Single Molecule Characterization of the Varkud Satellite Ribozyme and Bulk Native Purification of Non-coding RNA

by

Miguel J. Pereira

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2009

Doctoral Committee:

Associate Professor Nils G. Walter, Chair
Professor Roseanne J. Sension
Associate Professor Anna K. Mapp
Assistant Professor Bruce A. Palfey

ACKNOWLEGEMENTS

I firmly believe that the best science is not done in solitude but rather in a supportive and intellectually challenging environment. The time spent in the laboratory of my advisor Nils Walter has truly been such a place. His support is embodied in his open door policy and led to me knocking on his door at completely random times. Each time he answered, I could count on him being happy to take some time out to talk to me, no matter how busy he was. His sage insights were instrumental in getting me to where I am today. Thank you for being a great mentor.

The other current members and alumni of the Walter lab were also invaluable in making this endeavor a fulfilling one. To that group thanks for all the help with my experiments, thanks for all the good conversation, and all the best in your years to come.

To the love of my life, Jen Pereira, thank you for your kind listening and support each and every time that day's experiment laughed at me. To my brother and mom, thank you for being there for me, as you have always been.
And finally to my deceased father, I have no idea if you can read this or not, but if you can, this one is for you☺.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

Non-coding RNAs of complex tertiary structure are involved in numerous aspects of the replication and processing of genetic information in all organisms; however, an understanding of the complex relationship between their structural dynamics and function is only slowly emerging. The *Neurospora* Varkud Satellite (VS) ribozyme provides a model system to address this relationship. First, it adopts a tertiary structure assembled from common elements, a kissing loop and two three-way junctions. Second, catalytic activity of the ribozyme is essential for replication of VS RNA *in vivo* and can be assayed *in vitro*. Here we utilize single molecule fluorescence resonance energy transfer (FRET) microscopy to show that the VS ribozyme exhibits previously unobserved dynamic and hierarchical folding into an active structure. Readily reversible kissing loop formation combined with the slow observed cleavage of the VS RNA lead to the discovery of an additional downstream barrier that must be overcome before the VS RNA can cleave its immediate upstream substrate.

An alternative substrate available to the VS ribozyme is the distal downstream substrate that is part of the next monomer of the multimeric replication intermediate of VS RNA. Recent evidence suggests this may actually be the preferred substrate *in vivo*. The role that structural dynamics play in this preference was unknown and is investigated here. Single molecule FRET studies of an RNA that models this alternate substrate selection show that in contrast to the previously observed global dynamics, these VS RNA molecules exhibit comparably little global dynamics. These differences in folding characteristics are correlated to the large difference in cleavage rates. These observed differences in structural dynamics, in turn, are consistent with a proposed pathway *in vivo*

that minimizes exposure of the ends of linear VS RNA to the exonucleases of the host cell.

The VS ribozyme investigated here is but one example of a plethora of non-coding RNAs that play a central role in a multitude of essential cellular processes. A deeper understanding of how these other RNAs function when in their active states requires the design of experiments that accurately replicate what those RNA molecules are doing in nature. A prerequisite to these experiments is an RNA preparation procedure whose end-product is the target RNA in the same structure as found in the cell. To this end, a novel native RNA purification procedure has been developed. This protocol yields a pure target RNA whose structure more closely resembles what is found in nature than what the current standard RNA purification methods achieve.

CHAPTER 1:

EXPLORATION OF HOW RNA FUNCTIONS IN THE CHEMISTRY OF LIFE

1.1 Introduction

At the most fundamental level, all biological reactions occur through the function of single biomolecules. Our quest to better understand how that function is achieved lies at the heart of much of the research carried out over the past century. This endeavor started out with the development of ensemble assays where the experimental observable arose from the contribution of countless numbers of the macromolecule being investigated. Proper interpretation of this observable required an assessment of whether or not all of those molecules behaved in the same way. This introduces the paradoxical question of how can one confidently ascertain the heterogeneity or lack thereof in a sample of interest if the motivation for investigation was a lack of understanding it? This conundrum is circumvented if the experiment probes the behavior of individual molecules.

The development of single-molecule spectroscopy techniques has made such experiments a reality. Through the time of this writing, a large portion of these techniques has centered on the use of fluorescence as the observable to report on the behavior of the target. Figure 1-1 illustrates this growth of FRET based single-molecule publications. This trend also illustrates how the perceived role that RNA plays in the cell has evolved. This role that was postulated for RNA in the central dogma of molecular biology was initially limited to that of a passive

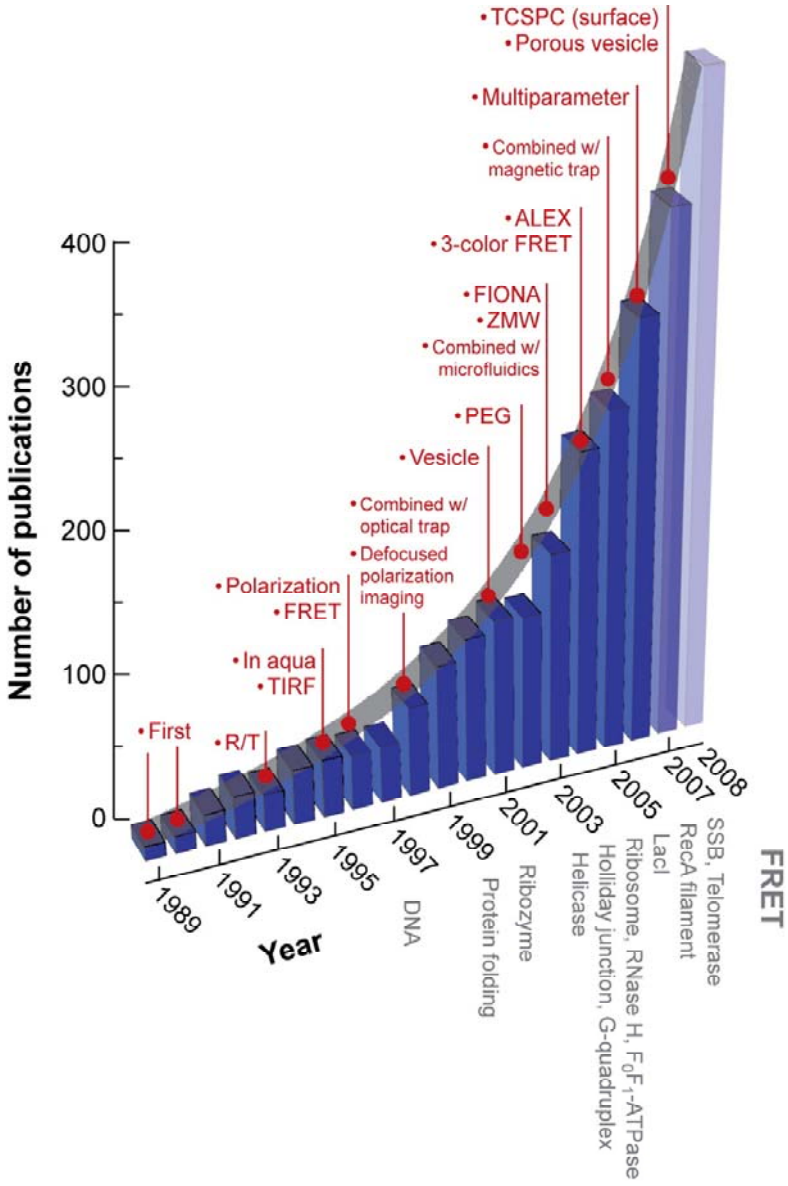


Figure 1-1. The exponential growth of single molecule work. Shown is the number of publications that come up each year when the keyword "single-molecule fluorescence" is searched in PUBMED. Major technical advances are marked in red. Reprinted, with permission, from the Annual Review of Biochemistry, Volume 77 © 2008 by Annual Reviews www.annualreviews.org

intermediary, allowing the information stored in the DNA genome to be converted into specific protein sequences. Today this view has completely transformed. It is now recognized that non-coding RNA's play a central role in many aspects of the replication, processing, modification, and regulation of genetic information (2-4).

1.2 Single Molecule FRET elucidates the structural dynamics of ribozymes

Ribozymes are RNA enzymes that catalyze the breakage/formation of covalent bonds with high specificity (5). Some of these ribozymes catalyze the site-specific cleavage of their own phosphodiester backbone (6). Their catalytic activity made these fascinating non-coding RNA's an active area of research partly because of the relative ease to assay for function. As with protein enzymes, there is strong motivation to develop single molecule methodologies to gain further insight into how these ribozymes work (7).

One of the pioneering single-molecule studies in this area was the investigation of the self-splicing *Tetrahymena thermophila* group I intron ribozyme (8). The authors employed total internal reflection fluorescence (TIRF) microscopy (9) to monitor the simultaneous emission intensities of a Cy3 and a Cy5 fluorophore, each of which was attached to a single *Tetrahymena* ribozyme molecule. These two fluorophores undergo distance-dependent fluorescence resonance energy transfer (FRET) (10) and, as a consequence, global distance changes between them result in relative changes in their respective emission intensities. The individual *Tetrahymena* ribozyme molecule is immobilized on a quartz slide where an evanescent light field generated by total internal reflection at the quartz-solution interface excites the Cy3 and both emission intensities are measured (8).

This ability to monitor global changes in structure in real-time led to a deeper understanding of how these RNA enzymes actually behave in nature as well as what fundamental principles drive such behavior. No other ribozyme is more exemplary of this than the hairpin ribozyme (11-16). Its smaller size allowed

its synthesis from standard phosphoramidite chemistry techniques. By this approach, modified bases can be site-specifically incorporated anywhere in the sequence, greatly expanding the types of questions that can be answered by single-molecule TIRF microscopy. Previously hidden from view in ensemble fluorescence assays, it was discovered that the hairpin ribozyme adopts four distinct structural folds when in the catalytically active docked state (11). These four folds revealed themselves through the observation of four distinct rates of the hairpin ribozyme's undocking from its active structure into an extended, inactive state that places the cleavage and active site far apart. Surprisingly, there was no inter-conversion between these four structures. It turns out that the individual hairpin molecules "remember" which docked structure they last adopted because it is that same structure that is formed once the next docking event happens (17). In addition, the introduction of targeted mutations has led to unexpected and insightful results (12). Modifications of essential residues distant from the site of catalysis altered the rate constants not only of docking and undocking but also of catalytic chemistry. This supported the idea that coupled molecular motions connect remote parts of the hairpin ribozyme fold with its functional core. Similar observations have been made for protein enzymes (18).

The inner workings of other ribozymes have begun to be uncovered as well through single-molecule fluorescence experiments. This includes a ribozyme whose presence is a prerequisite for every single life form on earth, the ribosome (19, 20). The complex kinetics of the multitude of conformational rearrangements that occur during each round of amino acid incorporation in translation have begun to be teased out using similar single molecule experiments as were done earlier on the hairpin ribozyme.

The hairpin ribozyme, along with the hepatitis delta virus (HDV), hammerhead, and Varkud Satellite (VS) ribozymes, are categorized as the group of "small ribozymes" (21-23). Of these, the hairpin ribozyme is arguable the best understood due in part to the significant number of single molecule publications devoted to it (11-16). By comparison, the VS ribozyme has posed many challenges to the RNA research community. Study of this ribozyme has been

hampered because its larger size precludes the possibility of studying it with just synthetic RNA. The use of transcription reactions has been the predominant means of producing it. This method, however, does limit the options available as far as non-natural site specific modifications are concerned.

Nevertheless, a fair amount is known about the VS ribozyme. The global architecture of the naturally occurring VS ribozyme and its relation to catalytic activity have so far been studied by ensemble activity and FRET assays, mutagenesis, native gel electrophoresis, hydroxyl radical footprinting, and UV-induced crosslinking (24-28). In combination, these studies have yielded two similar, if incomplete models of how the six helical elements self-assemble to juxtapose helices I and VI in the catalytic core (26, 28). These models depict the global architecture of the active VS ribozyme as built from the II-III-VI and III-IV-V three-way junctions and the I-V kissing loop interaction. While it is known that ribozymes commonly undergo dynamic structural changes along their reaction pathways (8, 29, 30), no study had yet addressed whether alternate global folds of the VS ribozyme exist, what their rates of inter-conversion may be, and how they may relate to biological function. Single molecule FRET is a particularly suitable tool to identify even short-lived conformational isomers and report on their exchange kinetics and heterogeneities under equilibrium conditions (31). Figure 1-2 illustrates how of total internal reflection (TIRF) single-molecule spectroscopy is applied to the VS ribozyme. Its initial use in the study of ribozymes began in 2000 (Figure 1-1) (8). As mentioned earlier, those studies lead to a fundamentally better understanding of how the dynamics of *Tetrahymena* and hairpin ribozymes are intimately related to their function. The single molecule characterization of the VS ribozyme described in chapters 2 and 3 also tells a revealing story of how its conformational dynamics and folding pathway are a key component to its ability to function.

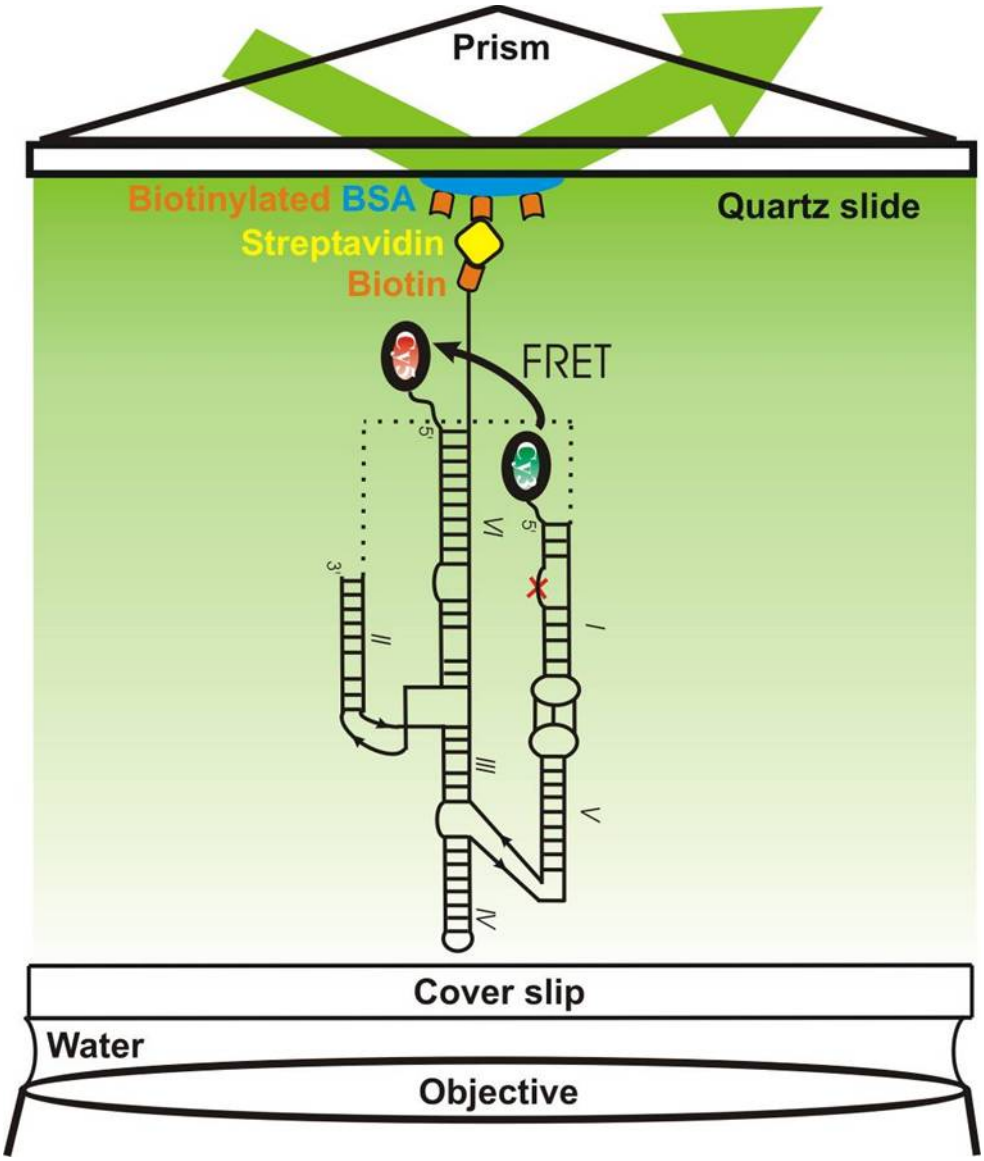


Figure 1-2. Single-molecule FRET microscopy diagram of a VS molecule. Surface immobilization of a VS RNA occurs when the biotin on the 3'end forms a stable complex with a streptavidin that is also bound to biotinylated BSA protein that is coated on the quartz surface. An evanescent field is generated to excite the Cy3 donor that can undergo FRET with the Cy5. The intensities of both emissions are sensitive to distance changes between the Cy3 and Cy5. This method of monitoring global structural changes over a time window of minutes has shown to be applicable to a diverse array of important macromolecules.

1.3 Using denaturing polyacrylamide gel electrophoresis to produce RNA: there is a better way

Despite repeated attempts by different groups, there is no atomic resolution structure of the VS ribozyme. This stands in stark contrast to the other small ribozymes (32-40), whose crystal structures have led to more penetrating and insightful conclusions from the variety of published biochemical studies. Several reasons have been proposed for why high-diffraction quality crystals so far have not been grown for the VS ribozyme. Perhaps it is too dynamic. However, the successfully crystallized hairpin ribozyme was shown to be a very dynamic RNA in solution (11), where its catalytically active docked conformation readily undocks into the vastly different undocked state. Both the successful HDV and hairpin ribozyme crystallization work utilized the binding of the positively charged U1A protein (41) to aid in crystal growth. This approach failed with the VS ribozyme, even though different U1A binding sites were tested.

These failed attempts began with the same protocol: carry out a large scale transcription reaction, separate by size all of the generated RNAs on a denaturing polyacrylamide gel, elute the desired band, and attempt to refold all of the eluted RNA back into its pre-denatured structure. The VS ribozyme did not evolve with selective pressure to refold back into its native structure once it has been denatured into a linear monomer. It is a reasonable proposition that in this case the denaturation is a somewhat irreversible process, much in the same manner that some proteins are permanently inactive once they are denatured *in vitro*. Compounding this issue is that the T7 RNA Polymerase protein that synthesizes the RNA off of the DNA template strand can introduce additional bases at the terminal 3' end (42) of the transcript and occasionally at the 5' end (43). Crystal structures of the hairpin and hammerhead ribozymes circumvented this issue through the use of synthetic RNA. This approach is untenable, however, for the VS ribozyme.

A new attempt at crystallizating the VS ribozyme would have to bypass both of these problems. The purified RNA would need to have sequence

homogeneity at both the 5' and 3' ends. In addition, once it was folded during the transcription reaction into its active conformation, it would have to remain that way throughout the purification process. At the same time all of the other reaction components need to be removed. The strategy to do this was built upon earlier published work (44-47) that had the same goal of achieving native RNA purification. Because the shortcomings of this published work appeared preventable, it was decided that development of a novel approach should be attempted. The methodology of native RNA purification that was born out of this development is described in chapter 4. It describes a protocol that is transferable to many research laboratories and yields an RNA sample in desirable quantities that is chemically pure and folded co-transcriptionally.

Prior to the completion of this dissertation, the body of knowledge of the folding behavior of non-coding RNA's was weighted towards those of the smallest size. The greater structural complexity of the VS ribozyme had so far precluded its study at the single molecule level. Nevertheless, such characterization of these more complex RNAs remains just as important. The body of knowledge accumulated in the prior single molecule studies did not contain any fundamental obstacle to studies like that illustrated in Figure 1-2. One goal of this dissertation was to gain a deeper understanding of the important role of folding and dynamics in the VS ribozyme.

The other major goal was the development of a purification assay that will enable the required preparation of target RNA that many biochemical assays demand. This goal evolved from the conclusion that the failed crystallization attempts of the VS ribozyme were due to shortcomings in its purification. These same shortcomings will continue to prevent other valuable insights from being realized as well. This fact was more than enough motivation to develop the protocol described in chapter 4.

CHAPTER 2:

SINGLE MOLECULE MICROSCOPY ELUCIDATES THE FUNDAMENTAL ROLE DYNAMICS PLAY IN THE FUNCTION OF THE VS RIBOZYME[a]

2.1 Introduction

Varkud Satellite (VS) RNA is 881 nucleotide long and is found in the mitochondria of genetically and geographically divergent species of the bread mold *Neurospora* (48). It is transcribed from a circular VS plasmid template by the host mitochondrial RNA polymerase to yield multimeric VS RNA copies (49). Embedded in each monomer of this repeating RNA is the VS ribozyme, minimally a 154-nucleotide sequence (50) that catalyzes a site-specific phosphodiester bond breakage to yield the 881-nucleotide monomers. These monomers in turn act as template for a reverse transcriptase that leads to the eventual regeneration of the VS plasmid. The action of the catalytic VS ribozyme RNA is essential to the replication of the VS plasmid in its fungal host.

The VS ribozyme is the least understood (6, 51-53) among the class of small pathogenic ribozymes that include the Hepatitis Delta Virus (HDV), Hammerhead, and Hairpin ribozymes. This lack of knowledge is due to the fact that the VS ribozyme is the largest of the small catalytic RNAs and less amenable to the application of modified synthetic RNAs in studies of its structure-function relationships. Site-directed mutagenesis, footprinting, and cross-linking studies (24-26) have revealed the secondary structure of the well characterized

[a] Reproduced in part from Pereira, M.J.P., Nikolova, E.N., Hiley, S.L., Jaikaran, D., Collins, R.A., and Walter, N.G. *Journal of Molecular Biology* **382**, 496-509. Copyright 2008 Elsevier Ltd.

G11 construct (54), featuring the six helices shown in Figure 2-1A. Furthermore, A756 has been implicated as a general acid/base during cleavage of the scissile phosphate located 3' of G620 (55). This requires that substrate stem-loop I and the internal bulge in helix VI come into close contact. The essential tertiary kissing interaction between the substrate stem-loop I and stem-loop V (56) is hypothesized to enable this juxtaposition. NMR-derived atomic resolution structures of isolated portions of the VS ribozyme (57-59) have led to reasonable hypotheses on how these distal yet crucial tertiary contacts occur sequentially. Bulk solution activity assays have shown that G11 has a comparatively slow cleavage rate of 0.14 min^{-1} at 37^{o}C (54).

The overall global architecture describing the relative orientation of the six helices has recently been further refined using hydroxyl radical footprinting (25), steady-state FRET measurements of portions of the VS ribozyme (27), and small angle X-ray scattering experiments (60). These data have yielded the incomplete tertiary structure model shown in Figure 2-1C. Using this available information we have developed a new, more representative secondary structure representation (Figure 2-1A), which now reflects the juxtaposition of the cleavage and active site as well as the stem-loop I-V kissing interaction. Despite repeated attempts in other laboratories, there is still no atomic resolution crystal structure of the VS ribozyme. Even with a complete static crystal structure, however, there would still be a large knowledge gap in our understanding of the role that global dynamics play in function of the VS ribozyme. As discussed in chapter 1, single molecule FRET microscopy is a suitable tool to probe deeper into RNA chemistry.

Here we employ single molecule FRET microscopy to measure the global folding dynamics of single VS ribozyme molecules and describe how those dynamics correlate to function. We combine the quantified heterogeneous population distributions and interconversion kinetics with ensemble activity assays and single-site as well as second-site revertant mutagenesis to show: (i) how the dynamics of the I-V kissing interaction affect catalytic activity; (ii) that a slow and hidden degree of freedom must be traversed before the ribozyme undergoes self-cleavage; and (iii) that the II-III-VI junction acts as a structural

scaffold for kissing loop formation.

2.2 Materials and Methods

Preparation of topWT, topΔI-V, toprescue 1-V, and topΔII-III-VI. QuickChange site-directed mutagenesis (Stratagene) on a plasmid encoding G11 (54) was used to introduce a transcription initiation site at G622 (2 nucleotides downstream of the cleavage site) and a termination site at the capping loop of helix VI through insertion of an EcoRI site. Based on this plasmid, defined as the WT sequence, further site-directed mutagenesis was performed to generate plasmids topΔI-V, toprescue 1-V, and topΔII-III-VI and confirmed by DNA sequencing.

Transcription on EcoRI-linearized plasmids was carried out with 20 mM GMP, 1 mM GTP, 4 mM ATP, UTP, CTP, 40 mM Tris-HCl, pH 8.0, 25 mM $MgCl_2$, 2 mM Spermidine, 5 mM DTT, 0.01% TritonX-100, 1 unit inorganic pyrophosphatase, and 0.2 mg T7 RNA Polymerase in a 2-ml reaction at 37 ^{o}C for 6 h. This yielded ~95% 5'-monophosphate initiated transcripts. The 3' ends of these transcripts were biotinylated after periodate oxidation as described (61).

Each of the 5'-monophosphorylated transcripts was ligated to ligS1 as follows. 3.7 μM ligS1, 2.45 μM splintS1 and 1.85 μM transcript (ratios taken from ref (62)) in 200 μl of 50 mM Tris-HCl, pH 7.5, 10 mM DTT, 1 mM ATP was heated to 90 ^{o}C for 2 min, then incubated for 10 min at room temperature. $MgCl_2$ was then added to 10 mM and recombinant T4 DNA ligase to 6 μM (final concentrations) and the reaction incubated at 30 ^{o}C for 6 h. The ligation reactions were desalted by gel filtration on NAP-10 columns (Pharmacia) and ethanol precipitated before purification on denaturing, 8 M urea, 10% (w/v) PAGE. Ligated RNA (yield typically ~40%) was detected based on its Cy3 fluorescence, cut out, eluted into 1 ml H_2O, 1 mM EDTA overnight at 4^{o}C, and ethanol precipitated. Typical total yields of ligated RNA were 10-40 pmol per reaction. Cleavable versions of topWT, topΔI-V, toprescue 1-V, and topΔII-III-VI were

prepared in the same way except that a ligS1 RNA variant was used without the 2'-O-methyl at G620, and the 3' ends of the transcripts were not biotinylated.

Preparation of synthetic oligonucleotides. RNA strands LigS1 (5'-Cy3-AAUUGCmGAA-3') and b1 (5'-NH_2-AAGUCAGUAUUGCAGCACAGCACAAGCCCGCUUGC-3'), where mG and NH_2 represent a 2'-O-methylated G and a 5'-Amino Modifier C6 (Glen Research), respectively, were purchased from the HHMI Keck Foundation Biotechnology Resource Center and deprotected as suggested by the manufacturer (http://info.med.yale.edu/wmkeck). The RNA was purified by denaturing, 8 M urea, 20% (w/v) polyacrylamide gel electrophoresis (PAGE) and C8-reverse-phase HPLC chromatography as described (63). Cy5 was attached postsynthetically to the 5' end of b1 via the succinimidyl ester of Cy5, as described (63) to generate strand Cy5b1. Oligodeoxynucleotide splintS1 (5'-CGGGGCGACGACG CCCTTCGCAATT-3') was purchased from Invitrogen.

Single-Molecule FRET measurements. A 10 µl solution of 100 nM top strand and 200 nM Cy5b1 in 40 mM Tris-HCl, pH 8.0, 50 mM KCl was heated to 70 ^{o}C for 2 min and cooled at room temperature for 10 min to anneal the VS ribozyme strands. The annealed ribozyme was diluted to 200 pM in 1× standard buffer (64) (40 mM Tris-HCl, pH 8.0, 50 mM KCl and 35 mM $MgCl_2$) and flowed into a microfluidic channel on a quartz slide, coated first with biotinylated BSA, then streptavidin as described (11, 31), to generate a surface density of ~0.1 molecules per μm^2 for optical resolution. Similar surface tethering has been extensively used and characterized as non-perturbing for the hairpin ribozyme (11, 13-15). After 2 min, the slide was washed with several volumes of 1× buffer. In parallel, 200 µl oxygen scavenger solution (OSS) was prepared that contained ~750 µg/ml glucose oxidase, 90 µg/ml catalase, 10% (w/v) glucose, and 1% (v/v) β-mercaptoethanol in 1× standard buffer. This solution was then flowed into the microfluidic channel or stored under nitrogen for later use. The donor and acceptor fluorescence signals of optically resolved single molecules were simultaneously monitored in real-time by prism-based total internal reflection

fluorescence video microscopy as described (11, 14, 15). There was little evidence of β-mercaptoethanol induced blinking of the Cy5 under our illumination conditions of 9 W/cm^2; we observed few excursions to a low-FRET state, which may suggest acceptor blinking, and none of the traces were analyzed after they had reached zero signal, as would indicate donor blinking. Our illumination conditions lie well below the excitation intensities used in experiments that characterized β-mercaptoethanol induced blinking of Cy5 (65). The optimal I-CCD frame rate was found to be two frames per second (fps), and all time traces were taken at 2 fps resolution unless otherwise stated.

Single Molecule Data Analysis. The aggregate FRET histograms were derived from molecules where both Cy3 and Cy5 were confirmed to be fluorescent. This criterion introduces a possible bias into the calculation of 4% H state for construct ΔII-III-VI, as the FRET ratio of the L state is indistinguishable from molecules with photobleached Cy5 (data not shown), making the terminal state ambiguous for traces that end in low FRET. In addition, observed dwell times in terminal states are biased towards smaller values due to premature Cy3 photobleaching and our limited observation window (see also below). The average time ΔII-III-VI molecules spend in their longer lived M and L states as based on confirmed dwell times is ~30 s each. To correct for the above effects, ambiguous terminal states were assigned to be of this length. The H state was still found to be 4% abundant, suggesting that our histogram analysis is robust against possible bias.

For dwell-time analysis, noise in time traces was reduced using a previously described non-linear filter with m = 3 and p = 20 as parameters (66, 67). Hidden Markov modeling (HMM) with two to five FRET states was then used to determine the most likely traversed path through each well defined trace as described, with each result undergoing visual inspection to ensure the best possible fit; a two or three-state description sufficed for all traces. The HMM algorithm assigns a single-exponentially distributed probability to each transition. All individual probabilities for each type of transition (H→M, M→H etc.) were

averaged to yield the rate constants reported in Figures 2-2D and 2-5D. The raw dwell-times that arose from the HMM analysis were also used to calculate normalized probability densities as described (68), and fit using Marquardt-Levenberg nonlinear least-squares regression in Origin 7.5. These density plots present an independent test for the single-exponential distribution of dwell-times that is assumed in the HMM analysis. All rate constants reported here from single molecule FRET analysis were corrected for the bias towards smaller values due to premature Cy3 photobleaching and our limited observation window by subtracting 0.2 min^{-1} as described (14).

*Distance estimates.*To obtain fluorophore distance estimates R for the various FRET states the apparent FRET efficiency was calculated from $E_{app} = \frac{I_{Cy5}}{I_{Cy5} + I_{Cy3} \times \frac{(\phi_{Cy5} \times \eta_{Cy5})}{(\phi_{Cy3} \times \eta_{Cy3})}}$, (69)where Φ and η signify the fluorophore quantum yields and detector channel efficiencies, respectively, and the donor and acceptor intensities I_{Cy3} and I_{Cy5}, respectively, were corrected for leakage of 20% of donor photons from the donor into the acceptor channel. R was then calculated from $E_{app} = c[1 + (R / R_0)^6]^{-1}$, where c = 0.69 and R_0 = 54 Å were used as described (70).

Ensemble cleavage assays. Cleavable versions of topWT, topΔI-V, toprescue I-V, and topΔII-III-VI were 3' end labeled with [^{32}P]-pCp using T4 RNA ligase. Trace amounts of 3' end labeled top strand were combined with 500 nM Cy5b1 in 40 µl of 40 mM Tris-HCl, pH 8.0, 50 mM KCl, heated to 70 ^{o}C for 2 min, equilibrated in a 25 ^{o}C bath for 10 min, and a 4-µl aliquot taken for a zero time point. The reaction was initiated by adding 4 µl of the appropriate stock solution of $MgCl_2$. At defined times 4-µl aliquots were taken and quenched by the addition of 10 µl 80% (w/v) formamide, 0.025% (w/v) xylene cyanol, 0.025% (w/v) bromophenol blue and 50 mM EDTA. The cleaved and uncleaved products were separated by denaturing, 8 M urea, 10% (w/v) PAGE. The gel was exposed to a

PhosphorImager screen overnight and quantified using ImageQuant software (Molecular Dynamics). The fraction cleaved versus time was fit with the single exponential first-order rate equation $y(t) = y_0 + A\left(1 - e^{(-k_{obs}t)}\right)$ to determine the observed cleavage rate (k_{obs}) and extent of cleavage (A). The Mg^{2+}-dependence of the cleavage rate constant was fit with a hyperbolic, non-cooperative binding equation of the form:

$$k_{obs} = k_{\max} \frac{[Mg^{2+}]}{[Mg^{2+}] + Mg_{1/2}}$$

where k_{max} is the cleavage rate under saturating Mg^{2+} conditions and $Mg_{1/2}$ is the magnesium half-titration point.

2.3 Results

Single wild-type VS ribozyme molecules display dynamic and heterogeneous three-state folding. Our design for single molecule FRET experiments takes advantage of the well-studied construct G11, derived from the natural VS RNA sequence from four nucleotides upstream to 163 nucleotides downstream of the cleavage site (71). This sequence was only slightly adapted to optimize the construct for FRET studies (Figure 2-1A). In particular, we labeled the 5' end of G11 with the FRET donor Cy3 by an enzymatic ligation approach that also allowed us to introduce a chemistry-blocking 2'-O-methyl modification at G620 for structural studies (Materials and Methods) . In addition, we opened the non-essential (72) (data not shown) closing loop of helix VI to attach a 5'-Cy5 FRET acceptor and a 3'-biotin for surface immobilization to take advantage of prism-based total internal reflection excitation (Materials and Methods). Our construct is thus designed to specifically report on distance changes between helices I and VI that are distal in primary sequence but juxtaposed in the tertiary structure models and compose the catalytic core (Figure 2-1A).

First, we characterized the catalytic activity of our labeled wild-type (WT) construct by standard ensemble assays (Figure 2-1B). The cleavage time course

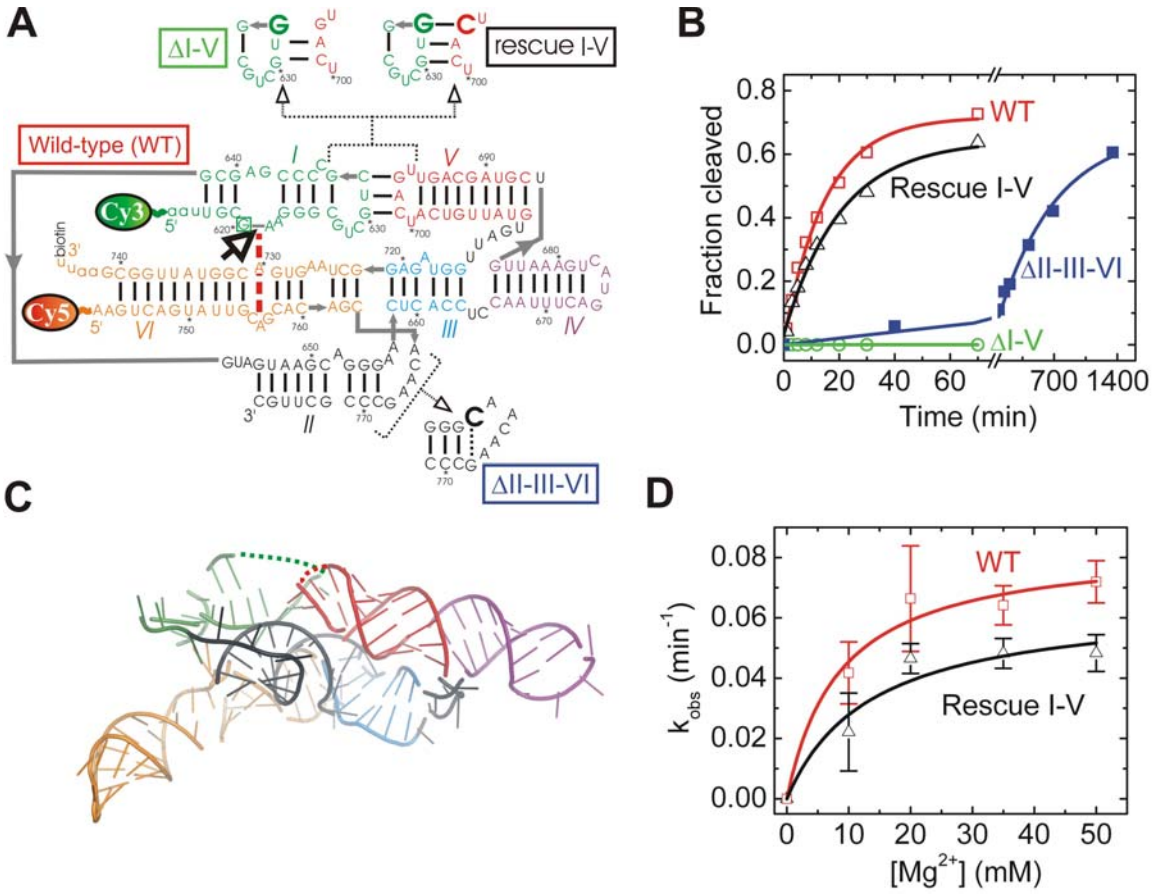


Figure 2-1. The G11 based WT construct that preserves native sequence and variants A) Secondary structure depiction of the four VS ribozyme variants studied here. Helices are color-coded, the bold open arrow denotes the cleavage site, the red dashed line a key interaction in the active site, gray lines denote backbone connectivities, and the Cy3, Cy5 and biotin labeling sites are indicated. Mutations in the ribozyme variants are shown next to their names. All single molecule FRET measurements employed a 2'-O-methyl modification to G620 to render the substrate non-cleavable. B) Cleavage time courses (data points) for each of the four ribozyme variants, with single exponential fits (lines) yielding the observed rate constants reported in the text. C) Tertiary structure model of the VS ribozyme (Zamel and Collins, unpublished data) using the same helix color code as in panel a. Some strand connectivities are incomplete in the model D) Magnesium dependence of the observed cleavage rate constants of WT and Rescue I-V.

is fit well with a single exponential, yielding an observed rate constant of 0.065 ± 0.006 min^{-1} (total cleavage extent of ~75%) under our standard conditions (40 mM Tris-HCl, pH 8.0, 50 mM KCl, and 35 mM $MgCl_2$ at 25 ^oC, see Material and Methods). The unmodified G11 construct has a very similar cleavage rate constant (0.14 min^{-1} in the same buffer (up to ~95% cleavage extent), but at 37 ^oC instead of 25 ^oC (71), which likely accounts for these differences) suggesting that our WT modifications minimally interfere with function. The cleavage rate constant saturated at 35 mM Mg^{2+} ($Mg_{1/2}$ = 9 mM when fit with a hyperbolic binding equation, Figure 2-1D), which was used in all subsequent single molecule experiments.

Immobilized single VS ribozyme molecules exhibit anticorrelated donor and acceptor fluorescence intensities as signature of global conformational changes (Figure 2-2A). (For a discussion of our evidence that fluorophore blinking is insignificant under our low-illumination conditions, please refer to Materials and methods.) Three distinct FRET states and thus distinguishable RNA folds are observed for WT. The particular ribozyme molecule in Figure 2-2A alternatively populates a high (H) FRET state with a FRET ratio (defined as $I_A/(I_A+I_D)$, where I_D and I_A are the donor and acceptor fluorescence intensities, respectively) of ~0.76, a mid (M) FRET state at ~0.50, and a low (L) FRET state at ~0.22. All analyzed time traces end with single-step donor photobleaching, confirming a single molecule as the source of the signal. A histogram of the FRET ratios of 94 WT molecules is well-described by a triple Gaussian that yields the relative population distributions of the three FRET states (Figure 2-2B). Donor-acceptor distances of 35 Å, 55 Å, and 96 Å (each ± 5 Å) were estimated for the H, M, and L states, respectively, after corrections for donor leakage into the acceptor channel, fluorophore photobleaching, and differential detector efficiencies for the donor and acceptor emissions (Materials and Methods) (69, 70).

Shot noise in the raw data was reduced by employing a non-linear filter as described (66). Hidden Markov modeling (HMM) was then used to determine for each molecule the most probable path traversed between the three FRET states

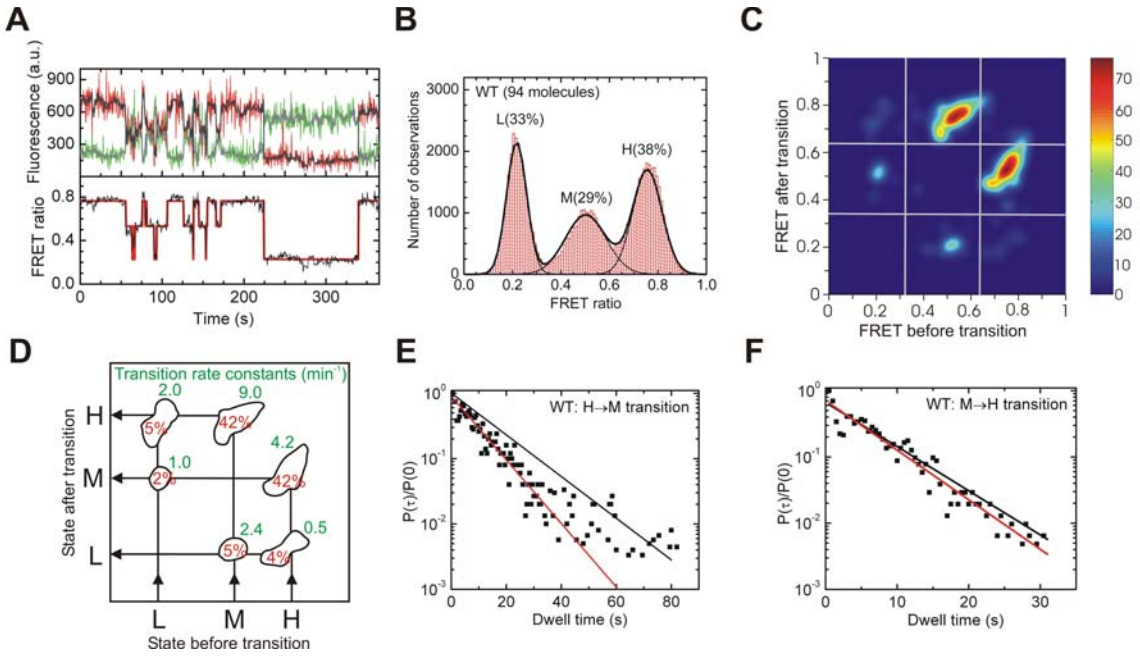


Figure 2-2. Single molecule FRET analysis of the WT VS ribozyme.

A) Representative FRET time trace of a WT molecule. The raw Cy3 donor and Cy5 acceptor fluorescence signals are green and red, respectively. Superimposed in gray are the data after applying a non-linear filter as described (66). The FRET ratio (black trace below) is calculated from the filtered data as $I_A/(I_A + I_D)$ and reveals three distinct states in Hidden Markov modeling (HMM, red line). B) An aggregate FRET histogram of 94 WT molecules shows Gaussian distributions for three states with FRET values of ~0.76 (H), ~0.50 (M) and ~0.22 (L) and the indicated abundances. C) Transition density contour map from HMM analysis (73) of all transitions for 77 well-defined WT molecules. D) Relative contributions and mean rate constants of the transitions illustrated in panel c. E,F) The normalized probability densities (68) of all dwell times preceding an $H^{\rightarrow}M$ (E) and $M^{\rightarrow}H$ (F) transition are fit with single exponential decay function (red lines) and are compared to decay functions (black lines) using the rate constant from the HMM analysis in panel d.

(Figure 2-2A) (67). The transition density contour map in 2-2C, 2-2D summarizes all six possible observed transitions between initial and final FRET states for the analyzed WT molecules. Single-exponential transition rate constants for the H to the M state of 4.2 min^{-1} and for the M to the H state of 9.0 min^{-1} were derived from the hidden Markov fitting procedure (67) (see Materials and methods) (Figure 2-2D). Together, these two conformational changes account for 84% of all observed transitions. Despite the overall predominance of these transitions, 10% of the analyzed molecules never sample the H state. These molecules are instead characterized by long dwell times in the M and L states (Figure 2-3). This small subset of molecules and the generally low transition rate constants for leaving the L state (Figure 2-2D) lead to the observation of 33% L state in the aggregate histogram (Figure 2-2B). These findings provide evidence for heterogeneity at the single molecule level, as previously noticed for other RNAs such as the hairpin ribozyme (11, 14, 15, 74), and suggest that the L state is a less relevant (rarer) structural state of the VS ribozyme.

Markov processes as assumed in HMM consist of a combination of state-to-state transitions with single-exponential decay kinetics (67), which may not always reflect the full complexity of single molecule behavior. To further characterize the nature of the kinetics of the H→M and M→H transitions, we therefore plotted the normalized probability densities of the corresponding dwell times obtained from the HMM path analysis for all WT molecules that accessed the H state at least once (Figure 2-2E, 2-2F). A single-exponential decay fit yields a rate constant of 6.5 min^{-1} for the H→M transition, which is significantly faster than the previously described, averaged Markov rate constant of 4.2 min^{-1} (compare red and black lines in Figure 2-2E). By contrast, both analysis methods yield similar rate constants (10.1 min^{-1} and 9.0 min^{-1}, respectively) for the M→H transition (Figure 2-2F). The discrepancy observed for the H→M transition is due to a higher occurrence of long dwell times than is predicted for a single-exponential distribution, providing evidence for heterogeneity also in the kinetic behavior of the H→M transition.

What is the structural nature of the H state and what role, if any, does it

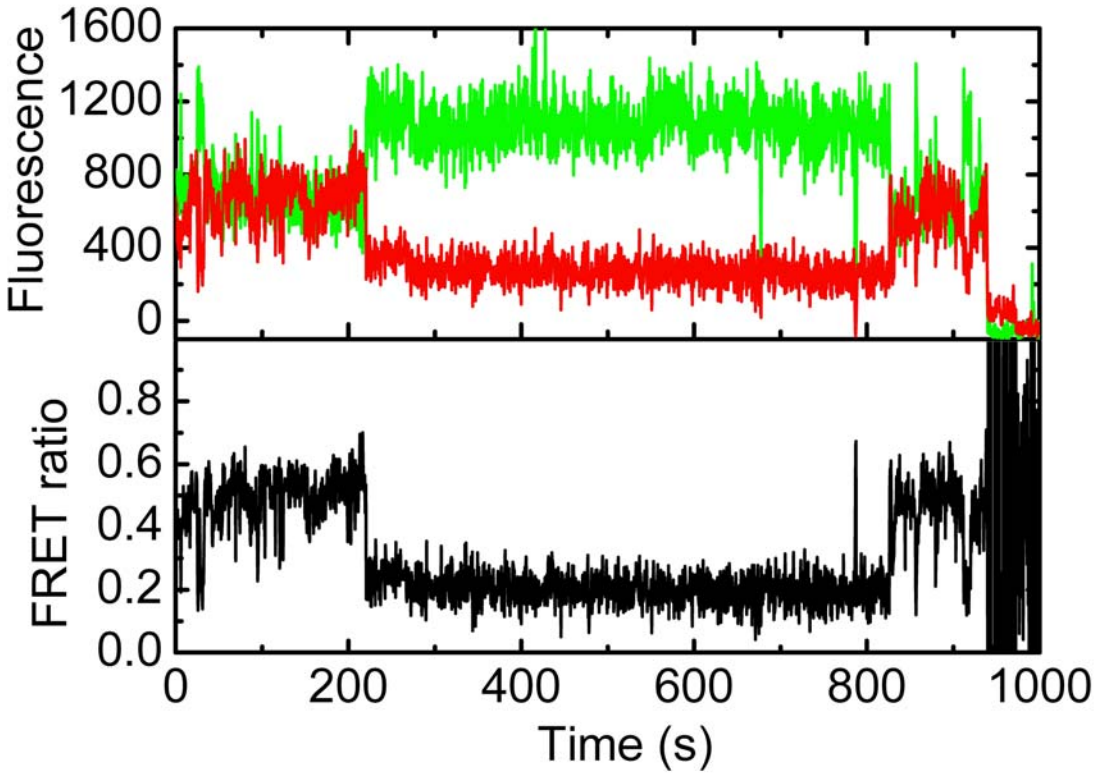


Figure 2-3: A misfolded WT VS molecule. Raw fluorescence and FRET signals of a misfolded wild-type VS ribozyme molecule; the catalytically competent H state of FRET ≈ 0.76 is never occupied. 10% of all WT time traces show this behavior, at least partially explaining the typically ~30% of inactive ribozyme in our ensemble cleavage assays (Figure 2-1B).

play in function? The previously established essential I-V kissing loop interaction has been shown to help dock the cleavage site in helix I with the residues in helix VI that are involved in catalysis (25, 28, 56, 72, 75). Given that our FRET probes are located in helices I and VI, we are able to test the hypothesis that the observed high FRET state reports on the catalytically active docked structure.

Mutation of the kissing loop links the H state to the active docked structure

To ask whether the I-V kissing loop interaction is present in our H state, we designed a ΔI-V construct with a C632→G mutation (Figure 2-1A) that is predicted to weaken the interaction by eliminating one of three base pairs. This mutant was previously shown to lack self-cleavage activity (56) and accordingly our construct yields no detectable product (Figure 2-1B). Significantly, the folding behavior of ΔI-V is markedly different from that of WT (Figure 2-4A). ΔI-V molecules spend long periods of time in global folds characterized by FRET ratios of ~0.48 (M) and ~0.27 (L) (the M state is typically interrupted by brief excursions into the L state, Figure 2-4A). A FRET histogram obtained from 25 molecules shows that the H state was not detected (Figure 2-4B). Due to the long lifetime of the L state (the molecule in Figure 2-4A, for example, dwells in it for ~8 min), its dwell times were ill-defined due to either premature donor photobleaching or the fact that the molecule already inhabits the L state at the beginning of the measurement. Separate measurements were therefore conducted over one hour where the laser was intermittently shuttered for an overall exposure time of only 9 min to limit photobleaching. The normalized probability densities of all observed M→L and L→M transition dwell times for ΔI-V (Figure 2-4E, 2-4F) show a clear distinction between short- and long-lived M and L states. The short lived dwell times were fit to single exponential decay functions yielding rate constants of 5.0 min^{-1} for M→L and 40 min^{-1} for L→M. The average values for the long dwell times are 17 min and 14 min for the M and the L state, respectively, which must be viewed as lower bounds given the natural limitation of our observation window.

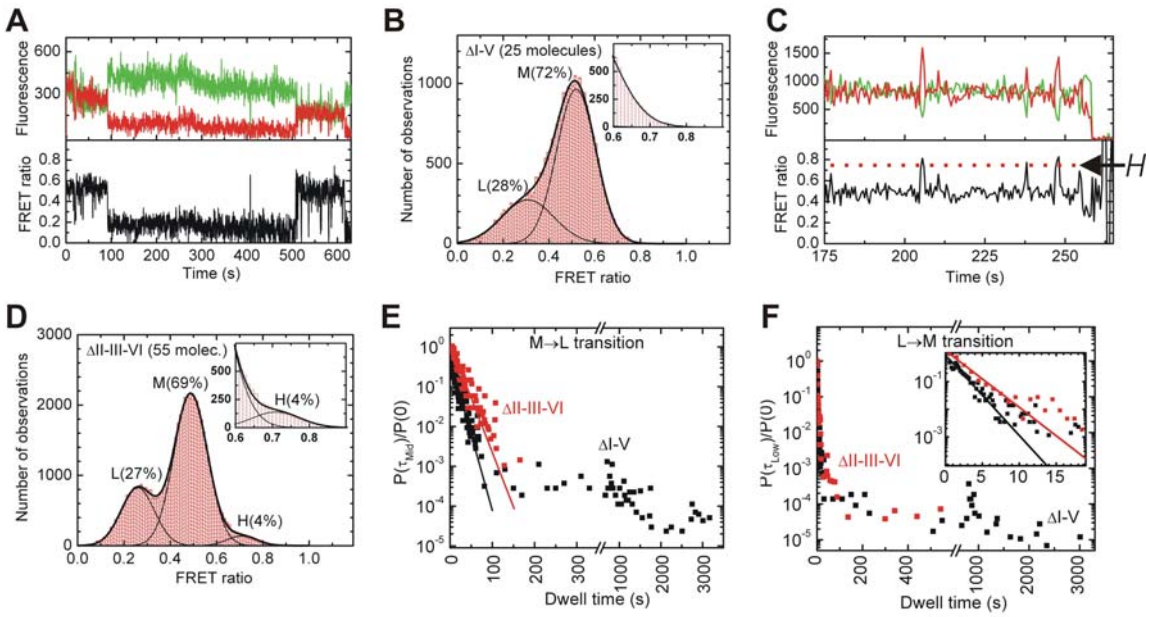


Figure 2-4. Single molecule FRET analysis of the ΔI-V and ΔII-III-VI variants. A) Representative FRET time trace of a ΔI-V molecule. Compared to WT, the H state is not observed in ΔI-V and is replaced by long lived L and M states. The M state shows brief excursions to the L state. B) An aggregate FRET histogram of the raw data for the 25 ΔI-V molecules is well described by Gaussian distributions for only two states with FRET ratios of ~0.48 (M) and ~0.27 (L), but no H state. C) Representative FRET time trace of a ΔII-III-VI molecule, showing brief excursions into the H state, as well as an early excursion into the L state. D) An aggregate FRET histogram of the raw data for 55 ΔII-III-VI molecules reports a ~0.69 (H) state that is populated ~4% of the time in addition to the ~0.49 (M) and ~0.25 (L) states (see inset). Please note that not all molecules visit all three states due to molecular heterogeneity so that these aggregate histograms do not reflect equilibrium distributions. E,F) The short dwell times of the normalized dwell time probability densities preceding an M→L (E) and L→M (F) transition in the ΔI-V and ΔII-III-VI variants are fit with single exponential decay functions (lines), revealing discrepancies from homogeneous kinetics. Please note that the signal-to-noise ratio in the M state due to shot noise did not allow for reliable HMM analysis of the short-lived H state of the ΔII-III-VI ribozyme.

The failure to observe the H state upon weakening of the kissing loop interaction suggests that adoption of the H state requires an intact kissing loop. The accompanying loss of cleavage activity links the H state with biological function of the VS ribozyme. Consistent with these observations, previous hydroxyl radical footprinting studies have observed a decrease in solvent protection of the riboses of several nucleotides involved in the kissing loop interaction when all three interacting bases of either loop I or V were mutated (25). An NMR study of stem-loop V predicts that the three kissing base pairs can form a short A-type helix (76), which we find to be disrupted by the single G-G mismatch in ΔI-V (Figure 2-1A). The time resolution of our data collection (500 ms) allows the possibility that stems I and V may transiently (for <1 s) come into close contact, but their interaction is not sufficiently stabilized to be observed by either single molecule FRET or catalytic activity.

The II-III-VI junction plays the role of an architectural scaffold. Correct juxtaposition of helices II, III, and VI has been implicated in playing an important role in VS ribozyme function. In particular, an A656C mutation at this junction has been shown to substantially reduce the cleavage rate constant in a trans-cleaving variant of the VS ribozyme, presumably by forming a structurally detrimental Watson-Crick base pair with G768 (27). To further explore the structural basis for this large reduction in activity we introduced the same mutation, generating construct ΔII-III-VI (Figure 2-1A), and assayed its cleavage activity (Figure 2-1B). The ΔII-III-VI mutant cleaves with a substantially reduced rate constant of 0.002 min^{-1}, as expected (27), further validating the relevance of our single molecule FRET construct.

Figure 2-4C shows a representative section of a time trace of a single ΔII-III-VI molecule. It predominates in the M state but also shows brief excursions into a high-FRET state similar to the H state of WT. These transient excursions are also evident from the aggregate FRET histogram, which includes a 4% contribution of the H state (Figure 2-4D). By contrast, the H state is never observed to be accessed by the ΔI-V construct (Figure 2-4B).

In conjunction with our earlier observation that the H state reports on formation of the kissing loop, our single molecule studies of the A656C mutant provide evidence that a correct architecture of the II-III-VI junction is also a prerequisite for kissing loop formation. This finding implies that the A656C mutation induces local changes in the II-III-VI junction that have a cascading, distal effect preventing loops I and V from docking. Previous steady-state FRET measurements on the isolated II-III-VI junction have posited a change in the tilt angle between helix II and the III/VI helical stack as a consequence of the local change in sequence (27). Such a rearrangement in three-way junction architecture may then disfavor kissing loop formation as observed here.

Disruption of the II-III-VI junction also has a profound effect on the transition kinetics between the M and L states. While many single ΔI-V molecules with an unaltered II-III-VI junction exhibit very long dwell times in the M state (~17 min, see above and Figure 2-4E, black symbols), the normalized probability density plot for the M→L transition of ΔII-III-VI instead is well described by a single rate constant of 3.6 min^{-1} with few, if any molecules exhibiting long dwell times in the M state (Figure 2-4E, red symbols). The normalized probability distribution of L→M transition dwell times for ΔII-III-VI, by contrast, remains multi-exponential including relatively long dwell times (Figure 2-4F). Fitting the dwell times of the rapid L→M transition yields a rate constant of 26 min^{-1} (Figure 2-4F), while the average of all longer dwell times is 3.2 min, somewhat shorter than the 14 min observed for ΔI-V.

Taken together, these observations are consistent with the notion that a properly folded II-III-VI junction, in the absence of the I-V kissing loop interaction, is associated with a very stable M state that forms an architectural scaffold for the long-range kissing loop interaction to build on. This stable scaffold is disrupted in the A656C mutant, consistent with the general lack of long M dwell times for ΔII-III-VI in Fig. 3e. By contrast, the long fluorophore distance in the L state (96 Å) together with the fact that both the ΔI-V and ΔII-III-VI constructs similarly show long L dwell times leads us to propose that L states generally do not have a properly folded II-III-VI junction, thus unraveling the central

architectural scaffold of the VS ribozyme in a way that dramatically increases the fluorophore distance. This model thus associates the L state with a misfolded II-III-VI junction.

Second-site reversion mutagenesis rescues both H state and ribozyme activity. Our ΔI-V construct completely loses its H state, presumably because of an inability to form a stable kissing loop interaction. If so, recovery of this interaction should restore the structural dynamics observed for WT. To test this hypothesis we designed a Rescue I-V construct with a second-site reversion mutation that flips the WT C:G to a G:C base pair (Figure 2-1A). The representative time trace in Figure 2-5A shows that the H state is indeed recovered, further providing evidence that the H state reports on the formation of the catalytically essential kissing loop interaction. Rescue I-V molecules spend 36% of their time in the H state (Figure 2-5B) versus 38% for WT (Figure 2-2B), suggesting that kissing loop rescue is nearly complete. This notion is further supported by the very similar transition density contour maps of Rescue I-V and WT (Figure 2-5D versus Figure 2-2D).

The transition density contour map of Rescue I-V is also similar to that of WT with the predominant transitions occurring between the H and M states. Comparison of the averaged transition rate constants from the contour map with those derived from analysis of the normalized dwell time probability densities shows good agreement (Figure 2-5E, 2-5F). This homogeneity of the transition kinetics contrasts with the significant heterogeneity observed for the H→M transition of WT (Figure 2-2E), thus illustrating how the simple reversal of a G:C basepair leads to subtle differences in dynamic behavior detectable by single molecule probing.

Rescue of the kissing loop interaction and reemergence of the H state coincide with substantial recovery of catalytic activity in Rescue I-V (Figure 2-1B). The Mg^{2+} dependence curve of Rescue I-V is predicted to saturate at 0.065 min^{-1} with a Mg^{2+} half-titration point of 14 mM, just slightly below and above the WT values of 0.084 min^{-1} and 9 mM, respectively (Figure 2-1D). These

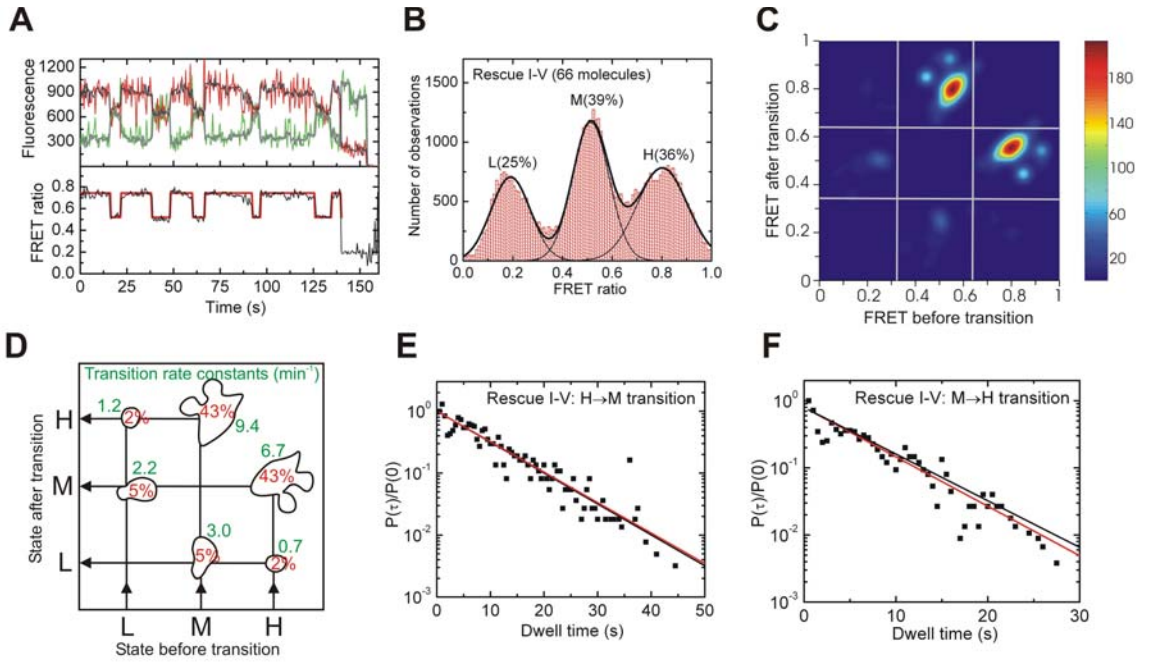


Figure 2-5. Single molecule FRET analysis of the Rescue I-V ribozyme.

A) Representative FRET time trace of a Rescue I-V molecule. The raw Cy3 donor and Cy5 acceptor fluorescence signals are green and red, respectively. Superimposed in gray are the data after applying a non-linear filter as described. The FRET ratio (black trace below) is calculated from the filtered data as $I_A/(I_A + I_D)$ and reveals three distinct states in Hidden Markov modeling (HMM, red line). B) An aggregate FRET histogram of 66 Rescue I-V molecules shows Gaussian distributions for three states with FRET values of ~0.80 (H), ~0.51 (M) and ~0.19 (L) and the indicated abundances. C) Transition density contour map from HMM analysis of all transitions for 62 well defined Rescue I-V molecules. D) Relative contributions and mean rate constants of the transitions illustrated in panel c. E,F) The normalized probability densities of all dwell times preceding an H→M (E) and M→H (F) transition are fit with single exponential decay functions (red lines) that are compared to decay functions (black lines) using the rate constant from the HMM analysis in panel d.

findings provide additional strong evidence that link the H state observed in single molecule FRET experiments to on-pathway folding of the kissing loop interaction and to catalytic activity. Consequently, the fact that a lower estimate of 10% of WT molecules never sample the active H state but instead dwell in the M and L states (Figure 2-3) at least partially explains why a significant fraction of ribozymes remain inactive in ensemble cleavage assays (Figure 2-1B), a common observation for the VS ribozyme independent of fluorophore labeling (54, 75), as well as for ribozymes in general (11, 14).

2.4 Discussion

The structural dynamics of tertiary structured RNA and their links to biological function are still not well understood, despite recent advances in our appreciation of the breadth of non-coding RNA function (2-4). Here we have exploited single molecule FRET to dissect the global structural dynamics of the VS ribozyme. This complex non-coding RNA contains combinations of tertiary structure elements commonly found in large RNAs, such as three-way junctions and a kissing loop, and is endowed with catalytic activity that reports on the biologically relevant fold.

Our findings support the notion that complex RNAs often fold into their active states by traversing discrete structural intermediates, similar to recent observations on other complex ribozymes such as group I (8, 77) and group II introns (78). Each newly formed intermediate acquires additional tertiary interactions whose formation is contingent upon the proper folding of the previous state. In the case of the VS ribozyme, the correctly folded II-III-VI junction serves as the structural scaffold upon which the kissing loop interaction builds. In turn, the kissing loop serves as a scaffold to juxtapose the internal loops of helices I and VI for cleavage of the former. We also observe evidence

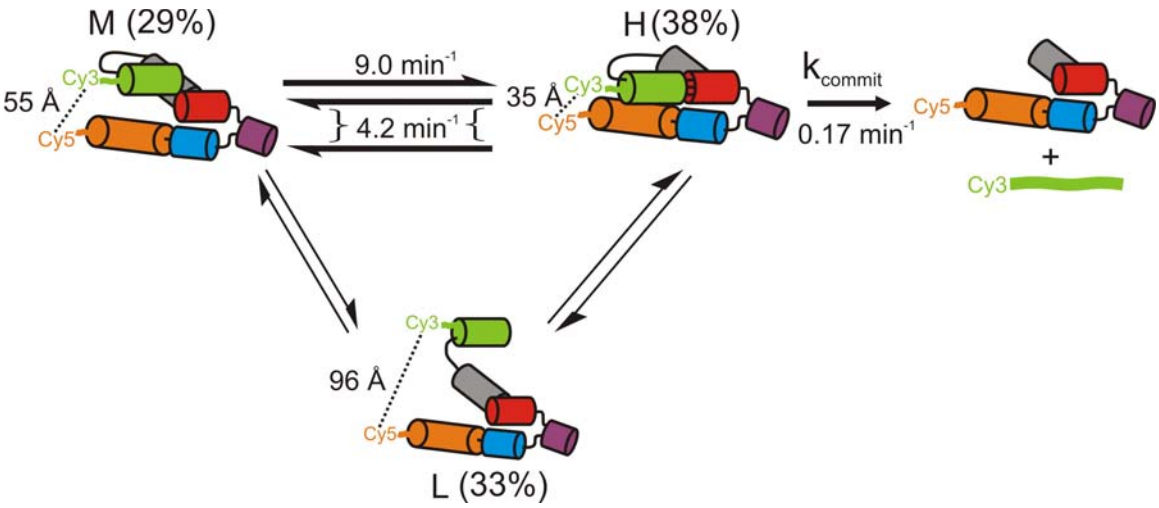


Figure 2-6. Structural and kinetic model of the reaction pathway of the WT VS ribozyme. The helix colors match those in Figure 2-1A. The reported H to M transition rate constant is the mean value obtained from the HMM analysis; the dual arrows represent heterogeneity (either static, i.e., never-changing or possibly dynamic, i.e., time-dependent in nature) in this transition as illustrated in Figure 2-2E. The fluorophore distances and relative abundances of the H, M and L states were obtained from single molecule FRET, while k_{commit} was modeled. L states, while often long-lived, are only infrequently accessed by a sub-population of molecules (static heterogeneity, thinner arrows).

for static (never-changing) and possibly dynamic (time-dependent) heterogeneity between individual molecules of the VS ribozyme, suggestive of a deeply furrowed folding free energy landscape with independent pathways separated by barriers too high for crossing by thermal motion, consistent with observations for other ribozymes such as group I introns (8, 77, 79) and the hairpin ribozyme (11, 14, 15, 74, 80).

The minimal mechanistic model in Figure 2-6 summarizes the structural dynamics of the full-length G11 VS ribozyme as directly observed and quantified here. We observe for WT altogether three states of distinct FRET values and thus fluorophore distances that are likely to lie on the folding pathway towards catalysis: L, M, and H states, which are about equally populated in the time and molecule average, although transitions into the long-lived L state are relatively rare (thin arrows in Figure 2-6). Our data support a model wherein population of the H state, attainable from both the L and M states, is necessary but not sufficient for catalysis. This on-pathway H state requires a properly folded II-III-VI junction, whose disruption in ΔII-III-VI greatly reduces both the probability of kissing loop formation (compare Figure 2-2B and Figure 2-4D) and catalytic activity (Figure 2-1B). The correctly folded II-III-VI junction thus serves as a structural scaffold for the formation the kissing loop interaction which we observe as the H state.

Our mechanistic model allows us to further evaluate the catalytic relevance of the H state. All observed transition rate constants between the three FRET states are at least an order of magnitude faster than the cleavage rate constant of 0.065 min^{-1} under the same conditions; the M→H and mean H→M conformational transitions are 138- and 65-fold faster, respectively (Figure 2-6). This rate difference indicates that the downstream pathway towards phosphodiester cleavage and product release minimally perturbs the equilibrium distribution between the L, M, and H states. We therefore define k_{commit} as the rate constant that describes the probability of cleavage and product dissociation from the H state (5'-product binds only weakly and is thus assumed to be irreversibly released under our conditions (J. Olive and R.A.C, unpublished)).

Straightforward analytical modeling of the reaction scheme in Figure 2-6 predicts that the rate constant observed in our cleavage assays is simply the product of k_{commit} and the fraction of time the RNA spends in the H state as reported in our FRET histograms. This simplest possible model assumes that the 2'-O-methyl modification does not disrupt the ground state interaction formed by the natural 2'-OH. k_{commit} is then calculated to be $(0.065\ min^{-1})/(0.38) = 0.17\ min^{-1}$ for WT, $(0.002\ min^{-1})/(0.04) = 0.05\ min^{-1}$ for ΔII-III-VI, and $(0.05\ min^{-1})/(0.36) = 0.14\ min^{-1}$ for Rescue I-V. The reasonable agreement among the three k_{commit} values as calculated from our experimental results lends further support to the minimal reaction scheme in Figure 2-6. It also suggests that WT and mutants must traverse similar local energy barriers for catalysis to occur after formation of the catalytically competent H FRET state.

What is the nature of this barrier: conformation or chemical? In our G11 based ribozyme the substrate stem-loop I is connected via its 3' end to the 5' end of helix 2 by just 3 unpaired bases. This barrier is apparently alleviated when a >25 nucleotide linker is instead used. VS ribozyme constructs that have such linkers that connect the substrate instead to the 3' end of the ribozyme have observed cleavage rates that are several hundred fold faster (81) than that of G11. We speculate that this alternate connectivity affords the substrate additional conformational flexibility that bypasses the energetic barrier(s) responsible for our observed k_{commit} value. This additional conformational flexibility optimizes the catalytic core concomitantly with kissing loop formation, leading to rapid cleavage. Further experimental work that explores such an alternate connectivity is presented in chapter 3.

CHAPTER 3

AN ALTERNATE CONNECTIVITY OF THE VS RIBOZYME EXHIBITS SIGNIFICANT DIFFERENCES IN FOLDING DYNAMICS AND CATALYSIS[b]

3.1 Introduction

The minimal contiguous sequence of natural VS RNA that carries its self-cleavage activity consists of 154 nucleotides (50), commonly referred to as the VS ribozyme. This small catalytic RNA minimally contains a single nucleotide immediately upstream (5') followed by 153 nucleotides downstream (3') of the cleavage site. Many studies of the VS ribozyme have revolved around a similar construct, G11, which is extended to 9 nucleotides upstream (54). The single molecule characterization described in chapter 2 was performed on such a construct. Later experiments discovered, however, alternate constructs that gave several hundred-fold higher cleavage rates (81). The fundamental property of these new VS ribozymes is an alternate connectivity between the ribozyme and the substrate stem-loop with the cleavage site. The cleavage site is now located at the 3' end of the ribozyme, connected via a longer, for example, 23-nucleotide linker. One of these constructs, RS19ΔL, has an observed rate constant for producing cleaved RNA of 1.9 s^{-1}, several orders of magnitude greater than for G11 (0.002 s^{-1}) (54). This rate consists of the sum of the forward cleavage rate constant (0.21 s^{-1}) and the reverse ligation rate constant (1.7 s^{-1}), so RS19ΔL exists ~90% of the time in its ligated form (81).

[b] Alex Buck-Johnson was a major contributor to the results shown in figures 3-2, 3-3, and 3-4.

The question arises as to the biological relevance of these fast rates. In its natural context, VS RNA is transcribed from the VS plasmid inside the mitochondrion of *Neurospora* to yield a multimeric series of head-to-tail tandem repeats (48, 82). This gives the VS ribozyme motif access to the 3' distal downstream cleavage site that is hundreds of nucleotides away. This access is expected to result in the ribozyme motif preferentially cleaving this 3' distal cleavage site (k_{cleav} ~ 0.2 s^{-1}) (81) instead of its neighboring upstream site (k_{cleav} ~0.002 s^{-1}) (54), as evidenced by comparison of the cleavage rate constants of RS19ΔL (0.2 s^{-1}) and G11 (0.002 s^{-1}). The subsequent finding that *in vitro* assays of multimeric VS RNA show preferential cleavage of downstream substrate (83) further supports this hypothesis.

There are many questions pertaining to RS19ΔL function that remain unanswered. How does its alternate connectivity confer such enhanced catalysis? What role do its structural states and the dynamics between them play in its function? Single molecule FRET microscopy and ensemble activity assays of a variant of RS19ΔL, termed FR3 (Figure 3-1A) were performed to gain insight into these questions. We show here that FR3 does not exhibit significant structural dynamics. This observation stands in sharp contrast to the dynamic behavior of the WT construct described in chapter 2. The FRET ratio associated with the observed stable FR3 states increases with higher magnesium concentrations. A parallel enhancement of the cleavage rate is also observed as the magnesium concentration is increased. Validation of these results was achieved by probing for and comparing FR3's catalytic activity both in bulk solution and during single molecule FRET measurements. The cleavage/ligation equilibrium of FR3 directly observed on the microscope slide is consistent with standard pre-steady-state cleavage assays performed on FR3, as well as the data previously reported for the biologically relevant, related RS19ΔL construct.

3.2 Materials and Methods

Preparation of FR3. FR3 is derived from the RS19 family of VS ribozymes (81), in which stem-loop I, which contains the site of cleavage, is attached via a linker of arbitrary sequence to the 3' end of the ribozyme core (Figure 3-1A). The sequence of stem-loop I was altered to introduce a 5-amino-allyl uridine at a position predicted from molecular modeling (not shown) to be on the exterior face of helix Ia for Cy3 labeling. Synthesis of FR3 was accomplished by VS ribozyme-mediated ligation using a strategy similar to that used previously to incorporate a site-specific 4-thio-uridine nucleotide into stem-loop I (26). The ribozyme portion was obtained by in vitro transcription by T7 RNA polymerase of a linearized plasmid template, followed by self-cleavage and gel-purification of the upstream cleavage product which ends at G620 and contains a 2'3' cyclic phosphate terminus; the transcription mixture included 4 mM ApG dinucleotide in which the adenosine contained an amino group at the end of a six-carbon linker attached to the 5' phosphate (Dharmacon, Inc) to allow for subsequent labeling of the 5' end of the RNA with Cy5 mono-reactive dye (GE Healthcare). A second RNA beginning at position 621 (with a 5' hydroxyl) was chemically-synthesized (Dharmacon, Inc.) and contains a 5-amino-allyl-uridine for subsequent labeling with Cy3 mono-reactive dye (GE Healthcare) at the position indicated in Figure 3-1A and a 3' terminal biotin. Incubation of these two Cy-labeled RNAs in the presence of 200 mM Mg^{2+} results in ligation to form the full-length FR3 RNA, which was gel-purified and ethanol-precipitated.

FR3 Histograms of the Magnesium Titration: A solution of 3.3 µl 3 nM FR3 + 97 µl of 1X (20 mM HEPES pH 7.4 + 100 mM KCl) was heated in the dark for 2 min at 70^{o}C and cooled to room temperature. This 100 pM FR3 solution was flowed onto a clean quartz slide that was first coated with biotinylated BSA, then streptavidin as described (11, 31). After 2 min the microfluidic channel was flushed with scavenging 1X (1X + 1mM Trolox (65) + 2.5 mM protocatechuic acid + 25 nM protocatechuate-3,4-dioxygenase + appropriate concentration of Mg^{2+})

to efficiently consume the dissolved oxygen (84). The donor Cy3 was excited with a 532nm laser using total internal reflection microscopy as described at a frame rate of 2 frames per second. Around 50 frames prior to the end of each movie, a 640nm laser was used (except for 5mM Mg^{2+}) to probe for the presence of an active Cy5. This unambiguously determined whether a low FRET ratio is due to low FRET efficiency or the absence of an active Cy5 acceptor. Noise in the time traces was reduced using a previously reported (66) non-linear filter with the parameters m = 3 and p = 20 prior to the appropriate binning and Gaussian fitting using Origin 7.5. Sample preparation for Figure 3-1B was similar except the background buffer used was 40 mM Tris pH 8.0 + 50 mM KCl and the glucose oxidase/catalase oxygen scavenging solution (31) was employed.

Single Molecule Probing of FR3 catalysis: A solution of 1.65 µl 3 nM FR3 + 97 µl of 1X (20 mM HEPES pH 7.4 + 100 mM KCl + 10 mM Mg^{2+}) was heated in the dark for 2 min at 70^{o}C and cooled to room temperature. This 50 pM FR3 solution was flowed onto a clean quartz slide that was first coated with biotinylated BSA, then streptavidin as described (11, 31). After 2 min the microfluidic channel was flushed with scavenging 1X. The donor Cy3 was excited with a 532nm laser using total internal reflection microscopy as described (9) at 2 frames per second. Approximately 30 frames into the trace, the 640nm red laser was also activated to probe for the presence of photoactive Cy5. The shutter was then closed, 6M urea was flowed into the microfluidic channel at 1 ml/min for 20 seconds using the Harvard Apparatus Pump-33 dual syringe pump, followed by the 1 ml/min flow of scavenging 1X and subsequent incubation for 5 min. The shutter was opened with both lasers still on to report on the presence (no dissociation) or absence (dissociation) of the already confirmed photoactive Cy5.

Bulk Solution Activity Assay of FR3: A 150 µl solution containing 45 picomoles of FR3 in 20 mM HEPES pH 7.4 + 100 mM KCl was heated to 70^{o}C for 2 min, then cooled to room temperature for 5 min. Then removed 25 µl of it

into 25 µl stop mix (formamide + 200 mM EDTA), added 25 µl Mg^{2+} solution to achieve either a 10mM or 200mM final Mg^{2+} concentration. At time points of 15 sec, 30 sec, 1 min, 2 min, 5 min, 10 min removed 25 µl into 25 µl of stop mix. The cleaved and uncleaved material were separated on a 10% denaturing PAGE gel and the intensity of Cy5 fluorescence on the gel was quantified for each band and analyzed using the Amersham Biosciences Typhoon 9410 Variable Mode Imager instrument and Image Quant 5.2 analysis software.

3.3 Results

The FR3 RNA (Figure 3-1A) is a single-stranded, catalytically active molecule that has the substrate portion linked to the downstream 3' end of the ribozyme motif by a 26-nucleotide linker (highlighted in bold in Materials and Methods). This connectivity is markedly different from that of the WT described in chapter 2 (Figure 2-1A), which instead has only a 3-nucleotide linker attaching the 5' upstream end of the ribozyme to the substrate stem-loop. The dynamic conformational behavior of FR3 differs greatly from that of WT described in chapter 2. Addition of 35 mM Mg^{2+} to immobilized FR3 molecules results in a non-dynamic, long-lived high FRET state (Figure 3-1B). The underlying structure is hypothesized to entail the catalytically essential kissing interaction since current models (25, 28) of the global structure predict that docking of the cleavage site into the active site brings the Cy3 and Cy5 fluorophores into close proximity. Histogram analysis of the trajectories shows that the FR3 molecules spend 94% of their time in the high FRET state. The globally static behavior in the example trace (Figure 3-1B) is representative of the molecules analyzed. These data reveal that the nature of the connectivity between the ribozyme and substrate has a profound effect on the presence/absence of global structural dynamics on the second timescale.

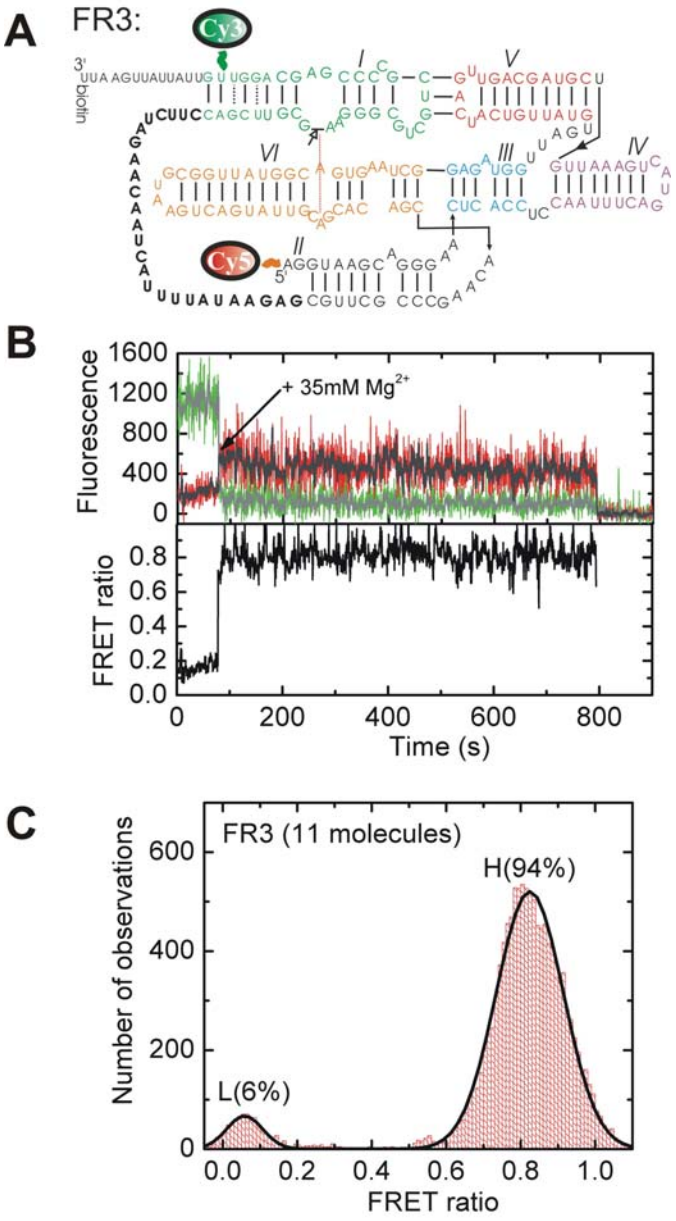


Figure 3-1. Connectivity of FR3 and its behavior upon addition of magnesium. A) The primary sequence and secondary structure of FR3. The substrate (green) is linked to the 3'end of the ribozyme by the nucleotides shown in bold. A cleavage (arrow) and denaturant induced dissociation event leaves behind the 3' biotin and Cy3 containing portion of the RNA. B) Addition of 35 mM Mg^{2+} induces a conformational state from a low to a high FRET state that is very stable and non-dynamic. C) Histogram analysis of multiple molecules shows a strong tendency of FR3 in 35 mM Mg^{2+} to populate a high FRET state.

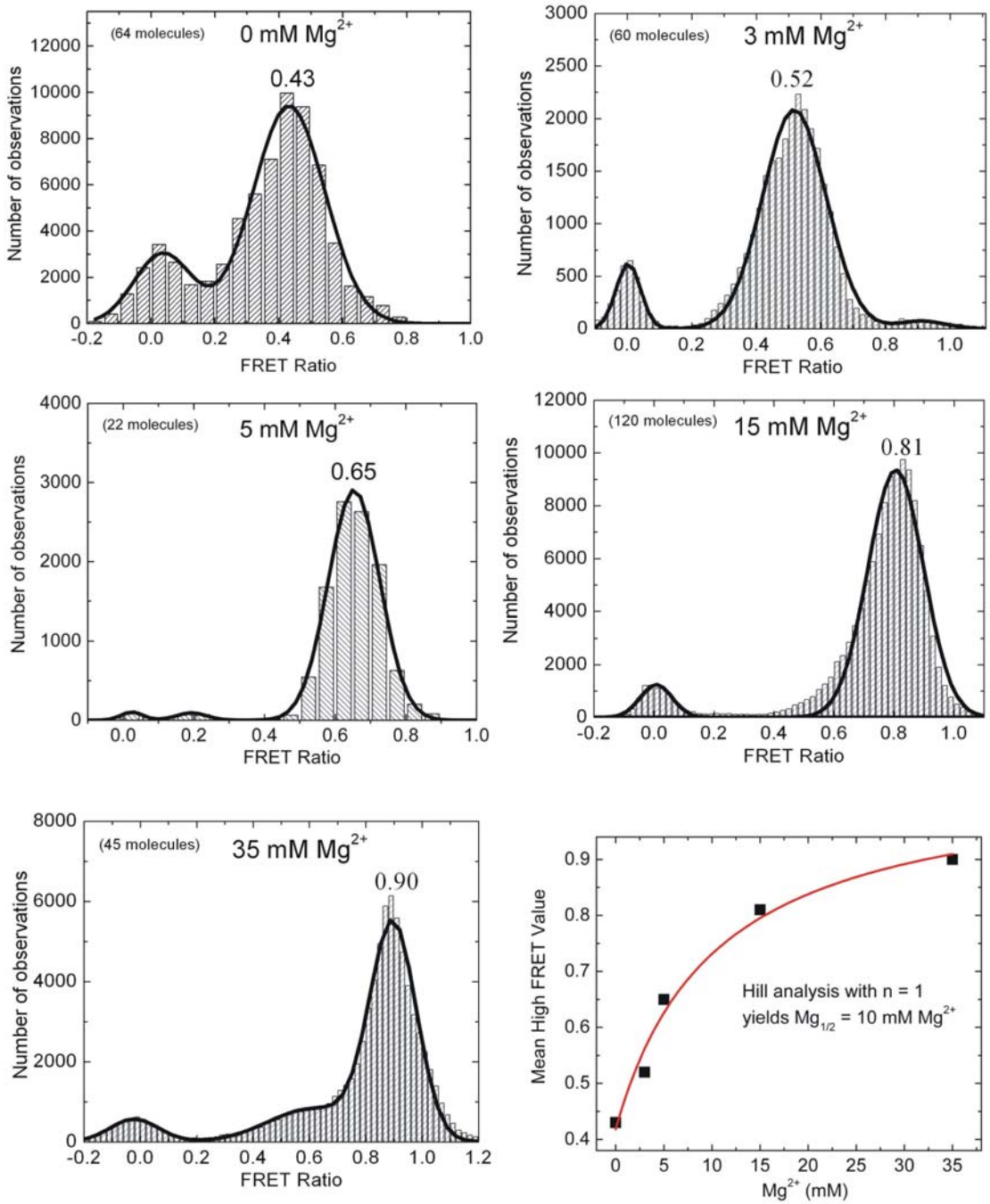


Figure 3-2. The effect of magnesium on the observed mean high FRET value. The predominant FRET ratio of FR3 molecules shifts upwards as the concentration of magnesium increases. A Hill analysis of the mean high FRET value versus [Mg^{2+}] yields an apparent magnesium half-titration point of $Mg_{1/2}$ = 10 mM.

The concentration of magnesium ions is known to play an important role in the dynamics or absence thereof in RNA (85). The effect of Mg^{2+} on the behavior of single FR3 molecules was investigated by conducting similar studies over a range of smaller magnesium concentrations. The background buffer consisted of 20 mM HEPES-KOH, pH 7.4, 100 mM KCl to more closely represent the pH and ionic strength conditions present *in vivo*. Figure 3-2 shows a histogram analysis of the resulting FR3 FRET ratios. These data reveal a gradual shift in predominant FRET ratio towards higher values as more Mg^{2+} is added. Visual inspection of these traces shows a consistent profile of single molecule trajectories with predominantly a single state. Fitting the data to a Hill equation with n=1 yields an approximate half maximal saturation point of ~10 mM Mg^{2+} (Figure 3-2). The quality of this fit with n = 1 argues against the existence of cooperative behavior with respect to the association of Mg^{2+} with FR3.

There is no chemical modification in FR3 that prevents cleavage and re-ligation of the scissile phosphate bond. The same construct used in the FRET histogram analysis can therefore be tested for its catalytic properties as well. The expanded length of the single molecule time trajectories is achieved in part because product dissociation, which would lead to disappearance of the donor fluorophore, is expected to only occasionally occur. There are 11 base pairs that connect the 5' product with the ribozyme even after bond breakage (Figure 3-1A), and dissociation of these base pairs is necessary for a characterization of FR3 catalysis.

Bulk solution activity assays were performed at the half-maximal saturation point of 10 mM Mg^{2+} as well as at 200 mM Mg^{2+} (Figure 3-3). This latter high concentration was previously used in the characterization of RS19ΔL, the VS construct from which FR3 was derived. The accumulation of cleaved FR3 after addition of magnesium to a final concentration of 200 mM Mg^{2+} is inadequately described by a single-exponential fit due to the biphasic nature of the data. Fits with a double-exponential, however, failed to converge because there are only two data points to describe the fast phase. A single-exponential fit at the lower 10 mM Mg^{2+} concentration does reasonably describe the data with a

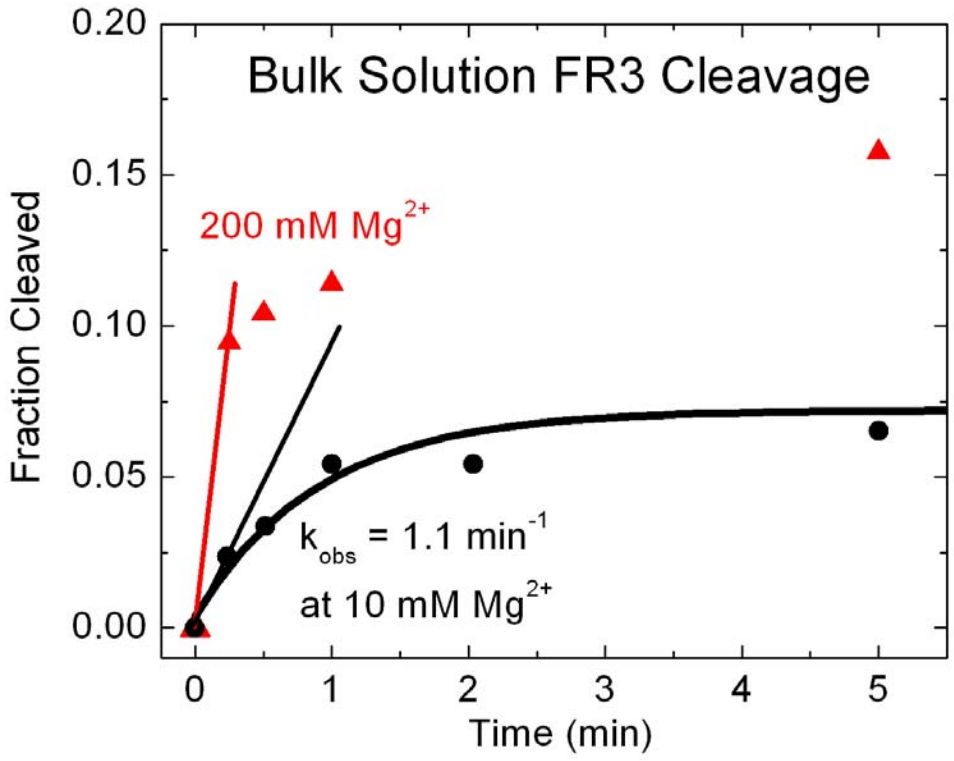


Figure 3-3. Accumulation of cleaved FR3. Ensemble activity assays done at 10 mM and 200 mM Mg^{2+} show differential rates of accumulation of cleaved FR3 over time. The slope between the 0 and 15 seconds time points for both assays are compared and a single-exponential fit to the 10 mM Mg^{2+} data shown.

k_{obs} = 1.1 min^{-1} (Figure 3-3). The total extent of cleaved material reached 8%. The slope of the burst phase for 200 mM Mg^{2+} is 7-fold greater than that at 10 mM Mg^{2+}. This gives an extreme lower estimate of k_{obs} for the burst phase at 200 mM Mg^{2+} of (1.1 min^{-1}) $\times$ (7) = 8 min^{-1}. Concomitantly with the increase in $k_{obs,}$ the total extent of observed cleavage reached is 16% at 200 mM Mg^{2+} compared to 8% at 10 mM Mg^{2+}. The observed 8% extent of cleavage for FR3 in 10 mM Mg^{2+} implies that k_{lig} is about 10-fold faster than k_{cleav} in the bulk solution experiment (assuming that most molecules are active), consistent with earlier observations (81). To validate our single molecule FRET approach the next questions we sought to address was: Are the same cleavage kinetics and cleavage extent observed for FR3 molecules that are immobilized on a streptavidin coated quartz slide?

Experimental conditions were optimized to allow for the direct determination of whether a single immobilized FR3 molecule existed in the cleaved or ligated format at a given time point. The magnesium half-titration point $Mg_{1/2}$ of 10 mM was chosen along with the same background buffer conditions (20 mM HEPES-KOH, pH 7.4, 100 mM KCl) used for the bulk solution activity assays (Figure 3-3) and the single molecule FRET histograms (Figure 3-2). Excitation of both Cy3 and Cy5 was used to confirm their presence, followed by closing of the excitation shutter and flowing of 6 M urea onto the slide. This secondary structure denaturant is expected to dissociate and flush away the non-biotin containing Cy5 portion of FR3 if that molecule exists in its cleaved form. In the ligated form the FR3 RNA will still denature, but the Cy5 remains covalently attached. As the 6 M urea is washed away, the original buffer with oxygen scavenger system is reintroduced and the presence or absence of Cy3 and Cy5 is again determined by direct excitation. Figure 3-4 shows the fingerprints of single FR3 molecules that upon injection of 6 M urea were either in the cleaved or ligated form. This process can be repeatedly applied to the same molecule, as is illustrated in Figure 3-4C and 3-4D.

Out of a total of 72 single molecule observations, 7 cleaved molecules were observed. This value of ~10% cleaved molecules is consistent with the

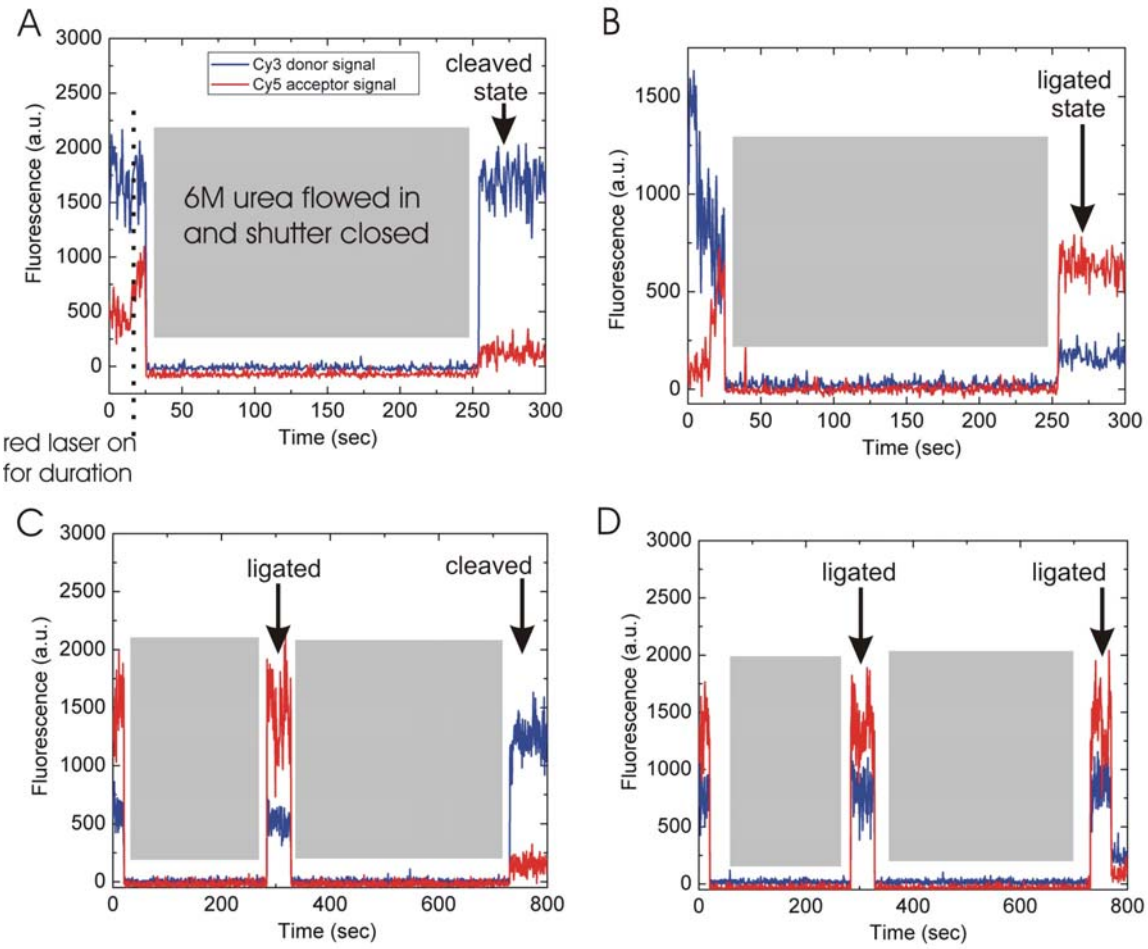


Figure 3-4. Direct single molecule detection of FR3 catalysis. Cy3 emission is shown in blue, Cy5 emission in red. Cy3 is initially excited, then both are excited when a red laser directly excites the Cy5. The shutter is closed and 6 M urea is flowed in, then removed by the original oxygen-free buffer and the single FR3 molecule is once again exposed to both lasers. This process was repeated. A) An FR3 molecule in its cleaved form loses its Cy5 as the 6 M urea dissociates it away. B) Another FR3 molecule is instead in its ligated form, not allowing the Cy5 to leave. C) The first 6 M urea flow reveals the ligated form, the second time the cleaved form. D) Both times reveal the ligated form.

amplitude of 8% obtained for the FR3 bulk solution activity assay at 10 mM Mg^{2+} shown in Figure 3-3. This consistency indicates that the underlying behavior of FR3 is both independent of and accurately probed by these different experimental methodologies. It also validates the relevance of the histogram results from Figure 3-2.

3.4 Discussion

We here show that changing the connectivity between the VS substrate stem-loop and the ribozyme motif has a profound effect on large-scale distance changes between the cleavage and active sites. Chapter 2 described large-scale and frequent structural changes between those two sites when the substrate is connected to the 5 'end of the ribozyme through a 3-nucleotide linker. Here we find that when the substrate is instead connected to the 3' end by a 26-nucleotide linker the VS ribozyme predominates in a stable global fold characterized by a FRET ratio that gradually shifts from 0.52 to 0.90 (Figure 3-2) upon increasing the magnesium concentration.

This upward shifting of the high FRET peak with increasing Mg^{2+} concentration is consistent with a gradual structural compaction of a stable fold of FR3. An alternate interpretation is that there are two rapidly inter-converting structures with constant fluorophore distance that are shifted towards the shorter distance species upon magnesium titration. The timescale of these dynamics would need to be faster than 0.5 s^{-1} so that our 2 frames per second data collection frequency would report only on the equilibrium distribution of these two structures. This possibility would require either the rapid inter-conversion of two globally disparate structures that both contain the stem-loop IV kissing interaction, or the breaking/reformation of this well-characterized triple base-pair tertiary contact on a time scale of more than twice a second. For the well characterized WT construct from chapter 2 it took on average 14 seconds for this tertiary contact to break. The longer FR3 connectivity would then somehow lead

to a destabilization of the kissing-loop interaction that would then confer much faster cleavage rates. While both scenarios cannot be ruled out, a magnesium induced overall structural compaction is more plausible. The compaction of RNA structure through stabilization by diffuse ions has been proposed for other large RNAs such as the *Tetrahymena* ribozyme (86)

How catalytically competent are these globally static folds of FR3? They are expected to have much faster observed cleavage rate constants than the reported 0.065 min^{-1} values for WT described in chapter 2, because fast rates for variants of FR3 have been reported previously (81). The high FRET values observed in Figure 3-2 are consistent with the idea that the essential loop I-V kissing interaction is maintained throughout the time traces. The FR3 construct that was used to generate the data in Figure 3-2 was also characterized kinetically to corroborate the relevance of the histogram data. The rate of reaching the ligation/cleavage equilibrium and the equilibrium position were experimentally measured using both bulk solution and single molecule techniques. The 8% amplitude of the 10 mM Mg^{2+} bulk solution cleavage assay in Figure 3-3 is in good agreement with the 10% of single FR3 molecules that were observed to exist in the cleaved form. This implies that the distinct conditions of FR3 molecules during single molecule versus bulk solution experiments do not impact their function. The histogram results can thus be confidently interpreted as data collected on single FR3 molecules whose catalytic ability is unperturbed relative to those in bulk solution. As has been described above, the bulk solution FR3 molecules that exhibit proficient catalytic activity also spend a large majority of their time in the high FRET state. This proficient catalytic activity also requires the formation of the loop I-V kissing interaction, which implies that our observed high FRET state structure contain the kissing interaction.

The FR3 bulk solution results are also consistent with the previously characterized RS19ΔL construct which FR3 was designed to model. RS19ΔL yielded a k_{obs} = 120 min^{-1} and a total fraction cleaved of 11% when measured using rapid quenched-flow instrumentation at 200 mM Mg^{2+} (81). The manual

mixing data (due to limited sample amount) shown in Figure 3-3 for FR3 at 200 mM Mg^{2+} are consistent with the published RS19ΔL data. In particular, the extent of cleavage is in good agreement with 11% for RS19ΔL after 5 s (no data were collected over a longer time frame) versus 9.5% for FR3 after 15 s. Visual inspection of the longer FR3 time points at 200 mM Mg^{2+} suggests biphasic behavior: 9.5% of the molecules had already reached their ligation/cleavage equilibrium within 15 seconds, as expected from the reported RS19ΔL rate constant of 120 min^{-1}. There is also a smaller subpopulation that shows much slower kinetics.

The data reported here for FR3 lead to further insights into how VS RNA functions in its biological context. The discovery of VS variants with greatly enhanced cleavage rates led to the hypothesis that the downstream substrate is preferentially cleaved, which was confirmed when multimeric VS RNA was shown to preferentially cleave downstream (83). These findings, however, shed little light on how the 881-nucleotide VS RNA selects against the upstream site. A comparison of the single molecule characterization of WT in chapter 2 with the results from the RS19ΔL-related FR3 construct described here offers new insight into this question. The readily reversible formation of the I-V kissing-loop interaction with high FRET in WT stands in sharp contrast to the much more stably formed high FRET conformer observed for FR3. This difference in docking stability is expected to partially contribute to the difference in cleavage rate constants between the WT (k_{obs} = 0.04 min^{-1} at 10 mM Mg^{2+}) and FR3 (k_{obs} = 1.1 min^{-1} at 10 mM Mg^{2+}), where now the essential kissing-loop interaction is formed a vast majority of the time, thus enhancing the rate at which downstream steps can occur that lead to catalysis.

Is there an evolutionary rationale for the lack of observed dynamics in our biologically relevant FR3 construct? Inside the mitochondria of the fungus *Neurospora*, one expects selective pressure on the replication cycle of VS RNA to minimize the population of vulnerable linear monomer. By contrast, the circular monomeric VS RNA is exonuclease resistant while still serving as a suitable substrate for the accompanying reverse transcriptase (49). Multimeric VS RNA

may have adapted to such selective pressures by fine-tuning its ribozyme motif to minimize the lifetime of the exonuclease vulnerable 2',3'-cyclic phosphate and 5'-OH termini required for the ligation into the circular VS RNA. Such self-protection of the RNA first requires slow cleavage of the proximal stem-loop I substrate upstream of the ribozyme motif to allow for transcription of the distal downstream substrate. What follows is dynamic docking and undocking of the proximal (upstream) stem-loop substrate with the ribozyme motif and a distal (downstream) substrate that now not only has access but also high affinity to that same ribozyme motif. The formation of this globally stable structure now leads to eventual dissociation of the cleaved distal (downstream) stem-loop I product. This event leaves a 5'-OH in the immediate vicinity of the corresponding proximal (upstream 2',3'-cyclic phosphate) product that was generated by its distal VS ribozyme partner, thus allowing for ligation of the VS RNA with minimal exposure of both ends to the exonucleases in the mitochondria of *Neurospora*. The FR3 data presented are consistent with this model of how the VS ribozyme may function in nature.

CHAPTER 4

A NOVEL APPROACH FOR THE NATIVE PURIFICATION OF RNA[c]

4.1 Introduction

Over the past few years a broad consensus has emerged that the typical cell contains a multitude of diverse and essential non-coding RNA's. These RNA molecules play a central role in the chemistry of life and as a consequence a large body of research is aimed at gaining a fundamental understanding of how they function. This research involves the design of *in vitro* experiments whose results are then argued to be biologically relevant. The purification of the RNA being studied for these experiments, however, has a variety of shortcomings (87). The predominant method of generating RNA molecules that are too long to be made synthetically utilizes T7 RNA polymerase and a DNA template strand (42, 88). This generates shorter aborted transcripts, transcripts to which a few extra nucleotides have been added by the polymerase (42), and a plethora of salts and unused nucleotide triphosphates alongside the RNA of interest. The predominant purification procedure for the desired RNA involves a desalting step prior to the separation by size using denaturing polyacrylamide gel electrophoresis (PAGE) and elution from the gel. This requires a subsequent step aimed at homogeneously folding the RNA into its functional structure.

[c] The final portion of this work was done with significant assistance from Vivek Behera, an undergraduate student.

The drawbacks of this methodology are well known (87). Separation by size on a denaturing PAGE gel is unreliable at removing the transcripts that contain few additional bases on the 3' end when the overall transcript is quite long. The refolding protocol can result in the generation of non-trivial amounts of misfolded species. This diversity in populated structures can be difficult to detect yet can give adverse results if structure probing is pursued, for example, by NMR, or can lead to the inability to grow crystals that are amenable to high-resolution X-ray diffraction and subsequent structure determination (88).

Several different methods and protocols have been previously developed to overcome these challenges (44-47). They aim to purify the desired RNA as folded during the transcription reaction, while avoiding denaturation and/or sequence heterogeneity in the final sample. This motivation comes from the fact that RNA sequences in nature have evolved to fold into their functional structure in the context of the transcription process (89). Size-exclusion chromatographic RNA purification has shown some promise, however, this technique does not yield RNA's with homogeneous 3' ends (45). The use of an immobilized DNA oligonucleotide and a DNAzyme has been shown to generate homogeneous 3' ends and bypass the need for denaturing gel electrophoresis. This protocol does, however, involve heating the RNA to 95^{o}C for 3 min (44) which re-introduces the problem of denaturing the desired RNA structure folded during transcription. The use of an MS2 coat protein to achieve immobilization combined with the *glmS* catalytic riboswitch as a means of removing the target RNA off a commercially available nickel-affinity resin has been shown to achieve native RNA purification with homogeneous 3' ends. Implementation does, however, require the user to purify large amounts of the HMM (His-tagged MBP-MS2 coat fusion) protein. This involves IPTG-induced expression, purification on a Nickel affinity column, and the use of Fast Protein Liquid Chromatography (FPLC) instrumentation to apply a gradient of 0-1 M NaCl to a Hi-Prep 16/10 SP-Sepharose column binding the protein (46). This can limit its applicability if the laboratory has limited expertise in protein purification and/or proper equipment available.

We present here a novel protocol that does not require protein purification and is expected to extend the ability of researchers to carry out native RNA purification with minimal difficulty. The purchase of commercially available streptavidin-coated magnetic beads, a 30-mer biotinylated DNA oligo, and the cloning of the user-specific double-stranded DNA template into our available plasmid is all that is required to implement the straightforward procedure described here and attain homogeneous, co-transcriptionally folded RNA .

4.2 Materials and Methods

Cloning of pVS1, pHDV1, pMCGL. The topWT plasmid (see materials and methods in chapter 2) was as the parental plasmid that lead to the eventual generation of pVS1. This involved minimal sequence modification to the 5' end of G620 to include a HindIII site and the addition of 250 bases 3' of C740 to introduce sequence that would encode the *glmS* ribozyme (90), followed by the helix motif that terminated with a EcoRI site. This was accomplished through multiple rounds of PCR whereby dangling 5' ends of the primers (5' – 3': CGCGCGCGCG CCCCTTTTTTAAAAAAAAAACGATCGATCGATCGATCGATA AAAAACTTG; CCTTACGGCTGTGATCTGCACACTCAGCCGGGAG GCATCC GCCGAAAAT TCGGTAACCTC; CAAGGAGTCACCCCCTTGGTTTGAAGAAA TCCTTACGGC TGTGATCTGCACACTCAGCCG; GCGTTGAATTCAGATCA TGTGATTTCTCT TTGTTCAAGGAGTCACCCCCTTGGTTTGAAGAAAT) served to add in the desired sequence. In each round the completed PCR reaction was separated by non-denaturing 10% (w/v) PAGE and the larger desired product was cut out and eluted into 1 ml of H_2O overnight at 4^oC. This eluant was used to seed the next round of PCR that added on more of the desired sequence. This was repeated until a PCR product was generated with HindIII and EcoRI sites on each end. Both sites were simultaneously digested with HindIII and EcoRI. The commercially available pGEM4Z plasmid was simultaneously digested and dephosphorylated with HindIII, EcoRI, and alkaline phosphatase. Both reactions

were passed through two microcon-EZ (Millipore) columns to remove the enzymes. The insert was then ligated into the plasmid using the standard conditions as recommended by the manufacturer of the T4 DNA Ligase (Invitrogen # 15224-017).

The pVS1 plasmid served as the parental plasmid to make pHDV1. The pVS1 plasmid was simultaneously digested with HindIII and AflII, heat inactivated at 70^{o}C for 20 min, phenol chloroform extracted, and run on a 1% (w/v) agarose gel. The doubly digested plasmid band was cut out and extracted using a commercial extraction kit (Qiagen #28706). The insert containing the HDV ribozyme sequence flanked by HindIII and AflII restriction sites was generated by primer extension PCR from two partially overlapping DNA oligonucleotides (5'-3': GCGAGCAAGCTTCCACTCCTTTTCGGGTCGGCATGGCTCTCCACCTCCTCG CGGTCCGACCTGGGCACCATT; CGCCTCTCTTAAGTTCGGCTCTCTGGCTC TCCCTTAGCCACGGGAGTGCAATGGTGCCCAGGTCGGACCGCGA). This PCR product was simultaneously digested by both enzymes and as with pVS1, heat inactivated at 70^{o}C for 20 min and phenol chloroform extracted to unambiguously remove all protein prior to ligation. Standard T4 DNA ligase reaction conditions (Invitrogen) were used to generate the pHDV1 plasmid.

The same doubly digested pVS1 plasmid used to make pHDV1 was also used to make pMCGL. The insert consisted of two synthetic DNA oligonucleotides with 5' phosphates whose ends when annealed together mimic the product of the HindIII and AflII digestions. A 20 µl reaction containing 33 fmols of this insert, 11 fmols of doubly digested pVS1, and 4 U of T4 DNA ligase (Invitrogen # 15224-017) was incubated at 16^{o}C overnight. An aliquot was then transformed into XL10-Gold Ultracompetent cells as suggested by the manufacturer (Stratagene # 200315) to yield ~50 colonies on a single LB-amp plate. DNA sequencing confirmed all plasmid sequences described above. All synthetic DNA oligonucleotides used in this study were purchased from Invitrogen.

Capture1 binding assay. Ten transcription reactions were run, each containing 4 mM NTP, 40 mM HEPES-KOH ph 8.0, 5 mM DTT, 25 mM $MgCl_2$, 200 mM NaCl , 50 nM EcoRI-linearized pHDV1 plasmid, 0.008U/µl inorganic pyrophosphatase, 0.1 mg/ml T7 RNA Polymerase, and variable (0, 25, 50, 75, 100, 125, 150, 175, 200, 250) pmols of capture1 in a 50 µl reaction volume for 2 hrs at 37^{o}C. Control reactions (shown on each end in Figure 4-2A) consisted of transcriptions with 100 or 250 pmols capture1 prepared identically as above except for the absence of the linearized pHDV1 transcription template. An equal volume of non-denaturing loading buffer was added and 5 µl samples were run on 6% ND-PAGE at 5W for 4 hrs. The gel was stained using a 1:100,000 dilution of Cybr-Gold (Invitrogen # S11494) in 1X TBE for 5 min. Fluorescence intensity of the gel was measured using a fluorimager (Amersham Biosciences Typhoon Variable Mode Imager 9410) with an excitation of 488 nm, an emission bandpass of 500 – 540 nm, and a PMT of 500V. The results were quantified using ImageQuant software (Molecular Dynamics).

The fluorescence intensities indicated in Figure 4-2A that arise from free capture1 were converted into pmol amounts using the control standards at both ends of the gel. These calculated free picomoles of capture1 were subtracted from the known amounts included at the beginning of the transcription to give the bound pmols of capture1 shown in Figure 4-2B. The numbers reported in 4-2 were scaled by a factor of 2 to correctly report the results for total transcription volumes of 100 µl.

Activity Assay of the glmS Catalytic Riboswitch. This assay began with a 150 µl transcription containing using the same standard conditions as for the capture1 binding assay: 4 mM NTP, 40 mM HEPES-KOH, pH 8.0, 5 mM DTT, 25 mM $MgCl_2$, 200 mM NaCl, 50 nM EcoRI-linearized pHDV1 plasmid, 0.008 U/µl inorganic pyrophosphatase, 0.1 mg/ml T7 RNA Polymerase, 350 pmol capture1, and a trace amount of ^{32}Pα-GTP. An incubation at 37^{o}C for 2 h followed by a phenol/chloroform extraction gave a final volume of 100 µl. Then 25 µl of 5 M NaCl was added to bring the final Na^{+} concentration in the transcription mixture

to ~1 M. The transcription reaction was then combined with 100 µl (1 mg) of Streptavidin C1 Dynabeads (Invitrogen # 650.01) that had been equilibrated in 40 mM HEPES-KOH pH 7.4, 1 M NaCl. The streptavidin-coated magnetic beads form a complex with the biotinylated-capture1•RNA transcript hybrid as the sample is rotated at room temperature for 30 min. First the supernatant was removed using a magnetic particle concentrator (Invitrogen # 120.20D), then the beads were washed 3 times with 100 µl of WB (wash buffer: 40 mM HEPES-KOH, pH 7.4, 1 M NaCl). A wash step entails the resuspension of concentrated beads with a buffer, followed by removal of buffer through the use of the magnetic particle concentrator. The buffer in this removal stage is referred to as the supernatant. After the third wash the beads were once more resuspended in 100 µl of WB. A 10 µl aliquot is collected to represent the overnight time point and added to 30 µl of SM (stop mix: 90% formamide in 1X TBE). The remaining supernatant was removed and 100 µl of WB supplemented with 10 mM $MgCl_2$ and 5, 50, 100, 200 or 500 µM glucosamine-6-phosphate (GlcN6P) was added to initiate the *glmS* ribozyme reaction. Progression of the reaction was monitored by taking out 10 µl aliquots into 30 µl of SM after 0.25, 0.5, 0.75, 1, 2, 5, and 10 min. The different populations present in the time points were separated on a 10% D-PAGE gel run for 2 h at 20 W. This gel was exposed to a phosphorimager screen overnight and quantified using ImageQuant software. The intensities of the separate bands were corrected for the number of guanines in each sequence. This directly correlates the intensity values to the relative counts of RNA molecules present in each band in the gel. The fraction cleaved was then calculated as $H_{intensity}/(H_{intensity} + HG_{intensity})$ using the notation in Figure 4-3A. The 200 µM GlcN6P data in Figure 4-3A were fit with the single-exponential first order rate equation $y(t) = y_0 + A\left(1 - e^{(-k_{obs}t)}\right)$.

Purification of Native HDV and VS Ribozymes. The same standard transcription reaction conditions used in the capture1 binding and *glmS* activity assays were used for the purification of the HDV (Figure 4-4) and VS (Figure 4-6B) ribozymes. A 150 µl transcription reaction with 4 mM NTP, 40 mM HEPES-

KOH, pH 8.0, 5 mM DTT, 25 mM $MgCl_2$, 200 mM NaCl, 50 nM EcoRI-linearized pHDV1 or pVS1 plasmid, 0.008 U/µl inorganic pyrophosphatase, 0.1 mg/ml T7 RNA Polymerase, 350 pmol capture1, and trace amount of ^{32}Pα-GTP was run at 37^{o}C for 2 h. This was followed by a phenol/chloroform extraction that yielded a 100 µl final volume, and then addition of 25 µl of 5 M NaCl to bring the final Na^+ conentration to ~1 M. This transcribed material was then added to 100 µl (1 mg) of Streptavidin C1 Dynabeads (Invitrogen # 650.01) that was equilibrated in WB (wash buffer: 40 mM HEPES-KOH, pH 7.4, 1 M NaCl). This mixture was rotated at room temperature for 30 min to form the complex between the biotinylated-capture1•RNA transcript hybrid and the streptavidin-coated magnetic beads. Using the magnetic particle concentrator removed the supernatant (transcription supernatant lane in Figure 4-4), washed the beads 4 times with 100 µl of WB while collected the supernatant each time (wash lanes 1-4). Added 100 µl of WB supplemented with 10 mM $MgCl_2$ and 200 µM GlcN6P to concentrated beads to initiate elution of the HDV or VS RNA. The sample was incubated for one min, then the supernatant was collected (H elution in Figure 4-4; V elution in Figure 4-6B). Then 100 µl of WB supplemented with 10 mM $MgCl_2$ was added and the supernatant collected (H elution lane 7 in Figure 4-4 or V elution lane 7 in Figure 4-6B). The beads were washed 4 times with 100 µl of EB (elution buffer: 10 mM Tris, pH 8.0). During the fourth time the beads were added to a 70^{o}C water bath for 5 min, then incubated at RT for 5 min, after which the supernatant was collected (HG, G elution lane 8 in Figure 4-4; VG, G elution lane 8 in Figure 4-6B). 100 µl of EB was added and the supernatant collected (HG, G elution lane 9 in Figure 4-4; VG, G elution lane 9 in Figure 4-6B). another 100 µl EB were added, the mixture placed back in the 70^{o}C water bath for 5 min, then incubated at RT for 5 min, and the supernatant again collected (repetition of elution lane 10 in Figure 4-4). 100 µl of EB was added and the supernatant collected (repetition of elution lane 11 in Figure 4-4). All of the collected supernatants were placed on the magnetic particle concentrator for a second time to remove any residual magnetic beads, each supernatant was removed and added to an equal volume of denaturing loading buffer (composition). 40 µl of each sample was loaded onto

a denaturing, 7 M urea, 10% (w/v) PAGE gel that was run at 20 W for 2 h. The gel was exposed to a phosphorimager screen overnight and quantified using ImageQuant software. The band intensities were corrected for the number of guanines in each sequence so that the intensity values report on the relative counts of RNA molecules that make up each gel. Analysis of these counts yielded the calculated >99% purity of desired RNA in the elution supernatants. The picomoles of post cleavage HDV or VS ribozyme present in the elution supernatants was derived from the A_{260} values measured on a Beckman Du 640B UV/Vis spectrophotometer.

Regeneration of the beads. A 700 µl transcription under the standard conditions of 4 mM NTP, 40 mM HEPES-KOH ph 8.0, 5 mM DTT, 25 mM $MgCl_2$, 200 mM NaCl, 50 nM EcoRI-linearized pHDV1, 0.008 U/µl inorganic pyrophosphatase, 0.1 mg/ml T7 RNA Polymerase, 2,450 pmol capture1, and trace amount of ^{32}Pα-GTP was run at 37^{o}C for 2 h. This was followed by a phenol/chloroform extraction to a final volume of 600 µl, to which 150 µl of 5 M NaCl was added to bring the final [Na^{+}] to ~1 M. A 100-µl reaction aliquot was removed and subjected to the same protocol as described above for Figures 4-4 and 4-6B. The supernatants containing wash 4, the first elution of the desired RNA, and the first elution of 70^{o}C treated material were kept. Then another 100 µl of the transcription reaction was added to the same beads that had already undergone one heating/cooling round, together with an additional 5 µl (0.05 mg) of fresh Streptavidin C1 Dynabeads equilibrated in WB buffer. The same binding and elution protocol was again followed (cycle 2), and this process was repeated for a total of six cycles. The collected supernatants were run on a denaturing, 7 M urea, 10% (w/v) PAGE gel, exposed overnight to a phosphorimager screen, and the intensities of the H band were quantified using ImageQuant software and corrected for the number of guanines as described above. An analysis of these intensities is displayed in Figure 4-5B.

4.3 Results

General Strategy of Our Approach. The general methodology of our native RNA purification is shown in Figure 4-1A. The transcription occurs with a biotinylated DNA oligonucleotide, referred to as capture1, that hybridizes to the 3' end of the transcript. This complex is then bound to streptavidin coated magnetic beads through the stable biotin-streptavidin interaction and the other reaction components are washed off. The green portion of the transcript is derived from the *glmS* catalytic riboswitch of *Bacillus subtilis* (90). Addition of glucosamine-6-phosphate (GlcN6P) to this *glmS* RNA sequence induces a site-specific cleavage event that releases only the desired RNA of choice shown (black). The biotin-streptavidin interaction is then broken under conditions that retain the functionality of the beads for further use.

Figure 4-1B shows the sequence of the exemplary RNA construct used in this study. The first RNA of choice is the antigenomic hepatitis delta virus (HDV) (91) ribozyme. The cleavage site of this catalytic RNA is indicated by the top arrow. At the 5' end of the cleavage site is a short sequence (orange) that represents a HindIII restriction enzyme site in the corresponding double-stranded DNA template. The terminal loop in helix P4 was replaced with the binding loop of the U1A protein (92). Binding of the U1A protein has been shown to aid in the crystallization of other important RNA molecules, including the genomic HDV ribozyme (32). The *glmS* riboswitch sequence shown in green was modified to include another short sequence (orange) that represents a AflII restriction site in the DNA template. The orange nucleotides at the 3' end correspond to a sequence in the plasmid template that gets cleaved by the restriction enzyme EcoRI. This allows for run-off transcription from a linear plasmid to readily yield the entire sequence shown in Figure 4-1B.

The two ribozymes present in this transcript imply that there are two site-specific cleavage events possible. When both of these occur from the same molecule, three smaller RNA products are generated that are described by the

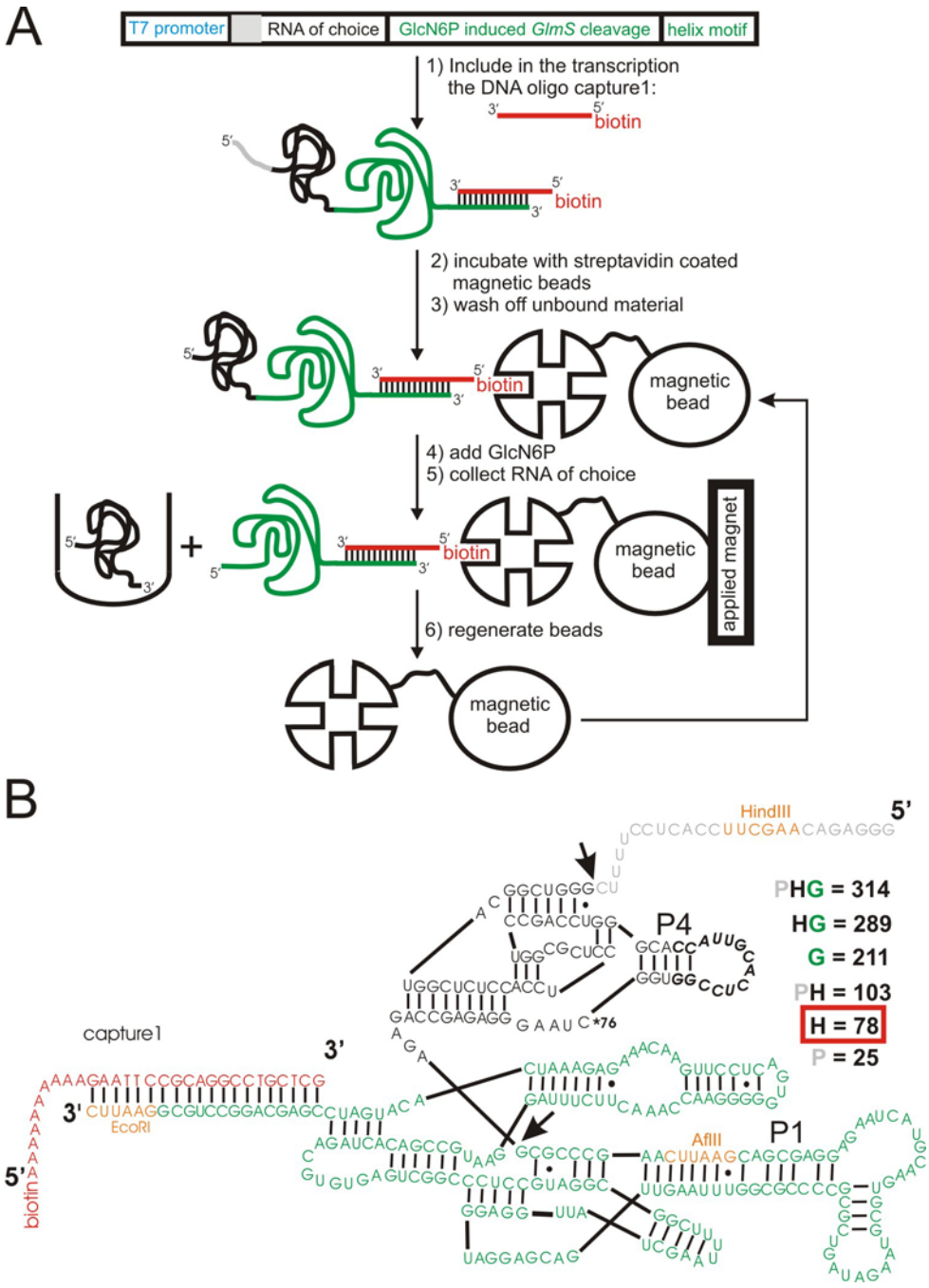


Figure 4-1. Schematic and sequence representation of the native RNA purification. A. Diagram of the strategy used to achieve the purification. B. The sequence of the transcript used in this study as an example of the application of the methodology. The H population of 78 nucleotides shown boxed in red are HDV molecules that underwent catalysis and are recovered to >99% purity using this procedure.

following nomenclature: “P” is the sequence upstream of the HDV cleavage site (upper arrow) shown in gray, “H” is the sequence downstream of the HDV cleavage site but upstream of the *glmS* cleavage site (lower arrow), and “G” is the sequence downstream of the *glmS* cleavage site. The sizes of these RNA’s and the other possible combinations of products are shown in Figure 4-1B.

Binding Capacity of capture1 for the Transcript. The application of the protocol outlined in Figure 4-1A is preferable if there is little unbound capture1 once the transcription is incubated with the magnetic beads. The regeneration of the beads involves heating the sample to 70^{o}C. Prior thermal stability studies have shown that ligand-free streptavidin has a T_m of 75^{o}C, however at full biotin saturation the T_m increases to 112^{o}C (93). This stabilization is proposed to arise from a closely packed β-barrel brought about by full biotin binding site saturation. There is no advantage to having free capture1 present when the transcription is incubated with the beads. This free capture1 could however saturate the 4 binding sites some streptavidins and raise its T_m. Free capture1 could also act as a competitive inhibitor for biotin binding sites accessible to the full complex.

Ten reactions under the standard transcription conditions containing different picomole amounts of capture1 were separated by non-denaturing polyacrylamide gel electrophoresis (PAGE). The gel was incubated with a sensitive fluorescent stain and the band intensities analyzed on a Fluorimager. The two important features of the gel are shown in Figure 4-2A. A slowly migrating population is observed, hypothesized to be HG (as the HDV ribozyme cleaves to completion, see below). Far faster migrating on the gel, bands of free capture1 begin to appear as larger amounts are included in the transcription. The control amounts of free capture1 shown on both ends of the gel in Figure 4-2A allowed for an estimate of the amount of bound picomoles of capture1 in each transcription, shown in Figure 4-2B. The observed increase is concomitant with the decrease of the HG population (Figure 4-2B). The binding of capture1 is saturated at around 350 picomoles, as indicated with the dashed line in Figure 4-2B. Additional amounts result only in more undesired free capture1 during the

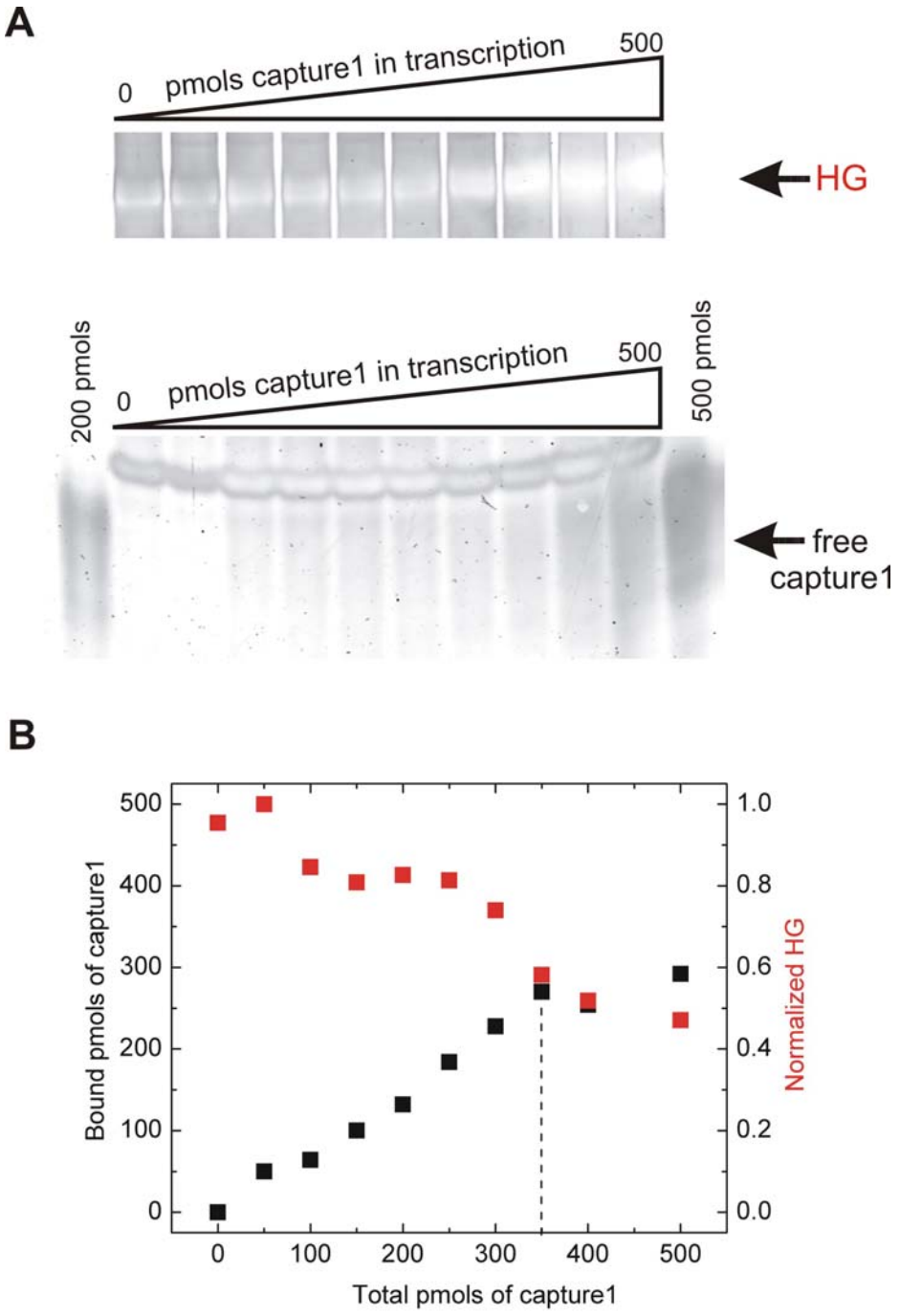


Figure 4-2. Binding conditions of capture1. A. Top and bottom portions of a non-denaturing PAGE gel shows the disappearance of the HG population along with the appearance of free capture1 as the total picomoles included in the transcription are increased. B. Quantification of the data from part A shows binding saturation at approximately 350 picomoles of capture1.

subsequent streptavidin binding step. This assay establishes 350 picomoles as the optimal amount of capture1 in the standardized transcription reaction conditions that are used for the rest of the study.

Activity Assay of the glmS Catalytic Riboswitch. The *glmS* catalytic riboswitch is a metabolite-responsive ribozyme that is activated upon the addition of glucosamine-6-phosphate (GlcN6P) (90). In the purification procedure described here it plays a crucial role by releasing the desired RNA only when GlcN6P is added. Previous kinetic characterizations of the *glmS* riboswitch revealed a cleavage rate of 1.5 min^{-1} upon addition of 200 μM GlcN6P and a total extent of cleavage of 70% with 100 μM GlcN6P (90). This efficient rate of catalysis combined with a low required concentration of metabolite makes the *glmS* riboswitch a good candidate for application in our purification. Previous native RNA purification work by Batey, et al. has also focused on employing this riboswitch to improve their method (46).

The catalytic properties of our version of the *glmS* riboswitch were investigated in the context of the finalized purification procedure. Building upon the results from Figure 4-2, a standard 2 h transcription reaction was run with 350 picomoles of capture1 to generate HG in complex with capture1. This transcription mixture was incubated with the streptavidin coated magnetic beads for 30 min while rotating at room temperature. Application of a magnet to the beads lead to removal of the supernatant. The beads were then washed 3 times with WB buffer. These beads were then resuspended with WB supplemented with 10 mM $MgCl_2$ and 200 μM GlcN6P to initiate cleavage of the bound HG RNA into the H and G components.

Figure 4-3A shows the accumulation of fraction of HG cleaved over time when 200 μM GlcN6P is added. The reaction is virtually completed after one min, as marked by the dashed line. Figure 4-3B summarizes the data gathered over a range of GlcN6P concentrations. The same total extent of cleavage is achieved after 10 min for concentrations ranging from 5 μM to 500 μM. The 200 μM data point is favored for further use in the purification protocol because the maximal

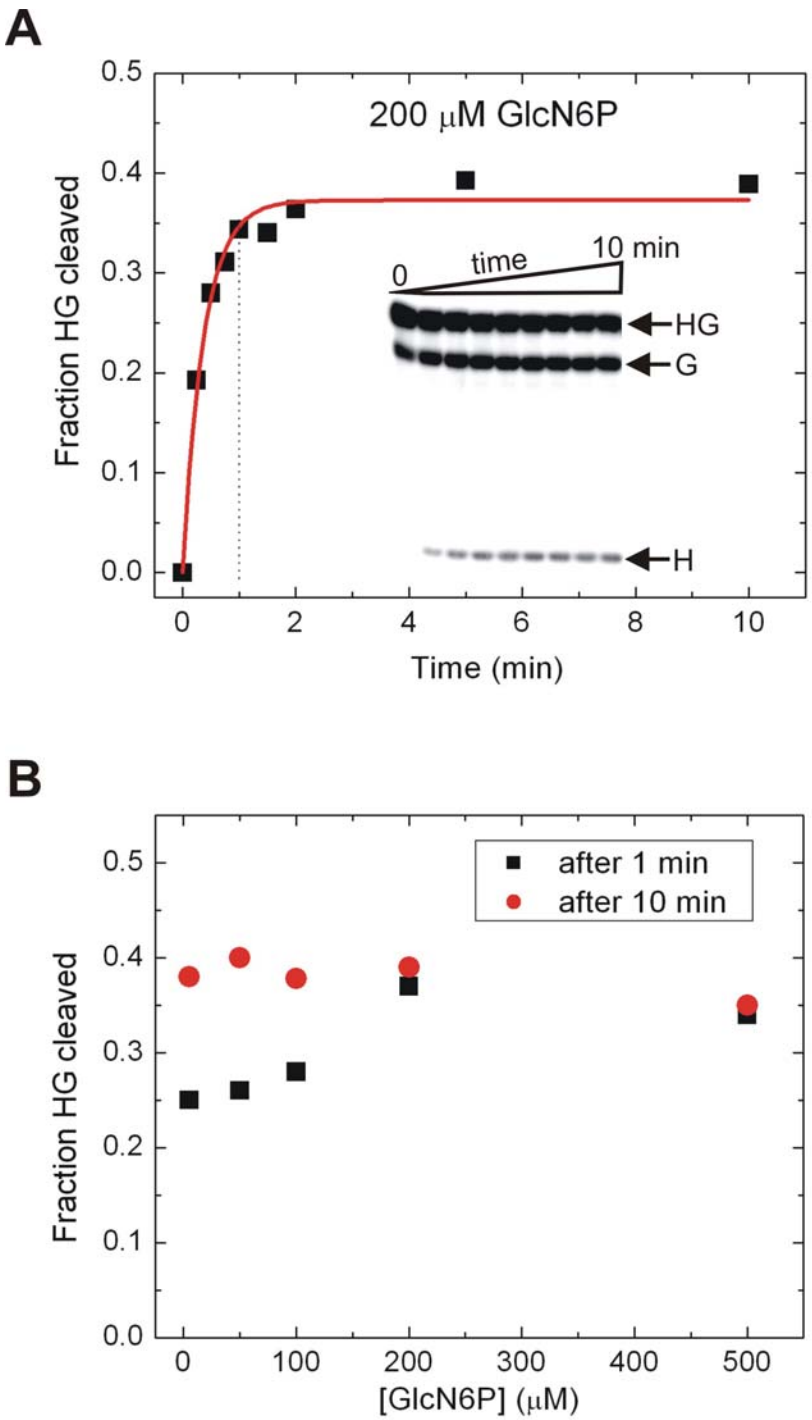


Figure 4-3. Kinetic characterization of the *glmS* riboswitch. A. An activity assay measures the cleavage of HG over time to yield H and G upon addition of 200 μM GlcN6P. These data points are derived from the denaturing PAGE gel shown in the insert. B. Activity assays over a range of GlcN6P concentrations show that 200 μM GlcN6P is the favorable concentration to use.

amplitude of the reaction is reached after only one min. A lower incubation time is preferable in this purification procedure because it minimizes the time over which the HG and G RNA molecules can potentially dissociate from their stable complex with the beads and get into the elution supernatant that contains the desired RNA.

Implementation of the native purification protocol. The optimized capture1 and *glmS* cleavage conditions are combined towards the purification scheme outlined in Figure 4-1A. The supernatants removed throughout the protocol were analyzed on a denaturing gel as shown in Figure 4-4A. The RNA transcript shown in Figure 4-1B was transcribed in the presence of capture1. The T7 RNA polymerase was then removed through a phenol/chloroform extraction because its presence resulted in small amounts of beads sticking to the sides of the microcentrifuge tubes. The NaCl concentration was increased to 1 M to per recommendation by the manufacturer of the magnetic beads (Invitrogen) to achieve optimal binding of biotin to the streptavidin. The reaction mixture was then incubated with the beads for 30 min. A magnetic separator was used to pull the magnetic beads to the side of the tube and allow for removal of the transcription supernatant. This supernatant contains the RNA that did not form a complex with the beads, as shown in Figure 4-4A. There is almost complete cleavage of the HDV ribozyme in the transcription reaction. This yields the P and HG populations from the original PHG transcript whose sequence is shown in Figure 4-1B. There is also some background cleavage of the *glmS* riboswitch in the HG population which gives rise to the observed H and G RNA's.

The beads were then washed four times with WB buffer (40 mM HEPES-KOH, pH7.4, and 1 M NaCl) and the supernatant removed each time (wash 1-4, Figure 4-4A). This high ionic strength is designed to maximize the stability of the helix between capture1 and the 3' end of the transcript (Figure 4-1B). Subsequent addition of WB supplemented with 10 mM $MgCl_2$ and 200 μM GlcN6P leads to a supernatant whose RNA content is dominated by the desired

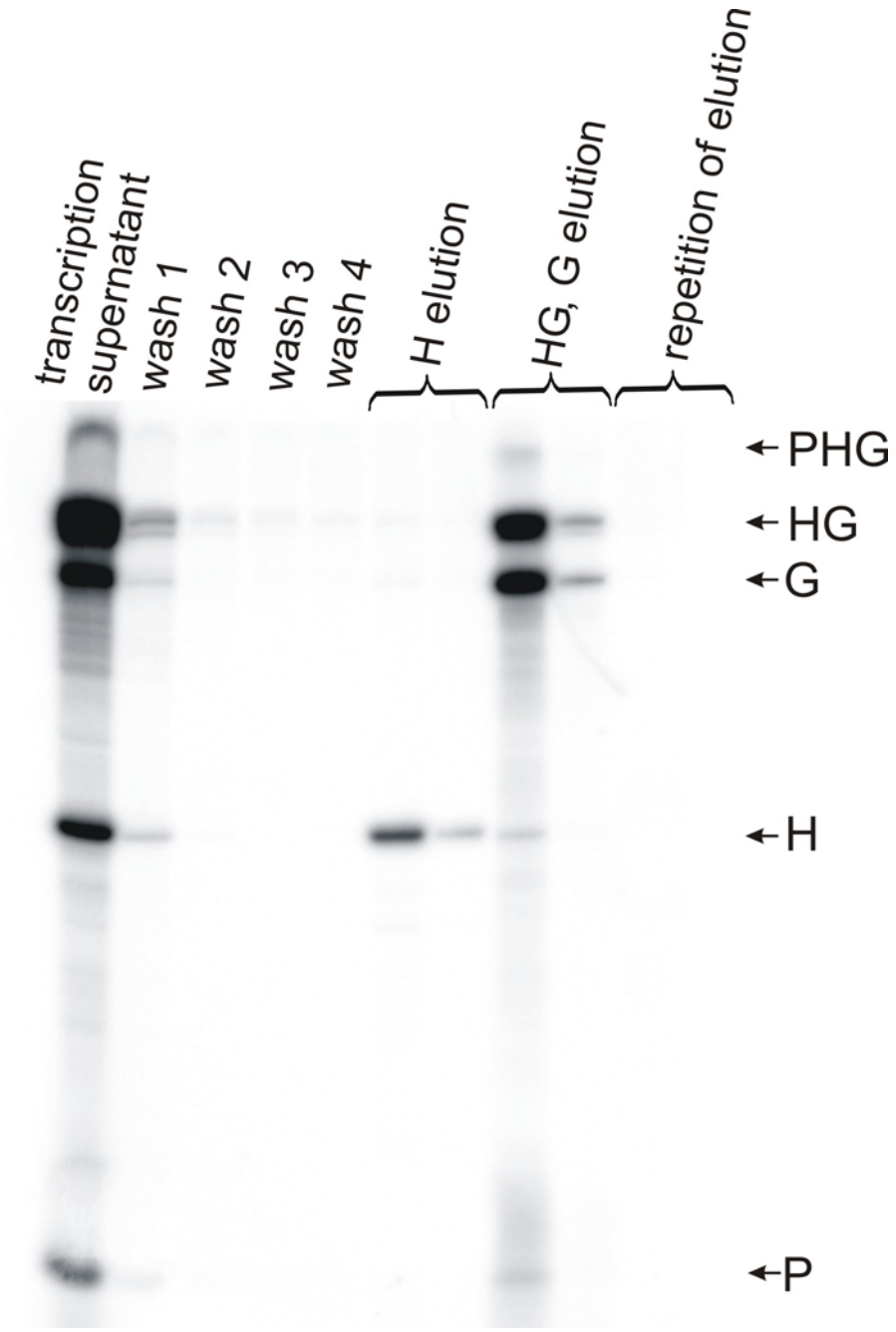


Figure 4-4. Analysis of the supernatants generated throughout the purification. A greater than 99% pure sample of cleaved HDV RNA (H population) is obtained in the elution step. A single heating/cooling cycle then removes all of the HG and G population that were complexed to the magnetic beads.

H population (see the 1st H elution lane in Figure 4-4). A further wash with WB supplemented with 10 mM $MgCl_2$ collects some additional H population (2nd H elution lane, Figure 4-4). Those H elution supernatants were obtained using 1 mg of beads and contained a total of 40 picomoles of cleaved HDV RNA that is >99% pure as calculated from the relative intensities of the bands in the "H elution" lane. The robustness of the capture1 helix formation in parallel with the efficiency of the GlcN6P induced cleavage enable this methodology to generate pure, cleaved HDV RNA that never underwent denaturation after having folded into its functional form during its transcription.

The final stage of the purification scheme outlined in Figure 4-1A is the regeneration of the beads to make the overall process cost-effective. Previous work (94) has shown that the biotin-streptavin interaction can be reversibly broken with heating under low ionic strength conditions. After GlcN6P induced cleavage the beads are washed with an elution buffer (EB) of 10 mM Tris-HCl, pH 8.0, then heated to 70°C in the same EB buffer for 5 min, then cooled to room temperature for 5 min. The supernatant now contains the HG and G RNA species that were previously bound to the beads. An additional wash with EB removes the remaining HG and G RNA's (Figure 4-4A, HG, G elution). A subsequent 90°C incubation for 5 min in EB buffer followed by immediate collection of the supernatant shows no additional stripping of nucleic acid material from the beads (data not shown).

The beads remain viable for additional rounds of purification after having undergone the entire procedure described above. Beads can be reintroduced repeatedly to another completed and prepared transcription reaction following the same sequence of steps described in Figure 4-4. The results are shown in Figure 4-5B. The purity is maintained above 96% even after six full purification cycles. The total amount of H RNA is also unaffected as the beads are subjected to repeated purification protocols. This greatly expands the production capacity of a single lot of purchased magnetic beads.

To demonstrate the broad applicability of this procedure, a native purification of the Varkud Satellite (VS) ribozyme was achieved. The sequence

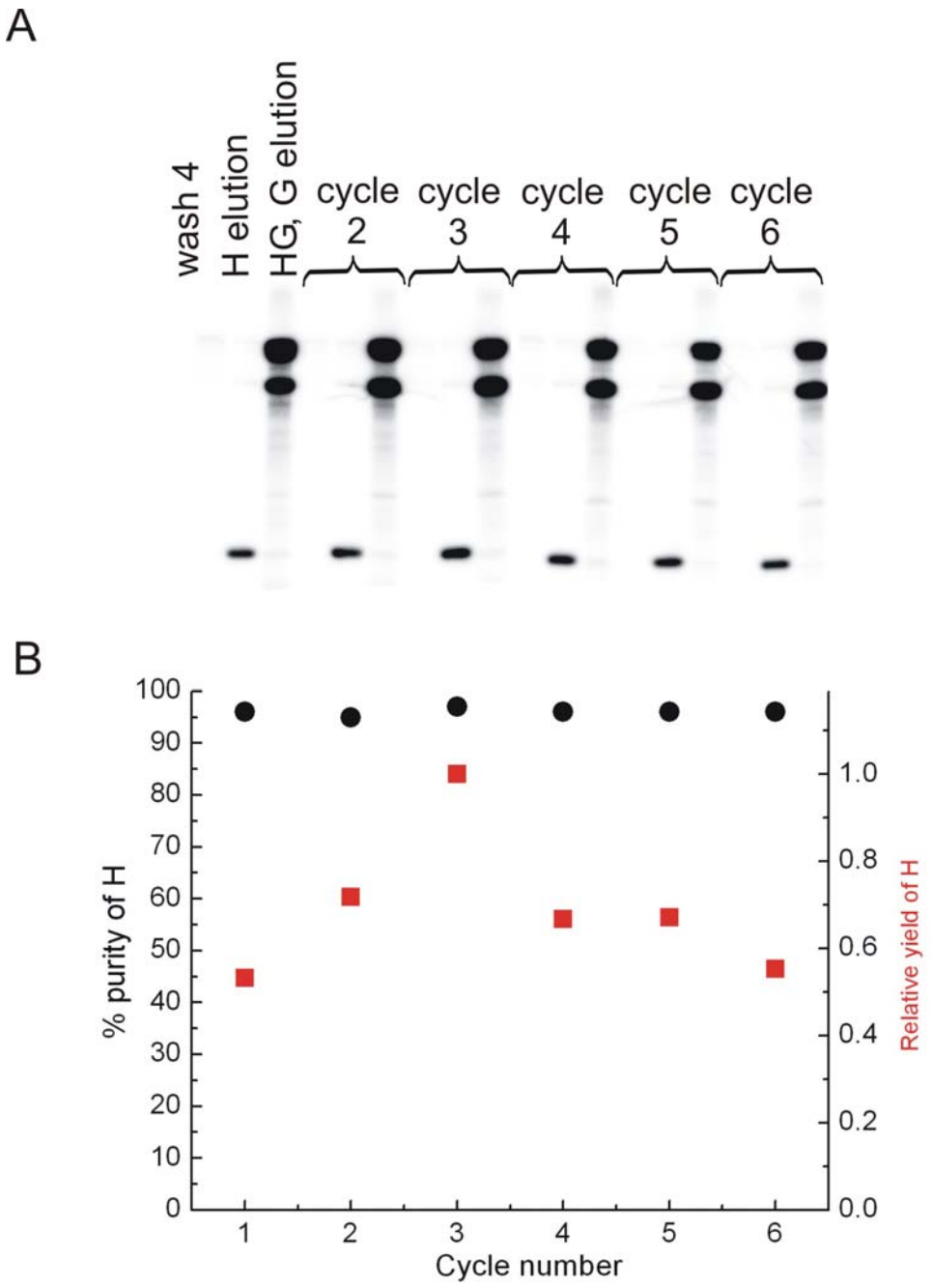


Figure 4-5. The repeated use of magnetic beads does not lower their performance. A) Using the same lane descriptions as in Figure 4-4, the wash 4, H elution, and HG + G elution supernatants of six successive rounds of purification are shown. The beads used in cycle 1 were re-used for the subsequent 5 cycles. B) The % purity of the H elution samples and the amount of H obtained are not compromised despite the recycling of the magnetic beads.

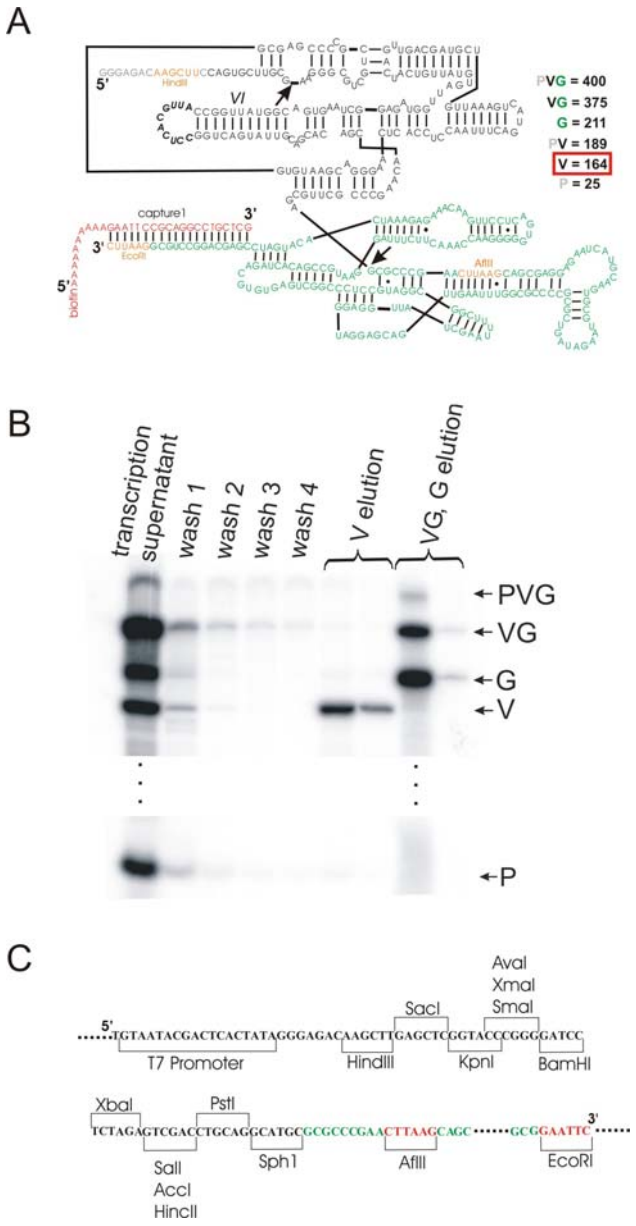


Figure 4-6. Successful application of the native RNA purification protocol to the VS ribozyme. A) Sequence of the VS ribozyme within the PVG transcript using the same color notation as Figure 4-1A. The sizes of the possible components of the PVG transcript are shown. B) Analysis of collected supernatants show the successful purification of cleaved VS RNA to >99% purity in the “V elution” lanes. C) Sequence of the multiple cloning site portion of the pMCGL plasmid.

and secondary structure of the VS ribozyme utilized is shown in Figure 4-6A. This VS construct is based on G11 (54) (Figure 2-1A) except that the end of helix VI was redesigned to incorporate the U1A protein binding loop (92); the upstream sequence of the cleavage site also differs due to insertion of a HindIII restriction site in the plasmid template. The same procedure was employed as detailed above for the HDV RNA. Once again 1 mg of beads yielded 40 picomoles of cleaved VS RNA at >99% purity (Figure 4-6B), attesting to the broad applicability of the technique.

A plasmid (4-6C) was engineered for the greater research community to allow for broad application of the procedure described here. The “RNA of choice” portion of the plasmid now contains a sequence similar to the universal multiple cloning site of the pGEM4Z vector for ease of cloning.

4.4 Discussion

We have presented a novel method for native RNA purification that produces the target RNA at greater than 99% purity when applied to the antigenomic HDV and VS ribozymes. In both examples the RNA was folded into its functional form during the transcription reaction, just as would be the case inside the cell. The functionality of these RNA structures is confirmed by the observation that a large majority of the transcribed ribozyme sequences has the catalytic activity to undergo self-cleavage. This observation supports the notion that the post-cleavage versions of the HDV and VS RNA’s that are collected using this protocol are functionally relevant structures.

The folded structures are maintained throughout the purification protocol. One particular detail worth noting is the absence of magnesium in the wash buffer. Earlier versions of the wash buffer contained a range of magnesium concentrations. This led, however, to contamination of both the HDV and VS ribozyme elution supernatants with some amount of *glmS* ribozyme. We have evidence to suggest that this side reaction arose from trans-interactions between

full-length transcripts (see Appendix 2). When this occurrs the GlcN6P induced cleavage event releases *glmS* ribozyme that was not attached to the beads through the biotin-streptavidin interaction mediated by capture1. Rather, it was complexed with another full length transcript since its immediate upstream target RNA forms a trans-interaction with another full length transcript that in turn is bound to the beads. The removal of magnesium abates this issue. While it is true that the transcribed target RNA experiences a removal of magnesium during the purification, this does not denature it as it is still maintained in a high ionic strength buffer. When the magnesium is re-introduced into the buffer solution containing 200 μM GlcN6P, the *glmS* riboswitch portion of the transcript exists in its catalytically active structure, as evidenced by the cleavage event itself. It is only from these transcripts that the target RNA is collected.

The HDV and VS ribozyme populations purified in this study also carry a high degree of primary sequence homogeneity. Their self-cleavage reaction results in homogeneous 5'-hydroxyl ends, and the eventual site-specific 3' end cleavage by the downstream *glmS* riboswitch ensures homogeneous 3' cyclic phosphate ends as well. This offers another advantage when the ultimate experimental goal places a high emphasis on sequence homogeneity, such as X-ray crystallography of the RNA. Such 3' end homogeneity is not achieved when the standard protocol of refolding of a transcribed denatured RNA is followed, as T7 RNA polymerase is prone to introduce heterogeneous 3' ends (42).

For future X-ray crystallography applications the new purification method described here will need to be significantly scaled up, which is expected to be straightforward. A robust screening of crystallization conditions requires at least on the order of 200,000 picomoles, or 400 μl of a 500 μM solution. In the studies described here 1 mg of beads yielded 40 picomoles from a 100 μl transcription of pure HDV or VS ribozyme RNA after one round of use. A direct scaling up of this protocol would yield the necessary 200,000 picomoles when the commercially available amount of 100 ml beads is used over six consecutive purification cycles. The procedure is also amenable towards NMR applications where the use of expensive ^{13}C and ^{15}N nucleoside triphosphates poses a limiting factor in

the generation of the transcribed RNA. In this case, the initially collected transcription supernatant can be re-incubated with regenerated beads. This can significantly add to the overall yield of the desired RNA from a given transcription volume if the amount of transcript made exceeds the initial binding capacity of the beads (data not shown).

Making this purification protocol available to as wide a research audience as possible was a central theme in its development. There is no protein purification expertise or equipment required. An investment in the magnetic beads and the magnetic particle collector is all that is required for a laboratory to utilize this procedure. The cloning of the appropriate sequence into our available pMCGL plasmid, and the implementation of this efficient, robust protocol should result in the successful production of natively folded, highly pure RNA of any desired sequence.

CHAPTER 5

SUMMARY AND FUTURE DIRECTIONS

It is now abundantly clear that non-coding RNA plays a pivotal role in a variety of cellular processes within the cell, including regulation, modification, splicing, and translation (2-4). It is critically important to understand how this central molecule of life is able to accomplish such a diverse array of tasks. This understanding will not occur unless the role that global structural dynamics play is thoroughly investigated for a variety RNAs. From synergistic observatins of a large number of investigators, a consistent set of themes is beginning to emerge. The new way of thinking can then be confidently applied to the other noncoding RNAs that will undoubtedly be discovered and studied in the future. The research presented here is a significant contribution towards such a collective of investigations highlighting the role of structural dynamics in non-coding RNA function.

5.1 Reaction pathway of the VS ribozyme

Two fundamentally different constructs of the VS ribozyme are presented here in chapters 2 and 3. The experiments described in chapter 2 expand on previous work on G11-based constructs (54). These experiments took investigation of G11 from the bulk solution realm into the world of single-molecule measurements. This new approach revealed G11 as a VS ribozyme that is

dynamic and regularly interchanges between 3 globally distinct populations. One of these structures (the H state, Figure 2-6) was shown to be an essential intermediate that must be traversed before the VS ribozyme can undergo catalysis. Although this structure is necessary, further downstream local structural rearrangements were found to be necessary before cleavage can take place.

New insight into how the H state is reached came from the results of single molecule studies of specific mutants described in chapter 2. It is a hierarchical process whereby each newly formed intermediate makes new tertiary contacts that are contingent on the proper folding of the previous state. The correctly folded II-III-VI junction is a prerequisite for the formation of the kissing loop interaction. This interaction is its own prerequisite for the correct juxtaposition of the cleavage and active site, which is itself a requirement for the biological function of the VS ribozyme, the site-specific cleavage of its RNA backbone. This theme of how the VS ribozyme folds into its catalytic active state in a hierarchical manner may apply to many other important RNAs (97).

The method of analyzing single molecule traces presented in chapter 2 (Figures 2-2 and 2-5) is broadly applicable to other systems as well. It involves the cogent combination of a number of previously known tools. An example trace of the raw data is shown. This is overlaid with an idealized traced obtained from the application of a non-linear filter (66) that effectively reduces the noise inherent in the raw data without losing valuable information. The path predicted by Hidden Markov Modeling (HMM) (67) is shown as well so the reader can visually see the conversion from raw data to the HMM fit that is used to derive all of the rate constants that describe the transitions between different states. Second, a histogram analysis of all the analyzed traces gives the reader the overall population distribution for all the analyzed single molecules. Third, a transition density plot (73) is shown to reveal which observed transitions from the HMM fits predominate as well as the rate constants associated with each of these transitions. Fourth, the dwell times for each transition are plotted as normalized probability density plots (68). The HMM analysis assumes *a priori*

that each transition is governed by a single exponential rate constant. Complementarily, the latter probability density plots reveal whether the dwell times are indeed sufficiently described by a single rate constant or whether more than one stochastic process is responsible for the generation of the observed dwell times. The combination of all the analyses described above gives a detailed and complete representation of the wealth of information that is available from single molecule experiments.

The results detailed in chapter 3 underscore how the connectivity of an RNA can have a large impact on its folding behavior. Here, a version of the VS ribozyme with an alternate connectivity was investigated. This rearrangement led to a pronounced absence of global dynamics compared to the G11-based construct. Such behavior is rationalized if one considers the model described in chapter 3 that operates on the assumption that VS RNA has evolved to defend itself against exonucleases present within the mitochondria of its *Neurospora* host. Our results also fit into the more general theme that not all non-coding RNAs can be assumed to be globally dynamic on timescales observable by single molecule experiments (98, 99). The *glmS* riboswitch that is applied in chapter 4 is another such example.

5.2 Native purification of RNA

A novel assay for the isolation of co-transcriptionally folded, pure RNA was developed. To illustrate the applicability of this assay, post-cleavage forms of both the antigenomic Hepatitis Delta Virus (HDV) and Varkud Satellite (VS) ribozymes were isolated. The assay is accessible to many research laboratories due to the availability of our developed pMCCL plasmid and lack of protein preparatory work required by the user. It is also cost effective due to the regenerative ability of the magnetic beads required by the user. This procedure is anticipated to be broadly applicable to the RNA field. For example, in her recent successful crystallization of the large group II intron (100), Prof. Anna Pyle was

fortuitously able to purify it from all the other components of the transcription reaction by using a 100-kDa ultrafiltration procedure. The success of this crystallization was attributed by Prof. Pyle in large part to the native purification of the group II intron, suggesting an important role of a more broadly applicable native purification protocol as developed here that does not require the target RNA to be much larger than any contaminant.

5.3 Future Directions

The successful characterization of the VS ribozyme construct presented in chapter 2 was in large part due to the use of different mutants. These were chosen to reveal as much information as possible. It was the combined analysis of four variants using complementary bulk solution and single molecule experiments that led to a deeper understanding of this RNA. An analogous strategy with the FR3 VS variant in chapter 3 is expected to produce a similar advance in knowledge.

The results shown in chapter 4 open the door to additional important structure probing techniques that are yet to be done on the antigenomic HDV ribozyme and the VS ribozyme. At the top of this list is X-ray crystallography. Both of the transcripts (Figure 4-1B and 4-6A) contain a U1A binding loop that is expected to aid in formation of crystal contacts. But most importantly, the assay described in chapter 4 will allow for the screening of crystallization conditions of VS or HDV RNA that was co-transcriptionally purified. The preliminary work done towards purification of the U1A protein described in Appendix 1 will aid in this future work as well.

This native purification assay can also be used to shed light on the question of heterogeneous folding in the hairpin ribozyme. Prior work has shown that the hairpin ribozyme separates into 2 bands on a non-denaturing gel (17). The source of this heterogeneity has turned out to be very difficult to ascertain. One possibility is the standard purification procedure that utilizes denaturing gel

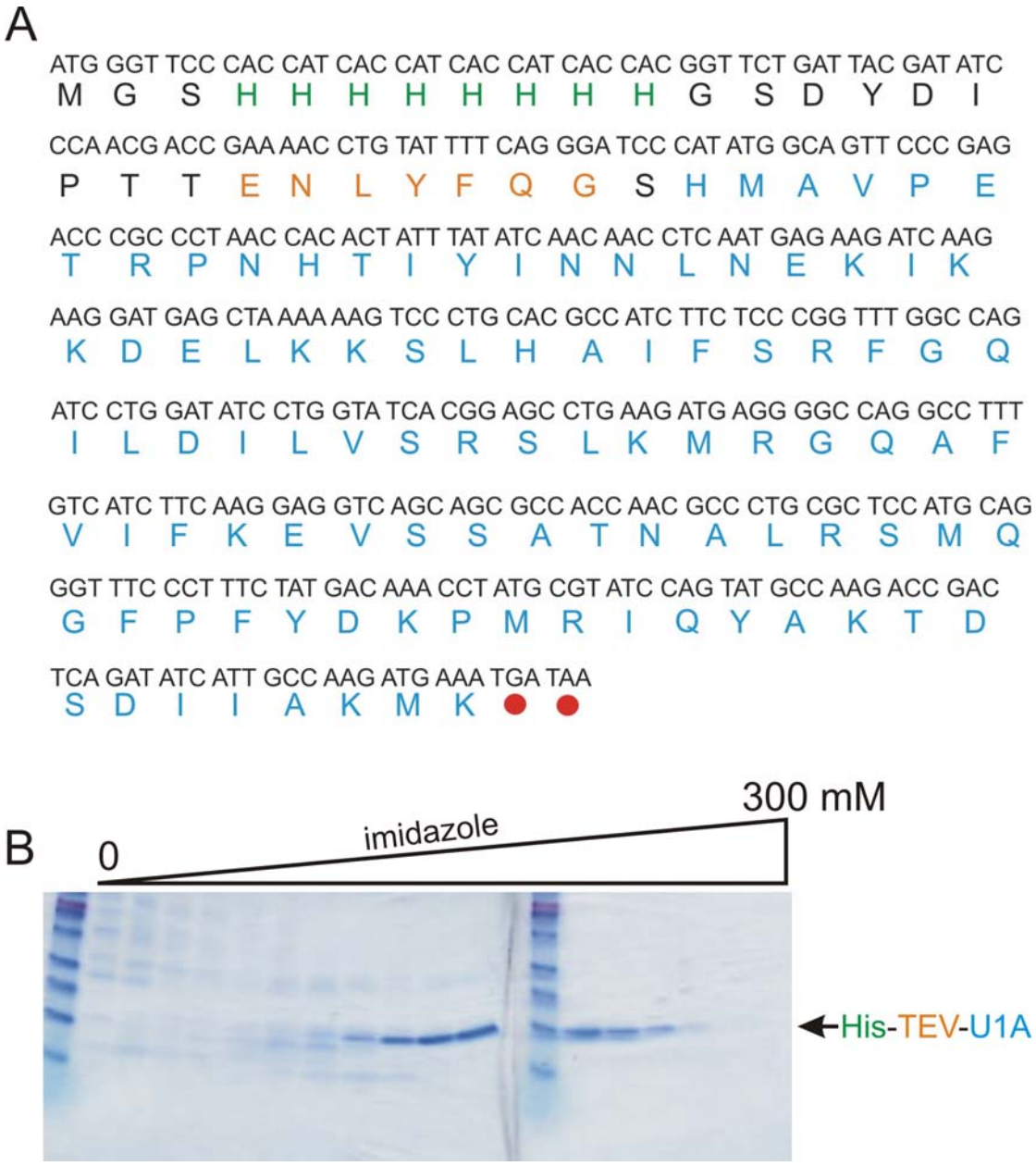
electrophoresis and may introduces a modification of some sort that then leads to the observed heterogeneous behavior on non-denaturing gels and in single molecule and functional assays (17). Purification of the hairpin ribozyme applying the protocol described in chapter 4, followed by analysis of the sample on a non-denaturing gel would conclusively determine whether or not the prior denaturation based purification protocol is responsible for this heterogeneity and can be avoided.

APPENDIX 1

PROGRESS TOWARDS PURIFICATION OF U1A PROTEIN

Purification of U1A protein was pursued as crystallization of the VS ribozyme was one of the longer-term aims for this dissertation. U1A has been instrumental in the attainment of high-resolution crystal structures of other RNAs. It forms a tight complex with a specific stem-loop sequence that can be engineered into the target RNA. Its isoelectric point lies at a higher pH than crystallizations conditions, which allows its net positive charge to interact with the negatively charged RNA backbone, thus facilitating crystal contacts and formation of an ordered unit cell (92). This requires the purification of a large amount of U1A that is free of any trace amount of RNases (32).

A strategy of attaching a histidine tag to U1A and employing an imidazole induced elution from a cobalt column was chosen (95). The first step was the creation of the psj4-U1A plasmid whose important features are shown in Appendix 1 Figure 1A. The U1A portion shown in blue came from a plasmid generously provided by Prof. Collins from the University of Toronto. This U1A sequence was inserted into the psj4 plasmid through the use of shared NdeI and HindIII restriction sites and standard molecular biology cloning techniques. The pET based psj4 vector makes the cell kanamycin resistant and is engineered for protein purification through the use of a cobalt column. Once the psj4 plasmid is inside of the BL21(DE3) strain of *E. coli*, addition of isopropyl-β-D-thiogalactopyranoside (IPTG) induces the expression of T7 RNA Polymerase within this strain of *E. coli*. This polymerase then transcribes starting from the T7



Appendix 1 Figure 1. Translated portion of the pjs4-U1A plasmid. A) Upon induction with IPTG, T7 RNA polymerase is expressed in the BL21(DE3) *E. coli* cells which then creates a transcript off of the pjs4 plasmid. The codon sequence portion of the transcript is shown along with its corresponding amino acid sequence. The histidine portion is shown in green. Orange denotes the TEV protease specific cleavage site. In blue is the amino acid sequence of the U1A protein. B) The amino acid peptide shown in A is eluted off of the cobalt column as the concentration of imidazole is continuously increased.

promoter on the psj4 plasmid until it reaches an RNA T7 terminator sequence. This transcribed RNA is then translated to create the polypeptide sequence (96) shown in Appendix 1 Figure 1A.

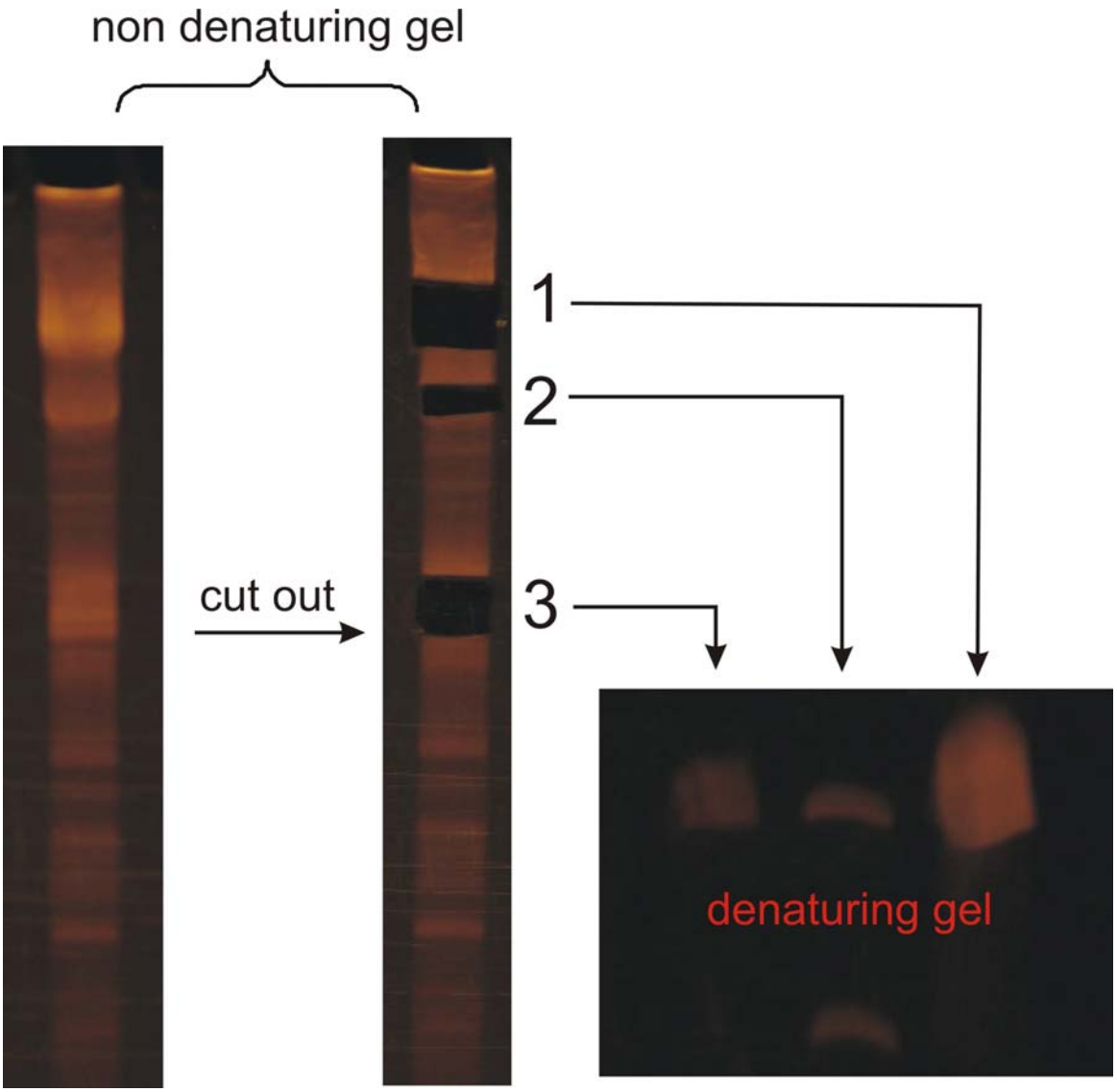
300 ml of BL21(DE3) cells containing the pjs4U1A plasmid were induced with 0.5 mM IPTG and shaken overnight at 37^{o}C. The cells were centrifuged at 4,000 rpm for 10 min and redissolved in 50 ml of cold phosphate saline buffer (PBS). These cells were sonicated using a large tip in a metal beaker on ice for 7 cycles of 30 s sonication, 30 s rest per cycle. The solution was spun down at 20,000 rpm, 4^{o}C for 30 min and the supernatant that contains the desired peptide was run through a 0.2 µm filter before being applied to a cobalt column. To this end, FPLC equipment was generously provided by the laboratory of Prof. Saper of the Biological Chemistry department at the University of Michigan. The psj4 plasmid was provided by Prof. Saper as well. The 8 ml Talon cobalt column was equilibrated with 1X PBS before loading of the sample at 1 ml per min. The column was then washed with 112 ml of 1X PBS, followed by a linear gradient in 1X PBS at 3 ml per minute from 0 to 300 mM imidazole over a total applied volume of 160 ml. Twenty 8-ml fractions were collected from which 10 µl were loaded onto 2 x 16% Tris-glycine gels and stained with Coomassie Blue. Analysis of these gels (shown in Appendix 1 Figure 1B) reveals that elution of the desired peptide begins at 165 mM imidazole and ends around 225 mM. This shows there is expression in the *E. coli* and the U1A-containing peptide can be extracted from the cells and eluted from a cobalt column in a relatively pure form. The actual extraction of the desired U1A protein from the column, however, requires the future steps of adding the TEV protease to cleave at the amino acid sequence shown in orange (Appendix 1 Figure 1A) to release the desired U1A protein. This approach allows for removal of the his-tag from the U1A protein.

APPENDIX 2

QUATERNARY INTERACTIONS PRESENT IN VS RIBOZYME TRANSCRIPTION REACTIONS

Early development of the native purification assay described in chapter 4 was hindered by the persistent contamination of the desired RNA with the *glmS* portion (abbreviated as “G” in chapter 4) of the transcript. This was puzzling since it was absent in the supernatants of the washes done prior to activation of the *glmS* riboswitch. It was the *glmS* cleavage event that led to its appearance. A hypothesis was that the transcripts were forming dimers in the transcription reaction, and once the beads were introduced only one of the two biotin molecules in the dimer bound to the streptavin-coated bead. This complex remained intact throughout the wash steps until addition of GlcN6P cleaved and released *glmS* RNA, if magnesium was included in the wash buffer. According to this hypothesis, the 3’ end of this *glmS* RNA would not be attached to the streptavidin-coated beads via capture1; its release upon washing steps would lead to contamination of the desired RNA.

The experiment shown in Appendix 2 Figure 2 confirms the presence of trans-interactions when PVG (Figure 4-6A) is transcribed. The standard transcription conditions of PVG were used. A 10 µl aliquot was then loaded onto a 6% (w/v) non-denaturing PAGE gel that was run at 5 W for 4 h, followed by staining with ethidium bromide. Three bands were cut out and placed in the lanes of a 10% denaturing PAGE gel that was run at 20 W for 2 h. Band 2 is shown to consist of 2 pieces of different sizes, consistent with the hypothesis that under non-denaturing conditions the proposed heteromultimers form.



Appendix 2 Figure 2. Confirmation of quaternary structures in a transcription reaction. Three separate bands from a non-denaturing gel are subsequently run on a denaturing gel. Band 2 is revealed to be a complex that is made up of at least 2 RNAs that differ in size. While the exact classification of these RNA strands is somewhat ambiguous, what is conclusive is the fact that there are multimer complexes present in the transcription reaction.

REFERENCES

1. Joo, C., H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha. 2008. Advances in single-molecule fluorescence methods for molecular biology. Annual review of biochemistry 77:51-76.
2. Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V. B. Bajic, S. E. Brenner, S. Batalov, A. R. Forrest, M. Zavolan, M. J. Davis, L. G. Wilming, V. Aidinis, J. E. Allen, A. Ambesi-Impiombato, R. Apweiler, R. N. Aturaliya, T. L. Bailey, M. Bansal, L. Baxter, K. W. Beisel, T. Bersano, H. Bono, A. M. Chalk, K. P. Chiu, V. Choudhary, A. Christoffels, D. R. Clutterbuck, M. L. Crowe, E. Dalla, B. P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C. F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T. R. Gingeras, T. Gojobori, R. E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T. K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S. P. Krishnan, A. Kruger, S. K. Kummerfeld, I. V. Kurochkin, L. F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K. C. Pang, W. J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J. F. Reid, B. Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S. L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C. A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S. L. Tan, S. Tang, M. S. Taylor, J. Tegner, S. A. Teichmann, H. R. Ueda, E. van Nimwegen, R. Verardo, C. L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S. M. Grimmond, R. D. Teasdale, E. T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J. S. Mattick, D. A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai, and Y. Hayashizaki. 2005. The transcriptional landscape of the mammalian genome. Science (New York, N.Y 309:1559-1563.
3. Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H. Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C. Wells, M. Frith, T. Ravasi, K. C. Pang, J.

Hallinan, J. Mattick, D. A. Hume, L. Lipovich, S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S. Mottagui-Tabar, Z. Liang, B. Lenhard, and C. Wahlestedt. 2005. Antisense transcription in the mammalian transcriptome. Science (New York, N.Y 309:1564-1566.
4. Pheasant, M., and J. S. Mattick. 2007. Raising the estimate of functional human sequences. Genome research 17:1245-1253.
5. Walter, N. G., and D. R. Engelke. 2002. Ribozymes: catalytic RNAs that cut things, make things, and do odd and useful jobs. Biologist (London, England) 49:199-203.
6. Doudna, J. A., and J. R. Lorsch. 2005. Ribozyme catalysis: not different, just worse. Nat. Struct. Mol. Biol. 12:395-402.
7. Cornish, P. V., and T. Ha. 2007. A survey of single-molecule techniques in chemical biology. ACS chemical biology 2:53-61.
8. Zhuang, X., L. E. Bartley, H. P. Babcock, R. Russell, T. Ha, D. Herschlag, and S. Chu. 2000. A single-molecule study of RNA catalysis and folding. Science (New York, N.Y 288:2048-2051.
9. Axelrod, D. 2003. Total internal reflection fluorescence microscopy in cell biology. Methods in enzymology 361:1-33.
10. Clegg, R. M. 1992. Fluorescence resonance energy transfer and nucleic acids. Methods in enzymology 211:353-388.
11. Zhuang, X., H. Kim, M. J. Pereira, H. P. Babcock, N. G. Walter, and S. Chu. 2002. Correlating structural dynamics and function in single ribozyme molecules. Science (New York, N.Y 296:1473-1476.
12. Rueda, D., G. Bokinsky, M. M. Rhodes, M. J. Rust, X. W. Zhuang, and N. G. Walter. 2004. Single-molecule enzymology of RNA: Essential functional groups impact catalysis from a distance. Proceedings of the National Academy of Sciences of the United States of America 101:10066-10071.
13. Liu, S., G. Bokinsky, N. G. Walter, and X. Zhuang. 2007. Dissecting the multistep reaction pathway of an RNA enzyme by single-molecule kinetic "fingerprinting". Proceedings of the National Academy of Sciences of the United States of America 104:12634-12639.
14. Rueda, D., G. Bokinsky, M. M. Rhodes, M. J. Rust, X. Zhuang, and N. G. Walter. 2004. Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance. Proceedings of the National Academy of Sciences of the United States of America 101:10066-10071.
15. Bokinsky, G., D. Rueda, V. K. Misra, M. M. Rhodes, A. Gordus, H. P. Babcock, N. G. Walter, and X. Zhuang. 2003. Single-molecule transition-state analysis of RNA folding. Proceedings of the National Academy of Sciences of the United States of America 100:9302-9307.
16. Walter, N. G. 2001. Structural dynamics of catalytic RNA highlighted by fluorescence resonance energy transfer. Methods (San Diego, Calif 25:19-30.
17. Ditzler, M. A., D. Rueda, J. Mo, K. Hakansson, and N. G. Walter. 2008. A rugged free energy landscape separates multiple functional RNA folds throughout denaturation. Nucleic acids research 36:7088-7099.

18. Benkovic, S. J., and S. Hammes-Schiffer. 2003. A perspective on enzyme catalysis. Science (New York, N.Y 301:1196-1202.
19. Blanchard, S. C., R. L. Gonzalez, H. D. Kim, S. Chu, and J. D. Puglisi. 2004. tRNA selection and kinetic proofreading in translation. Nature structural & molecular biology 11:1008-1014.
20. Blanchard, S. C., H. D. Kim, R. L. Gonzalez, Jr., J. D. Puglisi, and S. Chu. 2004. tRNA dynamics on the ribosome during translation. Proceedings of the National Academy of Sciences of the United States of America 101:12893-12898.
21. Fedor, M. J., and J. R. Williamson. 2005. The catalytic diversity of RNAs. Nature reviews 6:399-412.
22. Bevilacqua, P. C., and R. Yajima. 2006. Nucleobase catalysis in ribozyme mechanism. Current opinion in chemical biology 10:455-464.
23. Scott, W. G. 2007. Ribozymes. Current opinion in structural biology 17:280-286.
24. Beattie, T. L., J. E. Olive, and R. A. Collins. 1995. A secondary-structure model for the self-cleaving region of Neurospora VS RNA. Proc. Natl. Acad. Sci. USA 92:4686-4690.
25. Hiley, S. L., and R. A. Collins. 2001. Rapid formation of a solvent-inaccessible core in the Neurospora Varkud satellite ribozyme. EMBO J. 20:5461-5469.
26. Hiley, S. L., V. D. Sood, J. Fan, and R. A. Collins. 2002. 4-thio-U cross-linking identifies the active site of the VS ribozyme. EMBO J. 21:4691-4698.
27. Lafontaine, D. A., D. G. Norman, and D. M. Lilley. 2001. Structure, folding and activity of the VS ribozyme: importance of the 2-3-6 helical junction. EMBO J. 20:1415-1424.
28. Lafontaine, D. A., D. G. Norman, and D. M. Lilley. 2002. The global structure of the VS ribozyme. EMBO J. 21:2461-2471.
29. Pereira, M. J., D. A. Harris, D. Rueda, and N. G. Walter. 2002. Reaction pathway of the trans-acting hepatitis delta virus ribozyme: a conformational change accompanies catalysis. Biochemistry 41:730-740.
30. Walter, N. G., K. J. Hampel, K. M. Brown, and J. M. Burke. 1998. Tertiary structure formation in the hairpin ribozyme monitored by fluorescence resonance energy transfer. EMBO J. 17:2378-2391.
31. Ha, T. 2001. Single-molecule fluorescence resonance energy transfer. Methods (San Diego, Calif 25:78-86.
32. Ferre-D'Amare, A. R., K. Zhou, and J. A. Doudna. 1998. Crystal structure of a hepatitis delta virus ribozyme. Nature 395:567-574.
33. Ke, A., K. Zhou, F. Ding, J. H. Cate, and J. A. Doudna. 2004. A conformational switch controls hepatitis delta virus ribozyme catalysis. Nature 429:201-205.
34. Wedekind, J. E., and D. B. McKay. 1998. Crystallographic structures of the hammerhead ribozyme: relationship to ribozyme folding and catalysis. Annual review of biophysics and biomolecular structure 27:475-502.

35. Martick, M., and W. G. Scott. 2006. Tertiary contacts distant from the active site prime a ribozyme for catalysis. Cell 126:309-320.
36. Pley, H. W., K. M. Flaherty, and D. B. McKay. 1994. Three-dimensional structure of a hammerhead ribozyme. Nature 372:68-74.
37. Scott, W. G., J. B. Murray, J. R. Arnold, B. L. Stoddard, and A. Klug. 1996. Capturing the structure of a catalytic RNA intermediate: the hammerhead ribozyme. Science (New York, N.Y 274:2065-2069.
38. Klein, D. J., and A. R. Ferre-D'Amare. 2006. Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. Science (New York, N.Y 313:1752-1756.
39. Rupert, P. B., A. P. Massey, S. T. Sigurdsson, and A. R. Ferre-D'Amare. 2002. Transition state stabilization by a catalytic RNA. Science (New York, N.Y 298:1421-1424.
40. Rupert, P. B., and A. R. Ferre-D'Amare. 2001. Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. Nature 410:780-786.
41. Ferre-D'Amare, A. R., and J. A. Doudna. 2000. Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module. Journal of molecular biology 295:541-556.
42. Milligan, J. F., D. R. Groebe, G. W. Witherell, and O. C. Uhlenbeck. 1987. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic acids research 15:8783-8798.
43. Pleiss, J. A., M. L. Derrick, and O. C. Uhlenbeck. 1998. T7 RNA polymerase produces 5' end heterogeneity during in vitro transcription from certain templates. RNA (New York, N.Y 4:1313-1317.
44. Cheong, H. K., E. Hwang, C. Lee, B. S. Choi, and C. Cheong. 2004. Rapid preparation of RNA samples for NMR spectroscopy and X-ray crystallography. Nucleic acids research 32:e84.
45. Lukavsky, P. J., and J. D. Puglisi. 2004. Large-scale preparation and purification of polyacrylamide-free RNA oligonucleotides. RNA (New York, N.Y 10:889-893.
46. Batey, R. T., and J. S. Kieft. 2007. Improved native affinity purification of RNA. RNA (New York, N.Y 13:1384-1389.
47. Kieft, J. S., and R. T. Batey. 2004. A general method for rapid and nondenaturing purification of RNAs. RNA (New York, N.Y 10:988-995.
48. Saville, B. J., and R. A. Collins. 1990. A site-specific self-cleavage reaction performed by a novel RNA in Neurospora mitochondria. Cell 61:685-696.
49. Kennell, J. C., B. J. Saville, S. Mohr, M. T. Kuiper, J. R. Sabourin, R. A. Collins, and A. M. Lambowitz. 1995. The VS catalytic RNA replicates by reverse transcription as a satellite of a retroplasmid. Genes Dev 9:294-303.
50. Rastogi, T., and R. A. Collins. 1998. Smaller, faster ribozymes reveal the catalytic core of Neurospora VS RNA. Journal of molecular biology 277:215-224.

51. Fedor, M. J., and J. R. Williamson. 2005. The catalytic diversity of RNAs. Nature Rev. Mol. Cell Biol. 6:399-412.
52. Collins, R. A. 2002. The Neurospora Varkud satellite ribozyme. Biochem Soc Trans 30:1122-1126.
53. Lilley, D. M. 2004. The Varkud satellite ribozyme. RNA (New York, N.Y 10:151-158.
54. Collins, R. A., and J. E. Olive. 1993. Reaction conditions and kinetics of self-cleavage of a ribozyme derived from Neurospora VS RNA. Biochemistry 32:2795-2799.
55. Jones, F. D., and S. A. Strobel. 2003. Ionization of a critical adenosine residue in the neurospora Varkud Satellite ribozyme active site. Biochemistry 42:4265-4276.
56. Rastogi, T., T. L. Beattie, J. E. Olive, and R. A. Collins. 1996. A long-range pseudoknot is required for activity of the Neurospora VS ribozyme. EMBO J. 15:2820-2825.
57. Michiels, P. J., C. H. Schouten, C. W. Hilbers, and H. A. Heus. 2000. Structure of the ribozyme substrate hairpin of Neurospora VS RNA: a close look at the cleavage site. RNA (New York, N.Y 6:1821-1832.
58. Hoffmann, B., G. T. Mitchell, P. Gendron, F. Major, A. A. Andersen, R. A. Collins, and P. Legault. 2003. NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site. Proc. Natl. Acad. Sci. USA 100:7003-7008.
59. Flinders, J., and T. Dieckmann. 2004. The solution structure of the VS ribozyme active site loop reveals a dynamic "hot-spot". Journal of molecular biology 341:935-949.
60. Lipfert, J., J. Ouellet, D. G. Norman, S. Doniach, and D. M. Lilley. 2008. The complete VS ribozyme in solution studied by small-angle X-ray scattering. Structure 16:1357-1367.
61. Lambert, M. N., E. Vocker, S. Blumberg, S. Redemann, A. Gajraj, J. C. Meiners, and N. G. Walter. 2006. Mg2+-induced compaction of single RNA molecules monitored by tethered particle microscopy. Biophysical journal 90:3672-3685.
62. Ke, A., and J. A. Doudna. 2004. Crystallization of RNA and RNA-protein complexes. Methods (San Diego, Calif 34:408-414.
63. Walter, N. G. 2002. Probing RNA structural dynamics and function by fluorescence resonance energy transfer (FRET). Curr. Protocols Nucleic Acid Chem. 11.10:11.10.11-11.10.23.
64. Andersen, A. A., and R. A. Collins. 2001. Intramolecular secondary structure rearrangement by the kissing interaction of the Neurospora VS ribozyme. Proc. Natl. Acad. Sci. USA 98:7730-7735.
65. Rasnik, I., S. A. McKinney, and T. Ha. 2006. Nonblinking and longlasting single-molecule fluorescence imaging. Nature Methods 3:891-893.
66. Haran, G. 2004. Noise reduction in single-molecule fluorescence trajectories of folding proteins. Chemical Physics 307:137-145.

67. McKinney, S. A., C. Joo, and T. Ha. 2006. Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophysical journal 91:1941-1951.
68. Hodak, J. H., C. D. Downey, J. L. Fiore, A. Pardi, and D. J. Nesbitt. 2005. Docking kinetics and equilibrium of a GAAA tetraloop-receptor motif probed by single-molecule FRET. Proc. Natl. Acad. Sci. USA 102:10505-10510.
69. Cosa, G., E. J. Harbron, Y. Zeng, H. W. Liu, D. B. O'Connor, C. Eta-Hosokawa, K. Musier-Forsyth, and P. F. Barbara. 2004. Secondary structure and secondary structure dynamics of DNA hairpins complexed with HIV-1 NC protein. Biophysical journal 87:2759-2767.
70. Sabanayagam, C. R., J. S. Eid, and A. Meller. 2005. Using fluorescence resonance energy transfer to measure distances along individual DNA molecules: corrections due to nonideal transfer. J Chem Phys 122:061103.
71. Guo, H. C., D. M. De Abreu, E. R. Tillier, B. J. Saville, J. E. Olive, and R. A. Collins. 1993. Nucleotide sequence requirements for self-cleavage of Neurospora VS RNA. Journal of molecular biology 232:351-361.
72. Lafontaine, D. A., T. J. Wilson, Z. Y. Zhao, and D. M. Lilley. 2002. Functional group requirements in the probable active site of the VS ribozyme. Journal of molecular biology 323:23-34.
73. Joo, C., S. A. McKinney, M. Nakamura, I. Rasnik, S. Myong, and T. Ha. 2006. Real-time observation of RecA filament dynamics with single monomer resolution. Cell 126:515-527.
74. Tan, E., T. J. Wilson, M. K. Nahas, R. M. Clegg, D. M. Lilley, and T. Ha. 2003. A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. Proc. Natl. Acad. Sci. USA 100:9308-9313.
75. Wilson, T. J., A. C. McLeod, and D. M. Lilley. 2007. A guanine nucleobase important for catalysis by the VS ribozyme. EMBO J. 26:2489-2500.
76. Campbell, D. O., and P. Legault. 2005. Nuclear magnetic resonance structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping. Biochemistry 44:4157-4170.
77. Russell, R., X. Zhuang, H. P. Babcock, I. S. Millett, S. Doniach, S. Chu, and D. Herschlag. 2002. Exploring the folding landscape of a structured RNA. Proc. Natl. Acad. Sci. USA 99:155-160.
78. Pyle, A. M., O. Fedorova, and C. Waldsich. 2007. Folding of group II introns: a model system for large, multidomain RNAs? Trends Biochem Sci 32:138-145.
79. Bartley, L. E., X. Zhuang, R. Das, S. Chu, and D. Herschlag. 2003. Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. Journal of molecular biology 328:1011-1026.
80. Nahas, M. K., T. J. Wilson, S. Hohng, K. Jarvie, D. M. Lilley, and T. Ha. 2004. Observation of internal cleavage and ligation reactions of a ribozyme. Nature structural & molecular biology 11:1107-1113.

81. Zamel, R., A. Poon, D. Jaikaran, A. Andersen, J. Olive, D. De Abreu, and R. A. Collins. 2004. Exceptionally fast self-cleavage by a Neurospora Varkud satellite ribozyme. Proc. Natl. Acad. Sci. USA 101:1467-1472.
82. Collins, R. A., and B. J. Saville. 1990. Independent transfer of mitochondrial chromosomes and plasmids during unstable vegetative fusion in Neurospora. Nature 345:177-179.
83. Poon, A. H., J. E. Olive, M. McLaren, and R. A. Collins. 2006. Identification of separate structural features that affect rate and cation concentration dependence of self-cleavage by the Neurospora VS ribozyme. Biochemistry 45:13394-13400.
84. Aitken, C. E., R. A. Marshall, and J. D. Puglisi. 2008. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. Biophysical journal 94:1826-1835.
85. Draper, D. E., D. Grilley, and A. M. Soto. 2005. Ions and RNA folding. Annual review of biophysics and biomolecular structure 34:221-243.
86. Heilman-Miller, S. L., D. Thirumalai, and S. A. Woodson. 2001. Role of counterion condensation in folding of the Tetrahymena ribozyme. I. Equilibrium stabilization by cations. Journal of molecular biology 306:1157-1166.
87. Uhlenbeck, O. C. 1995. Keeping RNA happy. RNA (New York, N.Y 1:4-6.
88. Doudna, J. A. 1997. Preparation of homogeneous ribozyme RNA for crystallization. Methods in molecular biology (Clifton, N.J 74:365-370.
89. Wong, T. N., T. R. Sosnick, and T. Pan. 2007. Folding of noncoding RNAs during transcription facilitated by pausing-induced nonnative structures. Proceedings of the National Academy of Sciences of the United States of America 104:17995-18000.
90. Winkler, W. C., A. Nahvi, A. Roth, J. A. Collins, and R. R. Breaker. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. Nature 428:281-286.
91. Taylor, J. M. 2006. Structure and replication of hepatitis delta virus RNA. Current topics in microbiology and immunology 307:1-23.
92. Oubridge, C., N. Ito, C. H. Teo, I. Fearnley, and K. Nagai. 1995. Crystallisation of RNA-protein complexes. II. The application of protein engineering for crystallisation of the U1A protein-RNA complex. Journal of molecular biology 249:409-423.
93. Gonzalez, M., C. E. Argarana, and G. D. Fidelio. 1999. Extremely high thermal stability of streptavidin and avidin upon biotin binding. Biomolecular engineering 16:67-72.
94. Holmberg, A., A. Blomstergren, O. Nord, M. Lukacs, J. Lundeberg, and M. Uhlen. 2005. The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. Electrophoresis 26:501-510.
95. Bornhorst, J. A., and J. J. Falke. 2000. Purification of proteins using polyhistidine affinity tags. Methods in enzymology 326:245-254.
96. Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods in enzymology 185:60-89.

97. Greenleaf, W. J., K. L. Frieda, D. A. Foster, M. T. Woodside, and S. M. Block. 2008. Direct observation of hierarchical folding in single riboswitch aptamers. Science (New York, N.Y 319:630-633.
98. Hampel, K. J., and M. M. Tinsley. 2006. Evidence for preorganization of the glmS ribozyme ligand binding pocket. Biochemistry 45:7861-7871.
99. Tinsley, R. A., J. R. Furchak, and N. G. Walter. 2007. Trans-acting glmS catalytic riboswitch: locked and loaded. RNA (New York, N.Y 13:468-477.
100. Toor, N., K. S. Keating, S. D. Taylor, and A. M. Pyle. 2008. Crystal structure of a self-spliced group II intron. Science (New York, N.Y 320:77-82.