Submitted to Biochemistry

Characterization of the Secondary Structure and Melting of a Self-Cleaved RNA Hammerhead Domain by $^1$H NMR Spectroscopy

A.C. Pease and D.E. Wemmer

January 1990
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
Characterization of the Secondary Structure and Melting of a Self-Cleaved RNA Hammerhead Domain by $^1$H NMR Spectroscopy

Ann Caviani Pease and David E. Wemmer

Department of Chemistry
University of California

and

Chemical Biodynamics Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Running title: NMR of an RNA Hammerhead Domain

†This work was supported by the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098, and through instrumentation grants form the U.S. Department of Energy, DE FG05-86ER75281 and the National Science Foundation, DMB 86-09035.
ABSTRACT

INTRODUCTION

A site-specific self-cleavage reaction has been observed for the (+) and (-) strands of satellite tobacco ringspot virus RNA (STobRV RNA) which results in 5'-hydroxyl and 2',3'-cyclic phosphate termini (Prody et al. 1986; Buzayan et al. 1986a,b). Similar reactions have been observed for the (+) and (-) strands of avocado sunblotch virusoid (Hutchins, et al. 1986), and lucerne transient streak viroid (Forster & Symons 1987). Buzayan and coworkers have found that fewer than 100 nucleotides are necessary for efficient cleavage of (+)STobRV RNA (Buzayan, et al. 1986a). Symons and coworkers have proposed a "hammerhead" secondary folding model for the self-cleavage domain (Hutchins, et al. 1986; Forster & Symons 1987; Keese, et al. 1987). This domain consists of 13 conserved nucleotides and three variable length double-stranded helices (fig. 1). Cleavage occurs only at a specific phosphodiester linkage and requires no cofactor other than magnesium ions. Although the nucleotides directly at and near the cleavage site can vary, the position of cleavage within the hammerhead domain is conserved. Therefore it is clear that a specific three-dimensional folding of the hammerhead domain is required for cleavage. The nature of the interactions giving rise to this folding are as yet unknown.

We have chosen to study the three-dimensional structure of such an RNA by nuclear magnetic resonance (NMR). NMR has been used extensively to study the structure of transfer RNA in solution. Particular attention has been paid to the extreme low-
field (15-11 ppm) region of the spectrum. (For reviews see Kearns & Shulman 1974; Kearns 1976, 1977; Reid & Hurd 1977; Schimmel & Redfield 1980; Reid 1981.) Usually only hydrogen-bonded imino protons involved in secondary and tertiary base pairs resonate in this region, leaving it comparatively well resolved for a large nucleic acid (Reid, et al. 1975, 1977, 1979). One-dimensional nuclear Overhauser enhancements (NOE) have been used to assign this region of the spectrum for tRNA (Johnston & Redfield 1978, 1981; Sanchez, et al. 1980; Roy & Redfield 1978, 1981; Hare & Reid 1982a, 1982b), and the solution dynamics have been monitored under different solvent and temperature conditions (Reid 1981). We have used this NOE methodology to assign the low-field spectrum of a self-cleaving RNA, and have studied the stability of helical regions with respect to temperature in the absence and presence of magnesium ions.

Figure 1 shows the RNA sequence we have designed with the conserved nucleotides in outline print. The RNA is produced by in vitro T7 RNA polymerase run-off transcription, and because both this reaction and the cleavage reaction require magnesium ions, the RNA is cleaved at the site denoted by the arrow prior to purification. As an initial starting point we have made the assumption that cleavage does not result in a significant rearrangement of the tertiary structure, and therefore this species approximates a true "cleavage conformation". Although this specific RNA sequence does not occur in nature, cleavage at
the appropriate site has been demonstrated by showing that two RNA fragments of the correct sizes are formed, as analyzed by polyacrylamide gel electrophoresis under denaturing conditions (data not shown). In particular this sequence contains two GU wobble base pairs, one closing the hairpin loop of stem I, the other within stem II, that are not present in any of the naturally occurring hammerhead-type self-cleaving RNAs. They were introduced to facilitate the NMR assignments. The logic behind the use of GU wobble pairs is three-fold: (i) GU wobble pairs contain two hydrogen-bonded imino protons, allowing for two independent nearest-neighbor assignment pathways, (ii) the two imino protons of wobble pairs usually resonate slightly up-field of Watson-Crick type imino protons, reducing chemical shift degeneracies, and (iii) the two imino protons produce strong (30-40%) mutual NOEs due to their close spatial proximity (< 3Å), which facilitates their assignment. In fact, the GU4 of yeast tRNA\textsuperscript{Phe} (Johnston & Redfield 1978), the GU50 of E. Coli tRNA\textsuperscript{fMet} (Hurd & Reid 1979), and the GU50 of E. Coli tRNA\textsubscript{1Val} (Johnston & Redfield 1979) were the first reliably assigned resonances in their respective low-field NMR spectra (Reid 1981). In addition, x-ray studies of base pairing in DNA have shown that replacing a Watson-Crick base pair with a GT wobble base pair produces minimal perturbation in backbone torsion angles (Kneale, et al. 1985; Ho, et al. 1985).

MATERIALS AND METHODS

RNA Samples. Milligram quantities of RNA were produced by T7 RNA
polymerase run-off transcription of a linear plasmid DNA template (Lowary, et al. 1987). The proper RNA transcript was separated from plasmid DNA and short fall-off initiates on a 20x40x0.1 cm 12% native polyacrylamide slab gel. The band was visualized by shadowing with short wave-length UV light on a fluorescent TLC plate. The band was excised from the gel, and the RNA eluted by soaking the crushed gel slice in 0.1M Tris-HCl pH 7.5, 15mM EDTA, 0.15M NaCl, and 1%(w/v) SDS at 37°C for 24 hours. The aqueous phase was carefully removed from the acrylamide pieces using a drawn-out Pasteur pipette and the RNA was precipitated for 4 hours at -20°C in 3 volumes of ethanol. After centrifugation at 15K rpm and 4°C for 20 min., the pellet was resuspended in 2 mL of 5mM potassium phosphate buffer, pH 7.0, and dialyzed at 4°C for 24 hours against 4 liters of 5mM potassium phosphate buffer, pH 7.0, 0.025M NaCl, 0.05mM EDTA, and 0.005% NaN₃. After removal from the dialysis tubing the RNA was dried by lyophilization and stored at -20°C. A 20 mL transcription reaction containing 4 mg of linear plasmid DNA yielded 12 mg of pure RNA, representing a turnover of 400 copies of RNA per DNA template.

NMR Spectra. 12 mg of pure, lyophilized RNA was dissolved in 0.5 mL of 90% H₂O/10% D₂O to give a 1mM RNA, 20mM potassium phosphate, 0.1M NaCl, 0.2mM EDTA, 0.2% NaN₃ solution at pH 7. This was transferred directly to a Wilmad 528pp NMR tube. Between NMR experiments the solution was removed from the NMR tube and the RNA stored at -20°C as a lyophilized powder. For spectra recorded in the presence of magnesium ions, MgCl₂ was added to
Melting data was collected on a GN500 spectrometer using a 1-3-3-1 pulse (Hore, P. J. 1983) with the carrier frequency set at the resonance of water. 1024 scans of 8K data points with a recycle delay of 0.5 sec. were averaged for each temperature increment. The data were Fourier transformed, phase-corrected, and base-line corrected using a MicroVax III workstation and FTNMR software (Hare, D. unpublished). The spectra are indirectly referenced through the chemical shift of water at each temperature.

A complete set of NOE experiments was collected at both 15 and 20°C. For these experiments a peak of interest was selectively saturated using decoupler irradiation at that frequency. The decoupler power level was adjusted to saturate 90% of peak d in 0.8 sec. If one of a nearly degenerate pair of peaks was to be saturated, the decoupler frequency was set slightly off the exact top of the resonance in the direction away from the adjacent neighbor to reduce irradiation spill-over. The NOE data were collected as two experiments, one with the decoupler set on-resonance for a peak of interest, the other with the decoupler set off-resonance. The two experiments were collected in parallel, switching between on- and off-resonance at every scan. 1024 scans of 8K data points were averaged for each experiment. The off-resonance FID was subtracted from the on-resonance FID on the GN500 and the resultant FID transferred to a MicroVax workstation. Fourier transformation, phase-
correction, and baseline-correction using FTNMR software (Hare, D. unpublished) generated the difference spectra.

RESULTS

The imino proton spectrum of our self-cleaving RNA has been completely assigned. The base pair type was determined from intra-base pair imino to aromatic C2H (Watson-Crick AU pairs) or amino proton (Watson-Crick GC pairs) NOEs (Sanchez, et al. 1980), and inter-base pair imino-imino NOEs determined nearest neighbors (Johnston & Redfield 1981). In cases where near degeneracies in chemical shift occurred, second order NOEs, representing next-nearest neighbor connectivities, were used to aid the assignment procedure.

Figure 2 shows the imino-imino NOEs which establish the sequential connectivity of peaks r/t-m-a-i. Irradiation of resonance t (fig. 2F), which is broad even at 15°C, produces a strong NOE to r and a weak NOE to m. This NOE pattern is indicative of a GU wobble pair, and because both r and t have only one nearest neighbor NOE to m, this GU wobble must be a terminal base pair. The only terminal GU wobble is that which closes the hairpin loop of stem I (fig. 1), and therefore the sequential connectivity r/t-m-a-i must define stem I. This assignment is corroborated by the NOEs of peaks a and i. Irradiation of resonance a produces a strong NOE to an aromatic C2H at 7.8 ppm indicating that it is an AU base pair (fig. 2C), and weaker sequential NOEs to i and m. Irradiation of resonance i (fig. 2B) produces the expected NOE to a and no others,
identifying it as a terminal base pair.

Figure 3 shows the imino-imino NOEs sequentially connecting resonances e-d-k-n-b. The near degeneracy of resonances b and c (fig. 3A) and the lack of NOE to b upon irradiation of n (fig. 3E) makes the n-b connectivity slightly ambiguous. However, two lines of evidence support this assignment. The first is that although irradiation of peak n does not result in a sequential NOE to b, it does result in an NOE to the aromatic C2H associated with b (fig. 3E & F). Secondly, as the temperature is increased to 30°C (figure 4), resonance b broadens and disappears. Irradiation of peak c at 30°C results in NOEs only to peaks o and g (data not shown), meaning that it must be saturation of resonance b which results in an NOE to resonance n. The imino-imino NOE connecting e-d is also weak, but again NOEs to a common aromatic C2H confirm this assignment (fig. 3B & C). The sharp NOEs to aromatic C2H's identify resonances b, d, and e as AU base pairs. Because sequential AU pairs are unique to stem III (fig. 1), determination of the d-e connectivity assigns resonances e-d-k-n-b as those of stem III.

Continuing in a similar manner for all low-field resonances results in a complete assignment of the imino spectrum (fig. 1). There is no NMR evidence to substantiate the existence of the AU base pairs terminating stems II and III proposed by the hammerhead model. In similar studies of two other hammerhead RNA sequences, no imino resonances corresponding to these base pairs were observed (data not shown). Irradiation of resonance u at
both 20 and 15°C produced no aromatic or imino NOEs. It is most likely an imino proton in a hairpin loop or single stranded region of the molecule which is somewhat protected from rapid exchange with H₂O. It is possible that this is from a tertiary interaction, but further data are needed to determine if this is in fact the case. Because resonances o and p are nearly degenerate in chemical shift, the m'/p-o connectivity of stem II cannot be determined unambiguously (fig. 1). However, these base pairs are not necessary for cleavage, and this ambiguity in no way diminishes the usefulness of the overall assignment.

Figure 4 shows the imino region of the spectrum plotted as a function of temperature. An imino proton resonance will be observed only if the lifetime of the imino proton is long compared to the chemical shift difference between the imino form and the solvent. As the temperature is increased, base pairs begin to open allowing the imino proton to exchange with H₂O. This is observed in the NMR spectrum as a broadening, and eventual disappearance, of the imino resonance. Although the imino proton exchange rate is not limited by base pair opening (Leroy, J-L. et al. 1985), the order of resonance broadening generally does follow the order of melting. Treating the helical stems as individual sub-domains of the hammerhead, the opening of individual base pairs can be monitored as each helical stem melts. The melting behavior was used to corroborate the NOE assignments by looking for apparent aberrant stability of base pairs.
Based on helical length, stem I should melt before stem II and probably simultaneously with stem III. As the temperature is increased from 20 to 25°C, the terminal base pairs and the base pairs closing the hairpin loops of stems I and III (resonances b, e, r, and t) begin to open and exchange with H₂O. Upon increasing to 30°C, resonance e has completely disappeared and peaks a, i, r, and t of stem I have broadened significantly. In addition, the terminal base pairs of stem II (resonances f, l, m'/p) begin to open and exchange with H₂O. Increasing the temperature to 35°C results in a further broadening of these resonances, and finally, at 40°C only resonances m (the GC base pair of stem I), n and k (the two GC pairs of stem III), and c, o, g, h, q, and s (the center five base pairs of stem II) remain stable. This uniform "unzipping" from the ends of the individual helical sub-domains is in agreement with the assignments of figure 1, and suggests that there is no cooperativity between the helical stems upon denaturation.

If the RNA tumbles isotropically in solution, we can expect imino proton resonances corresponding to base pairs with long lifetimes on the NMR time scale to have similar linewidths. It is clear in figure 4 that this is not the case. The internal resonances of stem III (d, n, and k) exhibit slightly narrower linewidths than do similarly stable base pair resonances associated with stems I and II. Because narrow linewidths are indicative of increased molecular motion, this suggests that stem III possesses motion, possibly due to the nick introduced during
cleavage, which is independent of the global structure tumbling in solution. Stem III can, in essence, "flap around" independently of the rest of the structure. One would expect that as the temperature increases and the structure denatures the resonances of stem I and II would approach the linewidths of d, n, and k. Comparison of resonances j and k as the temperature increases from 35 to 40°C shows that this is indeed the case.

Because magnesium ions are required for the cleavage reaction, magnesium chloride was added to the sample with the rationale that it would stabilize tertiary base-pairing interactions necessary for cleavage. Figure 5 shows the imino spectrum as a function of temperature in 10mM MgCl₂. Comparing the 20°C spectrum with that taken in the absence of MgCl₂ (fig. 4), it is evident that addition of magnesium ions has not introduced tertiary base pairs, as there are no additional resonances in the imino proton spectrum. But this is not to say that addition of MgCl₂ has not perturbed the structure. There is a general loss of chemical shift dispersion, especially in the regions containing peaks a-f and l-n. Also, resonances b and c as well as t and u have become completely degenerate, suggesting that there is some structural rearrangement upon the addition of magnesium ions.

Addition of magnesium ions has also increased the stability of the helical stems. At 40°C resonances a, i, and m of stem I remain, whereas in the absence of magnesium only resonance m remains at this temperature. Also stabilized are the terminal
base pairs of stem II (resonances f, l, m', and p). However, this increased stability is probably due to the increased ionic strength of the solution rather than magnesium ion-specific interactions. As the temperature is decreased to 10°C, the spectrum broadens resulting in a loss of resolution. The most severe broadening occurs near the junction of the three helical stems (resonances e, l, and i). This effect is probably associated with Mg$^{2+}$ binding near fast exchange, but without further data we can't be much more specific. In fact, the cleaved form may not have a "good" Mg$^{2+}$ binding site.

It is interesting to note that the resonances of stem III in the presence of magnesium ions do not exhibit narrower linewidths than do the resonances of stems I and III as previously discussed. Although resonances n and d are no longer well enough resolved for linewidth comparison, it is clear that k, which is still well resolved in 10mM MgCl$_2$, is not narrower than resonance j. In fact, below 25°C resonance k is much broader than j. The addition of magnesium ions to the sample has apparently "tacked down" the once independent stem III, thereby eliminating its independent motion.

DISCUSSION

The focus of this study was two-fold. Ultimately our goal is to identify tertiary interactions, in the form of base pairing between conserved nucleotides, which drive the cleavage conformation of a hammerhead-type self-cleaving RNA. However, to achieve this goal it was necessary to design an RNA that not only
cleaved, but which also exhibited an imino proton spectrum well enough resolved to be completely, unambiguously assigned. This intermediate goal was achieved by replacing two Watson-Crick GC base pairs with GU wobble base pairs. While we have successfully assigned the imino proton spectrum, thereby substantiating, with the exception of two terminal AU base pairs, the hammerhead folding model, we have not reached our final goal. Addition of MgCl₂ to the RNA sample did not produce additional imino proton resonances as expected. It is difficult to rationalize the conservation of 13 nucleotides in such a small structure unless they interact in a sequence specific manner. Therefore we conclude that cleavage of the RNA results in a subsequent rearrangement of the structure, and this cleaved form does not represent a true cleavage conformation. To test this conclusion, it is necessary to produce RNA of essentially the same sequence which is slowly cleaving. Then biologically active uncleaved RNA can be purified and this imino proton spectrum compared to that of the cleaved form. We are currently investigating slowly cleaving forms of this RNA sequence to pursue this goal.

ACKNOWLEDGEMENTS

We would like to thank Mr. David Koh for synthesizing the necessary DNA sequences, and Dr. Rosalind Kim for the gift of the cloning vector and for sharing her cloning expertise.

†This work was supported by the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the U.S. Department of Energy under Contract
No. DE-AC03-76SF00098, and through Instrumentation Grants from the U.S. Department of Energy, DE FG05-86ER75281, and the National Science Foundation, DMB 86-09035.
Figure Legends

Figure 1. RNA sequence arranged in the hammerhead motif and the 500 MHz imino proton spectrum taken at 20°C in 90% H₂O/10% D₂O, 20mM phosphate buffer, pH 7, 0.1M NaCl, 0.2mM EDTA, and 0.02% NaN₃. The conserved nucleotides are shown in outline print and the cleavage site is denoted by the arrow. The proposed base pairs are connected with a dot. The assignments are discussed in the text.

Figure 2. (A) 500 MHz spectrum of Fig. 1 expanded to show the imino and aromatic proton region. (B-D) Difference spectra taken at 20°C. The large, unlabelled resonances denote the pre-irradiated peaks. NOEs are labelled in lower case letters according to the reference spectrum (A). Peaks due to spill-over irradiation and second-order NOEs are labelled with an "x". (E-F) Difference spectra taken at 15°C.

Figure 3. (A) 500 MHz spectrum of Fig. 1 expanded to show the imino and aromatic proton region. (B-F) Difference spectra taken at 20°C. The large, unlabelled resonances denote the pre-irradiated peaks. NOEs are labelled in lower case letters according to the reference spectrum (A). Peaks due to spill-over irradiation and second-order NOEs are labelled with an "x".
Figure 4. 500 MHz imino proton spectrum plotted as a function of temperature in 90% H₂O/10% D₂O, 20mM phosphate buffer, pH 7, 0.1M NaCl, 0.2mM EDTA, and 0.02% NaN₃. The temperature at which each spectrum was recorded is marked in the right-hand margin in degrees Celsius.

Figure 5. 500 MHz imino proton spectrum plotted as a function of temperature in 90% H₂O/10% D₂O, 20mM phosphate buffer, pH 7, 0.1M NaCl, 0.2mM EDTA, 0.02% NaN₃, and 10mM MgCl₂. The temperature at which each spectrum was recorded is marked in the right-hand margin in degrees Celsius.
REFERENCES


Hare, D. & Reid, B. R. (1982a) Biochemistry 21, 1835-1842

Hare, D. & Reid, B. R. (1982b) Biochemistry 21, 5129-5135


Hurd, R. E. & Reid, B. R. (1979) Biochemistry 18, 4017-4024


Johnston, P. D. & Redfield, A. G. (1978a) Nucleic Acids Res. 4, 3599-3615

Johnston, P. D. & Redfield, A. G. (1978b) Nucleic Acids Res. 5, 3913-3927


Reid, B. R. & Hurd, R. E. (1979) Biochemistry 18, 3996


