UNIVERSITY OF CALIFORNIA,
IRVINE

The influence of peripheral elements on the kinetics of HDV-like ribozymes

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Chemistry
with a concentration in Chemical Biology

By

Marie Ann-Juanita Myszka

Thesis Committee:
Professor Andrej Lupták, Chair
Professor Andy Borovik
Professor Robert Corn

2014

DEDICATION

To my family, the Matt, & Toby too

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my thanks and appreciation to my advisor and chair, Professor Andrej Lupták, for helping me decipher my spoken gibberish into scientific thought.

I would like to thank my committee members, Professor M. Andy Borovik and Professor Robert Corn, for allowing me to wrangle them into reading my summer sleeper. I tried not to eat all of your group's office snacks Rob and I claim fame to using that massive silica-column during my rotation with the Borovik lab.

I wanna thank everyone in the Lupták group for making lab an enjoyable place to be, thank you for putting up with my stream of conscious word vomit and weird antics (you know you enjoy it ;)).

ABSTRACT OF THE THESIS

The influence of peripheral elements on the kinetics of HDV-like ribozymes

By

Marie Ann-Juanita Myszka

Master of Science in Chemistry with a concentration in Chemical Biology

University of California, Irvine, 2014

Professor Andrej Lupták, Chair

Investigating the influence of peripheral domains on the kinetic rates of HDV-like ribozyme self-cleavage is important to discovering their biological relevance. The following work investigates the effects of peripheral elements on the self-scission of two HDV-like ribozymes, the mouse *CPEB3* ribozyme and the mosquito drz-Agam-2-1.

The *CPEB3* ribozyme, located in an intron of the mammalian *CPEB3* gene, is the first HDV-like ribozyme discovered in a genome other than the virus. ESTs from human cells suggest ribozyme activity *in vivo* and demonstrate tissue-specific expression[1]; the level of activity in human brain tissue suggests another factor, such as a chaperone, increases ribozyme activity by minimizing misfolding[2]. The rates of the co-transcriptional self-scission of the *CPEB3* ribozyme were determined using sequences varying by the number of nucleotides surrounding the ribozyme. Experimental rates range from 0.13 ± 0.029 to 0.46 ± 0.020 min^{-1} with the order of the constructs being -233/72, -197/72, -197/72/21, -233/72, -197/72/165, -233/72/165, -10/72, -49/72/165 and -257/72/328. The findings here demonstrate that *CPEB3*

ribozymes are kinetically fast-reacting and suggest a relationship between flanking sequences and rate.

The drz-Agam-2-1 ribozyme was the first naturally occurring ribozyme to contain a helix in the J1/2 domain[3], a domain shown to indirectly effect ribozyme self-cleaving[4]**.** The rate of self-cleavage increased with step-wise reintroduction of the helix into drz-Agam-2-1. Experiments correlated the stability derived from J1/2 to catalysis and provide an explanation for the ribozyme's Hill coefficient of n=1. The data shows the J1/2 structured element to impact ribozyme catalysis through stabilization of the catalytic core.

Chapter 1

General background

I. Small self-cleaving ribozymes

Ribonucleic acids (RNAs) demonstrate a range of biological functionality, including regulation[5] and catalysis[6,7], and these abilities support the "RNA world" hypothesis[8,9]. RNAs that catalyze chemical transformations are referred to as ribozymes or catalytic RNAs. The hammerhead[10], hairpin[11], hepatitis delta virus[12] (HDV), Varkud satellite[13] (VS), and the bacterial g*lmS*[5] ribozymes are a group of small, self-cleaving catalytic RNAs[14]. These small ribozymes catalyze the same transesterification reaction, but each ribozyme has a unique structure that results in site-specific catalysis[15,16]. Self-cleavage occurs by breaking the phosphodiester backbone of the RNA *via* nucleophilic attack of a 2′ hydroxyl on the neighboring 5′ phosphodiester, resulting in a 2′, 3′-cyclic phosphate and 5′-hydroxyl termini[17] (Figure 1-1).

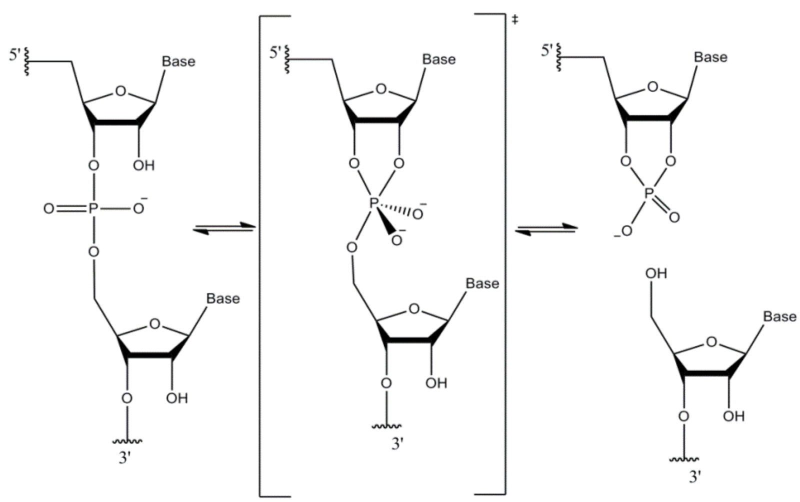


Figure 1-1: Ribozyme self-cleavage by nucleophilic attack of 2' hydroxyl on adjacent 5' phosphodiester, 2', 3'-cyclic phosphate and 5′-hydroxyl products shown.

Initially, ribozymes were only found in viroid satellite DNA; however, sequence-based searches recently revealed ribozymes in organisms of other phyla[18]. The underrepresentation of ribozymes in humans led to the search and discovery of four ribozymes in the human genome, including the cytoplasmic polyadenylation element-binding protein 3 (*CPEB3*) ribozyme[1]. Despite this, the biological roles and regulation of catalytic RNAs in higher eukaryotic species remains poorly defined[19].

II. HDV ribozyme

The hepatitis delta virus (HDV) is comprised of a single-stranded circular genome 1700 nt long[20], it was first discovered in human liver cells[21]. HDV carries out replication through a rolling-circle fashion[22] and contains an active ribozyme in both its genomic and antigenomic form[23]. The secondary structure of the HDV ribozyme consists of five paired helices: P1, P2, P3, P1.1 and P4 (Figure1-2). These regions stack to form two co-axial stacks, P1-P1.1-P4 and P2-P3 (Figure 1-2A) and are linked by joining strands J1/2 and J4/2[24]. The crystal structure of the cleaved genomic HDV ribozyme folds into a nested double pseudoknot, positioning the active- site cytosine (C75) near the site of cleavage and creating the catalytic core[25].

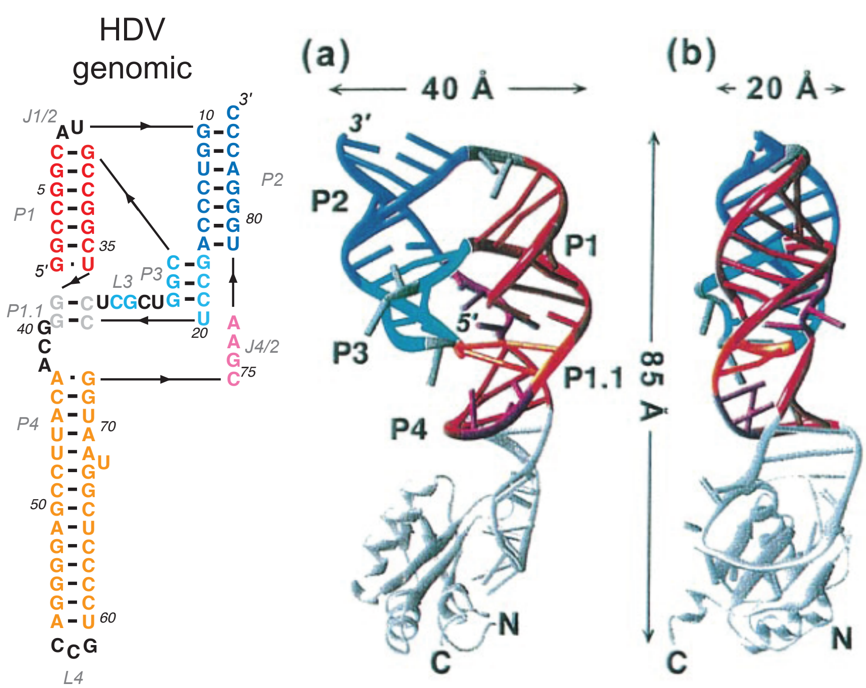


Figure 1-2: (Left) The genomic secondary structure of the HDV ribozyme. The crystal structure of the genomic HDV visualizing the co-axial stacking of the paired regions (A) and the crystal structure in profile[24].

III. *CPEB3* ribozyme

The CPEB3 protein itself can affect synaptic plasticity by regulating the translation of plasticity-related proteins (PRPs) for the long-term potentiation of neurons[26]. Local translation of mRNA at dendritic sites modulates the synaptic plasticity in the hippocampus[27,28]. The CPEB3 protein is a sequence-specific RNA-binding protein and represses translation of its target mRNA[29]. The ability of CPEB3 to regulate transcription of EGFR, a protein shown to modulate memory[30], along with the increased levels of CPEB3 mRNA in hippocampal tissue after kainate-induced seizures demonstrate a neuronal role of CPEB3[31].

The characterization of the human *CPEB3* ribozyme discovered through *in vitro* selection established that it is an HDV-like ribozyme[1]. The *CPEB3* ribozyme is located in the 2nd intron of the *CPEB3* gene, between the 2nd and 3rd exon, and is conserved throughout all mammalian species studied (Figure 1-3).

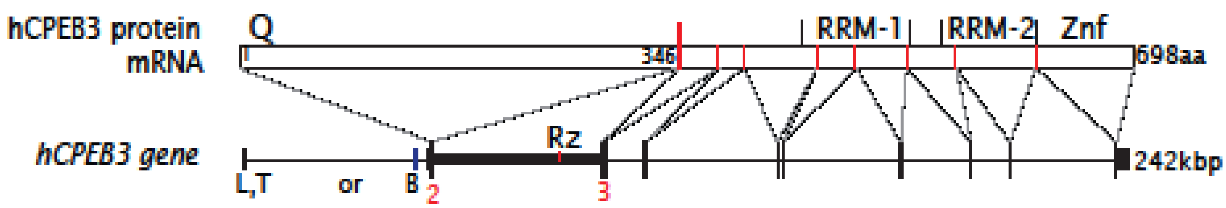


Figure 1-3: The human mRNA of the CPEB3 protein (top). The *CPEB3* gene displaying where exons composing the mRNA map onto the gene, Rz denotes the ribozyme location (bottom). The ribozyme sequence is conserved in all mammalian species studied and is located in the second intron between the 2nd and 3rd exon labeled 2 and 3, respectively; the size of the 2nd intron differs for each species[1].

The ribozyme was minimized to the region of highest conservation, yielding an 81 nt sequence capable of promoting self-cleavage, that results in a 9 nucleotide upstream fragment and a 72 nucleotide downstream fragment. The rate-limiting step of the *CPEB3* ribozyme self-scission was determined to involve at least one proton transfer by the constant kinetic pH profile and solvent isotope effect of the cleavage[1]. The ribozyme demonstrated a dependence on divalent metal ions for cleavage and data suggested that a hydrated divalent metal ion was required[1]. The hammerhead, hairpin, and VS small ribozymes cleave in the presence of either divalent metal ions or a high concentration of monovalent cations, however this is not true for the HDV ribozyme which cleaves only in the presence of divalent metal ions[32]. Since the *CPEB3* and HDV ribozymes both required the presence of divalent metal ions for catalytic activity, the structural similarities between the two were explored. Although the HDV and *CPEB3* ribozyme differ in sequence, *CPEB3* folds into a nested double pseudoknot forming the same secondary and tertiary structure as the HDV ribozymes[1] (Figure 1-4).

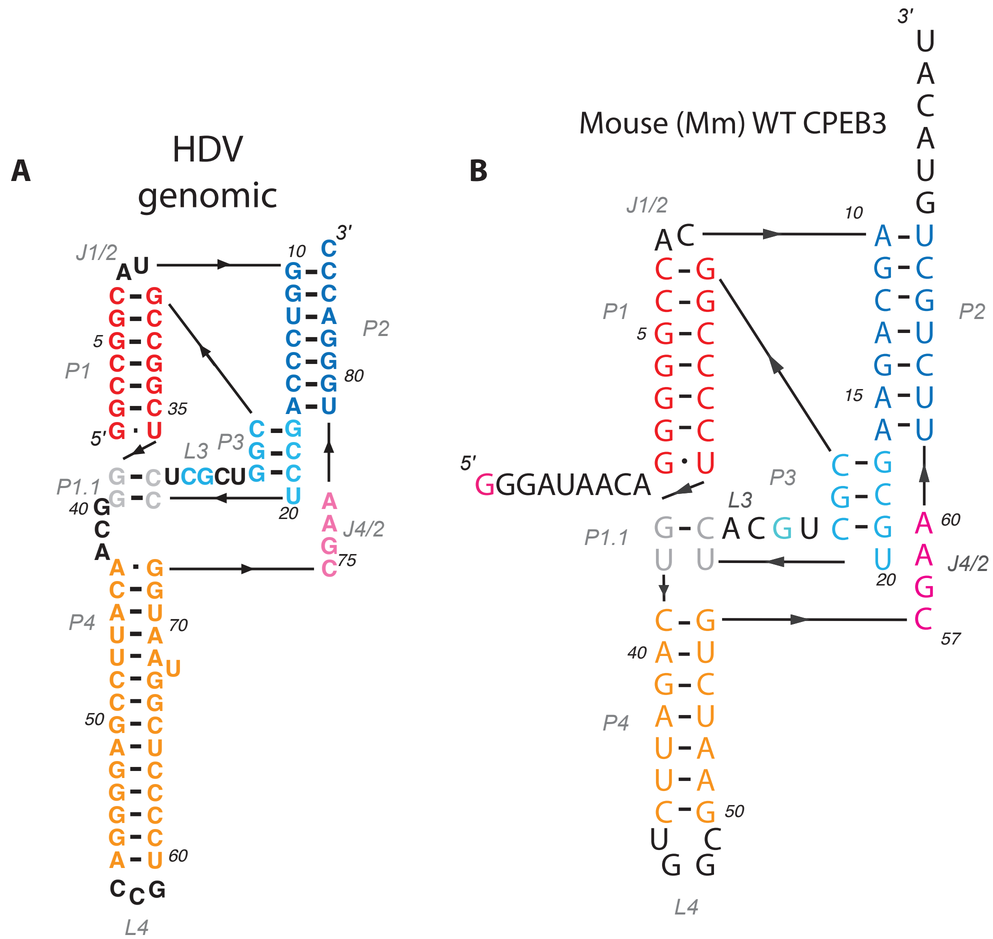


Figure 1-4: The genomic secondary structure of the self-cleaved HDV (A) and human *CPEB3* (B) ribozymes. The *CPEB3* active- site cytosine (C57) is in the same spatial position as that of HDV (C75).

IV. Drz-Agam-2-1

The drz-Agam-2-1 ribozyme was discovered through structured-based searches in the African mosquito *A. gambiae* and given the drz connotation to identify it as a HDV-like ribozyme[3]. It is the first naturally occurring HDV-like ribozyme to have a base-paired region in J1/2, also known as P1.2. The predicted secondary structure of the ribozyme in Figure 1-5 is similar to that of the HDV and *CPEB3* ribozymes.

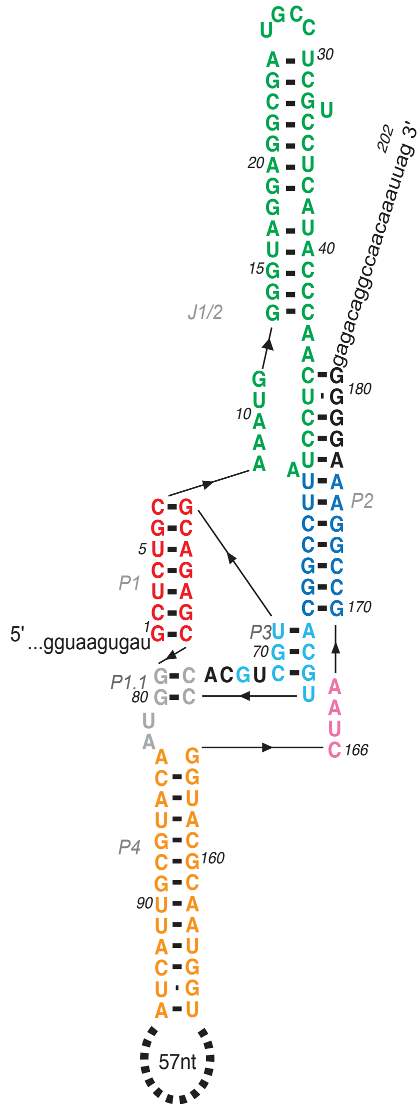


Figure 1-5: The drz-Agam-2-1 ribozyme secondary structure, similar to both HDV and *CPEB3* ribozymes. The leader sequence at the 5′ position of the structure is abbreviated in the figure. The defining feature of this ribozyme is the base-paired J1/2, located between the base-paired regions P1 and P2, which until its discovery was only a facet in engineered *trans*-acting ribozymes known as P1.2[3].

V. Objective of thesis

Expressed sequence tags (ESTs) from human liver and murine tissue suggest an active *CPEB3* ribozyme *in vivo*. However, the observed *in vitro* rate for a 209 nucleotide construct from the human genome (0.01 min^{-1}) did not support these findings. The *CPEB3* ribozyme is located roughly ten to fourteen thousand base pairs (bp) from the third exon[1] in a region that would take the polymerase (without pausing) about two and a half to four minutes to reach the third exon, and another five to ten minutes to splice the exons together[33]. Based on the rate and the secondary structure of the human *CPEB3* ribozyme it was hypothesized that the ribozyme was selectively hindered to allow normal splicing to occur a majority of the time[1], since splicing of pre-mRNA does not occur if an intron is cleaved by a fast ribozyme[34]. However, the ESTs gathered showed the efficiency of ribozyme cleavage occurred close to 50 % in human brain tissue[1], and this level of activity *in vivo* is not possible with the above mentioned *in vitro* rate. Previous studies of the HDV ribozyme have shown that flanking sequences beyond a determined length inhibit ribozyme activity by increased misfolding into inactive conformations or creating kinetic traps[35-37]. With the potential of flanking regions to act as kinetic traps, studies involving human *CPEB3* constructs with shortened flanking sequences were able to reveal a kinetically fast-reacting ribozyme (0.33-0.55 min^{-1})[2].

This result brings about the question of how nature unburdens *CPEB3* of its long flanking regions to increase catalytic rate. Group I and group II introns are large, self-splicing catalytic RNAs that require binding proteins, termed RNA chaperones, for activity[38]. RNA chaperones are proteins that associate with RNA molecules thereby destabilizing the RNA to avoid kinetic traps and forcing the RNA to fold into an active form[39,40]. The binding of an RNA chaperone to the flanking sequences of *CPEB3* could introduce cellular regulation by increasing ribozyme activity. If the *CPEB3* ribozyme does have a protein chaperone, it will be the first mammalian example[41,42].

The first objective of this thesis was to establish co-transcriptional kinetic rates of the mouse *CPEB3* with varying numbers of nucleotides bordering the ribozyme core. These rates can then be compared to later kinetics that will introduce the CPEB3 protein and determine if the protein demonstrates any RNA chaperone activity to infer its *in vivo* role. The co-transcriptional kinetics required optimization of the technique, after that was accomplished the average rate for several constructs were determined which revealed noticeable trends pertaining to flanking nucleotide length.

The second objective was to evaluate the role of the J1/2 peripheral element (P1.2) in the kinetics of drz-Agam-2-1. The drz-Agam-2-1 contains a helix in the J1/2 joining strand, the J1/2 was deleted from the ribozyme and reintroduced step-wise through base-pair addition. Kinetic rates for the constructs were established, showing increased activity with increased J1/2 structure. The manner in which the stability of the J1/2 affected the kinetic rates was monitored by changes in the activation energy versus the minimum free energy (MFE) of the ribozyme. A linear trend is emerging but more constructs are required to make a statement. The Mg^{2+} dependency of drz-Agam-2-1 and HDV constructs containing P1.2 generated Hill coefficients of n=1, but there was no difference between that and the wildtype's Hill coefficient for the HDV constructs. The presence of P1.2 in drz-Agam-2-1 and Hill coefficient of 1 supports the idea that the P1.2 is taking the place of the structural Mg^{2+} required for folding.

Understanding the impacts of neighboring sequences and structures on the ribozyme core and the resulting effects to activity can aide in recognizing environments that might utilize these functions/responses. Results presented demonstrate the changes in self-cleavage brought on by interactions with elements bordering the catalytic core of the ribozyme.

Chapter 2

Probing flanking nucleotide effects on the co-transcriptional kinetics of the *CPEB3* ribozyme

I. Introduction

To test and further explore the effects that the size of the flanking genomic sequences have on the *CPEB3* ribozyme, kinetic experiments were conducted using a collection of constructs* with varying sequences surrounding the minimum mouse *CPEB3* ribozyme. The mouse (Mm) *CPEB3* and human (Hs) *CPEB3* sequences differ at only two nucleotide positions and otherwise have identical secondary structures[1]. All kinetic studies on the self-cleavage of Mm *CPEB3* were done co-transcriptionally to avoid purification bias and more closely mimic the cleavage of the 2nd intron during *in vivo* transcription of Mm *CPEB3*. Optimization of co-transcriptional kinetics was a primary focus, with emphasis placed on ensuring transcriptional termination and reducing unwanted cleavage. Once scission has been induced, the transcriptional product must remain relatively constant for the remainder of the experiment for comparative purposes. Additionally, during transcription, minimal to no cleavage is necessary before initiating self-cleavage in order to gather an accurate portrayal of the ribozyme's rate of cleavage. Cleavage of the *CPEB3* ribozyme is triggered by the presence of

* Naming of constructs adopted from Chadalavada et al. (2010) as –*a*/72/*b*, where *a* is the number of nucleotides

Mg^{2+}; however, Mg^{2+} must also be present for transcription and any excess Mg^{2+} may catalyze self-cleavage of the newly transcribed product prior to the desired start time.

II. Co-transcriptional kinetics and optimization with Mm *CPEB3* constructs

Each of eight constructs maps to the Mm *CPEB3* gene (Figure 2-1) and cleaves the upstream flanking sequence from the ribozyme and downstream flanking sequence.

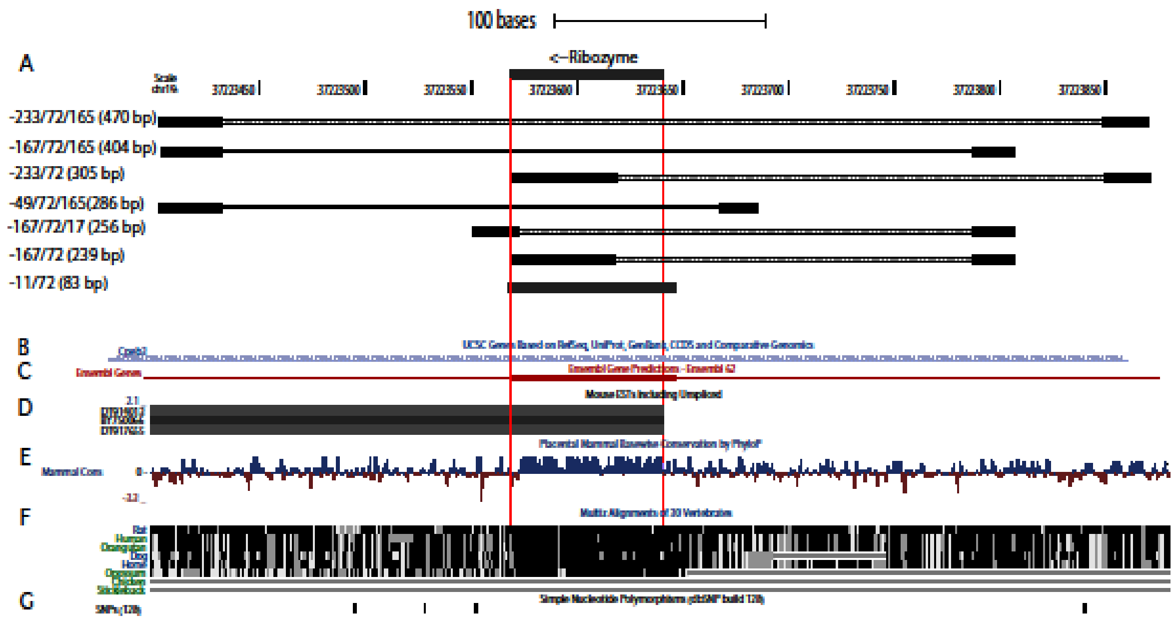


Figure 2-1: (A) *CPEB3* constructs mapped to the *CPEB3* gene with red, vertical lines highlighting the area of the ribozyme. Solid, black blocks on constructs denote position of the primers used for amplification of DNA. Construct -257/72/328 is not featured since it is outside the scale of the other constructs. (B) Shows the *CPEB3* gene relative to the constructs, (C) shows predicted gene (ribozyme is highlighted by red block), (D) are the expressed sequence tags (ESTs) of mouse *CPEB3*aligned to the gene, (E) shows the relative conservation of the *CPEB3* gene across sequences gathered from mammalian genomes, (F) shows the gene sequence alignment of eight mammals, and (G) are know the single nucleotide polymorphisms to occur for this gene segment.

Time points are taken from the transcription and quenched using an urea/EDTA stopping solution that both denatures the RNA and chelates the Mg^{2+} to eliminate further cleavage. A denaturing PAGE is then used to separate the 5′ and 3′ cleavage products of each time point and the relative fraction of cleaved and uncleaved product are quantified with the use of a phosporimager. Using the least squares method, the fraction of uncleaved product (F_{un}) at each time point (t) is modeled to both a mono-exponential (i) and bi-exponential decay (ii) function:

$$F_{un} = Ae^{-k_1 t} + C \qquad \textbf{(i)}$$
$$F_{un} = Ae^{-k_1 t} + Be^{-k_2 t} + C \qquad \textbf{(ii)}$$

where C is the fraction of uncleaved product at the end of the experiment, k_1 and k_2 are the rate constants, A and B are the amplitudes of the rate constants, respectively, and t is the time of the data point. Data are fit both mono- and bi-exponentially to determine which model better represents the Mm *CPEB3* cleavage. If all data fit closely to the mono-exponential function and no discernible differences between the mono- and bi-exponential fit are apparent, then the single rate equation appropriately describes the ribozyme[43]. The data from each construct were fit to equation (i) and (ii) with the resulting parameters and half-life displayed in Table 2-1 and any models with unrealistic fits were omitted. Results are from a multiple experiments with the average kinetic parameters listed in Table 2-1.

Table 2-1. Avgerage Mono- and Bi-Exponential Parameters of Co-Transcriptional Kinetics

Construct Name	**Mono-exponential**			**Bi-exponential**					**Half-life (min)**
	A	**k_1 (min^{-1})**	**C**	**A**	**k_1 (min^{-1})**	**B**	**k_2 (min^{-1})**	**C**	
-49/72/165				**0.88** ± 0.023	**0.42** ± 0.035	**0.013** ± 0.015	**0.11** ± 0.030	**0.039** ± 0.019	1.7
-10/72	**0.72** ± 0.090	**0.39** ± 0.091	**0.082** ± 0.026						1.8
-197/72				**0.54** ± 0.075	**0.13** ± 0.029	**0.21** ± 0.034	**0.25** ± 0.024	**0.013** ± 0.013	2.8
-197/72/21				**0.54** ± 0.13	**0.23** ± 0.11	**0.12** ± 0.0076	**0.026** ± 0.0081	**0.11** ± 0.024	3.0
-197/72/165				**0.80** ±	**0.32** ±	**0.013**±	**0.092**±	**0.036**±	2.2
-257/72/328				0.62	0.46	0.03	0.15	0.017	1.5
-233/72/165				**0.78** ± 0.044	**0.31** ± 0.044	**0.035** ± 0.0057	**0.17** ± 0.023	**0.029** ± 0.0054	2.2
-233/72				**0.62** ± 0.090	**0.16** ± 0.0020	**0.18** ± 0.12	**0.13** ± 0.11	**0.0051** ± 0.0062	4.3

Mono- and bi-exponential fits of each construct's co-transcription. Both k_1 and k_2 are rate constants; A and B are the amplitudes of k_1 and $k_{2,}$ respectively; and C is the fraction of uncleaved product at the end of the experiment. Half-life was calculated using the fastest rate. Each parameter is the average of three or more fits (except -197/72/21 which is the average of two).

Based on findings, all constructs react similarly to the human -8/68 (WT) *CPEB3* ribozyme (0.33 min^{-1}), making them fast-reacting[2]. Comparisons between the kinetic rates of each construct can be visualized in Figure 2-2, with the region kinetically relevant to the co-transcriptional splicing of *CPEB3* highlighted by gray vertical lines. Within the time it takes the polymerase to transcribe the region between the *CPEB3* ribozyme and the 3rd exon and splice the exons together (~seven to fourteen minutes)[33], the longest construct (-257/72/328), which is the most representative of the *in vivo* transcription product, would be completely cleaved with a half-life of 1.5 minutes. However, eliminating the downstream flanking sequence decreases the rate of the ribozyme by about three-fold (-257/72/328 vs. -233/72).

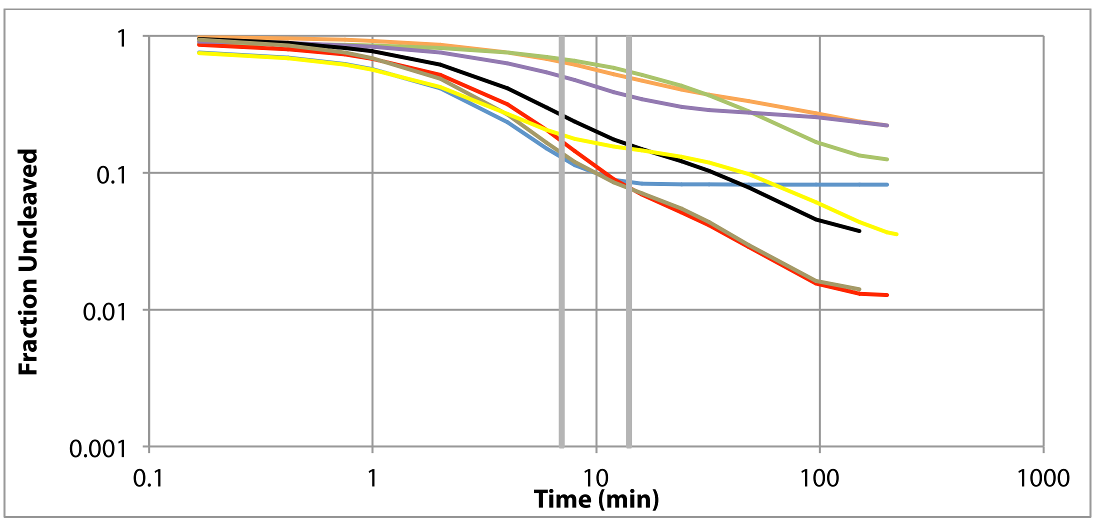


Figure 2-2: The average kinetic model of each construct is shown overlaid. The lines of corresponding color are the exponential fits. Shown are: color blue (-10/72 Co-transcription mono), color green (-195/72/21 Co-transcription, mono), color black (-233/72/165 Co-transcription, bi-exp), color orange (-195/72 Co-transcription, bi-exp), color red (-195/72/165 Co-transcription, bi-exp), color purple (-233/72 Co-transcription, bi-exp), color yellow (-257/72/328 Co-transcription, bi-exp), color brown (-49/72/165 Co-transcription, bi-exp). Highlighted vertically across the graph in gray is the time (~7-14 minutes) it takes the polymerase to reach the 3rd exon and splice the 2nd and 3rd exon together. Individual models in supplemental.

Among the constructs, it appears that constructs with short to no downstream flanking sequences react slower as seen by -233/72 vs. -233/72/165 and that increasing the upstream sequence has little effect, -197/72/165 vs. -233/72/165. Both -233/72 and -233/72/165 start similarly but quickly diverge, emphasizing the potential role of the downstream flanking sequence in ribozyme folding and, result in the increased activity of -233/72/165. This suggests that the downstream flanking sequence provides little to no alternative folding conformations, and could in fact stabilize the RNA of -49/72/165 and -233/72/165 by stabilizing the active fold, thus increasing the kinetic rate. The similar kinetic rates for -10/72 and -49/72/165 are further emphasized by the fact that the average kinetic models (Figure 2-2) overlap each other. Similar characterization of ribozyme activity was shown with reproduced experiments and presents a clear correlation between the ribozyme's activity and flanking sequences.

To determine the efficacy of optimization, kinetic data before and after optimization of co-transcriptional kinetics were compared (Table 2-2). Results from optimization revealed that construct -49/72/165 is indeed fast and that the reverse is true for -233/72.

Table 2-2. Comparison of Kinetics Before and After the Optimization of Co-Transcription

Construct Name	Mono-Exponential 6.75 mM [Mg] (Before)	Mono-Exponential 7.75 mM [Mg] (Before)	Mono-/Bi-Exponential (After)
	k_1 (min^{-1})	k_1 (min^{-1})	k_1 (min^{-1})
-49/72/165	0.29	0.16	0.42/0.46
-11/72	0.22	0.40	0.40
-233/72/165			0.19/0.24
-233/72	0.29	0.29	0.14

Comparison of the kinetic rates of Mm *CPEB3* constructs before (Before) optimization to the rates after optimization (After).

The concentration of Mg^{2+} used for transcription should be equimolar to the amount of ribonucleoside triphosphates (rNTPs) used. Initially constructs were co-transcribed at two different Mg^{2+} concentrations, the equimolar concentration of 7.75 mM, and a concentration slightly lower, 6.75 mM, to determine which would promote transcription without inducing self-scission; however, the actual concentration of rNTPs was never measured so both concentrations could be in excess. The *CPEB3* ribozyme reacts readily during transcription with excess Mg^{2+} and activity increases with temperature, therefore transcriptions were carried out below physiological temperature, at 25 °C instead of 37 °C, to reduce the chance of unwanted self-scission. Transcription involved either a five or ten µL sample that was incubated at 25 °C for five minutes. The transcription was then combined with a 95 or 90 µL solution, containing both physiological buffer and the Mg^{2+} to induce cleavage, and the resulting 100 µL reaction was heated to 37 °C. The mixture was held at 37 °C and aliquots of the reaction were taken at specified time points. When executing these experiments, the physiological buffer was added to the container holding the transcription at 25 °C, and then the container was heated and held at 37 °C. By performing the experiment in this order, the contents would be equilibrating to 37 °C during initial time points, misrepresenting the ribozyme's early behavior *in vitro* by possibly giving slower rates at a lower temperature. Additionally, experiments were only run for a maximum of 18 minutes, so the fraction of uncleaved product at the end of the experiment was skewed toward the high end in the models. By correcting these aspects of the co-transcription experiments, results should better represent *in vivo* conditions, and provide a better fit to the exponential model

III. Conclusions

Rates for the Mm *CPEB3* ribozyme range from 0.13-0.46 min^{-1}, and the primary determinate of the activity is the length of the flanking sequences surrounding the ribozyme. The hypothesis that a protein with RNA chaperone activity could modulate the rate of ribozyme cleavage of *CPEB3* by binding to the flanking sequences has yet to be studied. With the established kinetic rates, future directions are to introduce the CPEB3 protein at varying concentrations during transcription to determine if the protein has any inherent consequences on the rate of cleavage of Mm *CPEB3*. The CPEB3 protein can shuttle into the nucleus from the cytoplasm[30] making the simultaneous presence of both the CPEB3 protein and intron biologically relevant. If CPEB3 is an RNA chaperone for the *CPEB3* ribozyme the -233/72 and -233/72/165 constructs should show modulated kinetic activity corresponding to significantly faster or slower *in vitro* rates. The biological significance of ribozymes in higher eukaryotic species remains undetermined, but the high conservation of the *CPEB3* ribozyme in mammals suggests that it plays an important biological role. By probing the relationship between precursor and product, hopefully the biological relevance of mammalian *CPEB3* ribozymes can be elucidated.

Chapter 3

Investigating the influence of the J1/2 region on the kinetics of *HDV*-like ribozymes

I. Introduction

The HDV ribozyme consists of five base-paired regions that stack into two co-axial helices (P1 stacked with P1.1 and P4, P2 stacked with P3) and single-stranded linkers, J1/2 and J4/2, connect the based-paired regions[25]. These base-paired regions and the J4/2 strand are involved in the compact folding of the ribozyme core and contribute to catalysis[44,45]. Other elements peripheral to the ribozyme structure (J1/2, L4, and the 5′ leader and 3′ tail sequences) do not directly contribute to catalysis but their ability to stabilize/interact with the ribozyme core can affect the kinetic rate of the HDV ribozyme[4,46]. The antigenomic and genomic HDVs are kinetically fast but their trans-acting counterparts are slower in activity, lacking both the L4 and J1/2 joining strands which result in 10-100 fold lower activity[47]. The influence of J1/2 on activity is in contrast to earlier notions that it only joined the two helices together and had no substantial affect on kinetic activity[48]. The presence of the J1/2 results in faster self-scission by being able to bring the distal ends of the stacked P1-P1.1-P4 and stacked P2-P3 closer together and create faster folding and/or minimize the energy necessary for catalysis. A cis-ribozyme with a structured J1/2, drz-Agam-2-1, was discovered through structure-based searches[3] and prompted the exploration of the role a structured J1/2 played in the catalysis of HDV-like ribozymes.

II. Construction and kinetic rate determination of drz-Agam-2-1 ribozymes

Completing the kinetic picture

Constructs of drz-Agam-2-1 with varying lengths of the structured J1/2 domain, also referred to as P1.2, were previously designed to obtain kinetics of purified constructs. Rates of purified constructs involve the gel purification of the ribozyme RNA from the transcription mixture, the purified RNA is then introduced to a physiological buffer and held at 37 °C wherein self-scission is induced with the addition of Mg^{2+} and monitored. However the constructs did not show a consistent linear progression in cleavage rate leading up to the WT fully intact J1/2 helix and this required additional constructs to be designed and tested. In addition to getting rates of the new purified constructs at the previously established 1 mM Mg^{2+} concentration (rates in black), rates were also obtained at 300 μM Mg^{2+} (rates in green) and rates under co-transcriptional (immediately after transcription) conditions at 1 mM Mg^{2+} (rates in red) (Figure 3-1). The established constructs and rates of the purified constructs at 1 mM Mg^{2+} by Dr. Judy Webb are denoted with an asterisk and present a trend of increased rate of self-cleavage with elongation of the J1/2 (Figure 3-1A, B, D, G, H). The remaining three constructs align themselves with the well-established rate pattern with the exception of drz-Agam-2-1-9.1 (Figure 3-1C) showing an elevated rate relative to the J1/2 length. These later constructs stay consistent with the earlier constructs' design which lengthens the J1/2 to the naturally occurring structure but does the addition of each base-pair in a way that varies the loop among all the constructs. The combination of the varied loop and subsequent base-pair added to the J/12 of drz-Agam-2-1-9.1 could provide an alternative folding that promotes faster catalysis compared to its neighboring constructs, resulting in an increased rate. The depictions of these ribozymes are putative, based on the known secondary structure of the

HDV ribozyme[25] and predicted computational folding (DotKnot)[3]. The actual tertiary contacts of these constructs could deviate from expected folding patterns due to unpredicted interactions/folding.

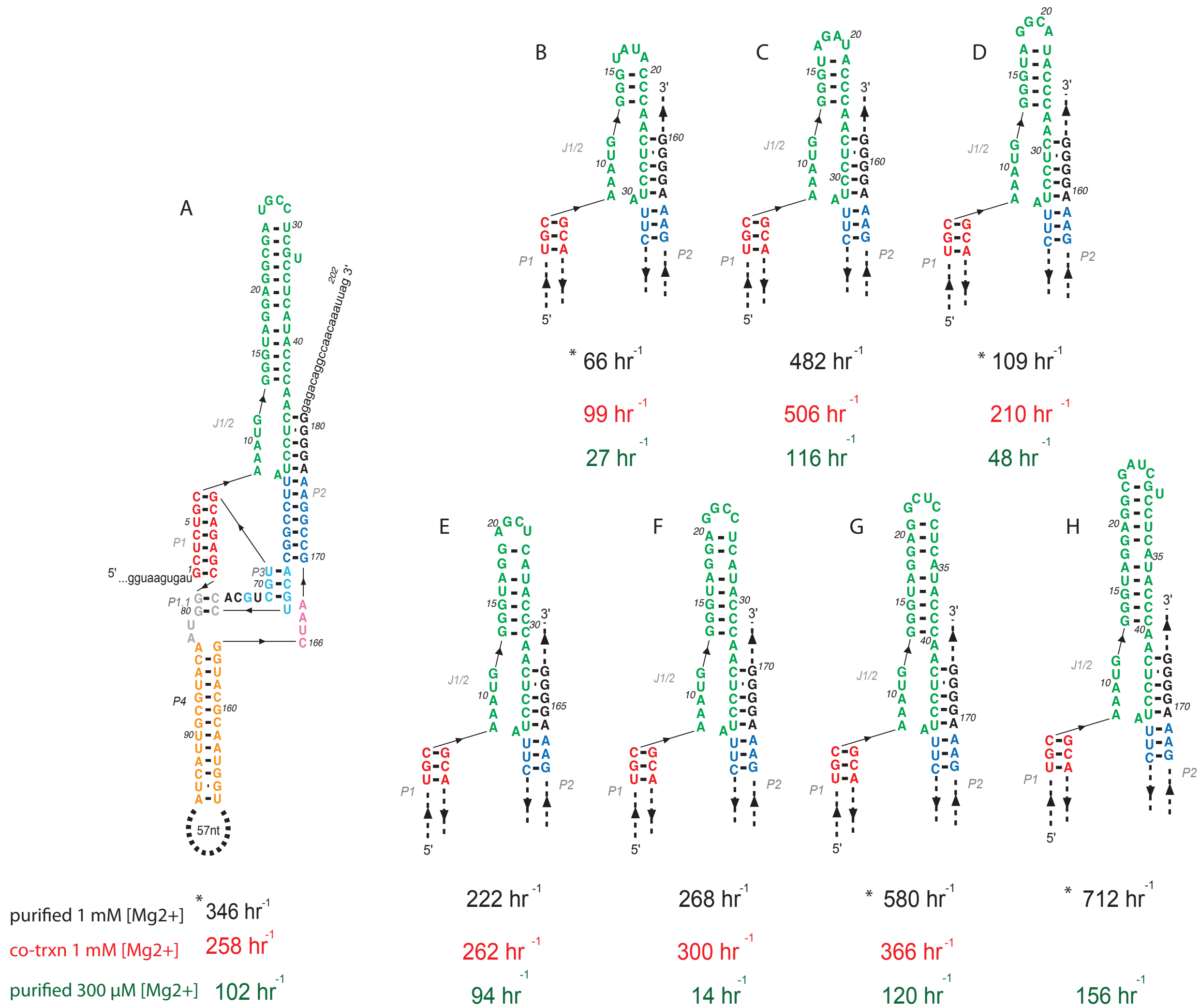


Figure 3-1: Secondary structures of (A) drz-Agam-2-1-1, (B) drz-Agam-2-1-9, (C) drz-Agam-2-1-9.1, (D) drz-Agam-2-1-10, (E) drz-Agam-2-1-10.1, (F) drz-Agam-2-1-10.2, (F) drz-Agam-2-1-11, and (H) drz-Agam-2-1-12. Kinetic rates measured at 37 °C are below each construct, shown in black are the kinetics of purified constructs measured at 1 mM Mg^{2+}, red for co-transcriptional kinetics at 1 mM Mg^{2+}, and green for the kinetics of purified constructs at 300 µM Mg^{2+}. * rate determined by Dr. Judy Webb. Constructs (A), (B), (D), (G), and (H) designed by Dr. Webb. All rates are the average of two-three experimental rates.

Co-transcriptional rates were determined and showed the same pattern of increasing rate with increasing length of J1/2 as observed in the purified form, including the aforementioned anomaly of drz-Agam-2-1-9.1. In some instances, the co-transcriptional and rates of purified constructs were in close agreement however the rate of drz-Agam-2-1-1 (Figure 3-1A) did not conform to the trend set by the rates of purified sequences. Monitoring the self-cleavage over time was not an automated process and done by hand, so after a certain rate it was not possible to observe the cleavage by the time the first time-point was pipetted. The observed rates of cleavage for the faster constructs (Figure 3-1A,E-G) were questionable since they exceeded the rate of the fastest time-point, this was true for both the co-transcriptional kinetics and the kinetics of purified sequences at 1 mM Mg^{2+}. Lowering the amount of Mg^{2+} available to the ribozyme for folding and catalysis reduces the kinetic rate and it was experimentally determined that 300 μM Mg^{2+} was low enough to catch the fastest constructs while still cleaving the slower constructs within a reasonable timeframe. In order to make a comparative study of the rates of purified sequences at 1 mM Mg^{2+}, kinetics at 300 μM Mg^{2+} were done under purified conditions. Rates of the purified constructs (Figure 3-1 green rates) under 300 μM Mg^{2+} retained the observed trend established at 1 mM with the exception of drz-Agam-2-1-10.2 (Figure 3-1F) whose repeated decrease in activity could not be recovered even after renaturing at 95 °C under monovalent ion conditions. Lowered activity is most likely the result of misfolding that hinders the formation of the catalytic core. The kinetic rates of purified sequences showed a consistent relationship with respect to the elongation of J1/2 leading indicating the more stable/structured the J1/2 (more base-pairs), the faster the rate of catalysis.

Filling the minimum free energy (MFE) landscape

Constructs of drz-Agam-2-1 within a specific MFE range were designed in order to fill a noted gap present when plotting the MFE of the existing constructs. The MFE constructs

(Figure 3-2) were engineered using the computational folding program DotKnot wherein nucleotide adjustments were made to create the desired MFE. The rates of the purified constructs were determined at 300 μM Mg^{2+}.

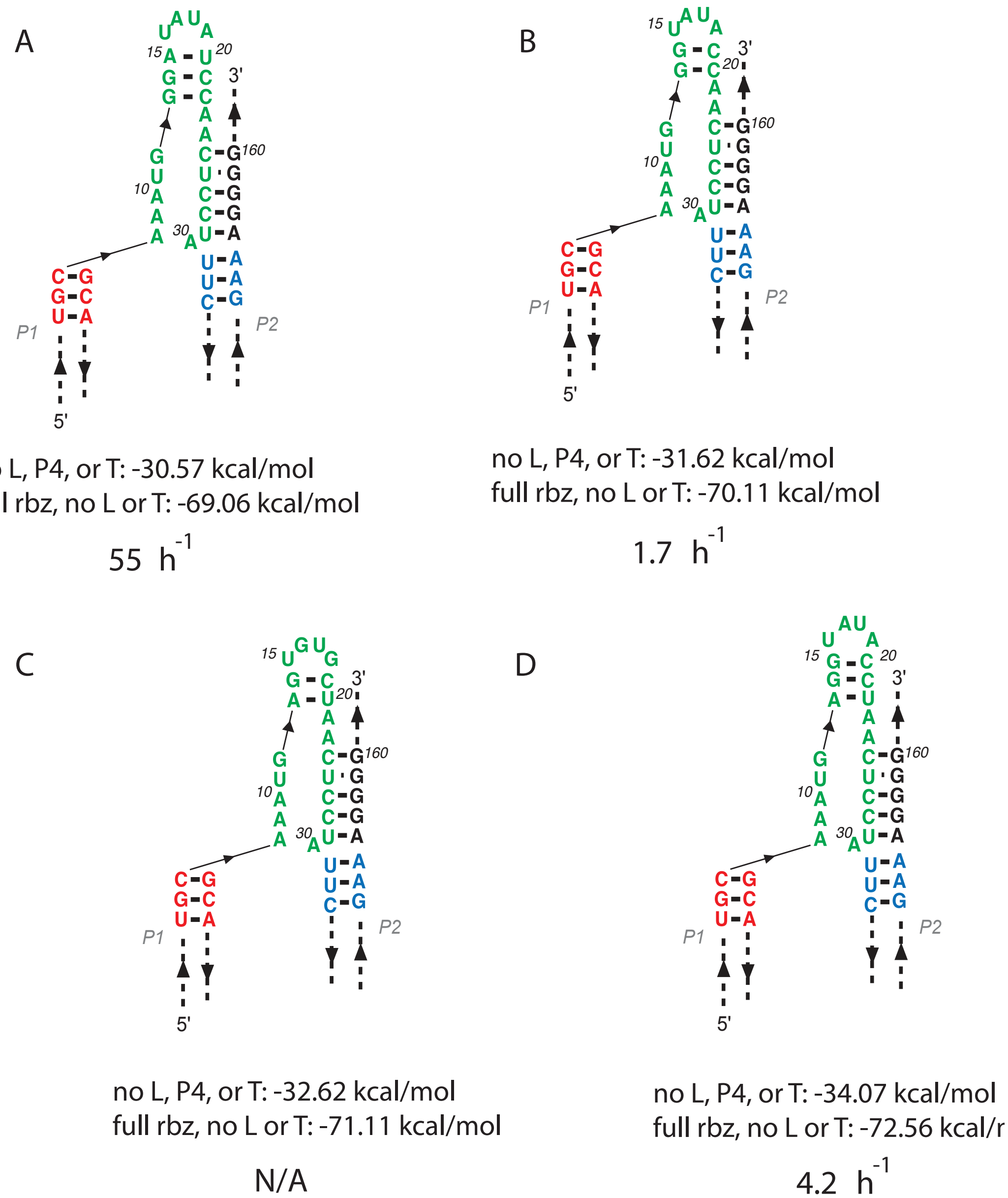


Figure 3-2: Secondary structures of (A) drz-Agam-2-1-8.1, (B) drz-Agam-2-1-8.2, (C) drz-Agam-2-1-8.3, and (D) drz-Agam-2-1-8.4. MFE values given for: ribozyme sequence containing the ribozyme (rbz) but no leader sequence (L), no P4, and no tail sequence (T); full ribozyme (rbz) which includes the P4 but no leader (L) and no tail (T) sequences. The kinetic rates of purified sequences at 300 μM Mg^{2+} was determined for each ribozyme except drz-Agam-2-1-8.3 which resulted in no reaction.

No self-scission was observed for drz-Agam-2-1-8.3 (Figure 3-2C) and drz-Agam-2-1-8.1 (Figure 3-2A) produced a relatively faster rate than the other two kinetically active constructs with lower MFE values. These constructs were used to further examine the effect of J1/2 on catalysis and discussed in the subsequent section.

III. Correlating J1/2 stabilization effects on HDV-like ribozyme catalysis

<u>Determining correlation</u>

In terms of a reaction coordinate diagram for RNA degradation, the intact RNA strand is the reactant and the product is the result of cleaving the RNA strand, the reaction only achieved when the activating energy of the reaction (E_a) is met. The activation energy required for RNA degradation can be lowered using a catalyst and can be done by introducing a ribozyme sequence into the RNA strand (Figure 3-3A). The Arrhenius equation can be solved for the activation energy generated from ribozyme base-pairing, the following derivation leads to equation (1) used to calculate the E_a from ribozyme kinetics:

$$k_{cat} = Ae^{-Ea/RT}$$

$$\frac{k_{cat}}{A} = e^{-Ea/RT}$$

$$\ln\frac{k_{cat}}{A} = -Ea/RT$$

$$-RT * \ln\frac{k_{cat}}{A} = Ea_{cat}$$

$$-RT * \ln\frac{k_{uncat}}{A} = Ea_{uncat}$$

$$Ea = Ea_{cat} - Ea_{uncat}$$

$$Ea = -RT * \ln\frac{k_{cat}}{A} + RT * \ln\frac{k_{uncat}}{A}$$

$$Ea = -RT * \ln\left(\frac{k_{cat}}{A} * \frac{A}{k_{uncat}}\right)$$

(1) $$Ea = -RT * \ln\left(\frac{k_{cat}}{k_{uncat}}\right)$$

where E_a is the activation energy, A the frequency factor, and k is the rate constant. The activation energy is calculated from the experimentally derived kinetic rate from self-cleavage (k_{cat}) relative to the uncatalyzed rate of cleavage (k_{uncat}). The uncatalyzed rate of RNA cleavage, obtained by equation (3), can be calculated from the following equations[49]:

(2) $k_{background} = 10^{\{-14.8+0.983(pH)\}}$ at pH 7.4 = 3 x10^{-8}

(3) $k_{projected} = k_{background} x69.3([Mg]^{0.8})x3.57([K]^{-0.419})$ at [K]of 1.4 M

where $k_{projected}$ is the uncatalyzed rate of degradation, [Mg] is the Mg^{2+} concentration, [K] is the K^+ concentration, $k_{background}$ is the pH adjusted component of the rate, and (pH) is the pH of the reaction. The minimum free energy (MFE) structure is a form or structure that minimizes internal energy at equilibrium. The energy required to form the core of the ribozyme and permit catalysis, determined by the base-pairing of the ribozyme, can be correlated to the stabilization of the ribozyme structure by the MFE (Figure 3-3B). If there is a cooperative relationship between the activation energy and MFE, then a positive linear trend should develop along the plot as the stability of the ribozyme structure decreases. This would mean that at higher ribozyme stability (lower MFE), less energy has to be used for folding and could be allocated to catalysis. Experiments correlating the activation energy to the structure stability of the ribozyme (MFE) were done by incrementally introducing the structured J1/2 element back into the drz-Agam-2-1 ribozyme and monitoring the changes this had on kinetic activity.

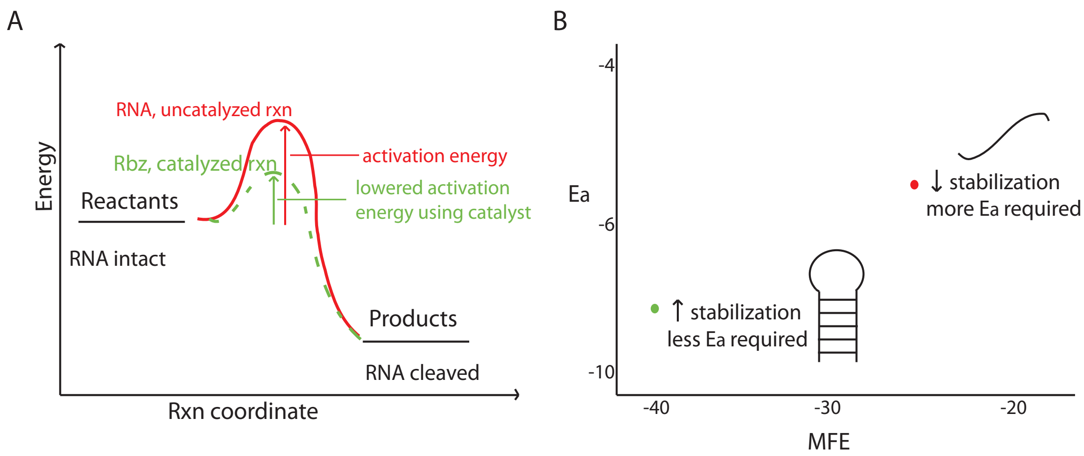


Figure 3-3: (A) Reaction coordinate diagram depicting the energy pathway of RNA degradation in both an uncatalyzed (red) and catalyzed (green) reaction. The catalyzed reaction is shown by including a ribozyme (rbz) sequence in the RNA strand as a method of catalyzing the reaction of RNA scission at a lower activation barrier. (B) Example of plotting the activation energy (E_a) versus the minimum free energy (MFE), showing the expected results in reference to RNA with structure/ribozyme giving rise to more stability/lower E_a (green dot) compared to an unstructured/no ribozyme RNA (red dot) that is unstable and requires more E_a in order to cleave.

Using the established kinetics of the purified drz-Agam-2-1 ribozyme constructs at 300 μM and 1 mM Mg^{2+} , equations (1) and (3), and the predicted MFE of each ribozyme structure calculated by DotKnot, the required activation energy and MFE for each construct were plotted (Figure 3-3). The construct names and associated values used in Figure 3-4 are listed in Table 3-1 (As a note, the rate produced from equation (3) is per minute and was converted to per hour for the Ea values reported.). The MFE for each construct is calculated in two ways: one where the sequence contains the ribozyme (rbz) but no leader sequence (L), no P4, and no tail sequence (T) (Figure 3-4A); the other has the full ribozyme (rbz) sequence which includes the P4 but does not include either the leader (L) nor the tail (T) sequences (Figure 3-4B). The rates of purified sequences at 1 mM (open symbols) and 300 μM Mg^{2+} (filled symbols) were compared. Some of the constructs present at 1 mM Mg^{2+} were solely tested by Dr. Webb and do not have comparisons at 300 μM.

Table 3-1: (A) MFE and E_a associated at the 300 µM Mg^{2+} kinetics for each construct. (B) MFE and E_a associated at the 1 mM Mg^{2+} kinetics for each construct. Data used to construct graphs in Figure 3-4.

A

		300 µM Mg			
No L, No P4	Full rbz, No T	$E_a = -RT*\ln(k_{cat}/k_{uncat})$			Construct
-27.8	-66.29				Agam-2-1-4
-33	-71.49				Agam-2-1-5
-29.3	-67.79				Agam-2-1-6
-29.1	-67.59				Agam-2-1-7
-30.57	-69.06	-9.21	-9.15	-8.91	Agam-2-1-8.1
-31.62	-70.11	-7.13	-6.7		Agam-2-1-8.2
-32.62	-71.11				Agam-2-1-8.3
-34.07	-72.56	-7.7	-7.27	-7.1	Agam-2-1-8.4
-34.87	-73.36		-8.51	-8.79	Agam-2-1-9
-34.52	-73.01	-9.48	-9.61	-9.58	Agam-2-1-9.1
-35.67	-74.16	-8.94	-9.15	-8.94	Agam-2-1-10
-35.58	-74.07	-9.55	-9.55	-9.08	Agam-2-1-10.1
-36.23	-74.72	-8.11	-8.34	-7.8	Agam-2-1-10.2
-40.69	-79.18	-9.67	-9.58	-9.48	Agam-2-1-11
-41.19	-79.68	-9.69	-9.89	-9.61	Agam-2-1-12
-45.62	-84.11	-9.72	-9.36	-9.27	Agam-2-1-1

B

		1 mM Mg			
No L, No P4	Full rbz, No T	$E_a = -RT*\ln(k_{cat}/k_{uncat})$			Construct
-27.8	-66.29	-1.98			Agam-2-1-4
-33	-71.49	-3.78			Agam-2-1-5
-29.3	-67.79	-6.06	-6.19		Agam-2-1-6
-29.1	-67.59	-6.33	-6.32		Agam-2-1-7
-34.87	-73.36	-8.66	-8.77	-8.81	Agam-2-1-9
-35.67	-74.16	-9.32	-9	-8.88	Agam-2-1-10
-40.69	-79.18	-10.2	-10		Agam-2-1-11
-41.19	-79.68	-10.2	-10.3		Agam-2-1-12
-45.62	-84.11	-9.8	-9.78	-9.63	Agam-2-1-1

As expected, the activation energies were higher for constructs with lower stability (higher MFE), which corresponded to constructs with a shorter structured J1/2 stem. Kinetics tested at both 1 mM and 300 µM Mg^{2+} were in agreement and with the exception of a few outliers a general trend can be discerned from the 300 µM kinetics (shown by the gray dash on graphs). Additional constructs are needed to sample the MFE range from -40 to -37 and -78 to -75 to insure that the activation energy stays linear with respect to ribozyme stability through that region.

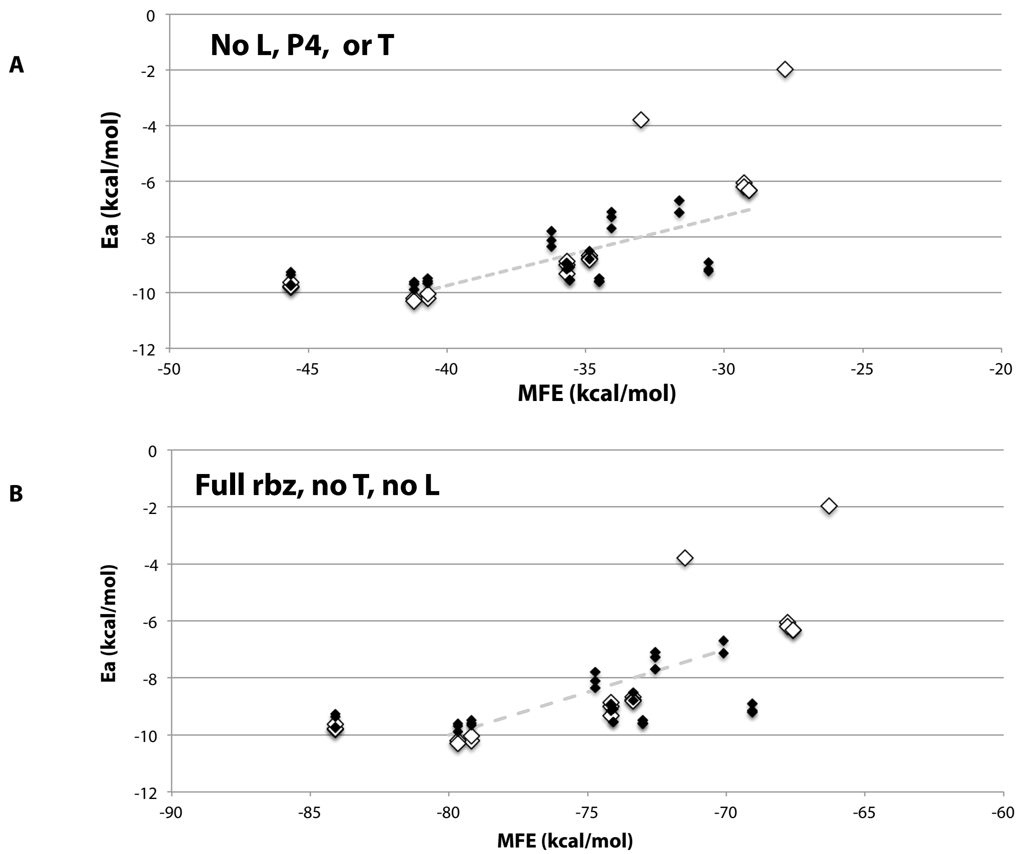
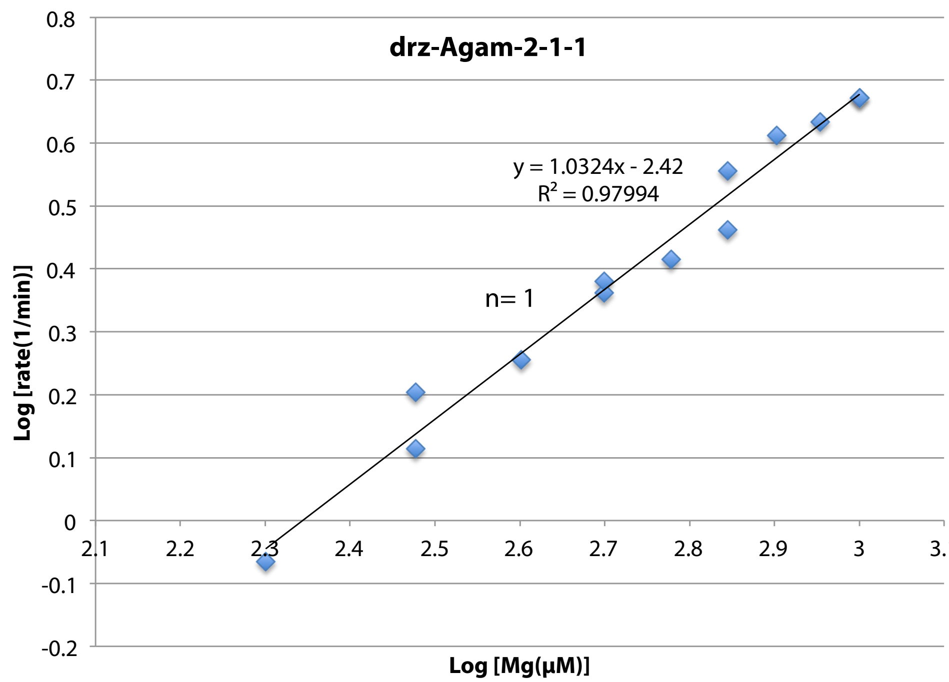


Figure 3-4: Graphs of activation energy vs MFE comparing the rates of purified sequences at 1 mM (open symbols) and 300 µM Mg2+ (filled symbols) of drz-Agam-2 -1 to drz-Agam-2-12 constructs listed in table 3-1. (A) MFE calculated using sequence containing the ribozyme (rbz) but no leader sequence (L), no P4, and no tail sequence (T). (B) MFE calculated using the full ribozyme (rbz) sequence that includes the P4 but does not include either the leader (L) nor the tail (T) sequences. Gray dashed line highlights emerging linear relationship amongst the constructs tested at 300 µM Mg^{2+}.

Hill coefficient of drz-Agam-2-1-1

The binding of Mg^{2+} ions is necessary to help fold and stabilize the tertiary structure of a ribozyme and perform catalysis[50]. The cooperativity of Mg^{2+} binding can be measured by the Hill constant/coefficient with respect to Mg^{2+} concentration. The Hill coefficient for the HDV ribozyme is n= ~1.5-1.7[51], a plot of the rates at different Mg^{2+} concentrations gave a Hill coefficient of n=1 (Figure 3-5) for drz-Agam2-1-1. The slightly lower Hill constant of drz-

Agam-2-1 as compared to HDV could be attributed to the presence of the structured J1/2, making the second Mg^{2+} unnecessary for folding but still requiring one for self-scission.



[Mg(µM)]	rate (1/min)
200	0.86
300	1.6
400	1.8
500	2.3
600	2.6
700	3.6
800	4.1
900	4.3
1000	4.7

Figure 3-5: (Right) Mg^{2+} curve of drz-Agam-2-1-1 plotting the log of the kinetic rate (per minute) versus the log of the Mg^{2+} concentration (µM). A linear regression was fit to the data giving a slope of ~1 meaning the Hill coefficient for drz-Agam-2-1-1 is n=1. (Left) Table of the Mg^{2+} concentrations and rates used to generate the Mg^{2+} curve.

IV. Insertion of J1/2 element into HDV and Hs *CPEB3*

<u>Construct design and initial kinetics</u>

To test whether the J1/2 helix could promote faster catalysis in similarly organized ribozymes, the J1/2 region from drz-Agam-2-1 was grafted onto both the antigenomic and genomic HDV (Figure 3-6) and Hs *CPEB3* ribozymes (Figure 3-7). Both HDV and *CPEB3* have two versions of the J1/2 insert: one where the J1/2 structure is flanked by the naturally occurring nucleotides in the ribozyme spanning between P1 and P2 (named J1/2.1), the other includes these nucleotides and additional nucleotides from the linking nucleotides from drz-Agam-2-1-1 to mimic the seven nucleotides linking J1/2 in drz-Agam-2-1 (named J1/2.2).

Additionally constructs included enough naturally occurring nucleotides upstream from the cleavage site to create a structured (loop/hairpin) leader shown as a MFE structure below its corresponding group of constructs in Figure 3-5G-H.

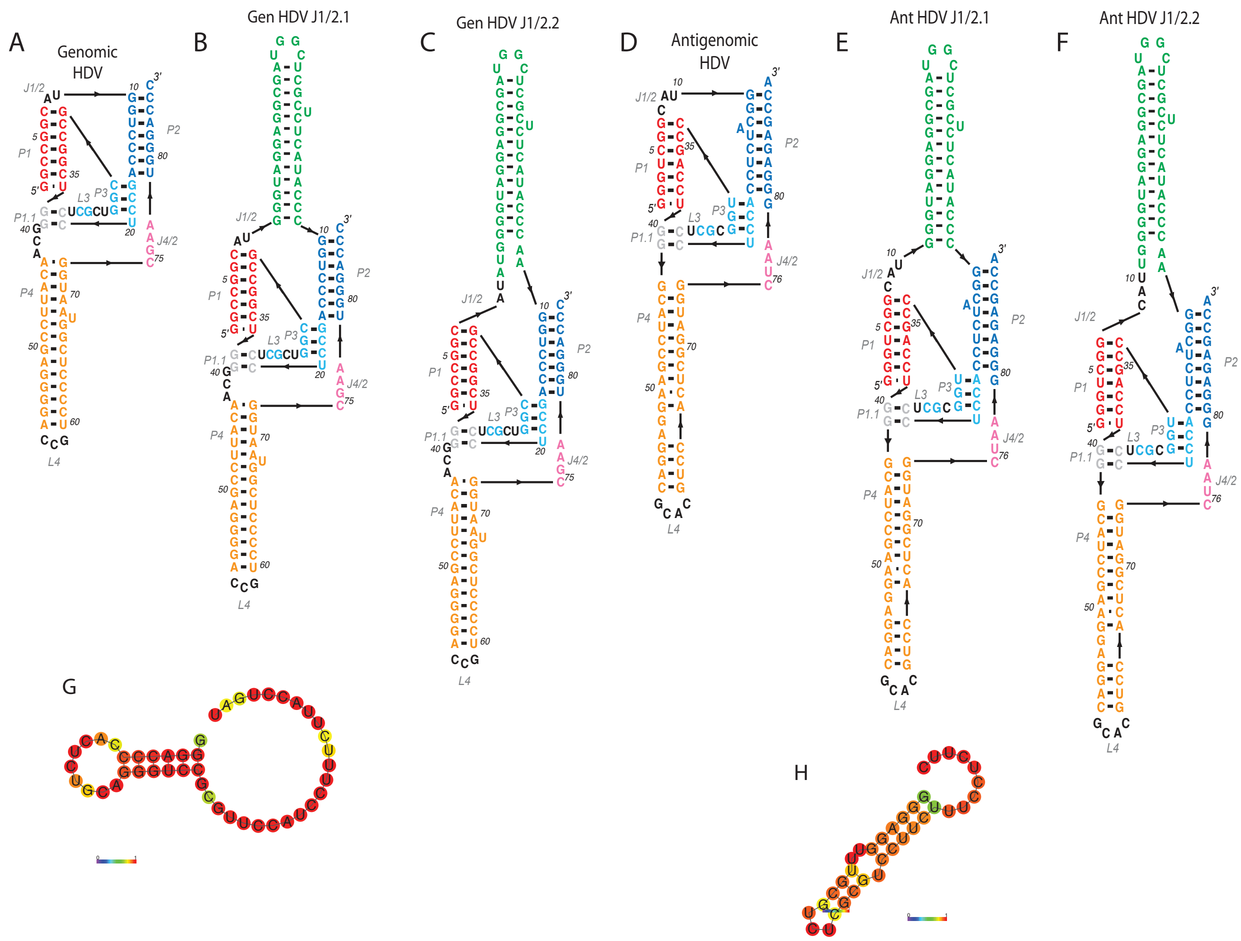


Figure 3-6: Secondary structures of (A) genomic HDV and (D) antigenomic HDV ribozymes. (B-C) Genomic HDV ribozyme constructs with drz-Agam-2-1 J1/2 structure, differing only by the number of nucleotides spanning P1 and P2. (E-F) Antigenomic HDV ribozyme constructs with drz-Agam-2-1 J1/2 structure, differing only by the number of nucleotides spanning P1 and P2. (G) MFE fold of the structured leader sequence present in the genomic HDV constructs (A-C) and not depicted in the above structures. (H) MFE fold of the structured leader sequence present in the antigenomic HDV constructs (D-F) and not depicted in the above structures.

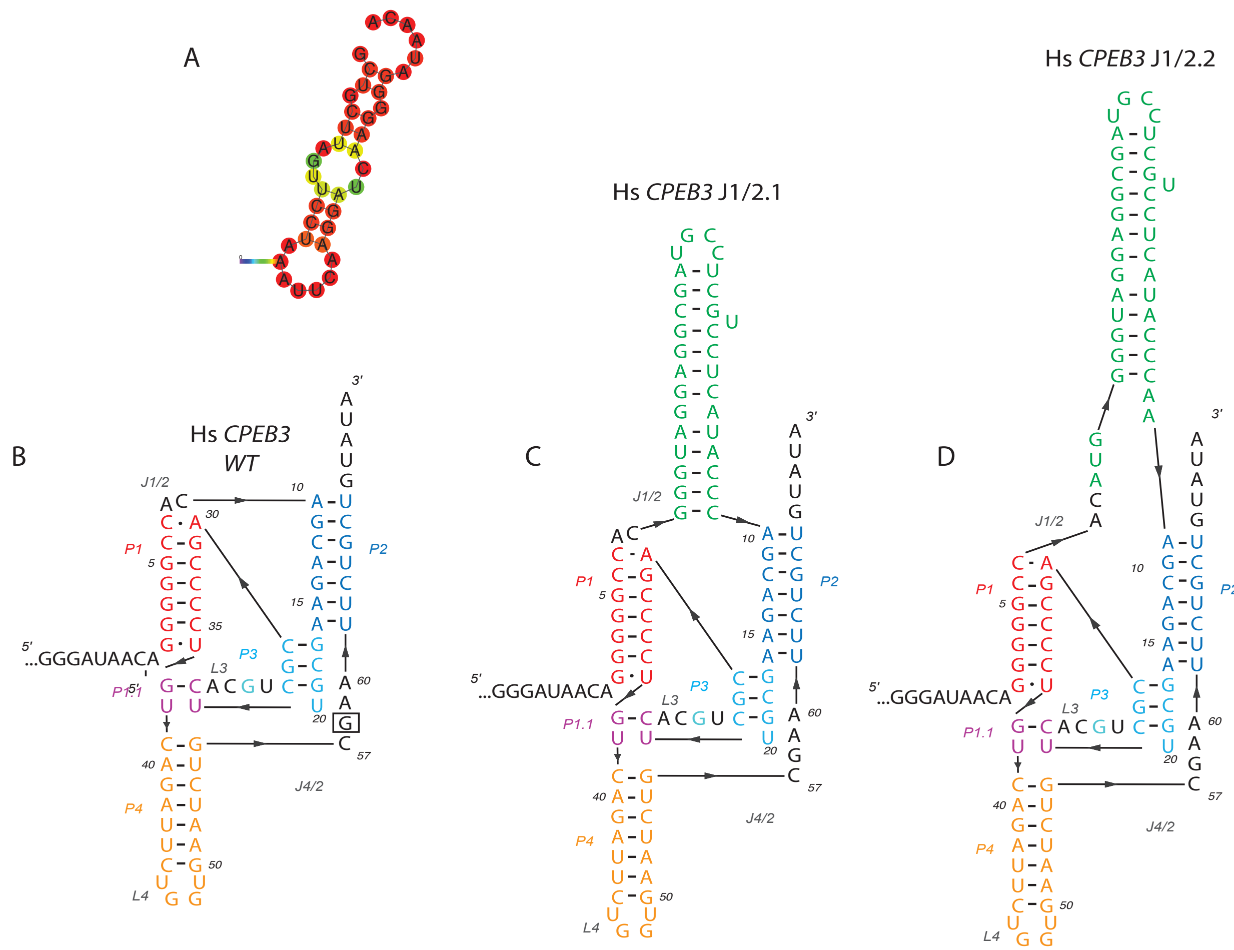


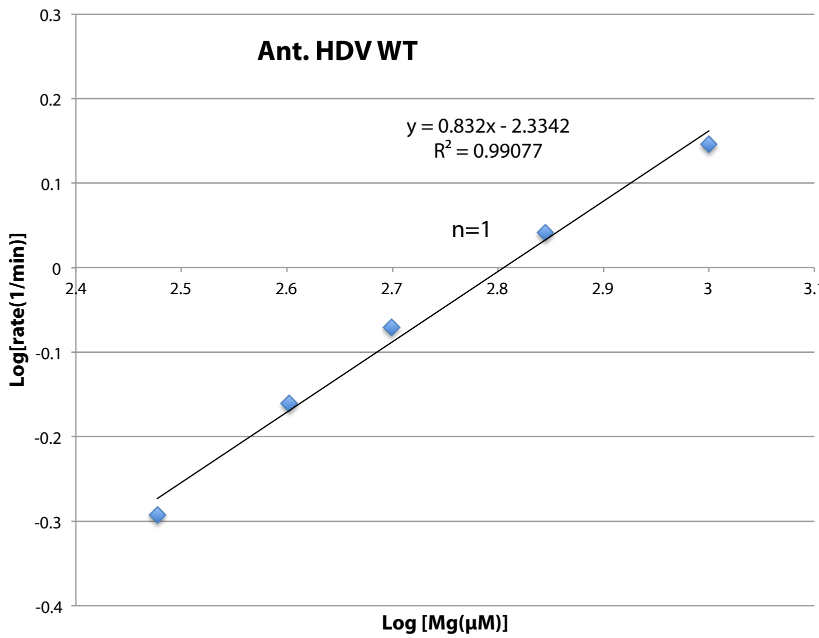
Figure 3-7: Secondary structure of (B) Hs *CPEB3* ribozyme. (C-D) Hs *CPEB3* constructs with J1/2 of drz-Agam-2-1 inserted between P1 and P2, structures differ only by the number of nucleotides spanning P1 and P2. (A) MFE fold of the structured leader upstream of the cleavage site and not depicted in the secondary structures of constructs (C-D).

Only the genomic and antigenomic HDV wildtype (WT) were made, design of the Hs CPEB3 WT from pre-existing oligos generated a construct of incorrect length making it unusable for comparisons. The HDV WTs are exactly like the J1/2.1 and J1/2.2 constructs but lack the J1/2 helix and only has a two nucleotide J1/2 linker for the genomic WT and a three nucleotide linker for the antigenomic WT. Kinetics of purified sequences were done on the HDV and Hs *CPEB3* J1/2 inserted ribozymes at 1 and 10 mM Mg^{2+}. The Hs *CPEB3* constructs showed low activity at both Mg^{2+} concentrations (less than 0.1 min^{-1}) and at such slow rates were not worth further investigation. The genomic HDV construct J1/2.1 also showed low activity but

J1/2.2 had an average rate of 0.57 min^{-1} at 1 mM Mg^{2+}. The antigenomic HDV ribozymes proved to be the fastest with J1/2.1 at ~1 min^{-1} and J1/2.2 at 0.5 min^{-1} in 1 mM Mg^{2+}. The role of the J1/2 helix in the constructs was unclear and led to experiments to determine the binding cooperativity of Mg^{2+} to the ribozymes. By adding the structured J1/2 to the ribozyme, the J1/2 could promote faster folding by bringing the ends of P1 and P2 closer together or in a quicker manner than they would normally with just a J1/2 joining strand. Normally Mg^{2+} ions are used to establish interactions in the tertiary structure of the ribozyme and used for catalytic activity, if the J1/2 helix could replace some of the Mg^{2+} necessary for folding then the Mg^{2+} cooperativity would decrease.

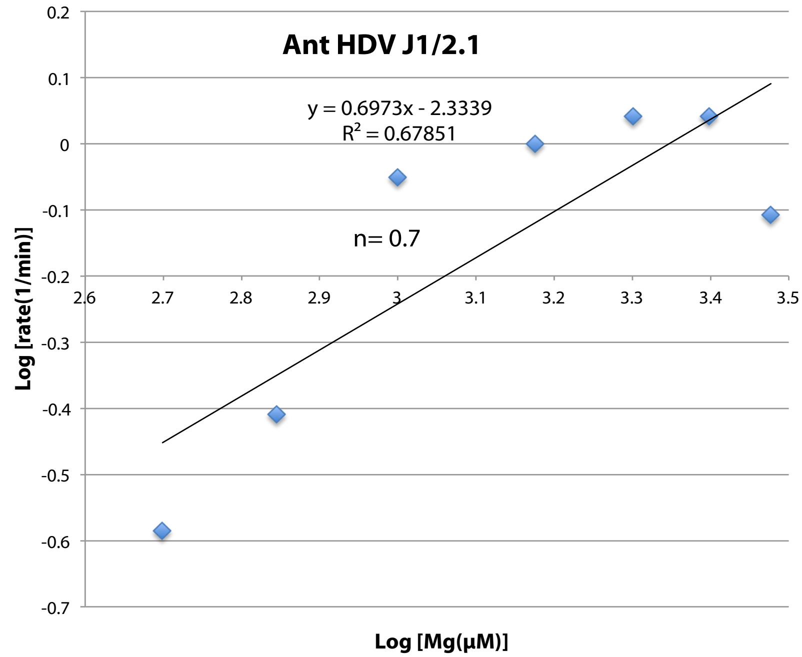
Determining the Hill constant of the J1/2 inserted HDV constructs

Kinetics of purified sequences were done at varying concentrations of Mg^{2+} on multiple constructs, the log of both the rate versus the concentration of the kinetics were graphed to create a Mg^{2+} curve. The linear regression of each curve was calculated with the slope of each line equivalent to the Hill constant for that construct. The Hill constant generated by the Mg^{2+} curve (not shown) for the genomic HDV WT demonstrated little to no cooperativity with a Hill of n=0.5 after generous fitting. The Hill coefficient for the antigenomic HDV WT was n=1 (Figure 3-8), the antigenomic HDV J1/2.1 was n=0.7 (Figure 3-9), and the antigenomic HDV J1/2.2 was n=1 (Figure 3-10). The fit of the regression of the antigenomic HDV J1/2.1 was not confident, speculation would guess the Hill is in the range of 1 (showing cooperativity) to ~0.5 which is too low to confirm cooperative binding. Repeating the kinetics at similar Mg^{2+} concentrations will hopefully give conclusive results. The antigenomic HDV WT had higher kinetic rates than the J1/2.2 at the same Mg^{2+} concentrations but both similar Mg^{2+} cooperativity with a Hill coefficient of n=1. The HDV ribozyme is made of a short RNA sequence and may not need the help of a structured J1/2 to fold quickly, the presence of the helix could be lowering the kinetic rate by interfering with tertiary contacts or misfolding.



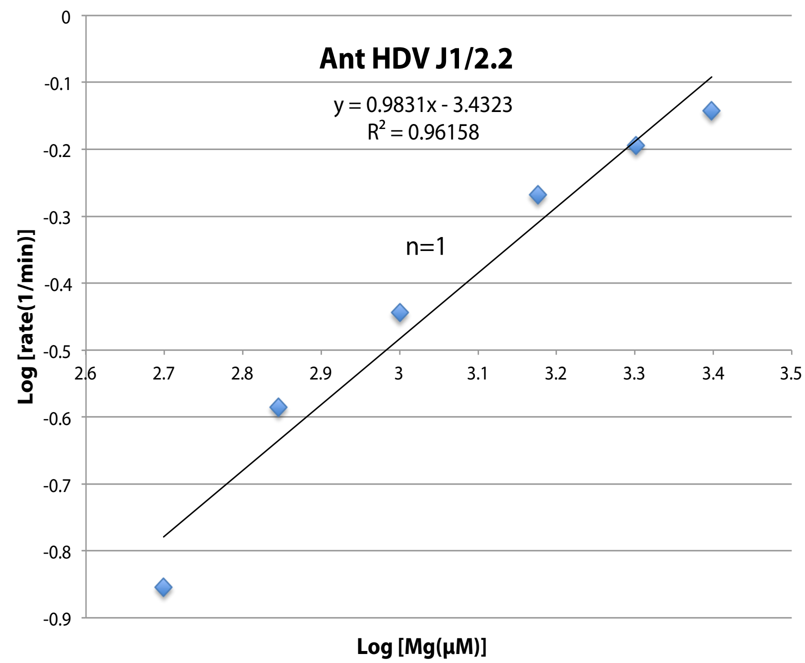
[Mg(μM)]	rate(1/min)
300	0.51
400	0.69
500	0.85
700	1.1
1000	1.4

Figure 3-8: (Right) Magnesium curve of antigenomic HDV WT plotting the log of the kinetic rate (per minute) versus the log of the Mg^{2+} concentration (μM). A linear regression was fit to the data giving a slope of ~1 meaning the Hill coefficient for antigenomic HDV WT is n=1. (Left) Table of the Mg^{2+} concentrations and rates used to generate the Mg^{2+} curve.



[Mg(µM)]	rate(1/min)
500	0.26
700	0.39
1000	0.89
2000	1.1
2500	1.1
3000	0.78

Figure 3-9: (Right) Mg^{2+} curve of antigenomic antigenomic HDV J1/2.1 plotting the log of the kinetic rate (per minute) versus the log of the Mg^{2+} concentration (µM). A linear regression was fit to the data giving a slope of ~0.7 meaning the Hill coefficient for antigenomic HDV J1/2.1 is n=0.7. (Left) Table of the Mg^{2+} concentrations and rates used to generate the Mg^{2+} curve.



[Mg(µM)]	rate(1/min)
500	0.14
700	0.26
1000	0.36
1500	0.54
2000	0.64
2500	0.72

Figure 3-10: (Right) Mg^{2+} curve of antigenomic antigenomic HDV J1/2.2 plotting the log of the kinetic rate (per minute) versus the log of the Mg^{2+} concentration (µM). A linear regression was fit to the data giving a slope of ~1 meaning the Hill coefficient for antigenomic HDV J1/2.1 is n=1. (Left) Table of the Mg^{2+} concentrations and rates used to generate the Mg^{2+} curve.

V. Conclusions

Peripheral elements in both the hammerhead and hairpin, though not directly involved in ribozyme catalysis, have been shown as an evolutionary necessity to enhance ribozyme scission by locking the RNA into a catalytically stable conformation[52,53]. The relevance of stabilizing domains was also emphasized in HDV wherein drastic decreases in activity resulted from the removal of either J1/2 or L4 and both resulted in less activity than either individual deletion[4]. Experiments showed an increase in activity with the growing stability of the J1/2 stem loop in the rates of the purified drz-Agam-2-1 constructs and co-transcriptional kinetics. The drz-Agam-2-1 ribozyme is relatively larger than HDV and might have developed a structured joining strand to reduce the number of Mg^{2+} ions necessary to fold and achieve catalysis, explaining the observed Hill coefficient of n=1 (compared to n=~1.5 for HDV[51]). The introduction of the structured J1/2 into ribozymes of similar secondary structures, HDV and Hs *CPEB3*, did not provide any further clarification after evaluating its effect on Mg^{2+} cooperativity. Once again, the smaller sizes of these ribozymes may not benefit from more than a simple linking strand at J1/2 to execute faster folding. Future directions of this project involve the additional design and testing of drz-Agam-2-1 constructs to fill in the MFE holes of the existing constructs and clarifying the influence of the J1/2 helix on catalysis through stabilizing core folding.

Chapter 4

Methods and Materials

I. Example of processing kinetic data and modeling for chapters 2 & 3

The data shown is the background subtracted percentage for each band based on the volume intensity within each gel lane, each lane represents an aliquot taken at a different time (Table 4-1).

Table 4-1: Data used to construct the kinetic graph from the ribozyme self-cleavage gel. Each time point has its own lane in the gel and the volume intensity of each band for the lane is given as a percentage of the total lane intensity along with the time. The amount of uncleaved product for each time point/lane is converted to a fraction and used along with the time in minutes to construct a graph of the fraction of uncleaved product versus the time in minutes in a log-log scale as seen in Figure 4-1. Minimizing the difference squared between the data and the model generates a best-fit model using a mono-exponential equation. The resulting variables of the modeled equation are listed as A, k, and C.

Ref band number	Lane 1 Band %	Lane 2 Band %	Lane 3 Band %	Lane 4 Band %	Lane 5 Band %	Lane 6 Band %	Lane 7 Band %
1	98.33	93.53	86.12	75.79	60.45	39.44	36.48
2	1.67	6.47	13.88	24.21	39.55	60.56	63.52
3	0	0	0	0	0	0	0
time (min)	0.0167	0.15	0.33	0.567	1.1	3.9	7.53
Fraction Uncleaved	0.9833	0.9353	0.8612	0.7579	0.6045	0.3944	0.3648
Model mono	0.996847473	0.928271813	0.847368082	0.758356201	0.613848947	0.389006708	0.368052474
Difference Squared	0.000184114	5.32122E-05	0.000225784	2.74435E-07	0.000142385	7.4774E-05	2.87421E-05
Sum of difference	0.000709286						
A	0.638905386						
k	0.864910581						
C	0.367104113						

The band percentage for the amount of uncleaved product for each aliquot is converted into a fraction (labeled fraction uncleaved in Table 4-1); a log-log graph is plotted using this and the time of each aliquot (shown as blue diamonds in Figure 4-1).

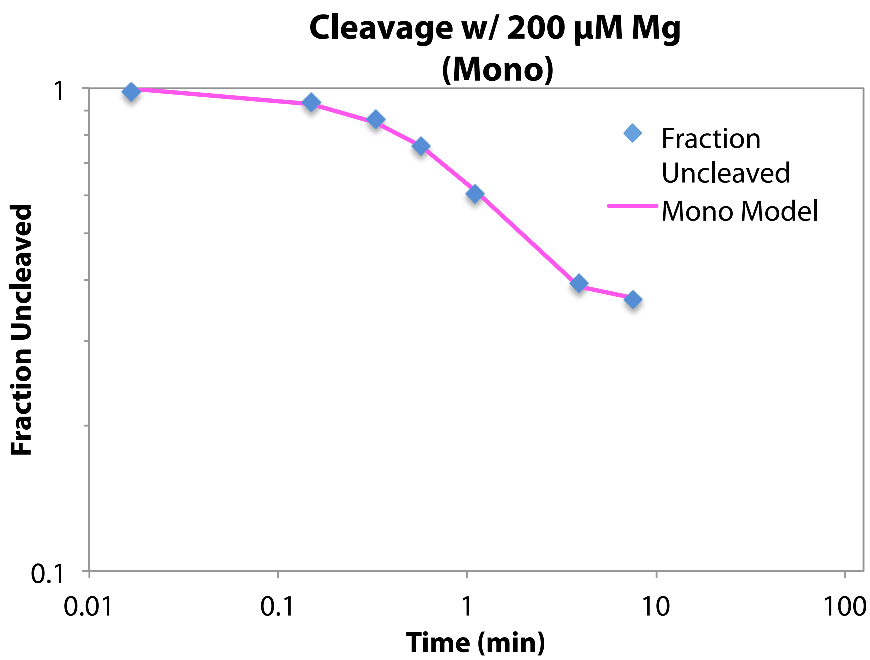


Figure 4-1: Graph of the time points from the kinetics of a purified HDV-like ribozyme RNA. The graph is the plot of the log of the fraction of uncleaved product versus the log of the time in minutes. The blue diamonds are the experimental time points and the pink line is the mono-exponential best-fit to those time points. The self-scission was induced using 200 µM Mg^{2+}.

Using the least squares method, the fraction of uncleaved product (F_{un}) at each time point (t) is modeled to both a mono-exponential (i) and bi-exponential decay (ii) function:

$$F_{un} = Ae^{-k_1 t} + C \qquad \textbf{(i)}$$

$$F_{un} = Ae^{-k_1 t} + Be^{-k_2 t} + C \qquad \textbf{(ii)}$$

where C is the fraction of uncleaved product at the end of the experiment, k_1 and k_2 are the rate constants, A and B are the amplitudes of the rate constants, respectively, and t is the time of the data point (shown as a pink line for the mono-exponential in Figure 4-1; mono-exponential solved for F_{un} listed as 'Model mono' in Table 4-1). Data are fit both mono- and bi-exponentially to determine which model better represents the cleavage, example only shown for kinetics of purified RNA with mono-exponential fitting but is similar for co-transcriptional kinetics of RNA and bi-exponential fitting.

II. Chapter 2 methods & materials.

General Methods. The concentration of oligonucleotides in solution was measured by UV spectroscopy on a Nanodrop ND-1000 spectrophotometer. Sample volumes were 2 μL; sample concentration was adjusted after initial reading and measured again until a consistent reading within the detection limits of the instrument was obtained. Measurements were taken at 260 nanometers.

Separation of radioactive co-transcriptional products was done on a denaturing 7.5% polyacrylamide gel containing 7 M urea. Gels were pre-run for 26-50 min then run anywhere from 30 min to ~2 h depending on construct length at 20 watts using a BioRAD model 3000 xi power source. Gel images were visualized using either a Molecular Dynamics or GE phosphor screen with exposure times of 2-2.5 h. Radioactive images were acquired using a GE Typhoon Trio variable mode imager.

General Method for PCR. All PCR reactions were carried out with DreamTaq Green PCR Master Mix (2x). PCR reactions included 5 μL of DNA template, 10 μL of each primer (all primers were 10 μM concentration), and 25 μL DreamTaq Green PCR Master Mix (2x). Initial concentration of DNA templates was unknown. The thermal cycling sequence was as follows: 95 °C for 30 seconds, then iterated cycles of 95 °C for 30 seconds, annealing temperature (different and specific to primers used) for 30 seconds, and 72 °C for 30 seconds. Upon completion of iterated cycles a final incubation of 72 °C for 30 seconds was performed. PCR experiments were performed using a BioRAD MJ Mini personal thermal cycler.

Method for Primer Extension. Primer extension of construct -11/72 was carried out with DreamTaq Green PCR Master Mix (2x). Primer extension included 1 μL of reverse primer (concentration ~ 40 μM), 6.41 μL of forward primer (concentration ~6.24 μM), 2.59 μL of double-deionized (dd) water, and 10 μL DreamTaq Green PCR Master Mix (2x). The thermal cycling sequence was as follows: 94 ºC for 1 min, then two iterated cycles of 50 ºC for 30 seconds and 72 ºC for 2 min. Primer extension experiments were performed using a BioRAD MJ Mini personal thermal cycler.

General Method of co-Transcription (optimized) and Gel Analysis. Transcriptions involved a 5 μL transcription that was allowed to transcribe for 10 min at 25 ºC using an Eppendorf AG mastercycler personal. Transcription contained: 1 μL of 5x transcription buffer (10 mM spermidine, 50 mM dithiothreitol, 120 mM trizma hydrochloride buffer pH 7.5, and 0.05% Triton), 1 μL of 5x ribonucleoside triphosphates (total concentration of 6.8 mM) and DTT mix, 1 μL of 5 mM Mg^{2+}, 1 μL DNA, 0.5 μL of 100% DMSO, 0.15 μL of DI water, 0.1 μL of BioLabs murine RNase inhibitor (40,000 units/mL), 0.125 μL of T7 polymerase, and 0.125 μL 32P rATP. Transcriptions were added and mixed with a 95 μL solution at 37 ºC that was a mixture of 50 μL of a physiological buffer with 2 mM Mg^{2+} (for induction of self-scission) and 45 μL of double-deionized water. The physiological buffer contained: 280 mM KCl, 20 mM NaCl, and 100 mM trizma hydrochloride (Tris-HCl) buffer pH 7.5. The 100 μL solution was then held at 37 ºC by a Perkin Elmer Cetus DNA thermal cycler for the reminder of the experiment while data points were taken. A 4 mM EDTA/7 M urea stopping solution was added to each aliquot collected. Aliquots were run on a denaturing PAGE (7.5% polyacrylamide, 7 M urea) at 20 W. The gel was exposed on a phophosimager plate for ~ 2 hours and then the plate was imaged using a Typhoon imager, allowing radioactivity to be visualized. Radioactive images were analyzed using Gel-Quant electrophoresis image analysis software.

III. Chapter 3 methods & materials

J 1/2 domain effects on Agam 2-1 information & methods

General Methods. The concentration of oligonucleotides in solution was measured by UV spectroscopy on a Nanodrop ND-1000 spectrophotometer. Sample volumes were 2 µL; sample concentration was adjusted after initial reading and measured again until a consistent reading within the detection limits of the instrument was obtained. Measurements were taken at 260 nanometers.

Separation of radioactive co-transcriptional products was done on a denaturing 7.5% polyacrylamide gel containing 7 M Urea. Gels were pre-run for 15 min then run anywhere from 45 min to ~1 h depending on construct length at 40 watts using a BioRAD model 3000 xi power source. Gel images were visualized using either a Molecular Dynamics or GE phosphor screen with exposure times of 12-36 h. Radioactive images were acquired using a GE Typhoon Trio variable mode imager.

General Method for PCR. All PCR reactions were carried out with the following reagents to make a 100 µL reaction: 4 µL of a mixture of 5mM dNTPs (A, T, C, & G NTPs were each present at 5mM in the mixture with water), 20 µL of 10 µM forward primer, 20 µL of 10 µM reverse primer, 1 µL (later increased to 2 because of decreased efficiency) of 5 unit/L inhouse Taq DNA polymerase, 2 µL of DNA template, 10 µL of New England Biolabs 10x reaction buffer for Taq DNA polymerase reactions (1x components: 10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, pH 8.3 @ 25 ^{o}C) and 43 µL of dd H_2O. Initial concentration of DNA templates was unknown, DNA template stocks were made for PCR by diluting some of the original DNA to a final concentration of 1:100 and used for PCR. The thermal cycling sequence was as follows: 95 °C for 30 seconds, then iterated cycles of 95 °C for 30 seconds, annealing temperature (different and specific to primers used) for 30 seconds, and 72 °C for 1 minute. Upon completion of iterated cycles a final incubation of 72 °C for 30 seconds was performed. PCR experiments were performed using a BioRAD MJ Mini personal thermal cycler. Presence of DNA verified by loading 3 µL of PCR aliquot with an equi-volume amount of loading buffer (mixture of 7 M urea and 5 mM trizma hydrochloride buffer pH 7.5 dissolved in water with xylene cyanol and bromophenol blue dyes) into a 2% agarose gel run at 200 volts for ~20 min.

General method of self-priming/primer extension. Primers used in method designed to anneal with one another and become the desired DNA product after extension. All self-priming reactions contained the following reagents: 5 µL of New England Biolabs 10x reaction buffer for Taq DNA polymerase reactions (1x components: 10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, pH 8.3 @ 25 °C), 2 µL of a mixture of 5mM dNTPs (A, T, C, & G NTPs were each present at 5mM in the mixture with water), 10 µL of 10 µM forward primer, 10 µL of 10 µM reverse primer, 1 µL of 5 unit/L inhouse Taq DNA polymerase, and 22 µL of double-deionized water. The thermal cycling sequence was as follows: 95 °C for 30 seconds, then one-two cycles of 95 °C for 30 seconds, annealing temperature (different and specific to primers used) for 30 seconds, and 72 °C for 1 minute. Products checked by agarose gel electrophoresis.

General Method of Transcription, purification, and Gel Analysis. Transcriptions involved a 20 µL transcription that was allowed to transcribe for 1 h at 37 °C using an Eppendorf AG mastercycler personal. A 1x transcription contained: 4 µL of 5x transcription buffer (10 mM spermidine, 50 mM dithiothreitol, 120 mM trizma hydrochloride buffer pH 7.5, and 0.05%

Triton), 1 µL of 200 mM DTT, 1 µL of 305 mM $MgCl_2$, 4 µL of rNTP mixture (rC, rG, & rU at 25 mM and rA at 1.25 mM), 4 µL of inhibitor oligo (later upped to 5 µL for the Agam 2-1 constructs, oligo varies since it is specific to the construct transcribed and is the reverse compliment, aka antisense stand, that spans the cleavage site of the ribozyme), 2 µL of 100% DMSO, 2 µL DNA, 0.5 µL inhouse T7 RNA polymerase, and 0.5 µL of alpha 32P radiolabeled rATP. Master mixes were made by multiplying by the number of transcriptions desired and an additional factor of 0.1 to account for pipet losses (i.e. desired number of transcriptions is 4, so multiply 1x transcription volumes by 4.1). Transcripts were purified by mixing the 20 µL transcription with 20 µL stop buffer (8 M urea, 20 mM EDTA, and 5 mM trizma hydrochloride buffer pH 7.5) and loading the 40 µL solution onto a thick-spaced denaturing PAGE (7.5 % polyacrylamide, 7 M urea) run at 20 watts for ~50 minutes. The gel was then exposed to a phosphorimage plate for ~15 minutes and the plate scanned by the typhoon. The radiolabelled RNA bands were cut from the gel by using a true-to-size printed image of the gel as a template. Each RNA band was placed in a Corning Costar Spin-X centrifuge filter (cellulose acetate membrane, pore size 0.45 µM) column with ~100-150 µL of dd H_2O (enough to cover the gel piece) and then shaken for 1 h to elute the RNA from the gel. The eluted RNA was further purified by filtering 50-75µL portions of the elution through Spin-X filter columns containing Sephadex G-25 beads, this process was done twice, each time with new Sephadex/column. The columns were prepared by adding 800 µL of Sephadex G-25 fine beads suspended in 10 mM trizma hydrochloride buffer pH 7.5 to Spin-X filter columns followed by spinning at 25,000 rcf for 1 minute and discarding the flow-through. The eluted RNA was added to the prepared Sephadex filter columns and spun at 25,000 rcf for 1 minute, the flow-through of purified radiolabeled RNA was stored at -80 ^{o}C and used in gel analysis.

Gel analysis was done by combining 5 µL of purified ribozyme RNA at 37 ºC to a 25 µL solution containing various concentrations of Mg^{2+} (for induction of self-scission) in a physiological buffer at 37 ºC, the mixture was held at 37 ºC by a Perkin Elmer Cetus DNA thermal cycler while aliquots were taken over time. 1 µL of the purified ribozyme RNA was removed before mixing with the Mg^{2+} solution to establish a zero time-point. The reaction was quenched by mixing each 3 µL aliquot to 3 µL of stop buffer (mentioned above). Aliquots were run on a denaturing PAGE (7.5% polyacrylamide, 7 M urea) at 40 W. The gel was exposed on a phophosimager plate for ~ 12-36 h and then the plate was imaged using a Typhoon imager, allowing radioactivity to be visualized. Radioactive images were analyzed using Gel-Quant electrophoresis image analysis software.

Method of making Agam-2-1 ribozyme constructs and oligos used. Agam-2-1-1 through 2-1-12 constructs were designed by former group member Judy Webb, additional agam-2-1 constructs made after Webb have a decimal followed by a number connotation (i.e. 2-1-10.1). Each construct has a different length J1/2 region linking the P1 and P2 helixes. The wild-type (WT) agam-2-1-1 (also known as just 2-1) was PCR amplified from drz-Agam-2-1 plasmid designed by Webb containing the full-length ribozyme including a 44 nt sequence upstream of the cleavage site and 24 nt sequence after the ribozyme. The WT ribozyme was PCRed from the plasmid using oligos AL 805 and AL 740, the construct produced known as AL 805-740. Subsequent agam constructs were made using 805-740 as the template DNA, oligo AL 740 as the reverse primer, and the oligo that would give the desired J1/2 portion of the construct as the forward primer. PCR product was checked for correct size by agarose and a final PCR done to add the T7 promoter and any missing leader sequence by using the PCR product as the DNA template, AL 805 as the forward primer w/ T7, and AL 740 as the reverse primer. Final PCR products checked on agarose. Naming system established by Webb was continued and explained as follows: construct naming uses the oligo name of the forward primer as the first number and the reverse primer oligo name as the second number from PCR to make a construct name (i.e. AL 805 F and AL 740 R used to make AL 805-740). If a third oligo is used to make the construct, specifically the J1/2 region, then that oligo name is inserted between the forward and reverse primers used in the final PCR (i.e. AL 805-1015-740).

Construct sequences and oligos used presented in 5′-3′ format below:

constructs presented in DNA form but kinetic data done using the RNA transcribed from the DNA construct & no longer contains the T7 promoter sequence

<u>Agam-2-1-1 (AL 805-740) w/ long leader & tail</u>
ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGGTAGGAGGCGA TGCCTCGTCCTCATACCCAACTCCTATTCCGGCACGTCCACGTCGTGCAGAGCGGTAACATGCGTTA CTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTGCATCAGATGGTAAC GCATGG CTAAGCCGGAAAGGGGgagacaggc caacaaatta g

<u>AL 805 forward primer w/ T7</u>
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat

<u>AL 740 reverse primer</u>
CTAATTTGTTGGCCTGTCTCCCCTTTCCGGCTTAGCCATG

<u>AL 2375 inhibitor oligo for Agam-2-1 constructs (technically a reverse primer)</u>
CCCATTTGCAGAGCATCACTTACCAAG

<u>Agam-2-1-4 (AL 806-740)</u>
TTCCCGCGAAATTAATACGACTCACTATAgggagttaacggcaaaaagttgaa tactcagcttggtaagtgatGCTCTGCCTCCTATTCCGGCACGTCCACGTCG TGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTG CACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

<u>AL 806 forward primer w/ T7</u>
TTCCCGCGAAATTAATACGACTCACTATAgggagttaacggcaaaaagttgaatactcagcttggtaagtgatgctc tgcctcctattccggcacgtcca

<u>Agam-2-1-5 (AL 805-1011-740)</u>

TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgatGCTCTGCACTCCTATTCCGGCACGTCCACGTCGTGCAGAGCGGTAACATGCGTTACTA GGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTGCATCAGATGGTAACGC ATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 1011 forward primer
gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGC A CTCCTA TTCCGGCAC GTCCACGTCG
Agam-2-1-6 (AL 805-1012-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAACTCCTATTCCGGCACGTCCACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGGTG CAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGCTA AGCCGGAAAGGGGgagacaggc caacaaatta g

AL 1012 forward primer
gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAA CTCCTA TTCCGGCAC GTCCACGTCG

Agam-2-1-7 (AL 805-1013-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAATGCTCCTATTCCGGCACGTCCACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGG TGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGC TAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 1013 forward primer
gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAA TG CTCCT A TTCCGGCAC GTCCACGTCG

Agam-2-1-8.1 (AL 805-2430-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaa GTTAACGGC AAAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAATG GGATATAT CCAACTCCTA TTCCGGCACG TCCACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTC TTTTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2430 forward primer
GTTAACGGC AAAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAATG GGATATAT CCAACTCCTA TTCCGGCACG TCCACGTCG

Agam-2-1-8.2 (AL 805-2431-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaa GTTAACGGCA AAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAA TGGGTATACC AACTCCTA TT CCGGCACGTC CACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTT TTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2431 forward primer
GTTAACGGCA AAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAA TGGGTATACC AACTCCTA TT CCGGCACGTC CACGTCG

Agam-2-1-8.3 (AL 805-2432-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaa GTTAACGGCA AAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAA TGAGTGTGCT AACTCCTA TT CCGGCACGTC CACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTT TTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2432 forward primer
GTTAACGGCA AAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAA TGAGTGTGCT AACTCCTA TT CCGGCACGTC CACGTCG

Agam-2-1-8.4 (AL 805-2433-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaa GTTAACGGCA AAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAATG GGGTATACCT AACTCCTA TTCCGGCACG TCCACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTC TTTTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2433 forward primer
GTTAACGGCA AAAAGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAATG GGGTATACCT AACTCCTA TTCCGGCACG TCCACGTCG

Agam-2-1-9 (AL 805-1015-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAATGGGGTATACCCAACTCCTATTCCGGCACGTCCACGTCGTGCAGAGCGGTAACATG CGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTGCATCAGATG GTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 1015 forward primer
gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAA TG GGGTA TA CCCAA CTCCT A TTCCGGCAC GTCCACGTCG

Agam-2-1-9.1 (AL 805-2356-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGG TATA CCC AACTCCTATTCCGGCAC GTCCACGTCGT GCAGAGC GGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTG CATCAGA TGGTAACGCATGG CTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2356 forward primer
gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGGT AGAT ACCC AACTCCTATTCCGGCAC GTCCACGTC

Agam-2-1-10 (AL 805-1663-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAATGGGGTAGGCATACCCAACTCCTATTCCGGCACGTCCACGTCGTGCAGAGCGGTAA CATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTAGTTGCATCA GATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 1663 forward primer
aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAA TGGGGTAGG CATA CCCAACTCCT A TTCCGGCAC GTCCACGTCG

Agam-2-1-10.1 (AL 805-2357-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGGTAGG AGCT CATACCC AACTCCTATTCCGGCAC GTCCACGTCGT GCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGC ACTAGTTGCATCAGA TGGTAACGCATGG CTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2357 forward primer
aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGGTAGG AGCT CATACCC AACTCCTATTCCGGCAC GTCCACGTCG

Agam-2-1-10.2 (AL 805-2358-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGGTAGGA GGCC TCATACCC AACTCCTATTCCGGCAC GTCCACGTCGT GCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGC ACTAGTTGCATCAGA TGGTAACGCATGG CTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 2358 forward primer
aaaagttgaa tactcagctt ggtaagtgat GCTCTGC AAATG GGGTAGGA GGCC TCATACCC AACTCCTATTCCGGCAC GTCCACGTCG

Agam-2-1-11 (AL 805-1664-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAATGGGGTAGGAGGCTCCTCATACCCAACTCCTATTCCGGCACGTCCACGTCGTGCAG AGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTGCACTA GTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

AL 1664 forward primer
aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAA TGGGGTAGGA GGC TCCTCATA CCCAACTCCT A TTCCGGCAC GTCCACGTCG

Agam-2-1-11.1 (AL 805-2434-740)
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaaGTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAA TGGGGTAGGA GGCGCGCCTC ATACCCAACT CCTA TTCCGG CACGTCCACGTCGTGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGA

GCTCTTTTTGCTGCACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

<u>AL 2434 forward primer</u>
GTTGAA TACTCAGCTT GGTAAGTGAT GCTCTGC AAA TGGGGTAGGA GGCGCGCCTC ATACCCAACT CCTA TTCCGG CACGTCCACG TCG

<u>Agam-2-1-12 (AL 805-1665-740)</u>
TTCCCGCGAAATTAATACGACTCACTATA ggga gttaacggca aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAATGGGGTAGGAGGCGATCGTCCTCATACCCAACTCCTATTCCGGCACGTCCACGTCG TGCAGAGCGGTAACATGCGTTACTAGGGGTGCAAGAGCTCTTTTTGAGGAGGAGCTCTTTTTGCTG CACTAGTTGCATCAGATGGTAACGCATGGCTAAGCCGGAAAGGGGgagacaggc caacaaatta g

<u>AL 1665 forward primer</u>
aaaagttgaa tactcagctt ggtaagtgat GCTCTGCAAA TGGGGTAGGA GGCGAT CGTCCTCATA CCCAACTCCT ATTCCGGCAC GTCCACGTCG

Method of J1/2 insertion into HDV and *CPEB3* constructs and oligos used. Naturally occurring antigenomic and genomic HDV sequences were engineered to contain the J1/2 region from agam-2-1-1. HDV constructs were made by self-priming, explained above, followed by PCR to add additional sequence using the self-priming DNA product as the template for PCR. Constructs included enough naturally occurring nucleotides, upstream from the cleavage site, to create a structured (loop/hairpin) leader. The Hs *CPEB3* sequence was also manipulated to contain the drz-Agam-2-1-1 J1/2 region and a naturally occurring structured leader. Both HDV and *CPEB3* have two versions of the J1/2 insert: one where the J 1/2 structure is flanked by the naturally occurring nucleotides in the ribozyme spanning between P1 and P2 (named J1/2.1) , the other includes these nucleotides and additional nucleotides from the linking nucleotides from agam-2-1-1 to mimic the seven nucleotides linking J1/2 in agam-2-1-1 (named J1/2.2). HDV construct naming via oligo names is as follows unless specified: forward PCR primer oligo, forward self-priming oligo, reverse self-priming oligo, reverse PCR primer oligo. Construct sequences and oligos used presented in 5′-3′ format below:

<u>Ant HDV WT (AL 2388-2384-2385-2387)</u>
TTCCCGCGAA ATTAATACGA CTCACTATA GGGAGGUUUG CGUCUCGCGU CCUUCUUUCC UCUUCGGGUC GGCAUGGCAU CUCCACCUCC UCGCGGUCCG ACCUGGGCAU CCGAAGGAGG ACGCACGUCC ACUCGGAUGG CUAAGGGAGA GCCA

<u>AL 2388 forward primer</u>
TTCCCGCGAA ATTAATACGA CTCACTATA GGGAGGTTTG CGTCT CGCGTCCTTC TTTCCTCTTC GGG

<u>AL 2387 reverse primer</u>
TGGCTCTCCC TTAGCCATCC GAG

<u>AL 2384 forward primer</u>
GGGAGGTTTG CGTCTCGCGT CCTTCTTTCC TCTTCGGGTC GGCATGGCAT CTCCACCTCC TCGCGGTCCG

<u>AL 2385 reverse primer</u>
TGGCTCTCCC TTAGCCATCC GAGTGGACGT GCGTCCTCCT TCGGATGCCC AGGTCGGACC GCGAGGAGGT GGAG

<u>AL 2391 inhibitor oligo for Ant HDV constructs (technically a reverse primer)</u>
GCCGACCCGA AGAGGAAAGA AGG

<u>Ant HDV J1/2.1 (2388-2422-2423-2387)</u>
TTCCCGCGAAATTAATACGACTCACTATAGGGAGGTTTGCGTCTCGCGUCCUUCUUUCCUCUUCG GGUCGGCAUGGGUAGGAGGCGAUGGCUCGCUCUCAUACCCGGCA UCUCCACCUC CUCGCGGUCC GACCUGGGCA UCCGAAGGAG GACGCACGUC CACUCGGAUG GCUAAGGGAG AGCCA

<u>AL 2422 forward primer</u>
GTCCTTCTTT CCTCTTCGGG TCGGCATGGG TAGGAGGCGA TGGCTCGCTC TCATACCCGG CATCTCCACC TC

<u>AL 2423 reverse primer</u>
TGGCTCTCCC TTAGCCATCC GAGTGGACGT GCGTCCTCCT TCGGATGCCC AGGTCGGACC GCGAGGAGGT GGAGATGCCG GGTAT

<u>Ant HDV J1/2.2 (2388-2424-2425-2387)</u>
TTCCCGCGAA ATTAATACGA CTCACTATA GGGAGGTTTG CGTCT C

GCGUCCUUCUUUCCUCUUCGGGUCGGCAUUGGGUAGGAGGCGAUGGCUCGCUCUCAUACCCA AGGCAUCUCCACCUCCUCGCGGUCCGACCUGGGCAUCCGAAGGAGGACGCACGUCCACUCGGA UGG CUAA GGGAGAGCCA

AL 2424 forward primer
GCGTCCTTCT TTCCTCTTCG GGTCGGCATT GGGTAGGAGG CGATGGCTCG CTCTCATACC CAAGGCATCT CCACCTCCTC G

AL 2425 reverse primer
TGGCTCTCCC TTAGCCATCC GAGTGGACGT GCGTCCTCCT TCGGATGCCC AGGTCGGACC GCGAGGAGGT GGAGATGCCT T

Gen HDV WT (2380-2376-2377-2379)
TTCCCGCGAA ATTAATACGA CTCACTATAGGGACCCCAC TCTGCAGGGT CCGCGTTCCA TCCTTTCTTA CCTGATGGCC GGCATGGTCC CAGCCTCCTC GCTGGCGCCG GCTGGGCAAC ATTCCGAGGG GACCGTCCCC TCGGTAATGG CGAATGGGAC CC

AL 2380 forward primer
TTCCCGCGAA ATTAATACGA CTCACTATAG GGACCCCACT CTGCAGGGTC CGCGTTCCAT CCTTTCTTAC CTG

AL 2379 reverse primer
GGGTCCCATT CGCCATTACC GAG

AL 2376 forward primer
GGGACCCCAC TCTGCAGGGT CCGCGTTCCA TCCTTTCTTA CCTGATGGCC GGCATGGTCC CAGCCTCC

AL 2377 reverse primer
GGGTCCCATT CGCCATTACC GAGGGGACGG TCCCCTCGGA ATGTTGCCCA GCCGGCGCCA GCGAGGAGGC TGGGACCATG CCGG

AL 2383 inhibitor oligo for Gen HDV constructs (technically a reverse primer)
ATGCCGGCCA TCAGGTAAGA AAGGA

Gen HDV J1/2.1 (2380-2426-2427-2428-2379)
*self-priming first with AL 2426 and AL 2427, then primer extension using 1 µL of the self-priming product as DNA template and AL 2426 and AL 2428 as primers, followed by PCR with AL 2380 and AL 2379.
TTCCCGCGAA ATTAATACGA CTCACTATA GGGACCCCACT CTGCAGGGT CCGCGTTCCA TCCTTTCTTA CCTGATGGCC GGCATGGGTA GGAGGCGATG GCTCGCTCTC ATACCCGGTC CCAGCCTCCT CGCTGGCGCC GGCTGGGCAA CATTCCGAGG GGACCGTCCC CTCGGTAATG GCGAATGGGA CCC

AL 2426 forward primer

CCGCGTTCCA TCCTTTCTTA CCTGATGGCC GGCATGGGTA GGAGGCGATG GCTCGCTCTC ATACCC

AL 2427 reverse primer
GACGGTCCCC TCGGAATGTT GCCCAGCCGG CGCCAGCGAG GAGGCTGGGA CCGGGTATGA GAGCGAGCCA TCG

AL 2428 reverse primer
GGGTCCCATT CGCCATTACC GAGGGGACGG TCCCCTCGGA ATGTT

Gen HDV J1/2.2 (2380-2382-2379)
*PCR was done using 2380-2376-2377-2379 as a DNA template and AL 2382 as the forward primer and AL 2379 as the reverse primer, then a final PCR using the diluted PCR product from the previous step and AL 2380 as the forward primer and AL 2379 as the reverse primer.
TTCCCGCGAA ATTAATACGA CTCACTATAGGGACCCCAC UCUGCAGGGU CCGCGUUCCA UCCUUUCUUA CCUGAU GGCCGGC AUAU GGGUAGGAGGCGAUGGCUCGCUCUCAUACCCAA GGUCCCAGCCUCCUCGCUGGCGCCGGCUGGGCAACAUUCCGAGGGGACCGUCCCCUCGGUAAU GG CGAA UGGGACCC

AL 2382 forward primer
TTCCCGCGAAATTAATACGACTCACTATAGGGCTGCTTAGTTCCTAAATTCAAGGATCATCCGCGTT CC ATCCTTTCTT ACCTGATGGC CGGCATATGG GTAGGAGGCG ATGGCTCGCT CTCATACCCA AGGTCCCAGC CTCCTCGCTG

CPEB3 J1/2.1 (2394-2451-2452-2395)
TTCCCGCGAA ATTAATACGA CTCACTATA GGGCTGCTTAGTTCCTAAATTCAAGGATCAAGGGGA TAACAGGGGG CCACGGGTAG GAGGCGATGC CTCGTCCTCA TACCCAGCAG AAGCGTTCAC GTCGCAGCCC CTGTCAGATT CTGGTGAATC TGCGAATTCT GCTGTATA

AL 2394 forward primer w/ T7 primer
TTCCCGCGAA ATTAATACGA CTCACTATA GGGCTGCTTAGTTCCTAAATTCAAGGATCA

AL 2395 reverse primer
TATACAGCAG AATTCGCAGA TTCACCAGAA T

AL 2451 forward primer
GGGCTGCTTAGTTCCTAAATTCAAGGATCAAGGGGATAACAGGGGGCCACGGGTAGGAGGCGATG C CTCGTCCTCA TACCCAGCAG AAGC

AL 2452 reverse primer
TATACAGCAG AATTCGCAGA TTCACCAGAA TCTGACAGGG GCTGCGACGT GAACGCTTCT GCTGGGTATG AGGACG

AL 2397 inhibitor oligo for *CPEB3* constructs (technically a reverse primer)
GTGGCCCCCT GTTATCCCCT TGAT

CPEB3 J1/2.2 (2394-2429-2450-2395)
TTCCCGCGAAATTAATACGACTCACTATAGGGCTGCTTAGTTCCTAAATTCAAGGATCAAG
GGGAUAACAGGGGGCCACAUGGGGUAGGAGGCGATGCCTCGTCCTCATACCCAAAGCAGAAGCG
UUCACGUCGCAGCCCCUGUCAGAUUCUGGUGAAUCUGCGAAUUCUGCUGUAUA

AL 2429 forward primer
GGGCTGCTTA GTTCCTAAAT TCAAGGATCA AGGGGATAAC AGGGGGCCAC ATGGGGTAGG
AGGCGATGCC TCGTCCTCAT A

AL 2450 reverse primer
TATACAGCAG AATTCGCAGA TTCACCAGAA TCTGACAGGG GCTGCGACGT GAACGCTTCT
GCTTTGGGTA TGAGGACGAG GCATCGCCT

References

(1) Salehi-Ashtiani, K.; Lupták, A.; Litovchick, A.; Szostak, J. W. *Science* **2006**, *313*, 1788.

(2) Chadalavada, D. M.; Gratton, E. A.; Bevilacqua, P. C. *Biochemistry* **2010**, *49*, 5321.

(3) Webb, C. H.; Riccitelli, N. J.; Ruminski, D. J.; Lupták, A. *Science* **2009**, *326*, 953.

(4) Tinsley, R. A.; Walter, N. G. *Biol Chem* **2007**, *388*, 705.

(5) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. A.; Breaker, R. R. *Nature* **2004**, *428*, 281.

(6) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147.

(7) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849.

(8) Doudna, J. A.; Lorsch, J. R. *Nat Struct Mol Biol* **2005**, *12*, 395.

(9) *The RNA World: The Nature of Modern RNA Suggests a Prebiotic RNA*; 3 ed.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 2006.

(10) Buzayan, J. M.; Gerlach, W. L.; Bruening, G. *Nature* **1986**, *323*, 349.

(11) Prody, G. A.; Bakos, J. T.; Buzayan, J. M.; Schneider, I. R.; Bruening, G. *Science* **1986**, *231*, 1577.

(12) Sharmeen, L.; Kuo, M. Y.; Dinter-Gottlieb, G.; Taylor, J. *Journal of Virology* **1988**, *62*, 2674.

(13) Saville, B. J.; Collins, J. A. *Cell* **1990**, *61*, 685.

(14) Fedor, M. J. *Annu Rev Biophys* **2009**, *38*, 271.

(15) Walter, N. G.; Perumal, S. *Springer series in biophysics* **2009**, *13*, 103.

(16) Ferre-D'Amare, A. R.; Scott, W. G. *Cold Spring Harb Perspect Biol* **2010**, *2*, a003574.

(17) Buzayan, J. M.; Bruening, G.; Harnpel, A. *Nucleic Acid Research* **1986**, *14*, 9729.

(18) Webb, C.-H. T.; Riccitelli, N. J.; Ruminski, D. J.; Luptak, A. *Science* **2009**, *326*, 953.

(19) Szostak, J. W.; Luptak, A. In *Ribozymes and RNA Catalysis*; Lilley, D. M. J., Eckstein, F., Eds.; The Royal Society of Chemistry: Cambridge, 2007, p 123.

(20) Taylor, J. M. *Hepatitis Delta Virus* **2006**, *307*, 1.

(21) Rizzetto, M.; Canese, M. G.; Arico, S.; Crivelli, O.; Trepo, C.; Bonino, F.; Verme, G. *Gut* **1977**, *18*, 997.

(22) Been, M. D. *Curr Top Microbiol Immunol* **2006**, *307*, 47.

(23) Kuo, M. Y.; Sharmeen, L.; Dinter-Gottlieb, G.; Taylor, J. *J Virol* **1988**, *62*, 4439.

(24) Ferre-D'Amare, A. R.; Doudna, J. A. *J Mol Biol* **2000**, *295*, 541.

(25) Ferre-D'Amare, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567.

(26) Huang, Y. S.; Richter, J. D. *Methods Enzymol* **2007**, *431*, 143.

(27) Schuman, E. M.; Kang, H. *Science* **1996**, *273*, 1402.

(28) Steward, O.; Schuman, E. M. *Neuron* **2003**, *40*, 347.

(29) Huang, Y.-S.; Kan, M.-C.; Lin, C.-L.; Richter, J. D. *The EMBO Journal* **2006**, *25*, 4865.

(30) Peng, S.-C.; Lai, Y.-T.; Huang, H.-Y.; Huang, H.-D.; Huang, Y.-S. *Nucleic Acid Research* **2010**, *38*, 7446.

(31) Theis, M.; Si, K.; Kandel, E. R. *Proc. Natl. Acad. Sci.* **2003**, *100*, 9602.

(32) Murray, J. B.; Seyhan, A. A.; Walter, N. G.; Burke, J. M.; Scott, W. G. *Chemistry & Biology* **1998**, *5*, 587.
(33) Singh, J.; Padgett, R. A. *Nature structural & molecular biology* **2009**, *16*, 1128.
(34) Fong, N.; Ohman, M.; Bentley, D. L. *Nature structural & molecular biology* **2009**, *16*, 916.
(35) Perrotta, A. T.; Been, M. D. *Nucleic Acid Research* **1990**, *18*, 6821.
(36) Perrotta, A. T.; Been, M. D. *Nature* **1991**, *350*, 434.
(37) Chadalavada, D. M.; Knudsen, S. M.; Nakano, S.; Bevilacqua, P. C. *J. Mol. Biol.* **2000**, *301*, 349.
(38) Lambowitz, A. M.; Perlman, P. S. *Trends Biochem. Sci.* **1990**, *11*, 440.
(39) Munroe, S. H.; Dong, X. F. *Proc. Natl. Acad. Sci.* **1992**, *89*, 895.
(40) Herschlag, D. *J. Biol. Chem.* **1995**, *270*, 20871.
(41) Michel, F.; Umesono, K.; Ozeki, H. *Gene* **1989**, *82*, 5.
(42) DePriest, P. T.; Been, M. D. *Journal of molecular biology* **1992**, *228*, 315.
(43) Uhlenbeck, O. C.; Stage-Zimmermann, T. K. *Rna* **1998**, *4*, 875.
(44) Perrotta, A. T.; Been, M. D. *Nature* **1991**, *350*, 434.
(45) Wadkins, T. S.; Perrotta, A. T.; Ferre-D'Amare, A. R.; Doudna, J. A.; Been, M. D. *RNA* **1999**, *5*, 720.
(46) Chadalavada, D. M.; Knudsen, S. M.; Nakano, S.; Bevilacqua, P. C. *J Mol Biol* **2000**, *301*, 349.
(47) Pereira, M. J.; Harris, D. A.; Rueda, D.; Walter, N. G. *Biochemistry* **2002**, *41*, 730.
(48) Perrotta, A. T.; Been, M. D. *Biochemistry* **1992**, *31*, 16.
(49) Li, Y.; Breaker, R. R. *J. Am. Chem. Soc.* **1999**, *121*, 5364.
(50) Pyle, A. M. *Science* **1993**, *261*, 709.
(51) Nakano, S.; Cerrone, A. L.; Bevilacqua, P. C. *Biochemistry* **2003**, *42*, 2982.
(52) Khvorova, A.; Lescoute, A.; Westhof, E.; Jayasena, S. D. *Nat Struct Biol* **2003**, *10*, 708.
(53) Walter, N. G.; Burke, J. M.; Millar, D. P. *Nature Structural Biology* **1999**, *6*, 544.