The EMBO Journal vol.12 no.6 pp.2567-2574, 1993

Essential nucleotide sequences and secondary structure elements of the hairpin ribozyme

Alfredo Berzal-Herranz, Simpson Joseph, Bharat M.Chowrira, Samuel E.Butcher and John M.Burke[1]

Department of Microbiology and Molecular Genetics, Markey Center for Molecular Genetics, The University of Vermont, Health Science Complex, Burlington, VT 05405, USA

[1]To whom correspondence should be addressed

Communicated by F.Eckstein

***In vitro* selection experiments have been used to isolate active variants of the 50 nt hairpin catalytic RNA motif following randomization of individual ribozyme domains and intensive mutagenesis of the ribozyme–substrate complex. Active and inactive variants were characterized by sequencing, analysis of RNA cleavage activity *in cis* and *in trans*, and by substrate binding studies. Results precisely define base-pairing requirements for ribozyme helices 3 and 4, and identify eight essential nucleotides (G_8, A_9, A_{10}, G_{21}, A_{22}, A_{23}, A_{24} and C_{25}) within the catalytic core of the ribozyme. Activity and substrate binding assays show that point mutations at these eight sites eliminate cleavage activity but do not significantly decrease substrate binding, demonstrating that these bases contribute to catalytic function. The mutation $U_{39}C$ has been isolated from different selection experiments as a second-site suppressor of the down mutants $G_{21}U$ and $A_{43}G$. Assays of the $U_{39}C$ mutation in the wild-type ribozyme and in a variety of mutant backgrounds show that this variant is a general up mutation. Results from selection experiments involving populations totaling more than 10^{10} variants are summarized, and consensus sequences including 16 essential nucleotides and a secondary structure model of four short helices, encompassing 18 bp for the ribozyme–substrate complex are derived.**

Key words: catalytic RNA/*in vitro* selection/ribozyme/RNA structure

Introduction

In vitro selection techniques represent powerful tools for the analysis of structure–function relationships of RNA molecules, greatly facilitating the elucidation of intramolecular and intermolecular interactions (for review, see Szostak, 1992; Burke and Berzal-Herranz, 1993). Such selection experiments (also termed directed molecular evolution) have been particularly useful for studying RNA enzymes. Initially, selection was applied to group I introns (Green *et al.*, 1990; Robertson and Joyce, 1990; Beaudry and Joyce, 1992). Recently, *in vitro* selection technology has led to the identification of catalytic RNA motifs not known to exist in nature (Pan and Uhlenbeck, 1992). We have developed a powerful *in vitro* selection scheme for analysis of the sequence and structural requirements of the hairpin ribozyme (Berzal-Herranz *et al.*, 1992).

The hairpin ribozyme is derived from the minus strand of the satellite RNA of tobacco ringspot virus, and catalyzes a sequence-specific and reversible RNA cleavage reaction, giving rise to products containing 5′-hydroxyl and 2′,3′-cyclic phosphate termini (Buzayan *et al.*, 1986). Minimum energy folding, limited mutational analysis and a very modest amount of phylogenetic data (three sequences) have been used to propose secondary structure models for the complex between the hairpin ribozyme and its substrate (Figure 1; Feldstein *et al.*, 1989; Hampel and Tritz, 1989; Haseloff and Gerlach, 1989; Hampel *et al.*, 1990; Rubino *et al.*, 1990).

In order to overcome the absence of extensive phylogenetic data, we chose to construct an artificial phylogenetic database through the development and application of an *in vitro* method that permits the selective replication of catalytically proficient hairpin ribozymes from a highly complex pool of variants, the great majority of which are catalytically incompetent (Berzal-Herranz *et al.*, 1992). To be selected as active, ribozyme variants are required to catalyze sequential cleavage and ligation reactions. These two RNA-catalyzed reactions result in the covalent attachment of a new primer binding sequence to the 3′ end of the ribozyme–substrate complex, and permit the selective replication of active molecules by a three-step RNA replication scheme involving reverse transcription,

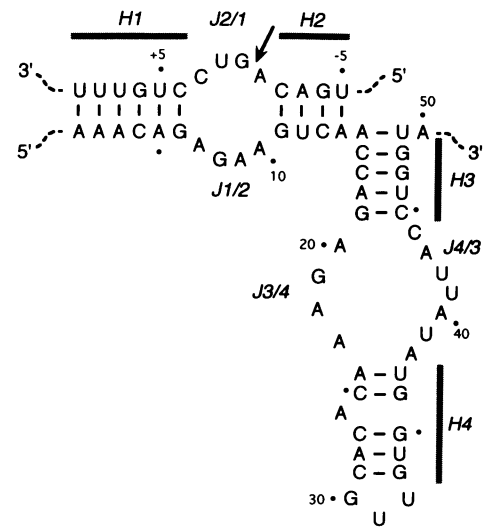

Fig. 1. Initial secondary structure model of the hairpin ribozyme. Ribozyme sequences are numbered 1–50; substrate sequences are −5 to +9. Arrow indicates cleavage–ligation site. Helices are numbered H1–H4. RNA segments joining helices are J1/2, J2/1, J3/4 and J4/3. For the self-cleaving molecules employed in *in vitro* selection, the 3′ end of the ribozyme is covalently linked to the 5′ end of the substrate with a short linker sequence (Berzal-Herranz *et al.*, 1992; Joseph *et al.*, 1993).

polymerase chain reaction, and transcription. In this selection procedure, the researcher has complete control over mutation sites, mutation frequency, and the stringency of the selection. We have used this procedure in the analysis of the substrate selection rules of the ribozyme (Berzal-Herranz *et al.*, 1992; Joseph *et al.*, 1993).

Here, we present the results of *in vitro* selection experiments aimed towards elucidating structure–function relationships within the catalytic domain of the hairpin ribozyme itself. This work completes the initial analysis of the entire ribozyme–substrate complex by *in vitro* selection. Together with our previously published work (Chowrira and Burke, 1991; Chowrira *et al.*, 1991; Berzal-Herranz *et al.*, 1992; Jospeh *et al.*, 1993), these results define a secondary structure containing four short helices, and serve to identify 16 nucleotides essential for efficient catalytic function.

Results

Strategies for in vitro selection

Three different strategies for *in vitro* selection were used in the experiments described here. First, to identify essential sequences and secondary structure elements in a region of interest, for example a putative helix, we randomized that region by introducing an equal frequency of G, A, U and C at every position. Although clones of active molecules are primarily of interest in this strategy, we also sampled the population of inactive molecules to confirm that all possible bases were introduced at each site that was randomized. Second, to identify point mutations that inactivate the ribozyme, we mutagenized all sites within the ribozyme–substrate complex at a low frequency (an average of one mutation per molecule) and selected inactive molecules. Third, to identify compensatory mutations within the complex, we mutagenized all sites in the ribozyme and substrate at a high frequency (an average of five mutations per molecule) and selected active molecules, or randomized several domains (for example all non-helical nucleotides) and selected active molecules.

Six essential bases within internal loop segments J1/2 and J3/4

J1/2 and J3/4 represent two internal loop segments separating helices 1 and 2, and helices 3 and 4, respectively (Figure 1). Interestingly, both J1/2 and J3/4 have the same sequence in the natural isolate, 5′-AGAA-3′. In independent experiments, J1/2 and J3/4 were randomized by introducing equal quantities of each of the four deoxyribonucleoside phosphoramidites during chemical synthesis of the DNA templates. Transcription of these templates gave two pools, each containing 256 sequences. For each selection experiment, both catalytically active and inactive molecules were selected, sequenced and characterized.

Nineteen clones obtained from the cDNA library of active J1/2 variants and 15 from the library of inactive variants were sequenced. Results show a strict sequence requirement in this region (Figure 2). Among the active clones, all possible variations at position 7 are found, while each of the active clones had the wild-type sequence G_8 A_9 A_{10} (Figure 2A). Sequencing of the inactive clones confirmed that each of the four nucleotides was present at every position in the initial population before selection (data not shown). Next, self-cleavage assays were carried out as described in Materials and methods. Each of the sequences selected in the active pool showed efficient self-cleavage activity (Figure 2B), while no activity was observed for any of the tested inactive clones. Note that a number of the inactive molecules contained single base substitutions at positions 8, 9 or 10. We confirmed the selection and self-cleavage results by synthesizing *trans*-acting ribozymes with the single base substitutions G_8U and G_8A, and then testing their activity in a standard *trans*-cleavage assay (Figure 2C). Consistent with the selection and self-cleavage results, the G_8U and G_8A mutations strongly inhibited RNA cleavage activity. From these results, we conclude that G_8, A_9 and A_{10} are required for ribozyme activity, while the identity of the base at position 7 is relatively unimportant.

In a similar manner, 23 clones from the J3/4 active library and 10 clones from the inactive library were characterized. The results are nearly identical to those obtained from the J1/2 region. Each of the 23 active clones has the wild-type sequence G_{21} A_{22} A_{23}, but all possible variations were identified at position 20 (Figure 2A). Active clones self-cleave efficiently, but no activity is observed for clones selected as inactive (Figure 2B). Thus, G_{21}, A_{22} and A_{23} are required for ribozyme activity. At position 20, all four nucleotides result in highly active ribozymes.

To complement these regional randomization experiments, *in vitro* selection was applied to pools of variant molecules resulting from scattering mutations randomly throughout the ribozyme–substrate complex at a frequency of 1–5 mutations per molecule (Materials and methods; Joseph *et al.*, 1993). Single base substitutions at each of the eight positions within J1/2 and J3/4 were isolated. Self-cleavage assays confirm that ribozymes containing point mutations at each of the invariant positions in J1/2 and J3/4 (G_8, A_9, A_{10}, G_{21}, A_{22} and A_{23}) are inactive, while point mutations at nonconserved positions (A_7 and A_{20}) are active (Figure 2 and data not shown).

Search for tertiary interactions between ribozyme internal loops

We examined the possibility of tertiary contacts between the bases of J1/2 and J3/4 by randomizing all eight nucleotides simultaneously and selecting active molecules from the resulting theoretical pool of 4^8 (65 000) variant ribozymes. Results from the sequence analysis of 12 active clones were identical to the results obtained from the independent selection of each region described above. No variation was observed at any of the six positions previously identified as invariant (data not shown). Similar results were obtained when each of the four internal loop segments in the ribozyme–substrate complexes (J1/2, J2/1, J3/4, J4/3) were randomized in all possible pairwise combinations. We conclude that these results do not support the existence of tertiary interactions between the bases studied. However, these experiments do not rule out the existence of such tertiary interactions, or other interactions involving these sites, for example base–sugar or base–phosphate contacts (see Discussion).

A mutation in J4/3 ($U_{39}C$) is a second-site suppressor of a mutation in J3/4 ($G_{21}U$)

Following the selection of active variants from a mutagenized pool containing an average of five mutations per molecule, we isolated a ribozyme clone containing two mutations

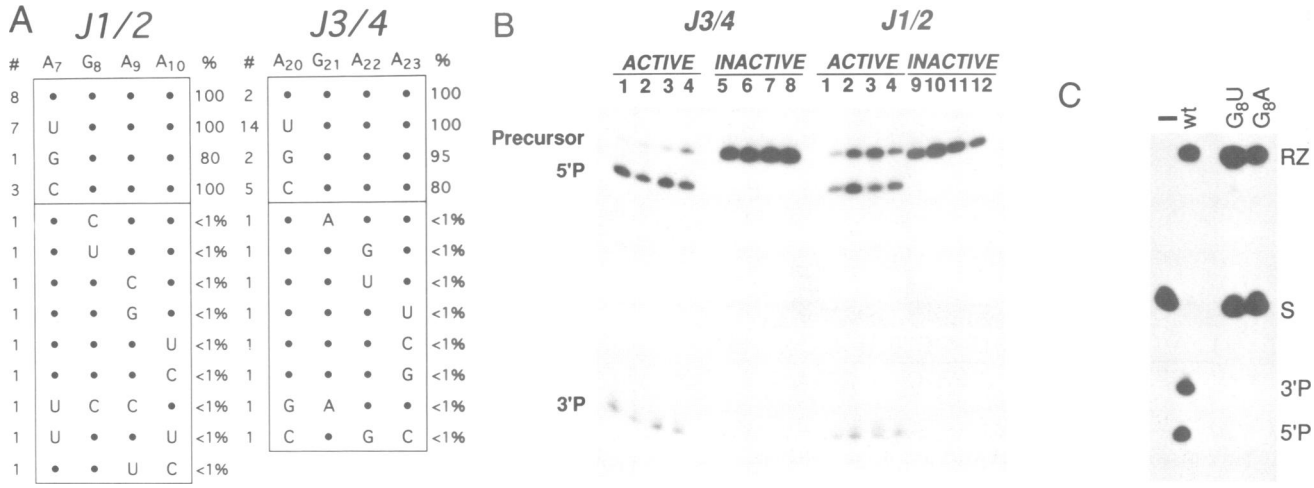

Fig. 2. Essential sequences in J1/2 and J3/4. (**A**) Variants obtained through *in vitro* selection of the J1/2 and J3/4 regions. In separate experiments, active and inactive molecules were selected after randomization of J1/2 and J3/4 and following mutagenesis of the entire ribozyme−substrate complex, as described in the text. Sequences from wild-type are shown at top. Sequences of all the active variants (above dividing line) and some of the inactive (below dividing line) variants obtained from the selection experiments are shown. Dot indicates presence of wild-type nucleotide at that position. Frequency of each sequence analyzed is shown at left, and relative self-cleavage activity is shown at right of each sequence. (**B**) Self-cleavage assay. Self-cleavage assay for individual active and inactive clones, performed as described (Berzal-Herranz *et al.*, 1992). Lane 1, wild-type sequence (AGAA). Lanes 2−4 are UGAA, GGAA and CGAA in J1/2 or J3/4 respectively. Lanes 5−8 are AAAA, GAAA, CGGC and AGAC, respectively, in J3/4. Lanes 9−12 are AGAC, UCCA, UGAU and AGUC, respectively, in J1/2. Precursor, full length uncut 79 nt RNA (Berzal-Herranz *et al.*, 1992). 5′P, 5′ cleavage product; 3′P, 3′ cleavage product. (**C**) *Trans*-cleavage assay of ribozymes containing the single substitution G_8U or G_8A. Internally labeled substrate RNA (240 nM) was incubated with internally labeled ribozyme (60 nM) at 37°C for 30 min as described in Materials and methods. RZ, ribozyme; S, 17 nt RNA substrate; minus indicates no ribozyme; wt, wild-type ribozyme and substrate.

$G_{21}U$ and $U_{39}C$. Self-cleavage assays showed that the double mutant $G_{21}U$:$U_{39}C$ had substantially greater activity than the single base mutant $G_{21}U$.

This result was confirmed and extended using *trans*-cleavage reactions using the two-piece ribozyme system previously described (Chowrira and Burke, 1992). To determine the pattern of suppression, all 16 possible dinucleotide combinations at positions 21 and 39 were tested (Figure 3). Results show that all three single base substitutions of G_{21} ($G_{21}A$, $G_{21}U$ and $G_{21}C$) are strong down mutations, significantly decreasing the extent of cleavage. Two of the three single base mutations of U_{39} ($U_{39}A$ and $U_{39}C$) slightly increase the extent of cleavage under these conditions, while the $U_{39}G$ change significantly decreases the extent of cleavage. Among the two-base substitutions, we found that the $U_{39}C$ mutation increases the extent of cleavage of ribozymes containing the $G_{21}A$, $G_{21}U$ and $G_{21}C$ down mutations by approximately 3-fold. However, suppression of the decreased activity of these G_{21} mutations by $U_{39}C$ was only partial; the cleavage extent did not approach the levels of the wild-type sequence. We have found that the activating effect of the $U_{39}C$ mutation is greatest at low $MgCl_2$ concentrations ($\leq$10 mM; data not shown). These results show that $U_{39}C$ is a nonspecific suppressor of G_{21} mutations.

Helix 3: five base pairs are important for optimal activity

Helix 3 (Figure 1), consisting of five proposed base pairs, was analyzed by randomizing the ribozyme sequence corresponding to nucleotides 15−19 and 45−50, to generate a theoretical initial pool of 4^{11} (4×10^6) sequence variants. In addition, a number of interesting variants were obtained from the mutagenesis experiments described above. Clones containing cDNAs from the active and inactive pools were characterized. Representative results are shown in Figure 4. All active molecules selected could form a stable helix 3, and most could form a 5 bp stem, as predicted for the native ribozyme. In contrast, molecules from the inactive pool selected following randomization show no significant base-pairing potential (see e.g. clone 22, Figure 4A).

At several positions within helix 3, single mismatches and G•U or A•C wobble pairs lead to a significant decrease in self-cleavage activity (Figure 4). However, point mutations in helix 3 do not generally inhibit ribozyme activity as strongly as do point mutations at essential sites within the internal loop sequences. These results confirm the importance of five base pairs in helix 3. Because of the absence of sequence conservation among the active clones, it is likely that the hairpin ribozyme simply requires five stable base pairs in helix 3. In addition, the results show no nucleotide preference at position 50, the unpaired 3′ terminal nucleotide of the ribozyme.

Helix 4 requires only three base pairs

Among the research groups studying the hairpin ribozyme, no clear consensus exists concerning the number of base pairs in helix 4. Indeed, some models do not show helix 4 at all (Haseloff and Gerlach, 1989; Feldstein *et al.*, 1990). In our own work, we have assumed that helix 4 consists of five base pairs, with A_{26} as a single base bulge (Figure 1). Helix 4 was analyzed by *in vitro* selection following randomization of positions 24−29 amd 33−37, and by mutagenesis as described above.

Analysis of 16 active clones and 10 inactive clones from the cDNA libraries indicates that stable base pairs at only three positions (between nucleotides 27•35, 28•34 and 29•33) are sufficient for efficient self-cleavage and ligation (Figure 5). The selection results provide no support for sequence conservation at any of these six positions. Contrary to our secondary structure model, there is no requirement for base pairing between nucleotides A_{24}•U_{37} and

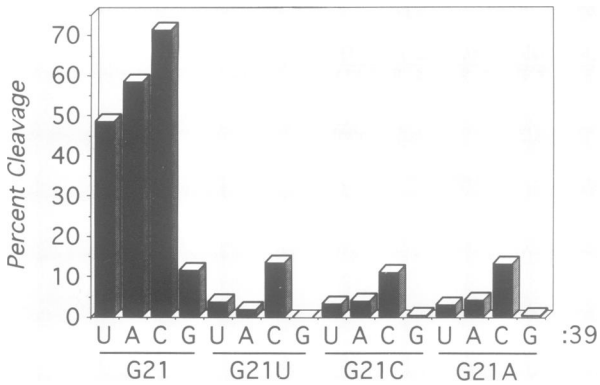

Fig. 3. $U_{39}C$ is an up mutation. Results of *trans*-cleavage assays using the two-piece ribozyme construct (Chowrira and Burke, 1992) are shown. In the 5′ segment of the ribozyme, G_{21} was changed to U, C or A. In the 3′ segment, U_{39} was changed to A, C or G. All 16 dinucleotide combinations were analyzed for substrate cleavage activity. Cleavage reactions were carried out by incubating 0.01 μM ribozyme with 0.05 μM substrate in a buffer containing 40 mM Tris−HCl (pH 7.5) and 5 mM $MgCl_2$ at 37°C for 30 min.

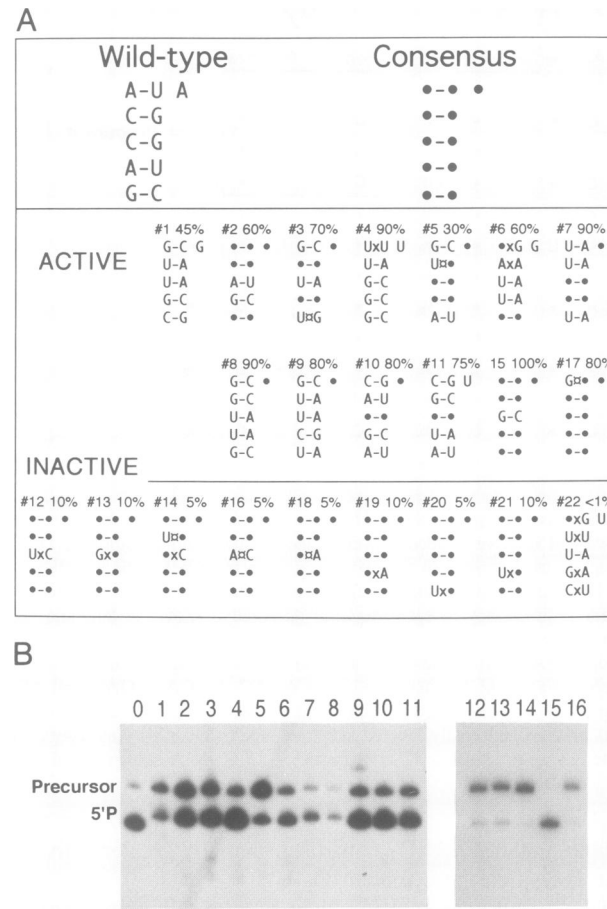

Fig. 4. Analysis of helix 3. (**A**) Sequences of active and inactive clones. All active helix 3 variants that were analyzed are shown, and a representative sample of the sequences obtained from the analysis of inactive clones is shown. Dots indicate wild-type sequence; letters denote sequence variations. Relative self-cleavage activity is indicated at the top of each individual sequence. Original sequence and derived consensus sequences are shown. (**B**) Self-cleavage assay of active and inactive molecules. Lane numbers correspond to sequences shown in (A). Lanes 12−16 are a series of variants showing single or double substitutions at the third base pair of helix 3, C_{17}●G_{47}. 0, wild-type self-cleaving molecule.

C_{25}●G_{36}. However, there is a very strong sequence bias for the wild-type sequence at A_{24} and C_{25} (Figure 5A). From random mutagenesis, we isolated single base substitutions at each of these positions in the inactive pool (clones 20 and 21; $A_{24}C$ and $C_{25}G$; Figure 5A), Self-cleavage assays showed that these two variants were completely inactive (data not shown). We further confirmed these results in a *trans*-cleavage assay after synthesizing ribozymes containing single base mutations ($A_{24}U$ and $C_{25}G$) and putative compensatory base substitutions ($A_{24}U$:$U_{37}A$ and $C_{25}G$:$G_{36}C$). Mutants $A_{24}U$ and $C_{25}G$ were catalytically inactive and the activity was not restored either by a complementary mutation to restore the putative pairing, or by increasing $MgCl_2$ concentration in the reaction (Figure 5C). Additionally, a clone containing a single base deletion of U_{37} (clone 1, Figure 5A) was isolated, and shows no reduction in self-cleavage activity (Figure 5B), while deletion of A_{24} (clone 19) completely inactivates the ribozyme (data not shown). These results indicate that A_{24} and C_{25} are essential for activity, but that base pairs between positions 24●37 or 25●36 are not necessary for ribozyme activity.

Two clones containing a deletion of A_{26} were isolated (clone 22, Figure 5A). The $A_{26}\Delta$ variant is inactive in a self-cleavage assay (data not shown), but a single base substitution at this position ($A_{26}C$) does not reduce activity (B.M.Chowrira, unpublished data), demonstrating that a nucleotide at position 26 is necessary for efficient ribozyme activity. We have also isolated several mutants in the ribozyme's loop (positions 30−32). These mutations do not affect self-cleavage activity (data not shown). Together with previous results involving a two-piece ribozyme in which this loop is eliminated (Chowrira and Burke, 1992), these results indicate that neither the sequence nor the presence of the loop itself is important for activity.

Mutations in essential bases do not affect formation of the ribozyme−substrate complex

The experiments above serve to identify eight essential nucleotides (G_8, A_9, A_{10}, G_{21}, A_{22}, A_{23}, A_{24} and C_{25}) within the ribozyme. Are these bases important for substrate binding, or do they function in an essential step that follows binding? We have shown that these eight bases do not form part of the secondary structure framework of the ribozyme, but could conceivably make important tertiary contacts necessary for substrate binding. To address this issue, we used a mobility shift assay that allowed us to directly visualize and quantitate complex formation using *trans*-acting ribozymes and the non-cleavable substrate analog $G_{+1}A$ (Chowrira *et al.*, 1991).

Ribozymes containing mutations that eliminated cleavage activity (point mutants G_8A, $A_{22}U$ and $C_{25}G$ and the double mutants $A_{24}U$:$U_{37}A$ and $C_{25}G$:$G_{36}C$) were tested for binding activity. Mutant $G_{21}U$, which shows very low activity, was also tested. Results show that all the mutant ribozymes have binding affinities similar to or slightly lower than that of the native ribozyme (Figure 6). In these mutants, decreases in binding activity range from 10 to 50%, while decreases in catalytic activity are very much more dramatic (see above). Therefore, we conclude that each of these mutations probably acts to inhibit an essential step that occurs after substrate binding during the catalytic cycle.

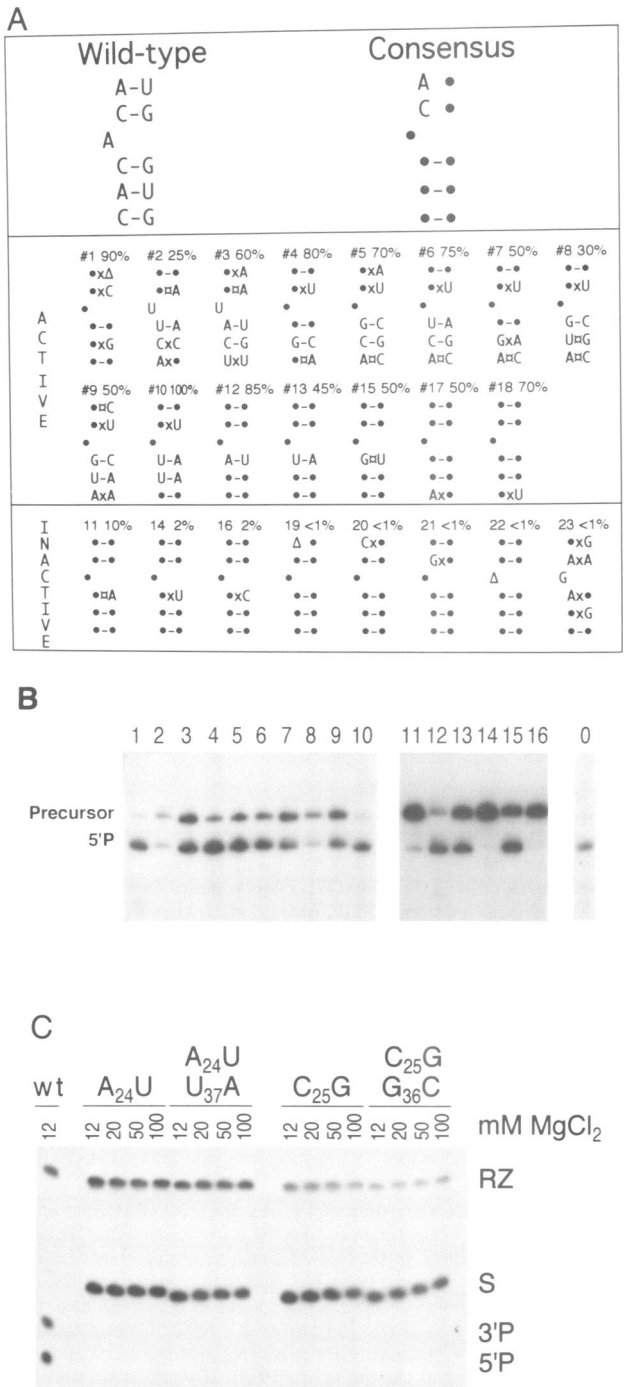

Fig. 5. Analysis of helix 4. (**A**) Sequences of active and inactive clones. All active helix 4 variants that were analyzed are shown, and a representative sample of the sequences obtained from the analysis of inactive clones is shown. Dots indicate wild-type sequence; letters denote sequence variations. Relative self-cleavage activity is indicated at the top of each individual sequence. Original sequence and derived consensus sequences are shown. (**B**) Self-cleavage assay of active and inactive molecules. Lane numbers correspond to sequences shown in panel A. Lanes 11–16 are a series of variants showing single or double substitutions at the third base pair of helix 4, C_{27}●G_{35}. 0, wild-type self-cleaving molecule. (**C**) Pairing between positions 24–37 and 25–36 is not required for activity. Results of a *trans*-cleavage assay using ribozymes containing single subsitutions ($A_{24}U$ and $C_{25}G$) or double base substitutions ($A_{24}U$:$U_{37}A$ and $C_{25}G$:$G_{36}C$) are shown. Internally labeled RNA substrate (0.3 μM) and ribozyme (0.1 μM) were incubated at 37°C for 1 h as described in Materials and methods. $MgCl_2$ concentrations were varied as indicated.

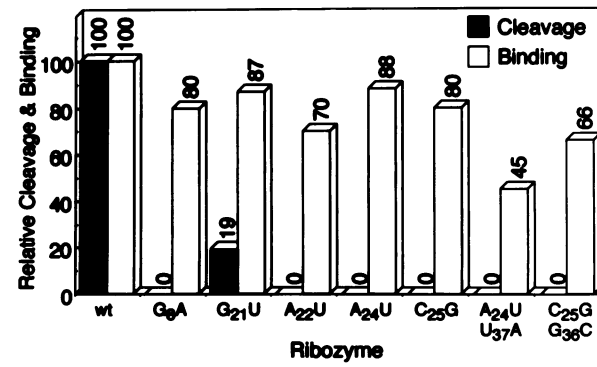

Fig. 6. Mutations that eliminate cleavage retain substrate binding activity. Results of cleavage assays (black bars) and binding assays (white bars) are compared for point mutations in J1/2 and J3/4. Binding assays use the noncleavable substrate analog $G_{+1}A$ (Chowrira *et al.*, 1991), and were carried out as described (Chowrira and Burke, 1991).

Discussion

The systematic *in vitro* analysis of the hairpin ribozyme described here and elsewhere (Chowrira *et al.*, 1991; Berzal-Herranz *et al.*, 1992; Joseph *et al.*, 1993) has led to the identification of nucleotides and structural motifs essential for cleavage and ligation. These results serve rigorously to define a secondary structure model for the ribozyme–substrate complex (Figure 7). In the course of the selection experiments, each of the base pairs shown in Figure 7 has been proven by testing all possible dinucleotide combinations and all possible nearest neighbors, and can be said to constitute an 'artificial phylogeny'. Because of the large number of active and inactive variants that we surveyed, our confidence in the model is high, and is comparable to that obtained by the several extensive natural phylogenies, i.e. transfer RNA (Levitt, 1969), ribosomal RNA (e.g. see Noller and Woese, 1981), group I introns (Michel *et al.*, 1989), and ribonuclease P RNA (Brown *et al.*, 1991). The fact that we have a positive selection for inactive as well as active ribozymes has allowed us to isolate and characterize point mutations that inactivate the ribozyme. In this regard, our *in vitro* selection results are in some ways comparable with those obtained from the combination of natural phylogenies and mutagenesis.

We have identified 15 essential nucleotides within the ribozyme (G_8, A_9, A_{10}, G_{11}, G_{21}, A_{22}, A_{23}, A_{24}, C_{25}, A_{38}, A_{40}, U_{41}, U_{42}, A_{43} and C_{44}), in addition to one essential nucleotide in the substrate (G_{+1}). At each of these 16 sites, single base substitutions bring about very significant reductions (at least 10-fold) in catalytic efficiency. In addition, we find sequence preferences at several sites within the substrate-binding domain, at positions 6 and 12 in the ribozyme, and +4, +3, +2, −2 and −3 in the substrate sequences (Chowrira *et al.*, 1991; Joseph *et al.*, 1993; Figure 7). At these sites, some but not all base substitutions reduce activity significantly. Taken together, the results obtained are consistent with the limited phylogenetic data available for this ribozyme (Rubino *et al.*, 1990), and with the limited mutational data (Hampel *et al.*, 1990).

With the sole exception of G_{11}, all of the essential bases lie within the nonhelical segments of the ribozyme–substrate complex, and their structure is, at this point, completely unknown. The biochemical function of the essential nucleotides within the ribozyme is not known. *A priori*, there

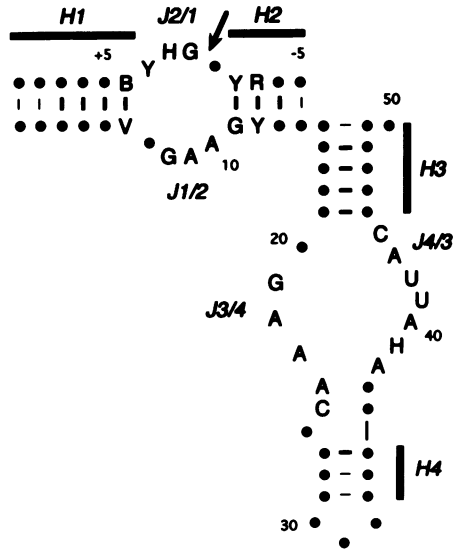

Fig. 7. Sequence and structural requirements for the hairpin ribozyme. Requirements for efficient RNA cleavage and ligation are shown, as determined by *in vitro* selection and by mutational analysis (Chowrira and Burke, 1991; Chowrira *et al.*, 1991; Berzal-Herranz *et al.*, 1992; Joseph *et al.*, 1993; this work). Key: ● represents A, C, G or U; **V** is A, C or G; **Y** is U or C; **R** is A or G; **B** is U, C or G; **H** is A,

are several possible reasons for the loss of activity associated with substitution of each of these essential bases.

First, it is possible that mutations could prevent substrate binding. However, the results of binding assays (Figure 6) show that each mutant that lacks catalytic activity retains most of its binding activity; the observed reductions in binding activity are much smaller than the decreases in catalytic activity. A second, more likely explanation for the loss of catalytic activity is that certain base substitutions might disrupt tertiary contacts that are required for catalysis but not the initial binding of substrate. These tertiary contacts could include interactions between the essential base and other bases, sugars, or phosphates within the ribozyme or substrate. Some mutations within critical sequences of the *Tetrahymena* group I ribozyme have been shown to act by this mechanism (Williamson *et al.*, 1987; Michel *et al.*, 1989, 1992; Michel and Westhof, 1990; Pyle *et al.*, 1992).

Third, substitution of the critical bases could disrupt the coordination of essential divalent cations within the complex. Recently, we have shown that the ribozyme contains at least two cation binding sites essential for catalysis (Chowrira *et al.*, 1993). If functional groups on a base are important for magnesium coordination, then base substitutions at these sites could strongly inhibit activity. Alternatively, a divalent cation could function to stabilize the tertiary structure of the ribozyme without directly participating in catalysis. Substitution of bases coordinating such a cation could, then, result in loss of activity due to disrupted tertiary contacts, as discussed above.

Fourth, it is possible that base substitution could result in activity loss if the base played a direct role in active site chemistry. Results from the use of guanosine analogs combined with molecular modeling work have led us to propose such a function for the 2-amino group of G_{+1}, the essential substrate guanosine immediately to the 3′ side of the scissile bond (Chowrira *et al.*, 1991).

In this work, we isolated the mutation $U_{39}C$ as a second-site suppressor of the mutation $G_{21}U$, and showed that $U_{39}C$ acts to increase the activity of ribozymes with any base, including the wild-type guanine, at position 21. Previously, we showed that $U_{39}C$ acts as a second-site suppressor of a different mutation in J4/3, $A_{43}G$ (Berzal-Herranz *et al.*, 1992). The most likely explanation of these results is that $U_{39}C$ functions as a general up mutation. It is possible that the $U_{39}C$ mutation may act generally to enhance the cleavage efficiency of all hairpin ribozyme constructs.

We are optimistic that our *in vitro* selection method will allow us to identify functionally important base–base contacts within the tertiary structure of the ribozyme. This information will be very useful for the formulation of specific tertiary structure models for the hairpin ribozyme. Our isolation of the $U_{39}C$ substitution as a second-site suppressor of $G_{21}U$ (this work) and, independently, as a suppressor of $A_{43}G$ (Berzal-Herranz *et al.*, 1992) demonstrates that the selection method has the potential for the identification of such tertiary contacts as non-Watson–Crick base pairs and base triples. To attempt to identify such contacts, we are currently performing selection experiments to identify second-site revertants of strongly inactivating point mutations, by starting the selection with highly mutagenized populations of molecules in which the strong down mutations are fixed.

Materials and methods

Plasmid and bacterial strains

Plasmid pGEM-3Zf(−) (Promega, WI, USA) was used for cDNA cloning. The *Escherichia coli* strain DH5α (F, *endA*1, *hsdR*17, *supE*44, *thy*-1, *recA*1, *gyrA*96, *relA*1, (Δ*argF-lacZYA*)U196, φf80Δ*lacZ*M15) was used as host for the individual clones.

Synthesis of oligonucleotides

All DNA and chimeric DNA/RNA oligonucleotides were synthesized using standard solid-phase phosphoramidite chemistry on an Applied Biosystems 392 oligonucleotide synthesizer (ABI; CA, USA), and were purified as described (Scaringe *et al.*, 1990; Chowrira and Burke, 1991). Ribonucleoside phosphoramidites were obtained from Glen Research (VA, USA).

In vitro selection

Selection experiments were performed as described (Berzal-Herranz *et al.*, 1992; Joseph *et al.*, 1993). Sequence variation is introduced during the solid-phase chemical synthesis of the transcriptional templates. Randomization at specific regions within the ribozyme–substrate complex was obtained by introducing equimolar quantities of all four bases at each of the desired positions, as described (Berzal-Herranz *et al.*, 1992). Contamination of phosphoramidites to mutagenize the ribozyme–substrate complex was carried out as described (Joseph *et al.*, 1993). For selection in inactive molecules, the ribozyme template was synthesized by mixing the DNA monomers such that 0.66% contamination with each of the three mutagenic nucleotides was obtained, and the substrate template such that 2.3% contamination was obtained. Using these mixes of phosphoramidites, one mutation per molecule was achieved. For selection of active molecules, molecules were mutagenized to an average of five mutations per molecule using phosphoramidites that were contaminated with 3.3% of each of the mutagenic nucleotides for synthesis of the ribozyme template and 12% for the substrate template.

Transcription of the DNA templates, RNA self-cleavage and purification of active and inactive variant molecules were performed as described (Berzal-Herranz *et al.*, 1992). Ligation reactions were done by incubating the purified 5′ cleavage product containing the 5′ fragment of the substrate linked to the 3′ end of the ribozyme (terminating with a 2′, 3′, cyclic phosphate) with 100 pmol of the ligation substrate (5′-*GUCCUGUUU*-AGATATCCGTCGACAAG-3′; ribonucleotides in italics, deoxyribonucleotides in plain type). Ligation products were gel purified, copied into cDNA and amplified by the polymerase chain reaction as described (Berzal-Herranz *et al.*, 1992).

Self-cleavage assays

Self-cleavage assays were carried out by runoff transcription of *Eco*RV or *Xho*I-digested plamsids, for clones selected as active or inactive respectively. Reactions were as described (Berzal-Herranz *et al.*, 1992).

Trans-cleavage reactions

Trans-cleavage reactions were carried out at 37°C in standard cleavage buffer containing 10 mM $MgCl_2$ (Chowrira and Burke, 1991) except when indicated. Reaction products were analyzed on 20% polyacrylamide–urea gels. Dried gels were quantitatively analyzed with a Betascan instrument (Betagen; MA, USA).

Ribozyme–substrate complex formation

The internally labeled noncleavable substrate analog, $G_{+1}A$, was synthesized by *in vitro* transcription in the presence of [α-^{32}P]CTP and used to measure formation of the ribozyme–substrate complex as described (Chowrira and Burke, 1991). Substrate binding by the ribozyme was allowed to proceed at 4°C in 12 mM $MgCl_2$ for at least 2 h. Complex formation was monitored following native gel electrophoresis in 12 mM $MgCl_2$, in such a manner that pre-existing complex is stabilized, but no additional binding occurs (Chowrira *et al.*, 1993).

Acknowledgements

We thank Charles Keller for synthesizing the many oligonucleotides used in this study. A.Berzal-Herranz was supported by a long-term EMBO postdoctoral fellowship. J.Burke is a member of the Vermont Cancer Center. This research was supported by grants from the NIH.

References

Beaudry,A.A. and Joyce,G.F. (1992) *Science*, **257**, 635–641.
Berzal-Herranz,A., Joseph,S. and Burke,J.M. (1992) *Genes Dev.*, **6**, 129–134.
Brown,J.W., Haas,E.S., James,B.D., Hunt,D.A. and Pace,N,R. (1991) *J. Bacteriol.*, **173**, 3855–3863.
Burke,J.M. and Berzal-Herranz,A. (1993) *FASEB J.*, **7**, 106–112.
Buzayan,J.M., Gerlach,W.Y. and Bruening,G. (1986) *Nature*, **323**, 349–353.
Chowrira,B.M. and Burke,J.M. (1991) *Biochemistry*, **30**, 8518–8522.
Chowrira,B.M. and Burke,J.M. (1992) *Nucleic Acids Res.*, **20**, 2835–2840.
Chowrira,B.M., Berzal-Herranz,A. and Burke,J.M. (1991) *Nature*, **354**, 320–322.
Chowrira,B.M., Berzal-Herranz,A. and Burke,J.M. (1993) *Biochemistry*, **32**, 1088–1095.
Feldstein,P.A., Buzayan,J.M. and Bruening,G. (1988) *Gene*, **82**, 53–61.
Feldstein,P.A., Buzayan,J.M., Van Tol,H., DeBear,J., Gough,G.R., Gilham,P.T. and Bruening,G. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 2623–2627.
Green,R., Ellington,A.D. and Szostak,J.W. (1990) *Nature*, **347**, 406–408.
Hampel,A. and Tritz,R. (1989) *Biochemistry*, **28**, 4929–4933.
Hampel,A., Tritz,R., Hicks,M. and Cruz,P. (1990) *Nucleic Acids Res.*, **18**, 299–304.
Haseloff,J. and Gerlach,W.L. (1989) *Gene*, **82**, 43–52.
Joseph,S., Berzal-Herranz,A., Chowrira,B.M., Butcher,S.E. and Burke,J.M. (1993) *Genes Dev.*, **7**, 130–138.
Levitt,M. (1969) *Nature*, **224**, 759–763.
Michel,F. and Westhof,E. (1990) *J. Mol. Biol.*, **216**, 585–610.
Michel,F., Hanna,M., Green,R., Bartel,D.P. and Szostak,J.W. (1989) *Nature*, **342**, 391–395.
Michel,F., Jaeger,L., Westhof,E., Kuras,R., Tihy,F., Xu,M.Q. and Shub,D.A. (1992) *Genes Dev.*, **6**, 1373–1385.
Noller,H.F. and Woese,C.R. (1981) *Science*, **212**, 403–411.
Pan,T. and Uhlenbeck,O.C. (1992) *Biochemistry*, **31**, 3887–3895.
Pyle,A.M., Murphy,F.L. and Cech,T.R. (1992) *Nature*, **358**, 123–128.
Robertson,D.L. and Joyce,G.F. (1990) *Nature*, **344**, 467–468.
Rubino,L., Tousignant,M.E., Steger,G. and Kaper,J.M. (1990) *J. Gen. Virol.*, **71**, 1897–1903.
Scaringe,S.A., Francklyn,C. and Usman,N. (1990) *Nucleic Acids Res.*, **18**, 5433–5441.
Szostak,J.W. (1992) *Trends Biochem. Sci.*, **17**, 89–93.
Williamson,C.L., Tierney,W.M., Kerker,B.J. and Burke,J.M. (1987) *J. Biol. Chem.*, **262**, 14672–14682.

Received on February 8, 1993