of the interaction between C5 protein and M1 RNA in the formation of the ribonuclease P holoenzyme from *Escherichia coli*. *Biochemistry 33*:1399–1405.

Vioque A, Arnez J, Altman S. 1988. Protein-RNA interactions in the RNase P holoenzyme from *Escherichia coli*. *J Mol Biol 202*:835–848.

Weeks KM, Cech TR. 1995. Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5′ splice site domain. *Cell 82*:221–230.

Weiner AM, Maizels N. 1987. tRNA-like structures tag the 3′ ends of genomic RNA molecules for replication: Implications for the origin of protein synthesis. *Proc Natl Acad Sci USA 84*:7383–7387.

Westhof E, Altman S. 1994. Three-dimensional working model of M1 RNA, the catalytic RNA subunit of ribonuclease P from *Escherichia coli*. *Proc Natl Acad Sci USA 91*:5133–5137.

Westhof E, Wesolowski D, Altman S. 1996. Mapping in three dimensions of regions in a catalytic RNA protected from attack by an Fe(II)-EDTA reagent. *J Mol Biol 258*:600–613.

Wilson C, Szostak JW. 1995. In vitro evolution of a self-alkylating ribozyme. *Nature 374*:777–782.

Yuan Y, Altman S. 1994. Selection of guide sequences that direct efficient cleavage of mRNA by human ribonuclease P. *Science 263*:1269–1273.