Sequence-dependent structural variations of hammerhead RNA enzymes

Hans A.Heus, Olke C.Uhlenbeck and Arthur Pardi*
Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, CO 80309-0215, USA

Received December 26, 1989; Revised and Accepted January 10, 1990

ABSTRACT

The discovery of *in vivo* catalytic activity for the hammerhead RNA self-cleaving domain has led to the development of a new class of sequence-specific RNA endonucleases. Two such ribozymes have been synthesized using *in vitro* transcription with T7 polymerase and their structures have been studied by optical spectroscopy, nuclear magnetic resonance and nondenaturing gel electrophoresis. These data show the presence of a stable hairpin consisting of a double helical stem and a tetranucleotide loop in both RNA enzymes. Additional structure, with different stabilities, is also observed in both RNA enzymes. The half-lives for cleavage of the complementary RNA substrates by these two RNA enzymes have been previously shown to differ by a factor of 50. The data presented here suggest that this rate difference may be a result of the formation of catalytically inactive conformations in the RNA enzyme which interfere with formation of the enzyme-substrate complex.

INTRODUCTION

The discovery of catalytic RNAs has broadened our view of biological catalysis (1–3). One class of such molecules exists within a variety of plant viroid and virusoid RNAs that are capable of undergoing site specific autocatalytic cleavage to generate 5′ hydroxyl and 2′,3′-cyclic phosphate termini (for a review see 4). These viral RNAs are believed to replicate by a rolling circle mechanism where a multigenome precursor undergoes a self-cleavage reaction to form linear single genomes which are then ligated to form circular RNAs. From comparison of the sequence of a number of plant viral RNAs, Forster and Symons (5) proposed a secondary structure model for this cleavage domain, which they termed a hammerhead. This hammerhead domain consists of three helices and contains 13 conserved nucleotides (6) and is schematically shown in Figure 1A. Although the *in vivo* cleavage reaction is intramolecular, it has been shown *in vitro* that an active hammerhead domain can be assembled from two molecules in several different ways (7–13). Since turnover is observed for the site-specific cleavage reaction in these systems, the RNA containing the cleavage site is termed the substrate and the other RNA is termed the enzyme. Haseloff and Gerlach (10) designed several ribozymes with different sequences for strands I and III that could base pair and cleave specific sites in a substrate mRNA coding for chloroamphenicol acetyl transferase.

All variants of the consensus hammerhead structure studied thus far *in vitro* show an absolute requirement for the three helices and ten of the thirteen conserved nucleotides (7,9,13, D.E. Ruffner & O.C. Uhlenbeck, unpublished results). An important conclusion from these studies is that whereas the presence of the helices I, II or III is essential for cleavage, the exact sequences of these helices can vary. However, different sequences in these helices can affect the half-lives of the cleavage reaction by over a factor of thousand (7). In this paper we present NMR data of two RNA enzymes aimed at correlating the observed rate differences with NMR structural data for these ribozymes. The data prove the existence of helix II as well as additional structure with different stabilities in both ribozymes. We suggest that these additional structures may be contributing to the observed rate differences.

MATERIALS AND METHODS

Preparation of RNA

The two RNAs, termed R8 and R9, following the numbering system of Ruffner *et al.* (7), with sequences given in Figure 1B were prepared by *in vitro* transcription with T7 RNA polymerase using synthetic DNA oligomers as templates, as previously described (14). A 100ml transcription reaction, containing 200nM DNA template, 4mM of each NTP and 10mg T7 polymerase in 40mM Tris-Cl, 25mM $MgCl_2$, 5mM dithiothreitol, 1mM spermidine, 0.01% Triton X-100 (v/v), pH 8.1 was incubated for 3 hours at 37°C. This crude RNA mixture was ethanol precipitated, then redissolved in 30ml of buffer A (10mM sodium phosphate pH 7.0) and applied to a 30×0.7cm fractogel DEAE column (Supelco). The column was washed with 50ml of buffer A to elute unreacted NTPs, after which the RNA transcripts were eluted with buffer B (10mM sodium phosphate pH 7.0, 0.6M NaCl) and ethanol precipitated. The RNA precipitate was dissolved in 6.4ml 50% water/50% formamide (v/v) and applied to eight 15% polyacrylamide gels containing 7M urea (40×60×0.3cm). Electrophoresis was performed for 10–12 hours at 800V, until the bromophenol blue dye marker reached the bottom of the gel. The band corresponding to full length

* To whom correspondence should be addressed

transcript RNA was detected by UV shadowing, and eluted with 50mM potassium acetate, 200mM KCl, 10mM EDTA. This eluted solution was diluted two fold with buffer A, applied to a 4×0.9cm fractogel DEAE column, washed with 50ml buffer A and eluted with buffer B. The RNA containing fractions were then lyophilized and desalted by passing them twice through a 20×0.5cm Sephadex G25 column, that was equilibrated with water. This procedure produced 5 to 7mg of purified RNA from a 100ml transcription reaction.

NMR Spectra

The NMR samples without magnesium were prepared by dissolving the purified RNA in 550μl of 10mM sodium phosphate, pH 7.0, 100mM NaCl, and 0.2mM sodium EDTA in 90% H_2O/10% ^2H_2O. The samples without magnesium were heated at 85°C for 2 minutes and slowly cooled to room temperature over 20 minutes. For the magnesium titration studies, a concentrated solution of $MgCl_2$ was added to the NMR sample. NMR spectra were recorded on a Varian VXR-500S spectrometer operating at 499.8 MHz. The 1D NMR spectra were collected with 8192 complex data points over a spectral width of 10 kHz using a binomial 1331 water suppression pulse sequence (15). The 2D NOE spectrum of R8 was collected at 30°C in the phase sensitive mode (16) with 96 scans and 4096 complex data points in t_2, using a 1331 acquisition pulse and 300 complex t_1 increments. The 2D spectrum was processed using the software package FTNMR, kindly provided by Dr. Dennis Hare (Hare Research, Inc., Woodinville, WA). The spectrum was apodized by a 65°shifted sinebell window function in both dimensions and a third degree polynomial baseline correction was performed in ω_2 after the first Fourier transform.

Thermal Denaturation Profiles

Optical melting curves were recorded on a Gilford 2400 spectrophotometer, equipped with a Gilford 2527 thermal programmer. Acquisition and curve fitting was performed using software provided by Dr. D. Turner (17).

Nondenaturing Gel Electrophoresis

^{32}P-labeled RNAs were prepared by adding 75μCi [α^{32}P]-CTP (3000Ci/mMole) to a 30μl transcription mix containing 1mM NTPs. Incubation was for 2 hours at 37°C, after which the transcription reaction was mixed with 30μl 7M urea, 0.2% xylene cyanol, 0.2% bromophenol blue and loaded directly on a 12% polyacrylamide, 7M urea gel. Full length RNA was detected by autoradiography, eluted with 50mM potassium acetate, 200mM KCl, 10mM EDTA and precipitated twice with ethanol. Samples for nondenaturing gels were prepared by mixing appropriate amounts of labeled and unlabeled RNA in 10μl gel running buffer (50mM Tris-borate, 1mM EDTA). The samples were denatured by heating to 95°C, slowly cooled to room temperature, mixed with 4μl 40% sucrose in gel running buffer and loaded on 15% polyacrylamide gels (40×16×0.15cm). Electrophoresis was performed for 12 hours at room temperature at 8V/cm.

RESULTS

As seen in Figure 1 both RNA enzymes contain the set of conserved nucleotides observed by Forster and Symons for hammerhead self-cleaving domains (5) as well as a hairpin structure consisting of a stem containing four CG base pairs with a 5'GAAA3' tetranucleotide loop for strand II. This loop

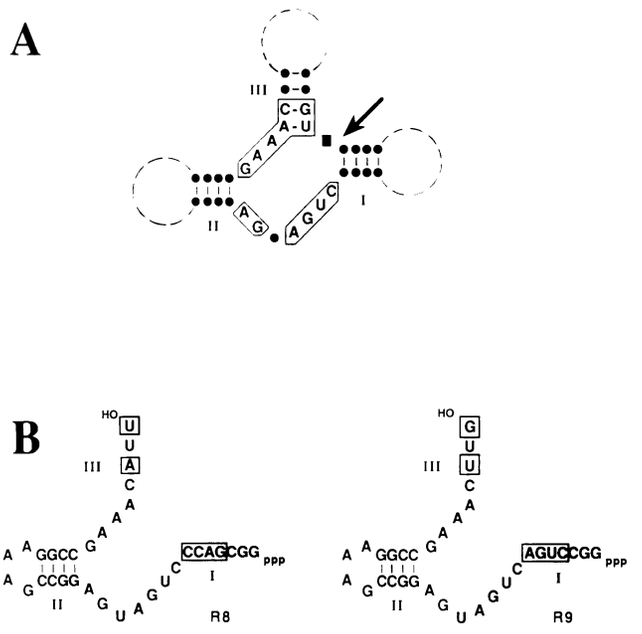


Figure 1. A) Consensus structure of the hammerhead RNA self cleaving domain (4,5). The arrow indicates the cleavage site, the three double stranded stems are labelled I–III, the dots indicate bases in standard Watson-Crick base pairs and the line between bases on two strands indicates base pair formation. B) The base sequences for the RNA enzymes R8 and R9. Boxed regions indicate sequence differences between the two enzymes.

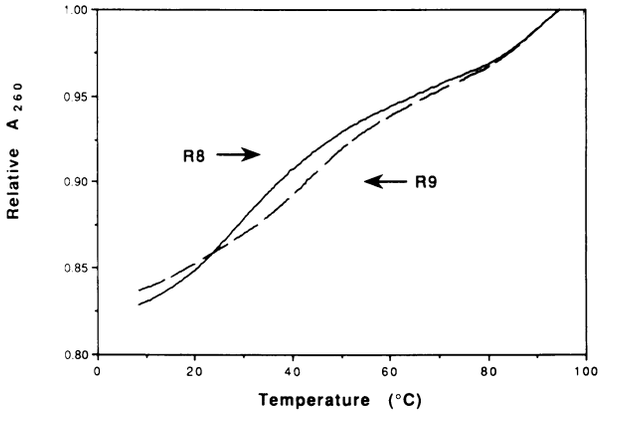


Figure 2. Absorbance-temperature profiles of 3μM R8(——) and R9(-----) in 10mM sodium phosphate, pH 7.0, 100mM NaCl, 0.1mM EDTA. The relative absorbance (normalized to 95°C) at 260nm in a 1cm path length cell is plotted versus the temperature.

sequence, when closed by a CG base pair, has been found to be exceptionally stable in both RNA (J. Haney & O.C. Uhlenbeck, unpublished results) and DNA (18) oligomers. Thus the stem and loop for strand II in both RNA enzymes were designed to be very stable in order to minimize alternate structures in this region of the molecule. The differences in the sequences for R8 and R9 occur in strand I and strand III and are boxed in Figure 1B.

From the design criteria discussed above, it was expected that hairpin II would form a very stable structure in solution, but additional structures could form as a result of interactions between the remaining 22 nucleotides in the molecules. Evidence for such additional structure is given by the thermal denaturation profiles

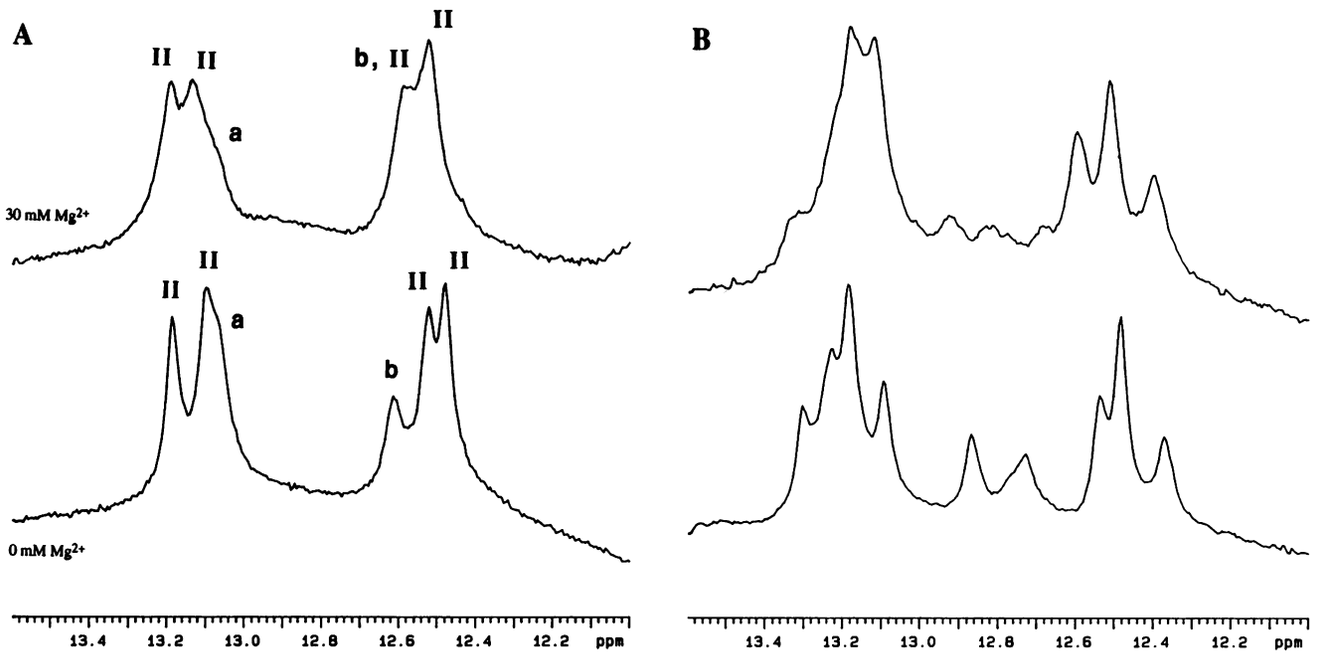


Figure 3. The imino proton region of the NMR spectra of A) R8 (0.2mM, 4000 scans) and B) R9 (0.2mM , 4000 scans) in 10mM sodium phosphate, pH 7.0, 100mM NaCl, 0.2mM sodium EDTA in the absence (lower traces) and presence (upper traces) of 30mM added $MgCl_2$. Spectra were recorded at 25°C. Resonances labeled II in A are assigned to helix II, resonances a and b are not assigned (see text).

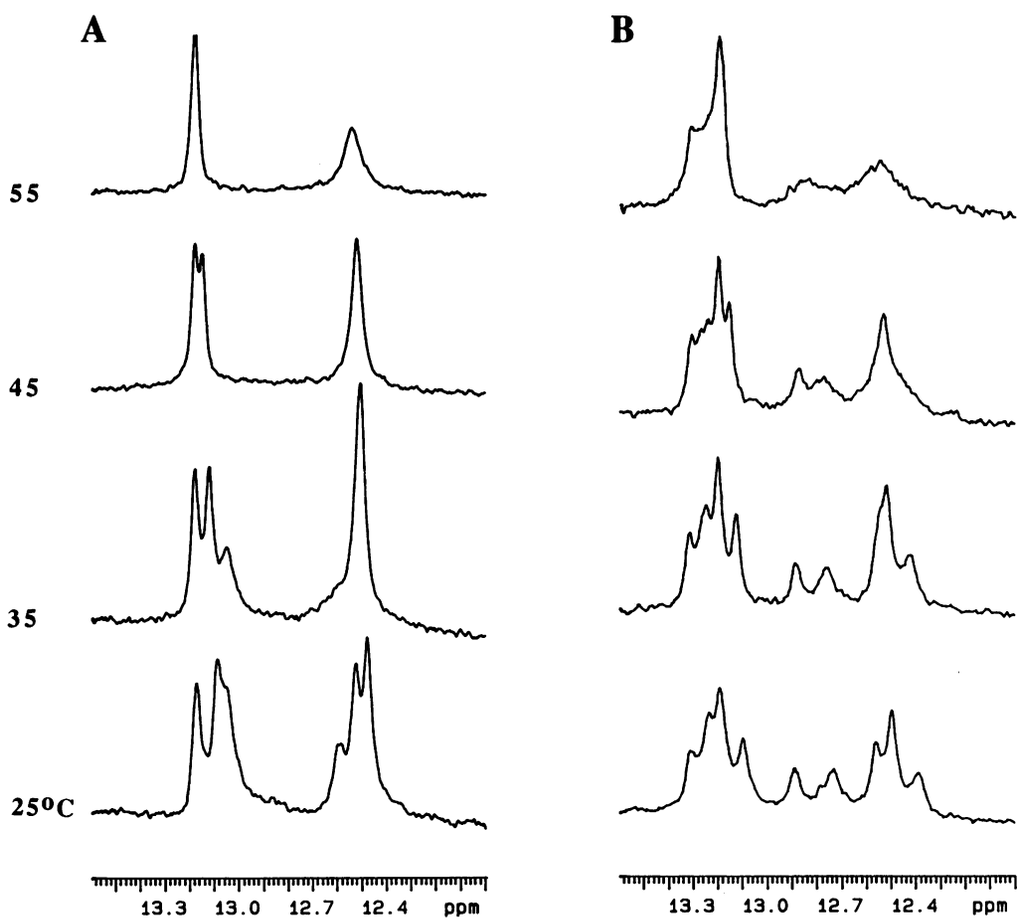


Figure 4. The imino proton region of the NMR spectra for A) 0.7mM R8 (64 scans) and B) 0.6mM R9 (256 scans) is plotted as a function of temperature in the magnesium free buffer given in Figure 3. These NMR spectra were recorded without magnesium in order to prevent the degradation of the RNA at high temperature.

for R8 and R9 shown in Figure 2. Instead of one transition arising from unfolding of hairpin II, two transitions are observed for both ribozymes. There is a broad transition around 40°C and a more cooperative transition at approximately 91°C. The second transition has a very high T_m and is identical for both ribozymes. This is consistent with melting of the CG rich stem of hairpin II since this stable hairpin is present in both RNAs. Additional structures in the ribozymes is suggested by the other transition at lower temperature. The low temperature transition in R9 consistently shows a slightly higher T_m than R8 which indicates that it has a more stable additional structure. To test for a concentration dependence of these transitions, thermal denaturation profiles were recorded at the ten fold higher concentration of 30μM. Within the range of error of these measurements, the profiles are the same at 3 and 30μM (results not shown), suggesting that the low temperature transition arises from an intramolecular interaction.

Figures 3 and 4 show the 1D NMR spectra of the ribozymes at various buffer and temperature conditions. In the absence of magnesium, the spectrum of R8 at 25°C has six resonances, corresponding to six base pairs in the imino proton region (Figure 3A, lower trace). Since there is an absolute requirement for a divalent cation for the *in vitro* cleavage reaction (8) the RNA enzymes were also studied in the presence of magnesium ion. The addition of 30mM Mg^{2+} to the R8 enzyme results in only a slight broadening and shifting of the imino proton resonances. Since no new resonances are observed, it appears that magnesium does not induce additional structure in the R8 ribozyme. The imino proton region of the NMR spectrum of R8 is plotted as a function of temperature in Figure 4A. Two of the six imino proton resonances broaden and disappear by 45°C, while the remaining four resonances only start to broaden at 55°C. Since broadening arises from chemical exchange of the imino protons with solvent H_2O, this indicates that two base pairs are more labile than the other four.

To assign the imino proton resonances in R8, a 2D NOE spectrum was run in H_2O. Figure 5 shows a plot of the region of the spectrum for crosspeaks involving imino, aromatic and amino protons. In the 2D NOE experiment crosspeaks arise from dipolar interactions for protons close in space (<4.5Å). Each of the four more stable imino proton resonances show two strong crosspeaks in this region, which is typical for a CG base pair. One crosspeak arises from a strong NOE between the imino proton and the nearby hydrogen bonded C amino proton (separated by 2.5Å in a standard Watson–Crick base pair). This NOE can be transferred by chemical exchange from the hydrogen bonded to the nonhydrogen bonded C amino proton by rotation around the carbon-nitrogen chemical bond, which generates the other crosspeak. Assignment of the two C amino protons is confirmed by a strong amino proton-amino proton crosspeak in the aromatic proton region (data not shown). In the case of an AU base pair only one strong imino to aromatic crosspeak is observed which arises from a NOE involving the U imino proton and the A2 proton. The crosspeak patterns in the 2D NOE spectrum of R8 assign the four stable imino protons to CG base pairs. The two more labile base pairs do not show up in the 2D NOE spectrum due to exchange of their imino protons with solvent H_2O. The chemical shifts however suggest that these imino proton resonances also originate from CG basepairs.

In addition to the strong intra base pair imino proton-amino proton NOEs, some weaker NOE crosspeaks appear in Figure 5. The crosspeaks, highlighted by arrows in Figure 5 arise from

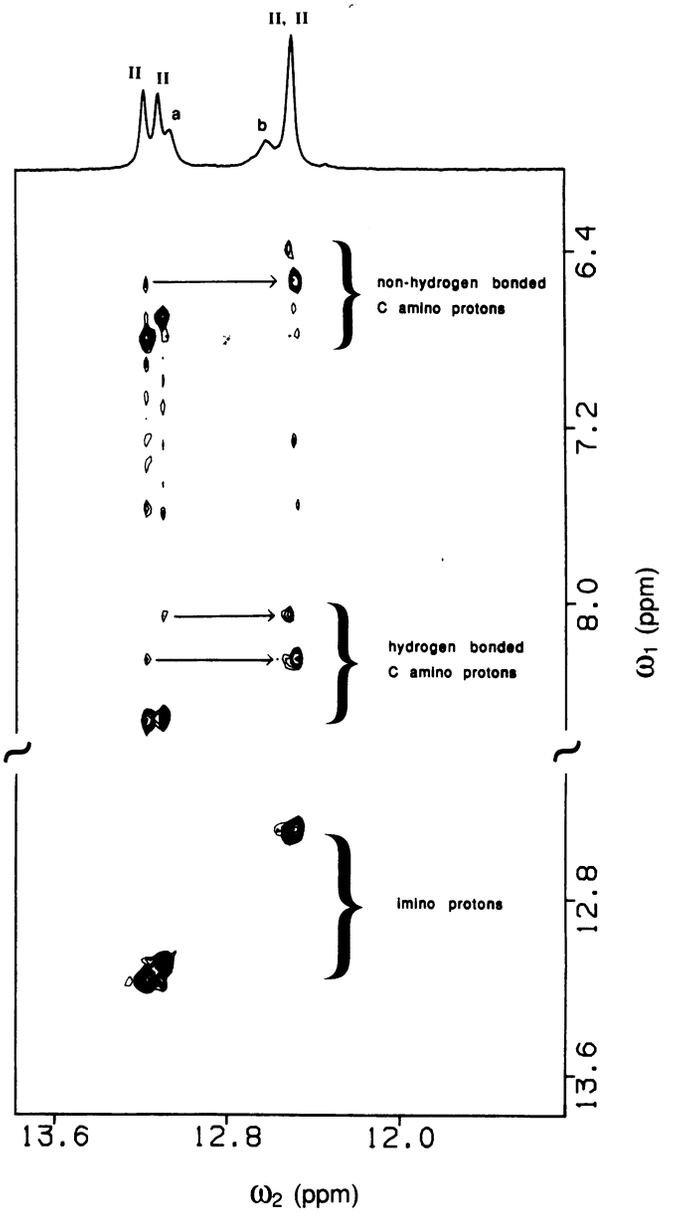


Figure 5. A contour plot of a portion of the phase sensitive 2D NOE spectrum of 0.7mM R8 in H_2O showing the imino to imino proton, and the imino to aromatic/amino proton regions. Sequential NOE connectivities are indicated by arrows and assignments to proton type are also shown. A 1D spectrum (512 scans) of the imino proton region recorded under identical conditions is plotted on top of the 2D spectrum.

NOEs between G imino and C amino protons on adjacent base pairs. These sequential NOE connectivities indicate that the four stable CG base pairs are adjacent in the RNA structure. A more extensive 2D NMR study of the full hammerhead domain, consisting of an RNA enzyme basepaired to a DNA substrate at helices I and III, confirms the sequential resonance assignments proposed here for helix II (H.A. Heus and A. Pardi, in preparation). Thus the NMR spectra on R8 confirm and extend the results from the thermal denaturation profiles proving the presence of the stable hairpin II, along with a less stable interaction involving two additional unassigned base pairs in the molecule.

The NMR properties of R9 are very different from R8. Figure

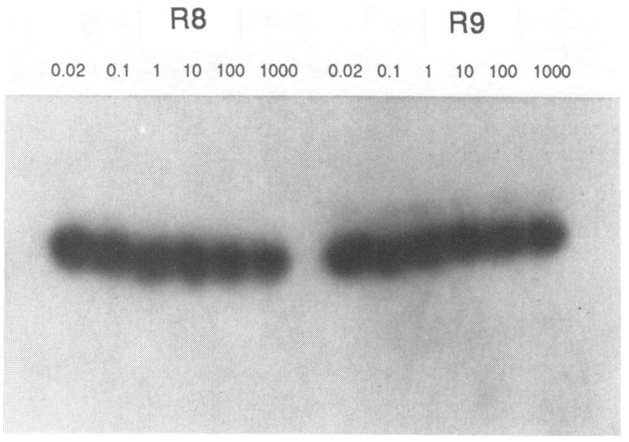


Figure 6. Nondenaturing polyacrylamide gel electrophoresis of R8 and R9 over a concentration range from 0.02μM to 1000μM. A fixed amount of ^{32}P labeled ribozyme was mixed with unlabeled enzyme to yield the given concentrations (see text).

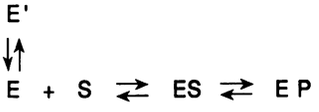
3B (lower trace) shows that in the absence of magnesium R9 has a much more complex imino proton spectrum with at least 3 more resonances. This indicates that R9 forms more base pairs than R8, which is consistent with the observed slightly higher T_m for the low temperature transition of R9. In the presence of 30mM magnesium, some line broadening and shifting occurs, but no new resonances appear in the spectrum (Figure 3B upper trace). Thus magnesium does not induce formation of additional base pairs in R9. The NMR spectra of R9 as a function of temperature are plotted in Figure 4B and show that the 3 additional resonances are quite stable and persist until high temperature. Although we do not have direct evidence from a 2D NOE experiment, the chemical shifts suggest that the additional resonances originate from CG basepairs.

These additional base pairs in R8 and R9 could arise from: i) An intermolecular interaction such as aggregation of the ribozyme or ii) an intramolecular interaction such as where a single conformation of the ribozyme has more base pairs formed or where different conformations of the ribozyme are in slow exchange on the NMR time scale (> milliseconds). The concentration independent melting behavior of R8 and R9 over a ten fold concentration range (3−30μM) argues against aggregation. To test for aggregation of the ribozymes at the NMR concentration, nondenaturing gels were run on R8 and R9 over a concentration range from 20nM to 1mM enzyme. Figure 6 shows the concentration dependence of the mobility of R8 and R9 in a nondenaturing gel where a single band is observed at all concentrations for both ribozymes. Nondenaturing gel electrophoresis has previously been used to probe conversion of DNA monomers into duplexes (19) and thus provides a direct probe of multimer formation. Multiple bands have also been observed on nondenaturing gels for another hammerhead RNA enzyme indicating the presence of aggregates (20). Our results show that the ribozymes do not form stable aggregates even at the high (0.7mM) concentrations used in the NMR experiments. Therefore the additional base pairs in R8 and R9 are due to an intramolecular (concentration independent) interaction.

DISCUSSION

The NMR spectral data and the thermal denaturation profiles on R8 and R9 indicate the existence of a stable hairpin structure consisting of a 4 CG base pair stem, and a GA_3 loop, which is in agreement with the consensus structure of the hammerhead RNA enzymes. However, these data also indicate the presence of additional, less stable, interactions which cannot be simply predicted from the primary structure of these ribozymes. The NMR data on the imino protons show that these interactions arise from formation of additional base pairs in the ribozymes. The data are consistent with a molecule that forms a single conformation in solution which contains some additional base pair interactions besides those predicted from Figure 1B. Another possibility is a molecule that has multiple conformations in solution which exchange slowly on the NMR time scale (>milliseconds), but not slowly enough to be separated on a nondenaturing gel (<hours). The gel data and the NMR experiments cannot distinguish between these two possibilities as the source of the additional base pairs in these ribozymes.

Under single turnover conditions, with equal amounts of ribozyme and RNA substrate, R8 cleaves its substrate 50 times faster than R9 (7). The slower enzyme is the one with the much more extensive additional structure which suggests a possible correlation between the structure and catalytic activity of the enzyme. The following scheme for the cleavage reaction is consistent with the data present here.



Where E, S are the 'active' conformations of the RNA enzyme and substrate, respectively, ES and EP are the enzyme-substrate and enzyme-product complexes, respectively and E′ represents an alternate conformation of the RNA enzyme which is unable to form a Michaelis complex. This scheme is analogous to the one proposed by Sheldon and Symons (12) for the intramolecular cleavage reaction. If assembly is the rate limiting step in the reaction, then any conformation of either the enzyme or substrate that interferes with complex formation can decrease the observed rate of cleavage. Using data from UV melting curves on a different hammerhead RNA enzyme and substrate, Ruffner et al. (7) proposed the existence of additional structure (not predicted from the consensus structure) from which they suggested that assembly could be a kinetically important step. However, even if assembly is not the rate limiting step, alternative structures of the ribozymes can still affect the observed cleavage rate by altering the K_m for the reaction. Steady state kinetic measurements of R8 and R9 and their RNA substrates showed that K_m was altered, while k_{cat} was nearly the same (20). The K_m of R9 is higher than R8, consistent with the difference in single turnover kinetics and in agreement with the mechanism proposed.

We note that for the plant viral RNA hammerhead domains, the *in vivo* cleavage reaction is an intramolecular process and the distinction between the ribozyme and the substrate may not be particularly meaningful. Thus the effect of alternate conformations for the substrate may be just as important as alternate structures of the ribozyme. In fact in some cases alternate structures for the substrate directly affect the catalytic efficiency of the cleavage reaction (20).

The effect of alternative conformations of virusoid RNA on self-cleaving activity has been well documented (5,6,21). Of particular interest to this study is the effect of heating followed by slow-cooling, versus heating followed by snap-cooling on ice.

For instance, plus and minus strands of vLTSV only selfcleave after snap-cooling and remain inactive after slow-cooling (6). Presumably slow cooling allows for the formation of the native, inactive conformation, while snap-cooling induces formation of an alternative, active conformation (6). Slow-cooled plus and minus RNA transcripts, containing 84% of the vLTSV genome have different mobilities from snap-cooled transcripts on nondenaturing gels (21), indicating the presence of alternative conformations. We have also recorded 1D NMR spectra of R8 and R9 after snap-cooling on ice (results not shown), and these spectra were identical to the spectra shown in Figure 3 (lower traces). Apparently the formation of additional or alternate structure in R9, correlated with inactive conformations is independent of prior treatment of the RNA. In the studies where inactive conformations were converted into active conformations by the heat-snap-cool method, self-cleavage does not go to completion (5,6,12,21), even in the case of small (43–44 nucleotide) molecules (12). This could very well arise from alternate conformations such as the ones described here.

RNA enzymes that form base pairs with an RNA substrate at helices I and III (Figure 1) have been designed and recently shown to be capable of cleaving a messenger RNA at specific sites (10). Such sequence-specific RNA endonucleases have important potential *in vivo* applications such as inactivation of specific mRNAs or RNA genomes. In order to optimize the cleavage reaction of these ribozymes it will be important to eliminate alternative inactive structures of the ribozyme (or substrate) that could interfere with catalysis. The combination of the structural and biochemical techniques discussed here give us a direct handle on probing alternate conformations of RNA oligomers in solution.

ACKNOWLEDGEMENTS

We thank S.C. Dahm, M.J. Fedor and D.E. Ruffner, for helpful discussions and M. P. Smith for preparation of the T7 RNA polymerase. This work was supported in part by grants from the Searle Scholars program of the Chicago Community Trust (85-C110) and NIH GM35807 to A.P. and NIH GM36944 to O.C.U. The 500 MHz NMR spectrometer was purchased with partial support from NIH Grant RR03283.

REFERENCES

1. Kruger, K., Grabowsky, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. & Cech, T.R. (1982) Cell **31**, 147–157.
2. Cech. T.R. & Bass, B.L. (1986) Annu Rev. Biochem. **55**, 599–629.
3. Cech, T.R. (1987) Science **236**, 1532–1539.
4. Keese, P. & Symons, R.H. (1987) in *Viroids and Viroid-like Pathogens* (ed. Semancik, J.S.) 1–47 (CRC Press, Boca Raton, Florida).
5. Forster, A.C. & Symons, R.H. (1987) Cell **49**, 211–220.
6. Forster, A.C. & Symons, R.H. (1987) Cell **50**, 9–16.
7. Ruffner, D.E., Dahm, S.C. & Uhlenbeck, O.C. (1989) Gene **82**, 31–41.
8. Uhlenbeck, O.C. (1987) Nature **328**, 596–600.
9. Koizumi, M., Iwai, S. & Ohtsuka, E. (1988) FEBS Lett. **228**, 228–230.
10. Haseloff, J. & Gerlach, W.L. (1988) Nature **334**, 585–591.
11. Jeffries, A.C. & Symons, R. H. (1989) Nucleic Acids Res. **17**, 1371- 1377.
12. Sheldon, C.C. & Symons, R.H. (1989) Nucleic Acids Res. **17**, 5665–5677.
13. Sheldon, C.C. & Symons, R.H. (1989) Nucleic Acids Res. **17**, 5679–5685.
14. Milligan, J.F., Groebe, D.R., Witherell, G.W. & Uhlenbeck, O.C. (1987) Nucleic Acids Res. **15**, 8783- 8798.
15. Hore, P.J. (1983) J. Magn. Reson. **55**, 283–300.
16. States, D.J., Haberkorn, R.A. & Ruben, D.J. (1982) J. Magn. Reson. **48**, 286–292.
17. Groebe, D.R. & Uhlenbeck, O.C. (1989) Biochem. **28**, 742–747.
18. Hirao, I., Nishimura, Y., Naraoka, T., Watanabe, K., Arata, Y. & Miura, K. (1989) Nucleic Acids Res. **17**, 2223–2231.
19. Kallenbach, N.R., Ma, R-I & Seeman, N.C. (1983) Nature **305**, 829–831.
20. Fedor, M.J. & Uhlenbeck, O.C. (1990) Proc. Natl. Acad. Sci. U.S.A. in the press.
21. Forster, A.C., Jeffries, A.C., Sheldon, C.C. & Symons, R.H.(1987) Cold Spring Harb.Symp.Quant.Biol. **52**,249–259.