

KINETICS, THERMODYNAMICS, AND DYNAMICS OF RIBOZYMES

By

Neil Andrew White

A DISSERTATION

Submitted to  
Michigan State University  
In partial fulfillment of requirements  
For the degree of

Biochemistry and Molecular Biology—Doctor of Philosophy

2016

## **ABSTRACT**

### **KINETICS, THERMODYNAMICS, AND DYNAMICS OF RIBOZYMES**

By

Neil Andrew White

RNA transcribed from DNA can be divided into two groups: RNA that codes for protein and RNA that does not code for protein, or so-called non-coding RNA. Non-coding RNA can be further divided into several classes based on function. Non-coding RNAs perform a wide array of functions in living organisms, from gene regulation, to scaffolding, to catalysis. It is amazing that despite RNA having only four, chemically-similar monomers it can have such important, wide-ranging functions. Proteins which also perform wide-ranging functions in organisms have twenty common monomers that are vastly more diverse in terms of chemical or functional groups and structure. How non-coding RNA, specifically catalytic RNA or ribozymes, overcome this inherent lack of chemical and structural diversity to have impressive, intricate structures and function is the focus of this thesis.

It is important to study how ribozymes are able to form intricate structure and execute function. They also have potential therapeutic applications, to control RNA viruses like HIV and oncogene transcripts, due to their ability to cleave RNA. Also, they provide a window back to a time described by the RNA World Hypothesis, a time before DNA and proteins, when RNA performed self-replication.

Ribozymes overcomes its lack of diversity in monomers by being a dynamic polymer. Conformational diversity or the ability to transition from one conformation to another is critical to function of ribozymes. Nuclear magnetic resonance is a tool that is unparalleled in its ability to provide site-specific insight on time ranges from pico-seconds to thousands of seconds. The

ribose dynamics of both the lead-dependent ribozyme or leadzyme and the hairpin ribozyme will be elucidated in chapters four and five, with their dynamics tied to the ribozymes' functions. These studies represent dynamics-function assays which are essential to going beyond a static view of molecules.

In this thesis, the first report of the binding kinetics of the junctionless hairpin ribozyme will be described in chapter two, which we published. The thermodynamic signature for the junctionless hairpin ribozyme will also be presented in chapter three, with important considerations of the commonly used cleavage-site modification. The kinetics and thermodynamics are essential in understanding how the junctionless hairpin ribozyme forms its active structure in a fundamental way.

Lastly, a RNA-protein interaction will be discussed in chapter six. The protein is present in *Trypanosoma brucei*, the parasitic protozoan that causes African Sleeping Sickness in humans. The key element of the protein's specificity for RNA was determined using *in vitro* selection. The specificity suggests that this protein may have a role in RNA editing. This is another case of specific interfaces being important to function.

In totality, this thesis examines the structure-function paradigm prevalent in molecular biology, in a RNA-centric manner. It also goes beyond static pictures of molecules and enters into the dynamics-function realm that is essential for a more complete picture of how RNA can function as a catalyst.

Copyright by  
NEIL ANDREW WHITE  
201

## ACKNOWLEDGMENTS

First and foremost I have to acknowledge my wife, Sara White. I am exceedingly grateful for her love and encouragement. Without her, my life simply would not be what it is today.

I have to thank a great personal-lineage of science teachers/mentors that have made my journey possible. First was Mrs. Kathy McDonald who sparked my scientific inquiry as an elementary student. Next, was Mr. G. Scott Gerstler, my high school biology teacher, who gave me a lot to enjoy and ponder. At that time, I would discuss what I was learning with my brother, Adam White, who was in his undergraduate studies at the time and it was a formative experience. In my undergraduate studies Dr. Neil Bowlby was exceptionally important to me. Without him, I don't see how I would have even gone to graduate school.

In graduate school I have be fortunate to have several great mentors. I first must thank Dr. Charles G. Hoogstraten, my mentor. I am fortunate to have been taken into his lab, despite not having much experience in general, and specifically none of real consequence with NMR. I have enjoyed the opportunity to learn and conducted experiments under his mentorship.

I also need to thank Dr. Cori L. Fata-Hartley my mentor for my teaching-as-research project. Not only did I learn about how students learn, I learned about my own learning. This was an important piece of development in graduate school for me.

I was fortunate have very good relationships with the Hoogstraten lab members. I spent a vast amount time learning from and discussing things with Dr. Mina Sumita. I was very grateful to have had her help and insight. I was also fortunate to work with Dr. Fadhiru Kamba and Dr. Patrick Ochieng. I wish to thank both of them for their friendship and being helpful and supportive.

I also wish to thank the members of my Guidance Committee: Dr. Leslie Kuhn, Dr. Honggao Yan, Dr. Erik Martinez-Hackert, and Dr. Lisa Lapidus for their discussion and direction.

I also am blessed to have a great support system comprised of family and friends. I have never forgotten this and it means very much to me. In particular, I would like to acknowledge a dear friend, Adam Flory, for his support.

## TABLE OF CONTENTS

LIST OF TABLES .....	x
LIST OF FIGURES .....	xi
KEY TO ABBREVIATIONS .....	xiii
CHAPTER 1 .....	1
RIBOZYMES AND THE DYNAMICS-FUNCTION RELATIONSHIP .....	1
1.1 INTRODUCTION TO RIBOZYMES .....	2
1.1.1 Ribozymes, a type of non-coding RNA .....	2
1.1.2 RNA as a catalyst .....	3
1.1.3 Hairpin ribozyme: biological role to therapeutic approaches .....	7
1.1.4 The RNA World .....	8
1.2 DYNAMICS-FUNCTION RELATIONSHIP .....	8
1.2.1 Necessity of conformational re-arrangement in the hairpin ribozyme .....	8
1.2.2 Double conformational capture .....	10
1.2.3 Kinetics, thermodynamics, and dynamic of ribozymes .....	14
1.2.4 Dynamics as compensation for RNA's inherent lack of chemical diversity .....	14
1.2.5 Surface plasmon resonance .....	15
1.2.6 NMR spin relaxation .....	16
REFERENCES .....	19
CHAPTER 2 .....	27
BINDING KINETICS OF INTERMOLECULAR DOMAIN DOCKING IN THE HAIRPIN RIBOZYME .....	27
ABSTRACT .....	28
2.1 INTRODUCTION .....	29
2.1.1 Intermolecular domain docking .....	29
2.1.2 Surface plasmon resonance .....	33
2.2 MATERIALS AND METHODS .....	33
2.2.1 RNA preparation .....	33
2.2.2 Surface plasmon resonance .....	35
2.2.3 Data analysis .....	36
2.3 RESULTS .....	36
2.3.1 Metal ion concentration associated with docking .....	36
2.3.2 Kinetics of docking for the junctionless hairpin ribozyme .....	37
2.4 DISCUSSION .....	39
2.4.1 Slow, but tight intermolecular docking .....	39
2.4.2 A biological role for the junction .....	40
2.5 CONCLUSION .....	41
APPENDIX .....	42

REFERENCES .....	57
CHAPTER 3 .....	61
THERMODYNAMICS OF RNA TERTIARY STRUCTURE FORMATION IN THE JUNCTIONLESS HAIRPIN RIBOZYME.....	61
ABSTRACT .....	62
3.1 INTRODUCTION .....	63
3.2 MATERIALS AND METHODS.....	64
3.2.1 RNA preparation.....	64
3.2.2 Temperature-dependent surface plasmon resonance .....	66
3.2.3 Data analysis .....	67
3.3 RESULTS .....	68
3.3.1 Rate and dissociation constants for the various hairpin ribozyme constructs .....	68
3.3.2 Temperature-dependent SPR studies .....	70
3.4 DISCUSSION .....	73
3.4.1 The effect of the cleavage-site modification.....	73
3.4.2 Thermodynamic effects of a junction on docking .....	75
3.4.3 Thermodynamic strategies of RNA tertiary structure formation.....	75
3.5 CONCLUSION.....	78
REFERENCES .....	79
CHAPTER 4 .....	83
THE RELATIONSHIP BETWEEN RIBOSE DYNAMICS AND CATALYSIS AT THE CLEAVAGE SITE OF THE LEAD-DEPENDENT RIBOZYME .....	83
ABSTRACT .....	84
4.1 INTRODUCTION .....	85
4.1.1 The lead-dependent ribozyme.....	85
4.1.2 Ribose dynamics and catalysis in the lead-dependent ribozyme .....	88
4.1.3 NMR studies of ribose dynamics .....	92
4.2 MATERIALS AND METHODS AND RESULTS.....	93
4.2.1 RNA preparation and verification.....	93
4.2.2 Ribose dynamics (Power-dependent $T_{1\rho}$ relaxation rates and CPMG) .....	94
4.2.3 Ribose dynamics ( $T_1$ relaxation rates and hNOE experiments).....	100
4.3 DISCUSSION .....	103
4.3.1 Conformational exchange on the $\mu$ s-ms timescale .....	103
4.3.2 Fast motions .....	108
4.4 CONCLUSION.....	108
APPENDIX.....	110
REFERENCES .....	116
CHAPTER 5 .....	121
RIBOSE DYNAMICS OF LOOP A OF HAIRIN RIBOZYME.....	121
ABSTRACT.....	122
5.1 INTRODUCTION .....	123

5.2 MATERIALS AND METHODS.....	126
5.2.1 Sample preparation and verification .....	126
5.2.2 $T_{1\rho}$ relaxation series .....	128
5.2.3 $T_1$ relaxation series.....	131
5.3 DISCUSSION.....	134
5.4 CONCLUSION.....	135
REFERENCES .....	136
CHAPTER 6 .....	139
RNA BINDING SPECIFICITY OF PPR27 .....	139
ABSTRACT.....	140
6.1 INTRODUCTION .....	141
6.2 MATERIALS AND METHODS.....	142
6.3 RESULTS .....	146
6.4 DISCUSSION.....	148
6.5 CONCLUSION.....	148
APPENDIX.....	149
REFERENCES .....	158
CHAPTER 7 .....	161
CONCLUSIONS AND FUTURE DIRECTIONS.....	161
7.1 CONCLUSIONS AND FUTURE DIRECTIONS .....	162
7.1.1 Exploring the kinetics of the hairpin ribozyme.....	162
7.1.2 The effect of cleavage site modifications on the thermodynamic signature.....	162
7.1.3 Ribose dynamics at the cleavage site of the lead-dependent ribozyme .....	163
7.1.4 Ribose dynamics of the hairpin ribozyme .....	163
REFERENCES .....	165

## LIST OF TABLES

Supplemental Table 2-1: SPR Ordering information.....	54
Table 3-1: Kinetic and dissociation constants determined by SPR .....	69
Table 3-2: Kinetic and dissociation constants determined by temperature-dependent SPR .....	71
Table 4-1: Effect of dynamics-function probes .....	90
Table 4-2: Power-dependent $T_{1\rho}$ series.....	99
Table 6-1: SELEX binding scheme .....	145
Table 2-1-1: SPR Ordering information .....	149
Supplemental Table 6-1: Round by round overview .....	151

## LIST OF FIGURES

Figure 1-1: Catalytic mechanism form the hairpin ribozyme.....	6
Figure 1-2: Necessity of conformational rearrangement for docking in the hairpin ribozyme .....	12
Figure 1-3: Conformational capture and induced fit.....	13
Figure 1-4: Timescales accessible using NMR.....	18
Figure 2-1: Junctionless hairpin ribozyme.....	31
Figure 2-2: Docking of the junctionless hairpin ribozyme .....	32
Figure 2-3: Docking of the junctionless hairpin ribozyme .....	38
Supplemental Figure 2-1: Typical streptavidin modification .....	51
Figure 3-1: Constructs utilized.....	65
Figure 3-2: Representative data of 2'-OH loop A and A38C loop B at 25° C.....	69
Figure 3-3: van't Hoff plots for docking.....	72
Figure 3-4 Arrhenius plots for the rates of docking and undocking.....	72
Figure 3-5: Various thermodynamics signatures of RNA tertiary structure formation .....	77
Figure 4-1: Secondary and tertiary structure of the leadzyme .....	87
Figure 4-2: Ribose conformations .....	90
Figure 4-3: Effect of locking the C6 nucleophilic ribose in the C3'-endo conformation ..	91
Figure 4-4: 2', 4' cytidine leadzyme.....	95
Figure 4-5: T <sub>1ρ</sub> relaxation series comparing C6 ribose and a typical helical residue .....	98

Figure 4-6: $T_{1\rho}$ relaxation rates for resolved sites in the $^{13}\text{C}$ , 2'-4' cytidine leadzyme .....	98
Figure 4-7: $T_1$ relaxation series comparing C6 ribose and a typical helical residue.....	101
Figure 4-8: $T_1$ relaxation rates for resolved sites in the $^{13}\text{C}$ , 2', 4' cytidine leadzyme.....	101
Figure 4-9: hNOE ratios for resolved sites in the $^{13}\text{C}$ , 2', 4' cytidine leadzyme.....	102
Figure 4-10 Comparing the C6 ribose to a typical helical ribose over the power-dependent series.....	105
Figure 4-11: Environment of the C10 ribose .....	107
Figure 5-1: Construct of loop A.....	125
Figure 5-2: $^{13}\text{C}$ 2', 4' adenosine loop A.....	127
Figure 5-3: $^{13}\text{C}$ 2', 4' guanosine loop A .....	127
Figure 5-4: Enhanced $T_{1\rho}$ relaxation rate for peak 10 compared to a typical site .....	129
Figure 5-5: $T_{1\rho}$ relaxation rates for $^{13}\text{C}$ 2', 4' adenosine loop A .....	129
Figure 5-6: $T_{1\rho}$ relaxation rates for $^{13}\text{C}$ 2', 4' guanosine loop A.....	130
Figure 5-7: $T_1$ relaxation rates for Peak 9 and 10 $^{13}\text{C}$ 2', 4' adenosine loop A.....	132
Figure 5-8: $T_1$ relaxation rates for $^{13}\text{C}$ 2', 4' adenosine loop A.....	132
Figure 5-9: $T_1$ relaxation rates for $^{13}\text{C}$ 2', 4' guanosine loop A .....	133
Figure 6-1: One round of <i>in vitro</i> selection .....	144
Figure 6-2: Ribogreen assay to access relative enrichment.....	147
Figure 6-3: WebLogo of RNA sequences in the pool after 11 cycles of selections with PPR27 .....	147

## KEY TO ABBREVIATIONS

$\mu$ RIU	Micro-refractive index units
2'-OMe	2'- <i>O</i> -methyl
2WJ	two-way junction or hinged
4WJ	four-way junction
AMP	adenosine monophosphate
CD	circular dichroism
CMP	cytosine monophosphate
CPMG	Carr-Purcell-Meiboom-Gill
ddH <sub>2</sub> O	double-deionized water
<i>E. coli</i>	<i>Escherichia coli</i>
EDC	N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride
EDTA	Edetic acid
FRET	fluorescence resonance energy transfer
GMP	guanosine monophosphate
HIV TAR	human immunodeficiency virus trans-activation response element
HSQC	heteronuclear single quantum coherence
LNA	locked nucleic acid
ncRNA	non-coding ribonucleic acid
NHS	N-Hydroxysuccinimide

NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
$R_1$	longitudinal relaxation rate
$R_{1\rho}$	rotating-frame transverse relaxation rate
RF	radio frequency
RNA	ribonucleic acid
SPR	surface plasmon resonance
$T_1$	longitudinal relaxation time
$T_{1\rho}$	rotating-frame transverse relaxation time
UMP	uridine monophosphate

# **CHAPTER 1**

## **RIBOZYMES AND THE DYNAMICS-FUNCTION RELATIONSHIP**

## 1.1 INTRODUCTION TO RIBOZYMES

### 1.1.1 Ribozymes, a type of non-coding RNA

Non-coding RNA (ncRNA) performs a vast array of biological functions. A recent review profiled over 50 classes of ncRNA on the basis of function and size (1). The extensive research to discover and characterize ncRNA has transitioned the view of RNA from a simple and transitory molecule between DNA and proteins to one of complexity and functional diversity and importance.

Catalytic RNA molecules, or ribozymes, are a type of non-coding RNA. Catalytic RNA was discovered by Thomas Cech and Sidney Altman (2,3). They were jointly awarded the Nobel Prize in Chemistry in 1989 for their discovery. This resulted in a paradigmatic shift, where RNA could be viewed as molecule capable of catalysis (4-6). That initial discovery of the ribozymes, has allowed for a wide range of catalytic RNA to be found and will be discussed below.

Considering catalytic RNA, in a broad sense, provides ubiquitous and essential examples. For instance, the yeast structure of the spliceosome shows a RNA active site for catalysis, for this reason, the Shi lab supported labelling the spliceosome a ribozyme (7,8). For another vital, protein-RNA mega-complex, the ribosome, the same conclusion has been drawn (9). For these systems, the widest definition of a ribozyme is being applied. Even though they are both integrated protein-RNA machines, the protein is serving a structural role and the RNA is involved in catalysis. In contrast, the narrowest definition for a ribozyme would be catalytic RNA that is found *in vivo*, without protein, to perform multiple turnover (10). There is not a ribozyme that has been discovered to date to have all this criteria. Irrespective of the definition one chooses, these are important criteria to consider.

A ubiquitous (bacteria to eukaryotes, including yeast and humans) example of catalytic RNA performing multiple turnover is RNaseP (3). Its function is to process the 5'ends of tRNA precursors (3,11). An excellent crystallographic study shows the enzyme to be comprised of RNA, with two metal ions at the active site and protein only serving a role as a co-factor, which is too far away to make direct contacts to the substrate (12). RNaseP quite closely resembles protein enzymes, with performing multiple turnover catalysis *in vivo* and using cofactors.

Another prevalent class of catalytic RNA is the self-splicing introns, termed, Group I and Group II. They catalyze the removal of the intron they contain and in this manner they are both enzyme and substrate (13-15). Since that class of RNA is excising their own intron there is no need for multiple turnover capacity. The small, self-cleaving enzymes (discussed in 1.1.2) are very similar to the self-splicing introns in that they are comprised of substrate and enzyme and have no need for multiple turnover catalysis *in vivo* (16).

### **1.1.2 RNA as a catalyst**

A very important question to examine is how can RNA, with only four monomers that are chemically and structurally similar, catalyze reactions? As illustrated by the spliceosome and ribosome examples, a RNA active site can exist where protein plays a supporting, structural role. However, RNase P and the Group I and Group II introns demonstrate a RNA active site can be supported by RNA playing a structural role, with protein only as a co-factor. The small, self-cleaving ribozymes go even further in that direction as they are comprised entirely of RNA (10,17-19).

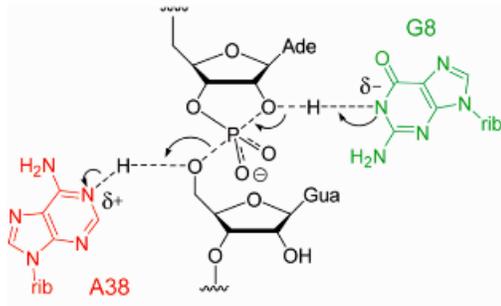
Small, self-cleaving ribozymes include the hairpin ribozyme, the hammerhead ribozyme, and the *Neurospora* Varkud satellite (VS) ribozyme and the hepatitis delta virus (HDV). They are excellent examples of how the structure-function relationship can be satisfied with only

RNA, both *in vivo* and *in vitro*. All of those ribozymes were discovered in the late 1980s to the early 1990s. Very recently (mid 2010s), four new self-cleaving motifs have been discovered by the Breaker lab; they are the twister, twister sister, pistol, and hatchet self-cleaving ribozymes (20,21). If the ribosome and spliceosome are considered ribozymes then it is very easy to conclude that ribozymes are wide-spread and integral in the some of the cell's most important reactions. The twister ribozyme has 2690 examples and is in the genomes of bacteria, fungi, plants, and animals including vertebrates (20). These innovative discoveries have dramatically increased the known *in vivo* frequency of small, self-cleaving ribozymes and allow for them to be thought of as being far more prevalent than initially thought.

The small (30-100 nucleotides), self-cleaving ribozymes catalyze sequence specific cleavage of their own backbone to yield 2', 3'-cyclic phosphate and 5'-OH terminus. They also catalyze the reverse of this reaction (or ligation) of their backbone. In order for this function to be achieved the ribozymes must adopt a specific structure. All of the self-cleaving ribozymes are understood to use a  $S_N2$  mechanism, which invokes in-line nucleophilic attack of the 2'-OH on the scissile phosphate, via a trigonal bipyrimidal transition state (22,23).

Determining the active site chemistry of these ribozymes is difficult, due in part to the challenge of implicating the bases of the nucleotides as the catalytic participants in general acid-base catalysis (16). Not only are the four nucleotides of RNA not very diverse, but their  $pK_a$ 's, free in solution, are not near neutrality. This is in contrast to the  $pK_a$  of the amino acid, histidine, which is commonly implicated in general acid-base catalysis. Due to this view, ribozymes were initially thought to be metalloenzymes, meaning they position metal ions (such as  $Mg^{2+}$ ) in the active site and those metal ions directly participate in the active site chemistry (24).

However, the hairpin ribozyme challenged this conception by being fully active in chemically and substitution inert cobalt hexamine and monovalent salts (25-27). It was proposed, by reexamining pH-profiles of catalysis, that the bases themselves of the hairpin ribozyme were participating in the active site chemistry (28). At present, the simplest catalytic mechanism that is consistent with biochemical data is that the bases of the hairpin ribozyme are catalytic participants in general acid-base catalysis (Figure 1-1) (28,29).



**Figure 1-1: Catalytic mechanism from the hairpin ribozyme.** This shows the active site chemistry, in which G8, of loop A, and A38, of loop B act as general base and a general acid respectively. The catalytic bases aid the nucleophilic attack of the 2'-hydroxyl for cleavage between what has been denoted A-1 and G+1 of loop A. This model has been supported with evidence from the Lilley lab, figure from reference (29).

The hairpin ribozyme provides an excellent opportunity to study the capabilities of RNA as it forms a functional structure both *in vitro* and *in vivo* that is unaided by protein. Also, diffuse metal ions assist in a passive, structural manner but do not participate in the active site chemistry. Understanding how RNA is capable of achieving catalysis in this manner is very important in its own right.

### **1.1.3 Hairpin ribozyme: biological role to therapeutic approaches**

The hairpin ribozyme was discovered in the minus strand of the satellite RNA associated with the Tobacco Ringspot Virus (30,31). There are now three known naturally occurring examples of this ribozyme (32). Its role is to promote rolling circle replication, which has been well reviewed (19). Simply, a circular RNA is replicated by a RNA-dependent, RNA polymerase forming a concatomer, of multiple genomes that are connected. The hairpin ribozyme cuts itself at genome-genome boundaries to yield individual units. It also performs self-ligation to re-circularize individual genomes for packaging. In the plus strand of the same satellite RNA, the hammerhead ribozyme provides the exact same functions.

When the hairpin ribozyme is functioning in its natural biological context it is, in essence, both substrate and enzyme, and multiple turnover is not done, nor is it needed. However, the hairpin ribozyme can be engineered to perform multiple, sequence specific, trans-cleavage reactions (33). This means an engineered hairpin ribozyme could target essential genes of RNA virus genomes, such as HIV, or detrimental RNA transcripts, such as those of oncogenes. This type of treatment has been applied against the hepatitis C virus and against HIV, in HIV positive humans, in successful Phase I clinical trials (34,35). The hairpin ribozyme was the first ribozyme to be used as a therapeutic agent in human clinical trials (36).

### **1.1.4 The RNA World**

The RNA-World Hypothesis is the idea that in a so-called RNA-World, RNA was the molecule responsible for both information storage and function in the primordial world, and this hypothesis has been well reviewed in the recent years (37-40). Subsequently, this period gave way to the era of DNA being favored for information storage and proteins being developed for performing a vast array of functions. However, RNA is still providing informational storage, e.g. mRNA and virus genomes, and ncRNA is still performing a wide-array of functions. For the RNA World to exist, self-replicating RNA would have been necessary (41). To this end, self-replicating RNA has been recently developed *in vitro*, making the RNA World Hypothesis even more plausible (42-46). With these significant advancements, it seems even more possible that in the RNA World, a ribozyme existed that was only RNA and could catalyze self-replication. The ribosome has been tied to such an enzyme, before transitioning to its modern state (9). Catalytic RNA provides a potentially powerful window back in time to the RNA World.

## **1.2 DYNAMICS-FUNCTION RELATIONSHIP**

### **1.2.1 Necessity of conformational re-arrangement in the hairpin ribozyme**

The hairpin ribozyme is a well-studied, self-cleaving RNA (18,19,36,47). Cleavage occurs between what has been denoted G+1 and A-1 in loop A, with G8 of loop A and A38 of loop B implicated as acting a general base and a general acid, respectively (Figure 1-1). The hairpin ribozyme exists naturally in a four-way junction with loop A and loop B on adjoining arms. Loop A and loop B come together in a minor-groove, minor-groove fashion, termed docking and there is an extensive interface between loop A and loop B and the docked structure has been studied extensively by x-ray crystallography (48-55). The crystal structures illustrate the important features of the docked structure, which include the in-line ( $S_N2$ ) geometry at the

cleavage site, the catalytic residues, G8 and A38, in close proximity to the cleavage site, and stabilizing interactions of the docked state. They also illustrate the extensive interface between loop A and loop B which includes the Watson-Crick base-pair between G+1 of loop A and C25 of loop B, a ribose zipper, and the pocket of U42 of loop B (these are illustrated and described further in Figure 2-2).

Also, the NMR solutions structures for loop A and loop B, in the absence of each molecules partner loop, have been determined (56,57). By studying the ground state structures of loop A and loop B and the docked complex of loop A and loop B, it is clear that extensive conformational change is necessary to form the docked complex, which is inherently necessary for function. Our lab has effectively detailed these features and they will be discussed in greater detail later in this thesis (Figure 2-2) (58). For now, the focus will be on one feature that has dynamics essential to function and is exploited in subsequent studies in this thesis.

As shown in Figure 1-2, the G+1 base of loop A is an important element in the docked structure of the hairpin ribozyme. In the ground state structure of loop A the G+1 base is incorporated into the noncanonical helix of loop A. In the docked structures, the G+1 base is incorporated into the noncanonical helix of loop B, forming a Watson-Crick base pair with the C25 base of loop B. This is an essential feature in the formation of the in-line geometry, which is necessary for cleavage. The G+1 residue of loop A adopting that conformation illustrates how structure relates to function, or the paradigmatic structure-function relationship. How the G+1 residue of loop A adopts this conformation, requires exploring how dynamics relate to function, or use an analogy, the dynamics-function relationship. This is one example of numerous necessary rearrangements in both loop A and loop B that is essential to forming the docked structure.

### 1.2.2 Double conformational capture

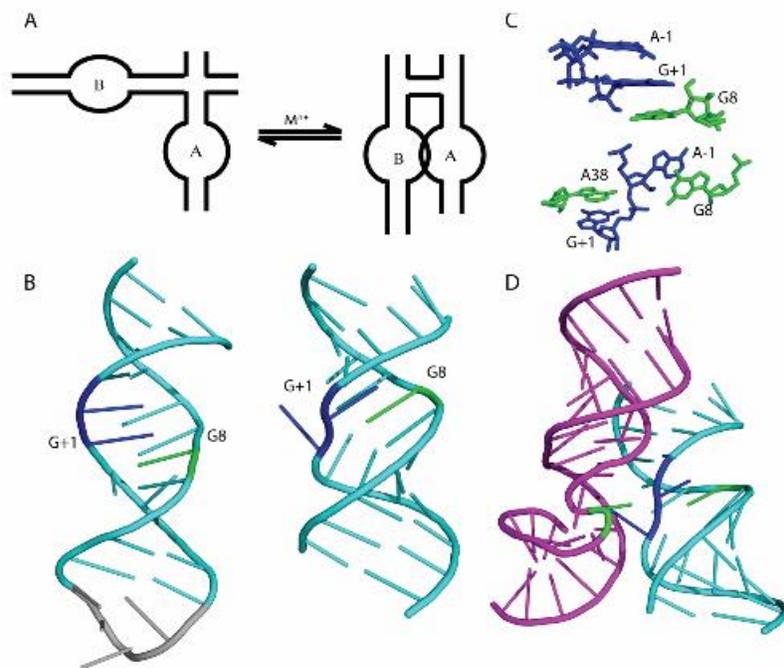
With the goal of describing how, from a molecular dynamics standpoint, the docked structure is formed there are two simple, but conceptually useful, paradigms for binding where conformational change is apparent (Figure 1-3). They are conformational capture (sometimes called tertiary capture or conformational selection) and induced fit (59). In the conformational capture model there is a ground state conformation that is present a majority of the time and an excited or “bound-like” state that exists a small minority of the time and resembles the bound structure. In this model, binding can only occur with the excited state and not the ground state. In the induced fit model, binding occurs with a ground state conformation and induces conformational change. Evidence for conformational capture involves finding the existence of these excited states, in the absence of the binding partner. These states have been found successfully in protein using NMR spin relaxation spectroscopy (60).

An outstanding example of how these excited states can be found by applying NMR spin relaxation spectroscopy, to study conformational dynamics, comes from the Wright Lab with their comprehensive study of dihydrofolate reductase (61). What they show is there is conformational flexibility (on the  $\mu$ s-ms regime) in regions that need to remodel to achieve the subsequent state in the catalytic cycle. They clearly demonstrate that there is a lowly populated, excited state that resembles the next structure in the enzymatic pathway and this allows the enzyme to proceed on its kinetic path. They tie the conformational transitions to limiting the global transitions in the pathway (62). Their work on conformational dynamics as it relates to function is demonstrative of obtainable insight with this strategy.

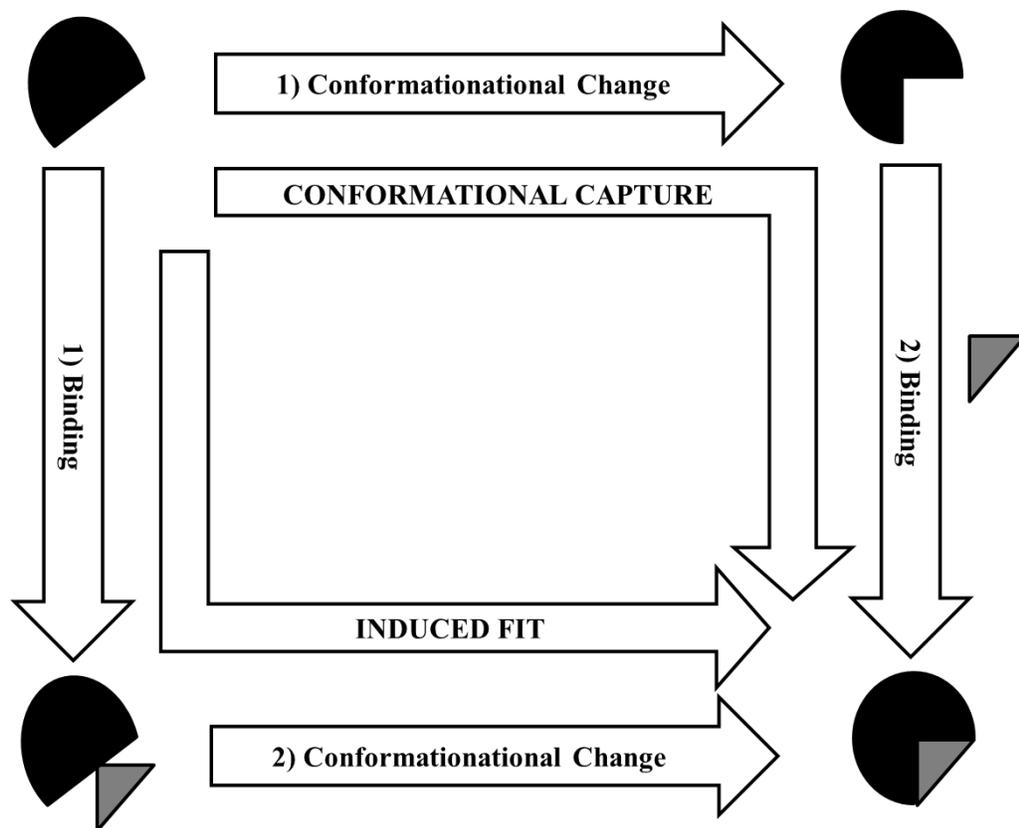
Ultimately, we are looking to test how conformational dynamics relate to function in the hairpin ribozyme. As pointed out, loop A and loop B undergo major rearrangements from their

undocked to docked state. We hypothesize that major rearrangements occur, in both loop A and loop B, prior to docking and thus we hypothesize that docking occurs via double conformational capture. This means that docking only occurs when both loop A and loop B are in their excited states, resembling the dock structure. Referring back to G+1 of loop A as an example, that base is incorporated into the non-canonical helix of loop B in the docked structure (56). One could imagine that base is flipped out of its non-canonical helix a minority of the time (as it is in the docked structure) and that could be a necessary event to allow docking to occur (48).

In order to test the hypothesis that loop A and loop B are sampling “docked-like” conformations it very useful to study the loops in the absence of each other. In this approach, the detection of “docked-like” conformations is due to conformational dynamics. As illustrated in the above example, evidence for conformational capture exists in finding the excited states, via conformational dynamics, in the absences of their binding partner. To this end, we are studying the hairpin ribozyme in its junctionless form, meaning loop A and loop B reside on separate molecules which allows for dynamics of loop A and loop B to be studied in the absence of each other. Our studies of conformational dynamics of the hairpin ribozyme will enable selective testing of the functional importance of the dynamics exhibited.



**Figure 1-2: Necessity of conformational rearrangement for docking in the hairpin ribozyme.** The cation-driven process of docking is illustrated using the four-way junction (panel A). This figure illustrates an important transition of the G+1 residue of loop A that is essential in forming the docked complex. On the left side of panel B, is the loop A construct used in recent molecular dynamic simulations, done by our lab in collaboration with the Feig lab. It is representative of the ground state of loop A. The conformation of loop A on right of panel B, is the conformation of loop A in the dock structure of loop A and loop B (loop B removed for clarity). Then loop A and loop B (purple) are show together (panel D), for reference (PDB: 2OUE). Panel C illustrates distinct differences between the ground state of loop A in the absence of loop B and the docked structure. Both the catalytic G8 residue and the cleavage site between A-1 and G+1 is highlighted here. The loop A construct for molecular dynamics was developed by Dr. Patrick Ochieng originating from coordinates from the NMR solution structure, graciously shared with us from the Tinoco lab (56). The crystal structure of the docked complex was solved by the Wedekind lab (52). This figure is from a recent publication from our lab (63).



**Figure 1-3: Conformational capture and induced fit.** The structure the black binding partner has a different conformation in its unbound state (upper left) and bound state (lower right). The two models induced fit or conformational capture, provide distinct paths. Either binding of the grey partner induces conformational change, termed induced fit. Or the grey partner only binds when the black partner resembles the bound structure, termed conformational capture.

### **1.2.3 Kinetics, thermodynamics, and dynamics of ribozymes**

The selection of the junctionless construct of the hairpin ribozyme enables spectroscopic studies of conformational dynamics (Chapter 5). (The constructs with junctions, studied extensively by others, are discussed in 1.2.6, Chapter 2, and Chapter 3.) Also, the kinetics of docking will be characterized (Chapter 2), which to our knowledge, has only been done by our lab (58). The thermodynamic signature for this interaction, with important considerations for the cleavage site modification will be presented (Chapter 3), which was also previously unknown. The kinetics and thermodynamic signature provides insight into the nature of docking. Conformational dynamics provide insight into how, from the standpoint of molecular motion, loop-loop docking is achieved. The functional importance of these dynamics is an important topic of future studies.

Another ribozyme that provides important insight into the dynamics-function relationship is the lead-dependent ribozyme or leadzyme (64). In both the NMR solution structure and the x-ray crystallographic structure, the cleavage site is not consistent with the proposed in-line geometry and also has poor agreement with biochemical studies (64-67). This necessitates that conformational dynamics are essential to formation of the active structure, and this is therefore a good model system to study how RNA dynamics relate to function. The cleavage site ribose dynamics using conformationally restricted probes and NMR spin relaxations will also illustrate the important link between dynamics and function (Chapter 4).

### **1.2.4 Dynamics as compensation for RNA's inherent lack of chemical diversity**

Both the hairpin ribozyme and leadzyme undergo major conformational rearrangement to obtain a catalytically active conformation. Importantly, RNA dynamics are critical to function as well (68). RNA has been described as having a rough or rugged energy landscape, where many

different conformations are accessible to a single molecule that have different energies and high intervening barriers to transition (69-71).

The rugged energy landscape, or potential dynamic nature, is likely the answer to how a molecule with four chemically similar monomers is capable of such a vast array of functions (68,72). An exemplary case is the HIV-1 TAR (transcriptional response element) studied by Al-Hashimi and co-workers (59). The HIV-1 TAR has been shown, in the absence of ligand, to have the ability to adopt seven distinctly different conformations in the presence of co-factors. This is a prime example of dynamics being essential to structure formation, which is critical to function.

### **1.2.5 Surface plasmon resonance**

The hairpin naturally exists as a four-way junction. The four-way junction and two-way junction (hinged) constructs of the hairpin ribozyme have been studied extensively (29,73-76). They are both unimolecular constructs, as loop A and loop B reside on adjoining arms and they will be discussed in more detail in Chapter 2 and 3. The junctionless system by comparison, is significantly less-studied (see Chapter 2) (77,78). Studying the junctionless form, in addition to making the previously mentioned dynamics studies possible, also allows for an investigation into the biological role of the junction and provides an excellent opportunity to study RNA tertiary structure formation in isolation.

Since, in the junctionless form, loop A and loop B are on separate molecules, this opens the possibility of using a variety of ensemble techniques unavailable to the unimolecular systems. In this thesis, surface plasmon resonance (SPR) was used to studying the docking interaction. This technique detects mass by exploiting a property of total internal reflection of light interacting with the molecule(s) (79). It also allows for bi-molecular interactions to be studied in real time (80).

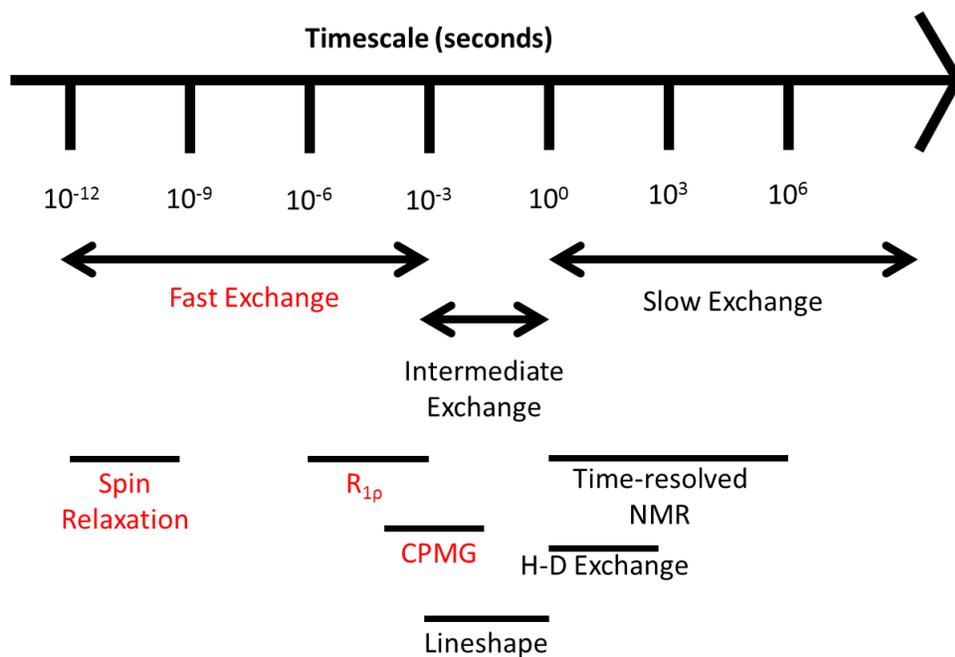
SPR is challenging, in that relies on the development of specific immobilization and regeneration methods for each interaction. However, it is an information rich technique because, unlike most ensemble techniques, one can determine association and dissociation rates as well as the equilibrium dissociation constant using small amounts of label-free molecules. Commonly, ensemble techniques rely on labels such as radioactive or florescent labels and do not allow determination of kinetic parameters (as discussed further in Chapter 2).

### **1.2.6 NMR spin relaxation**

RNA dynamics are essential in the systems being studied here, as well as being important to the functional diversity of RNA in general (81,82). There are a variety of experimental and simulation techniques available to study important RNA dynamics (68). Of these techniques, NMR is unrivaled in its ability to probe multiple timescales, in a site-specific manner (83). For instance, if we consider the ribose at the cleavage site of leadzyme, it could have two populations, one that is the inactive ground state and one that is the functionally relevant, excited state. This idea could also be true for loop A in the absence of loop B; there could be two populations for the ribose at the cleavage site, one corresponding to the ground state conformation and one that is representative of the “docked-like” conformation. If the two conformations have different chemical environments that yield different chemical shifts, then NMR is an excellent tool to study these populations (60,62). The rate at which the two populations are exchanging determines the appropriate technique, hence linking molecular motion to the available repertoire of NMR techniques (Figure 1-4) (84). If they are in the slow-exchange limit then the two populations can be seen in the NMR spectrum as two peaks. If the exchange rate falls within intermediate-exchange then a broadening of the two peaks can be

seen. If the exchange rate is within the fast-exchange limit, then a single peak will be seen and spin relaxation techniques must be utilized; this is the case with leadzyme.

Determining the conformational dynamics of a ribose within the fast-exchange limit is difficult, due to magnetic transfer within the ribose ring with uniformly labelled samples. Our lab has developed an innovative, site-specific labelling protocol that makes these spin-relaxation techniques possible (as discussed in greater detail in Chapter 4) (85,86). If there are two states for a ribose at the cleavage site, for example in these ribozymes, then this techniques reports on the rate of exchange between the populations, the percentage of each population (e.g. 95% ground state and 5% excited state), and the chemical shift difference between the two populations, which is effected by the chemical environment (60). Beyond studying the existence of conformational dynamics of ribose sugars on various timescales, the functional importance of these dynamics can be studied via conformationally restricted probes (64). How conformationally restricted probes relate dynamics to function in the leadzyme will be discussed in Chapter 4. Potential use of conformationally restricted probes to assess the relationship of ribose dynamics to docking in the hairpin ribozyme will be discussed in Chapter 5 and the Future Directions section. Incorporation of dynamics-function studies allow for an understanding that goes beyond static renderings of a given interaction or system and allow for the appreciation of how interactions are formed.



**Figure 1-4: Timescales accessible using NMR.** In this depiction of selected NMR experiments and the time scales they probe the fast exchange regime has been emphasized (red) as well as the experiments that were utilized in this thesis.

## **REFERENCES**

## REFERENCES

1. Cech, T.R. and Steitz, J.A. (2014) The Noncoding RNA Revolution-Trashing Old Rules to Forge New Ones. *Cell*, **157**, 77-94.
2. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, **31**, 147-157.
3. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849-857.
4. (2002) Ribozymes, the first 20 years. *Biochemical Society transactions*, **30**, 1162-1166.
5. Bevilacqua, P.C. (2015) The wonder of RNA: a personal reflection of the last 20 years. *RNA (New York, N.Y.)*, **21**, 515-516.
6. Doherty, E.A. and Doudna, J.A. (1999) Ribozyme structures and mechanisms. *Annual review of biochemistry*, **69**, 597-615.
7. Yan, C., Hang, J., Wan, R., Huang, M., Wong, C.C. and Shi, Y. (2015) Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science (New York, N.Y.)*, **349**, 1182-1191.
8. Hang, J., Wan, R., Yan, C. and Shi, Y. (2015) Structural basis of pre-mRNA splicing. *Science (New York, N.Y.)*, **349**, 1191-1198.
9. Steitz, T.A. and Moore, P.B. (2003) RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends in biochemical sciences*.
10. Hoogstraten, C.G. and Sumita, M. (2007) Structure-function relationships in RNA and RNP enzymes: recent advances. *Biopolymers*, **87**, 317-328.
11. Guerriertakada, C. and Altman, S. (1984) Catalytic Activity of an Rna Molecule Prepared by Transcription In vitro. *Science*, **223**, 285-286.
12. Reiter, N.J., Osterman, A., Torres-Larios, A., Swinger, K.K., Pan, T. and Mondragon, A. (2010) Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, **468**, 784-789.
13. Michel, F., Umesono, K. and Ozeki, H. (1989) Comparative and functional anatomy of group II catalytic introns--a review. *Gene*, **82**, 5-30.
14. Michel, F. and Ferat, J.L. (1995) Structure and activities of group II introns. *Annu Rev Biochem*, **64**, 435-461.

15. Cech, T.R. (1990) Self-splicing of group I introns. *Annu Rev Biochem*, **59**, 543-568.
16. Fedor, M.J. and Williamson, J.R. (2005) The catalytic diversity of RNAs. *Nat Rev Mol Cell Biol*, **6**, 399-412.
17. Fedor, M.J. (2008) Comparative enzymology and structural biology of RNA self-cleavage. *Annual review of biophysics*, **38**, 271-299.
18. Doherty, E. and Doudna, J. (2000) Ribozyme structures and mechanisms. *Annual review of biochemistry*, **69**, 597-615.
19. Ferré-D'Amaré, A.R. and Scott, W.G. (2010) Small self-cleaving ribozymes. *Cold Spring Harbor perspectives in biology*, **2**.
20. Roth, A., Weinberg, Z., Chen, A.G., Kim, P.B., Ames, T.D. and Breaker, R.R. (2014) A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nat Chem Biol*, **10**, 56-60.
21. Weinberg, Z., Kim, P.B., Chen, T.H., Li, S., Harris, K.A., Lünse, C.E. and Breaker, R.R. (2015) New classes of self-cleaving ribozymes revealed by comparative genomics analysis. *Nature chemical biology*, **11**, 606-610.
22. Jamal, M.B., Wayne, L.G. and George, B. (1986) Non-enzymatic cleavage and ligation of RNAs complementary to a plant virus satellite RNA. *Nature*, **323**.
23. van Tol, H., Buzayan, J.M., Feldstein, P.A., Eckstein, F. and Bruening, G. (1990) Two autolytic processing reactions of a satellite RNA proceed with inversion of configuration. *Nucleic Acids Res*, **18**, 1971-1975.
24. Cech, T. (2002) Ribozymes, the first 20 years. *Biochemical Society transactions*, **30**, 1162-1166.
25. Hampel, A. and Cowan, J.A. (1997) A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage. *Chemistry & biology*, **4**, 513-517.
26. Young, K.J., Gill, F. and Grasby, J.A. (1997) Metal ions play a passive role in the hairpin ribozyme catalysed reaction. *Nucleic acids research*, **25**, 3760-3766.
27. Nesbitt, S., Hegg, L.A. and Fedor, M.J. (1997) An unusual pH-independent and metal-ion-independent mechanism for hairpin ribozyme catalysis. *Chemistry & biology*, **4**, 619-630.
28. Bevilacqua, P.C. (2003) Mechanistic considerations for general acid-base catalysis by RNA: revisiting the mechanism of the hairpin ribozyme. *Biochemistry*, **42**, 2259-2265.

29. Kath-Schorr, S., Wilson, T.J., Li, N.-S.S., Lu, J., Piccirilli, J.A. and Lilley, D.M. (2012) General acid-base catalysis mediated by nucleobases in the hairpin ribozyme. *Journal of the American Chemical Society*, **134**, 16717-16724.
30. Gerlach, W.L., Buzayan, J.M., Schneider, I.R. and Bruening, G. (1986) Satellite tobacco ringspot virus RNA: biological activity of DNA clones and their in vitro transcripts. *Virology*, **151**, 172-185.
31. Buzayan, J.M., Gerlach, W.L. and Bruening, G. (1986) Satellite tobacco ringspot virus RNA: A subset of the RNA sequence is sufficient for autolytic processing. *Proc Natl Acad Sci U S A*, **83**, 8859-8862.
32. DeYoung, M., Siwkowski, A.M., Lian, Y. and Hampel, A. (1995) Catalytic properties of hairpin ribozymes derived from Chicory yellow mottle virus and arabis mosaic virus satellite RNAs. *Biochemistry*, **34**, 15785-15791.
33. Zeller, S.J. and Kumar, P. (2011) RNA-based gene therapy for the treatment and prevention of HIV: from bench to bedside. *The Yale journal of biology and medicine*, **84**, 301-309.
34. Jarczak, D., Korf, M., Beger, C., Manns, M.P. and Krüger, M. (2005) Hairpin ribozymes in combination with siRNAs against highly conserved hepatitis C virus sequence inhibit RNA replication and protein translation from hepatitis C virus subgenomic replicons. *The FEBS journal*, **272**, 5910-5922.
35. Wong-Staal, F., Poeschla, E.M. and Looney, D.J. (1998) A controlled, Phase 1 clinical trial to evaluate the safety and effects in HIV-1 infected humans of autologous lymphocytes transduced with a ribozyme that cleaves HIV-1 RNA. *Human gene therapy*, **9**, 2407-2425.
36. Fedor, M.J. (2000) Structure and function of the hairpin ribozyme. *Journal of molecular biology*, **297**, 269-291.
37. Cech, T.R. (2012) The RNA worlds in context. *Cold Spring Harb Perspect Biol*, **4**, a006742.
38. Yarus, M. (2011) Getting past the RNA world: the initial Darwinian ancestor. *Cold Spring Harb Perspect Biol*, **3**.
39. Breaker, R.R. (2012) Riboswitches and the RNA world. *Cold Spring Harb Perspect Biol*, **4**.
40. Robertson, M.P. and Joyce, G.F. (2012) The origins of the RNA world. *Cold Spring Harb Perspect Biol*, **4**.

41. Bartel, D.P. and Unrau, P.J. (1999) Constructing an RNA world. *Trends Cell Biol*, **9**, M9-M13.
42. Baskerville, S. and Bartel, D.P. (2002) A ribozyme that ligates RNA to protein. *Proc Natl Acad Sci U S A*, **99**, 9154-9159.
43. Beaudry, A.A. and Joyce, G.F. (1992) Directed evolution of an RNA enzyme. *Science*, **257**, 635-641.
44. Johnston, W.K., Unrau, P.J., Lawrence, M.S., Glasner, M.E. and Bartel, D.P. (2001) RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science*, **292**, 1319-1325.
45. Robertson, M.P. and Joyce, G.F. (2014) Highly efficient self-replicating RNA enzymes. *Chem Biol*, **21**, 238-245.
46. Horning, D.P. and Joyce, G.F. (2016) Amplification of RNA by an RNA polymerase ribozyme. *Proc Natl Acad Sci U S A*, **113**, 9786-9791.
47. Cochrane, J.C. and Strobel, S.A. (2008) Catalytic strategies of self-cleaving ribozymes. *Accounts of chemical research*, **41**, 1027-1035.
48. Rupert, P.B. and Ferré-D'Amaré, A.R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, **410**, 780-786.
49. Rupert, P.B., Massey, A.P., Sigurdsson, S.T. and Ferré-D'Amaré, A.R. (2002) Transition state stabilization by a catalytic RNA. *Science (New York, N.Y.)*, **298**, 1421-1424.
50. Ferré-D'amaré, A.R. and Rupert, P.B. (2002) The hairpin ribozyme: from crystal structure to function. *Biochemical Society transactions*, **30**, 1105-1109.
51. Alam, S., Grum-Tokars, V., Krucinska, J., Kundracik, M.L. and Wedekind, J.E. (2005) Conformational heterogeneity at position U37 of an all-RNA hairpin ribozyme with implications for metal binding and the catalytic structure of the S-turn. *Biochemistry*, **44**, 14396-14408.
52. Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. and Wedekind, J.E. (2006) Water in the active site of an all-RNA hairpin ribozyme and effects of Gaa8 base variants on the geometry of phosphoryl transfer. *Biochemistry*, **45**, 686-700.
53. MacElrevey, C., Salter, J.D., Krucinska, J. and Wedekind, J.E. (2008) Structural effects of nucleobase variations at key active site residue Ade38 in the hairpin ribozyme. *RNA (New York, N.Y.)*, **14**, 1600-1616.

54. Spitale, R.C., Volpini, R., Heller, M.G., Krucinska, J., Cristalli, G. and Wedekind, J.E. (2009) Identification of an imino group indispensable for cleavage by a small ribozyme. *Journal of the American Chemical Society*, **131**, 6093-6095.
55. Liberman, J.A., Guo, M., Jenkins, J.L., Krucinska, J., Chen, Y., Carey, P.R. and Wedekind, J.E. (2012) A transition-state interaction shifts nucleobase ionization toward neutrality to facilitate small ribozyme catalysis. *Journal of the American Chemical Society*, **134**, 16933-16936.
56. Cai, Z. and Tinoco, I. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.
57. Butcher, S.E., Allain, F.H. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature structural biology*, **6**, 212-216.
58. Sumita, M., White, N.A., Julien, K.R. and Hoogstraten, C.G. (2013) Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis. *RNA Biol*, **10**, 425-435.
59. Zhang, Q. and Al-Hashimi, H.M. (2009) Domain-elongation NMR spectroscopy yields new insights into RNA dynamics and adaptive recognition. *RNA (New York, N.Y.)*, **15**, 1941-1948.
60. Baldwin, A.J. and Kay, L.E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nature chemical biology*, **5**, 808-814.
61. Boehr, D.D., McElheny, D., Dyson, H.J. and Wright, P.E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis. *Science (New York, N.Y.)*, **313**, 1638-1642.
62. Boehr, D.D., Dyson, H.J. and Wright, P.E. (2008) Conformational relaxation following hydride transfer plays a limiting role in dihydrofolate reductase catalysis. *Biochemistry*, **47**, 9227-9233.
63. Ochieng, P.O., White, N.A., Feig, M. and Hoogstraten, C.G. (2016) Intrinsic Base-Pair Rearrangement in the Hairpin Ribozyme Directs RNA Conformational Sampling and Tertiary Interface Formation. *The journal of physical chemistry. B*.
64. Julien, K.R., Sumita, M., Chen, P.-H.H., Laird-Offringa, I.A. and Hoogstraten, C.G. (2008) Conformationally restricted nucleotides as a probe of structure-function relationships in RNA. *RNA (New York, N.Y.)*, **14**, 1632-1643.
65. Hoogstraten, C.G., Legault, P. and Pardi, A. (1998) NMR solution structure of the lead-dependent ribozyme: evidence for dynamics in RNA catalysis. *Journal of molecular biology*, **284**, 337-350.

66. Wedekind, J.E. and McKay, D.B. (1999) Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nature structural biology*, **6**, 261-268.
67. Wedekind, J.E. and McKay, D.B. (2003) Crystal structure of the leadzyme at 1.8 Å resolution: metal ion binding and the implications for catalytic mechanism and allo site ion regulation. *Biochemistry*, **42**, 9554-9563.
68. Al-Hashimi, H.M. and Walter, N.G. (2008) RNA dynamics: it is about time. *Curr Opin Struct Biol*, **18**, 321-329.
69. Pljevaljčić, G., Klostermeier, D. and Millar, D.P. (2005) The tertiary structure of the hairpin ribozyme is formed through a slow conformational search. *Biochemistry*, **44**, 4870-4876.
70. Pan, T. and Uhlenbeck, O.C. (1993) Circularly permuted DNA, RNA and proteins--a review. *Gene*, **125**, 111-114.
71. Mustoe, A.M., Brooks, C.L. and Al-Hashimi, H.M. (2014) Hierarchy of RNA functional dynamics. *Annu Rev Biochem*, **83**, 441-466.
72. Al-Hashimi, H.M. (2013) NMR studies of nucleic acid dynamics. *J Magn Reson*, **237**, 191-204.
73. Tan, E., Wilson, T.J., Nahas, M.K., Clegg, R.M., Lilley, D.M. and Ha, T. (2003) A four-way junction accelerates hairpin ribozyme folding via a discrete intermediate. *Proc Natl Acad Sci U S A*, **100**, 9308-9313.
74. Wilson, T., Nahas, M., Araki, L., Harusawa, S., Ha, T. and Lilley, D. (2007) RNA folding and the origins of catalytic activity in the hairpin ribozyme. *Blood cells, molecules & diseases*, **38**, 8-14.
75. Klostermeier, D. and Millar, D.P. (2000) Helical junctions as determinants for RNA folding: origin of tertiary structure stability of the hairpin ribozyme. *Biochemistry*, **39**, 12970-12978.
76. Pinard, R., Hampel, K.J., Heckman, J.E., Lambert, D., Chan, P.A., Major, F. and Burke, J.M. (2001) Functional involvement of G8 in the hairpin ribozyme cleavage mechanism. *The EMBO journal*, **20**, 6434-6442.
77. Hampel, K.J., Walter, N.G. and Burke, J.M. (1998) The solvent-protected core of the hairpin ribozyme-substrate complex. *Biochemistry*, **37**, 14672-14682.
78. Butcher, S.E., Heckman, J.E. and Burke, J.M. (1995) Reconstitution of hairpin ribozyme activity following separation of functional domains. *The Journal of biological chemistry*, **270**, 29648-29651.

79. de Mol, N.J. and Fischer, M.J. (2010) Surface plasmon resonance: a general introduction. *Methods in molecular biology (Clifton, N.J.)*, **627**, 1-14.
80. Myszka, D.G. (1998) Improving biosensor analysis. *Journal of molecular recognition : JMR*, **12**, 279-284.
81. Russell, R., Zhuang, X., Babcock, H.P., Millett, I.S., Doniach, S., Chu, S. and Herschlag, D. (2002) Exploring the folding landscape of a structured RNA. *Proc Natl Acad Sci U S A*, **99**, 155-160.
82. Russell, R., Millett, I.S., Tate, M.W., Kwok, L.W., Nakatani, B., Gruner, S.M., Mochrie, S.G., Pande, V., Doniach, S., Herschlag, D. *et al.* (2002) Rapid compaction during RNA folding. *Proc Natl Acad Sci U S A*, **99**, 4266-4271.
83. Rinnenthal, J., Buck, J., Ferner, J., Wacker, A., Fürtig, B. and Schwalbe, H. (2011) Mapping the landscape of RNA dynamics with NMR spectroscopy. *Accounts of chemical research*, **44**, 1292-1301.
84. Palmer, A.G. (2004) NMR characterization of the dynamics of biomacromolecules. *Chemical reviews*, **104**, 3623-3640.
85. Johnson, J.E., Jr., Julien, K.R. and Hoogstraten, C.G. (2006) Alternate-site isotopic labeling of ribonucleotides for NMR studies of ribose conformational dynamics in RNA. *J Biomol NMR*, **35**, 261-274.
86. Johnson, J.E. and Hoogstraten, C.G. (2008) Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using <sup>13</sup>C NMR spin relaxation and specific isotope labeling. *Journal of the American Chemical Society*, **130**, 16757-16769.

## CHAPTER 2

### BINDING KINETICS OF INTERMOLECULAR DOMAIN DOCKING IN THE HAIRPIN RIBOZYME

**Portions of this chapter have been published in the following manuscripts:**

Sumita, M., White, N. A., Julien, K. R., and Hoogstraten, C. G. (2013) Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis.

*RNA Biology 10*, 425–35.

Hoogstraten, C. G., Sumita, M., and White, N. A. (2014) Unraveling the thermodynamics and kinetics of RNA assembly: surface plasmon resonance, isothermal titration calorimetry, and circular dichroism.

*Methods in Enzymology 549*, 407–32.

## **ABSTRACT**

The hairpin ribozyme naturally exists in a RNA four-way junction, with loop A and loop B on adjacent arms. Here it is being studied in its trans or junctionless form, meaning loop A and loop B are residing on separate molecules. Significant structural rearrangements in the loop A and loop B are essential for interloop binding, or what has been termed docking. Docking is the obligatory step that precedes catalysis. In order to study the dynamics of loop A in the absence of loop B and vice-versa, via spectroscopic means, it is necessary to study the hairpin ribozyme in its junctionless form.

RNA tertiary structure formation is an intricate and specific process. Despite a distinct role not being easily deducible, metal ions are a very important factor in the tertiary structure that RNA adopts. To determine the appropriate metal ion concentration associated with docking, we developed a novel CD assay. It was determined that sub-millimolar concentrations of cobalt-hexamine are sufficient for driving the docking process.

This determination, of metal ion concentration, guided the surface plasmon resonance (SPR) assays that allowed for the reporting of the binding affinity of loop A and loop B and the first reporting of kinetic rates of association and dissociation of loop A and loop B. It was found that the docking process is slow but the affinity is relatively tight.

This slow association rate is likely due the significant rearrangements necessary for docking. This is hypothesized to occur via double conformational capture. The stability of the docked complex, as evident by a slow dissociation rate constant is most likely due to the extensive interface that is formed between loop A and loop B, even in the absence of the native junction.

## 2.1 INTRODUCTION

### 2.1.1 Intermolecular domain docking

When RNA adopts a tightly folded structure or specific tertiary structure there are negative charges, inherent to the phosphodiester backbone, which are juxtaposed. Positive charges, in the form of metal ions, are an important counterbalance to this unfavorable interaction (1-3). Despite dogma in the RNA structure field for years being that ions bound to a specific site are the essential ions in tertiary structure, emerging evidence supports a model that diffuse ions or atmospheric ions are actually more important in driving RNA tertiary structure formation (4).

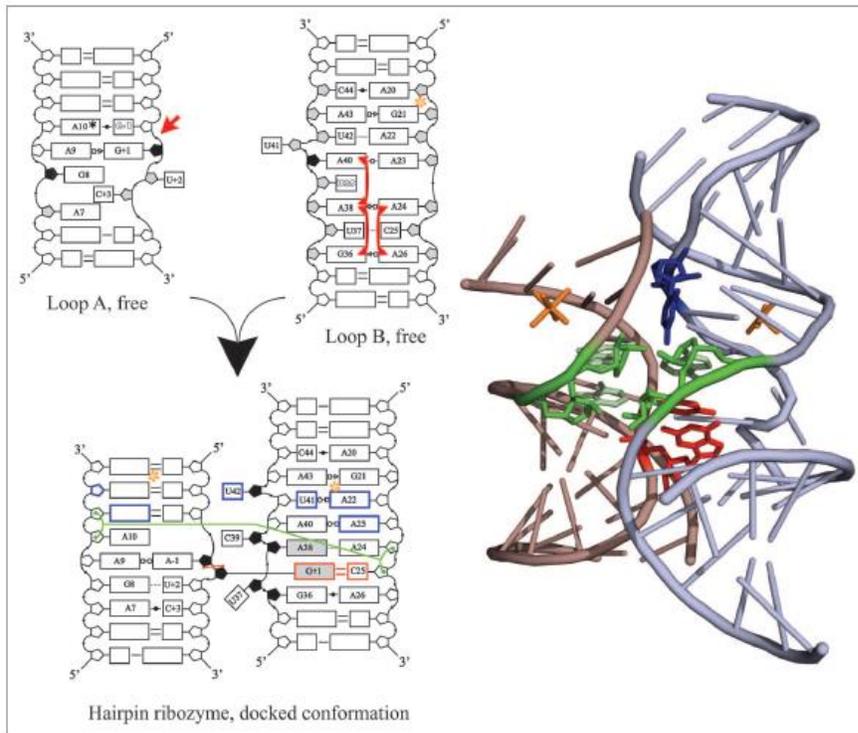
Loop A and loop B of the hairpin ribozyme dock, meaning that the two loops have come together in a minor-groove, minor-groove fashion with the catalytic participants in their proper positions to help facilitate  $SN_2$  in-line, nucleophilic attack (5,6). This structure has been termed the docked or pre-catalytic state. Extensive crystallographic studies have shown essentially no variance in the critical components of the hairpin ribozyme from the pre-catalytic state, to the transition state analogues, to the post-catalytic state (7,8). Also, whether investigating the four-junction or the junctionless form of the hairpin ribozyme by x-ray crystallography, the catalytic core is also structurally conserved (9,10).

In this chapter our studies of intermolecular domain docking of loop A and loop B, which have been published, will be presented in this chapter (11,12). The constructs used can be seen in Figure 2-1. The NMR structures of loop A and loop B solved independently show the necessity for considerable structural rearrangement of the loops to reach the docked structure solved by crystallography (Figure 2-2) (13-15).

Figure 2-2 shows there are significant rearrangements that can be observed between the undocked and docked states of the hairpin ribozyme. The G+1 base (red) transitions from being in non-conical helix of loop A in the undocked state. Then when docked, the G+1 base hydrogen-bonds with C25 base of loop B; this incorporates itself into non-conical helix of loop B. This critical residue is utilized in these docking studies. Mutation of this base (G+1A) abolishes docking and serves as our negative control. Another important structural element is the ribose zipper (green): a stabilizing element that utilizes hydrogen-bonding with the 2'-OH, inherent to RNA. The U42 residue of loop B is incorporated into the non-conical helix of loop B in the absence of loop B. In the docked form the U42 residue is flipped out of the non-conical helix of loop B and is in an elaborate pocket.

Other information represented in Figure 2-2 includes: the black pentagons indicate the C2'-endo conformation of the ribose, while white indicate C3'-endo, and shaded (grey) indicate an interconversion between the two major ribose conformations. It is worth pointing out that since the structures of loop A and loop B were solved using NMR spectroscopy interconversion is a possible answer. In contrast, the docked structure solved by x-ray crystallography results in two major ribose conformations being a possible obtainable outcome but not an interconversion between the two.





**Figure 2-2: Docking of the junctionless hairpin ribozyme.** Description of rearrangements and extensive loop-loop interface in text. Figure reproduced with permission from reference (11).

### **2.1.2 Surface plasmon resonance**

As mentioned above, the presence of metal ions is a critical component of RNA tertiary structure formation. Cobalt hexamine was used in these studies and it is chemically and exchange inert. It can support RNA tertiary structure formation as a diffuse metal ion. That choice was made because it is structurally similar to magnesium hexahydrate and the hairpin ribozyme has previously been shown to be catalytically active in cobalt hexamine (6,16-18).

Surface plasmon resonance (SPR) is an excellent assay for measuring binding. Inherent to the technique, the association and dissociation rates are determined, which is not typical using other binding assays. SPR measures bi-molecular interactions in real-time utilizing an optical technique that takes advantage of the phenomenon of attenuated total internal reflection (19). This phenomenon is used to detect mass at the surface via changes in the refractive index. The challenges of utilizing SPR mainly lie in development of specific methods, which can be quite laborious.

In this application loop A is immobilized and loop B is in the mobile phase. Increase in mass at the surface results from docking, and nonspecific interactions which are subtracted off. By measuring mass in real-time there is no need for fluorescent or radioactive labels that are required for other assays. This study provided the first kinetic data for the junctionless hairpin ribozyme which will provide important insight into the nature of the interaction (11,12).

## **2.2 MATERIALS AND METHODS**

### **2.2.1 RNA preparation**

The loop A constructs were synthesized by Dharmacon Inc. (Lafayette, CO) with a biotin-tag at the 5' end of the sequence. The RNA was deprotected using the protocol provided by the company. The 2'-OMe loop A sequence, of 26 nucleotides, is 5'-

GCGCAmGUCCUCGUAAGAGAGAAGCGC-3', where Am represents a 2'-*O*-mehtyl modification at the cleavage site. This prevents cleavage of the scissile phosphate. The negative control, G+1A loop A is the same, except the G, 3' to the Am residue, has an adenine base. This is known from previous studies to completely abolish docking (20). Both loop A species were lyophilized after deprotection and dissolved and exchanged into double-deionized, RNase-free water using Amicon Ultra 4 centrifugal units. They were then dialyzed against HPLS (HEPES low-salt) buffer which is comprised of 20 mM Hepes and 20  $\mu$ M Na<sub>2</sub>EDTA at pH 7.5. This dialysis step was done using a Spectra/Por microdialyzer with a cellulose membrane with a 2 kDa molecular weight cutoff. The purity was assessed by gel electrophoresis.

The loop B construct was synthesized by *in vitro* transcription, since no modifications are present with it (21). The sequence of the so-called wild-type loop B, of 42 nucleotides, is 5'-GCGAGAGAAACACACGACGAAAGUCGUGGUACAUAUACCUCGC-3'. The DNA template was purchased from Integrated DNA Technologies (Coralville, IA).

Following transcription, desalting occurred via ethanol precipitation. Then, the resuspension was purified on a Superdex 75 column (GE Healthcare) with an isocratic elution buffer of 10 mM phosphate, 100 mM sodium chloride, pH 6.5. Following purification the desired transcription product was concentrated with an Amicon Ultra 15 centrifugal unit with a 3 kDa molecular weight cut-off. The transcription product was exchanged into RNase-free, double-deionized water three times before concentrating to desired volume. Since RNA is very water-soluble, there was not an instance of the RNA being too concentrated.

Following concentration of the purified product, the loop B was dialyzed as described as above for loop A. The important distinction is that after dialysis into HPLS buffer a second round of dialysis occurred with HPLS buffer with 250  $\mu$ M cobalt hexamine in order to fix the

free concentration of metal ions. The typical time for dialysis into HPLS buffer was 0.5 hr. while the typical time for dialysis into HPLS with  $\text{Co}(\text{NH}_3)_6$  being 1 hr. The final concentration of loop B was determined with the extinction coefficient of  $354,000 \text{ M}^{-1}\text{cm}^{-1}$ .

### 2.2.2 Surface plasmon resonance

For a step-by-step protocol, developed by Dr. Minako Sumita and myself, refer to the appendix for this chapter. The SPR instrument used in these studies is a Reichert Technologies SR7000DC; it is a two-channel instrument. The sensor chip was a mixed self-assembled monolayer [10%  $\text{COOH}-(\text{EG})_6$ -Alkanethiol, 90%  $\text{HO}-(\text{EG})_3$ -Alkanethiol] surface. The chip was primed using 0.1% SDS (1 min), 50 mM glycine-HCl, pH 2.2 (10 min), 2 M NaCl (10 min), and 10 mM NaOH (22). The chip was modified to by covalently-attaching streptavidin (New England Biotechnologies) using standard EDC/NHS coupling (22).

Several challenges arose with streptavidin modification. First, although sensor chips with streptavidin are commercially available, they did not work in our hands. They could have been too old or not functional. We found modifying the sensor chips ourselves to be more effective. More specifics can be found within the in-house protocol (appendix for this chapter) but if the EDC/NHS coupling fails it is typically a problem with the EDC. EDC is very hydroscopic. It should be stored under  $\text{N}_2$  gas and only dissolved and mixed with NHS the day of its use.

After streptavidin modification, biotinylated 2'-OMe and the G+1A loop A were added to the two separate channels (100-125  $\mu\text{RIU}$ ) using the HPLS buffer. The buffer was switched to HPLS with  $250 \mu\text{M Co}(\text{NH}_3)_6^{3+}$ . This dialysis on the sensor chip typically took 1 hr. Then diluted loop B samples were prepared using the same buffer.

For the SPR experiments, concentrations of loop B ranging from 150-750 nM were used with multiple buffer-only blanks. An association phase of 15 minutes was used with a 10 minute

dissociation phase. The surface was regenerated with 1 M NaCl and ddH<sub>2</sub>O, each had a contact time of 10 minutes each. The flow rate was 10  $\mu$ L/min. Test scans demonstrated the interaction to be independent of flow rate. In general, the entire assay was able to be replicated 2-3 times.

### **2.2.3 Data analysis**

The data was analyzed using Scrubber 2 (Biologic, provided with Reichert instrument). The negative control, G+1A loop A interaction with loop B (reference channel) was subtracted from the 2'-OMe loop A interaction with loop B (experimental channel). Also, the average of multiple blanks was subtracted from all signals. This process is the so-called double referencing and is an important component in obtaining quality data (23).

The data were then fit globally using a simple 1:1 Langmuir interaction model which accounted for a bulk-shift. The rate constants and equilibrium constant reported here are the result of eight independent global fits. In practice that amount of replicates is not necessary but as this was the first successful data set on the Reichert instrument, the precaution of many replications was taken.

## **2.3 RESULTS**

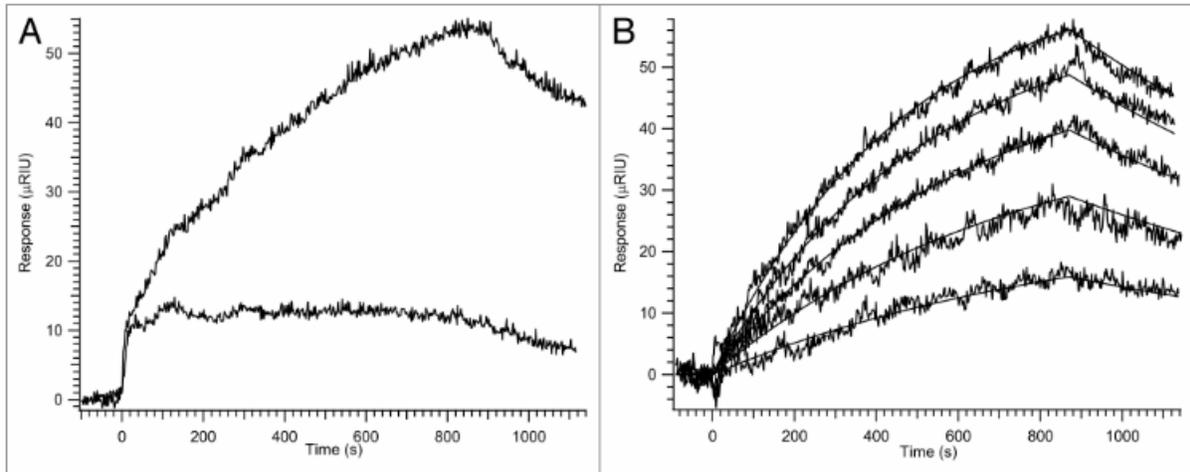
### **2.3.1 Metal ion concentration associated with docking**

The phenomenological [Cobalt]<sub>1/2</sub> of trans-docking is  $48.7 \pm 35.0$   $\mu$ M (11). This value was determined by a novel CD assay that was developed specifically for this system by Dr. Minako Sumita of our lab. The assay works by difference and thus gives fairly large error. For the surface plasmon resonance (SPR) assay that will be described in detail in this chapter 250  $\mu$ M cobalt hexamine was used, a value approximately five times the [Cobalt]<sub>1/2</sub> to ensure the presence of a saturating level of metal ions.

### 2.3.2 Kinetics of docking for the junctionless hairpin ribozyme

Surface plasmon resonance was used to determine the kinetics of docking at 250  $\mu\text{M}$   $\text{Co}(\text{NH}_3)_6^{3+}$ . The association and dissociation rates were very slow, with association rate of  $k_a = (1.97 \pm 0.29) 10^3 \text{ M}^{-1}\text{s}^{-1}$  and dissociation rate of  $k_d = (7.1 \pm 1.0) 10^{-4} \text{ s}^{-1}$ . These rates result in the dissociation constant  $K_d = 372 \pm 101 \text{ nM}$ . Representative data is displayed in Figure 2-3. The association rate is several orders of magnitude below the diffusion limit, suggesting significant conformational re-arrangements must be made before docking. These results have been published in reference (11).

Preliminary investigations of this specific interaction show that decreasing metal ion concentration results in an even slower association rate constant with the dissociation rate constant not being severely impacted. This suggests that the metal ions could be playing an important role in stabilizing docking-competent conformations.



**Figure 2-3: Docking of the junctionless hairpin ribozyme.** In panel A the upper trace is the raw data of the interaction between 2'-OMe loop A and loop B. The lower trace is the non-specific interaction between G+1A loop A and loop B. It is important to note that whether G1+A loop A was added to the reference channel or no RNA was on the reference channel the traces look fundamentally the same, implying they are due to bulk shifts. This clearly demonstrates that there is a specific interaction occurring between 2'-OMe loop A and loop B. In panel B, each trace is the result of subtracting the sample channel from the right channel (e.g. traces in panel A) and the average of multiple blanks have been subtracted as well. In other words the data has been double-referenced. The loop B concentrations in panel B range from 150-750 nM. Figure reproduced with permission from reference (11).

## 2.4 DISCUSSION

### 2.4.1 Slow, but tight intermolecular docking

The results in this chapter represented the first kinetic binding data on the junctionless hairpin ribozyme (11). We showed that despite a very slow association rate overall the docking is a tight interaction. The tightness of the interaction, kinetically, is due to a slow rate of dissociation. From previous structural studies of the undocked and docked junctionless hairpin ribozyme, it is clear that substantial rearrangements in the loop regions are necessary upon docking (Figure 2-2). The slow rate of docking is consistent with hypothesis of double conformational capture. In the double conformational capture hypothesis, docking only occurs when both loop A and loop B are in docking-competent conformations, which are present a minority of the time. For comparison with another RNA-RNA interaction has been previously studied by SPR, HIV TAR RNA and its RNA aptamer of complementary hairpin have an association rate approximately 100-fold faster than what is being reported here ( $K_d$  for that interaction is low nanomolar) (24,25). In that interaction an unwinding of five base pairs in a duplex are necessary for interface formation.

Previously, a  $K_d = 4.8 \pm 1.8 \mu\text{M}$  was determined for the junctionless hairpin ribozyme (26). This is approximately a 10-fold difference to what was found in our study. Those researchers used a slightly different RNA construct: it has a 2'-H at the cleavage site. The next chapter will focus on how what is present at the cleavage-site can have a major impact on loop-loop interactions. The prior result was obtained using Fe-EDTA fingerprinting, which is based on the formation of a solvent protected core. Also, different ionic conditions [1.5 mM  $\text{Co}(\text{NH}_3)_6^{3+}$ ] were used in that study.

### 2.4.2 A biological role for the junction

It is interesting that the junctionless hairpin ribozyme can even dock at all, given that it is easy to think of the junction aiding in proximity and orientation in the docking process of the two loops. Previously, the hairpin ribozyme has been studied with a native, four-way junction (4WJ) or a two-way junction (2WJ) by single-molecule FRET. This however means that in the 4WJ and 2WJ loop A and loop B are present in a unimolecular system. This is contrast to our ensemble technique where loop B is in great excess to loop A. The docking of the 4WJ is so fast a rate is not determinable (27). All one can conclude for the 4WJ is docking basically occurs upon addition of metal ions. The docking of 2WJ is also favorable, but rates are observable (28). Interestingly, the rate of undocking observed in that study of the 2WJ was  $4.5 \times 10^{-4} \text{ s}^{-1}$  and we found a rate of undocking  $7.1 \times 10^{-4} \text{ s}^{-1}$  for the junctionless form. Those similar rates suggest that a tight complex is formed in both the 2WJ and junctionless form. Perhaps the biological role for the junction is bringing loop A and loop B together, rather than is stabilizing the docked complex. It could be the extensive interface between loop A and loop B that provides that role, rather than a junction.

Also in published work from our group, we determined the cleavage rate ( $0.113 \pm 0.027 \text{ min}^{-1}$ ) for the junctionless hairpin ribozyme (11). This shows that the junctionless hairpin ribozyme is functional, as it can carry out self-cleavage (29). The junctionless hairpin ribozyme effectively cannot perform self-ligation as a strand of RNA is released after cleavage. This is different than the junction forms that can perform cleavage and ligation. The 4WJ favors ligation by a factor of 35 and the 2WJ favors ligation by a factor 13 (30). Another biological role for the junction is likely making ligation not only possible but favorable. Without a junction, the RNA

tertiary structure for ligation will simply not be formed, as the cleavage reaction goes to completion (11).

## **2.5 CONCLUSION**

In this chapter, the only reported binding kinetic data for the junctionless hairpin ribozyme was presented (11). The insights gained from this study are that intermolecular docking occurs in a slow, but tight manner. The slow rate of association supports the hypothesis of double conformational capture. It also provides evidence that biological role of the junction is to bring loop A and loop B together and the extensive interface is the major source of stability.

The metal ion concentration selected for this study was carefully chosen after a novel CD assay guided this decision. In the 4WJ, 2WJ, and junctionless forms of the hairpin ribozyme, docking is a metal-dependent process. None of the forms will dock in the absence of metal. A direct role of metal ions is difficult to ascertain. One possible role is the stabilization of docking-competent conformations. Molecular dynamics studies with NMR in the presence and absence of metal might provide further insight (31).

## **APPENDIX**

## **APPENDIX**

### **Surface Plasmon Resonance (SPR) Protocol**

Neil Andrew White and Mina Sumita

Fall, 2011

- 1) Before you attempt SPR
  - a. what you must know and do
- 2) Instrument cleaning and sanitization
  - a. what is necessary before your work
- 3) Installing a chip and connecting the Inter-fluid cartridge
- 4) Priming a chip
- 5) Modifying a chip with Streptavidin
- 6) Immobilizing and running an experiment
  - a. for LoopA/LoopB
- 7) Analyzing data
  - a. how to export your data
- 8) Ordering materials
- 9) Selected Readings
- 10) Additional Resources

# Before Starting (1)

## Preparing Solutions

The instrument used in this protocol is a Reichert SR7000DC.

**All solutions and samples** should be fresh and **HAVE TO BE 0.22  $\mu\text{m}$  filtered and degassed** to avoid any micro-organisms and bubbles. Also samples should be dialyzed against the buffer you are actually using as your running buffer for the experiment.

## Knowledge that will save you an immense amount of time

The samples inject 110 % of sample loop and **require extra 350  $\mu\text{L}$  as dead volume**. If the sample tube is 250  $\mu\text{L}$ , you have to prepare at least 250  $\mu\text{L}$  + 25  $\mu\text{L}$  (extra 10 %) + 350  $\mu\text{L}$  (dead volume) = **625  $\mu\text{L}$  of sample for one injection**. If you use the **insert** (100  $\mu\text{L}$  dead volume), make **400  $\mu\text{L}$  sample for one 250  $\mu\text{L}$  injection**. If you don't follow this guidelines you will get an **SPR error**, the machine then reads -1  $\mu\text{RIU}$  and the baseline comes back to a different point and makes that injection incomparable to your previous work.

Also you need to take the best care possible of EDC or it will not be usable, but there will be a section on this with reference material to help.

The pump draws from the reservoir of running buffer you put the tube in but it only fills when necessary. So, when you change solutions it will the pump generally has to draw three full times from the running buffer to reach a stable baseline.

The recommended warm-up time is an hour and a half, so wait at least a half hour before you claim a "stable baseline."

See the References section for general overviews of SPR experiments and theory.

Maintaining a temperature at least a few degrees different than room-temperature is ideal; too close and the baseline will oscillate noticeably with the heating and cooling of the instrument. You may need to adjust the temperature of the lab.

# Instrument Cleaning and Sanitization (2)

In short you do a cleaning procedure for the 250  $\mu\text{L}$  sample loop and then a cleaning procedure with the 1000  $\mu\text{L}$  sample loop and then sanitize the system. This is necessary to do when changing systems or **anytime before you start doing RNA work**.

## Instrument Cleaning

Starting an experiment with a clean instrument is very important. Change to the "cleaning chip", and the background flow as ddH<sub>2</sub>O. (See section 3) Set the flow rate to **100  $\mu\text{L}/\text{min}$**  and ensure the 250  $\mu\text{L}$  sample loop is attached. (If it is not remove the front of the auto-sampler by squeezing in the black tabs on the side at the top. Then change the loop. You should have a small plastic piece at each distal end in addition to the hand screws slightly further in on the loop.)

Dispense 2 mL each of the following solutions to an injection bottle, and place the bottle into the following place:

1A - ddH<sub>2</sub>O

1B - 20% ethanol (don't degas b/c of EtOH, just filter)

1C - 0.5% SDS (don't degas, just filter with syringe)

1D - 50 mM glycine, pH 9.5 (don't degas, just filter with syringe)

1E - ddH<sub>2</sub>O

Click "Instrument Cleaning". It is already programmed the cleaning procedure. Click "Run". The refractive index of ethanol is much higher than water. A large positive response will be observed for both channels.

Change the sample loop to the 1000 µL sample loop. Dispense 2 mL each of the following solutions to an injection bottle, and place the bottle into the following place:

1A - ddH<sub>2</sub>O

1B-3B - 20% ethanol (don't degas b/c of EtOH, just filter)

1C-3C - 0.5% SDS (don't degas, just filter with syringe)

1D-2D - 50 mM glycine, pH 9.5 (don't degas, just filter with syringe)

1E-3E - ddH<sub>2</sub>O

Click "1000cleaning". Click "Run".

### **Instrument Sanitization**

Change the flow rate to 200 µL/min. Fill 5 sample vials with 2 mL of ~1% Hypochlorite Solution and place in the following positions 1A, 1B, 1C, 1D, and 1E. Click "Sanitize." Click "Run."

Repeat to allow 10 mL of ~1% Hypochlorite Solution run through the system.

Then wash the loop with water. (There is a "1000LoopWash" procedure.)

# Chip Installation and Connecting IFC (3)

(the Userguide, pages 16-18)

## VERY IMPORTANT NOTE

If after installation the two channels seem out of sync (i.e. one is steady and the other one is drifting) after 15 min. and there is no good reason, it is likely do the tightness of the screws. We have had this problem quite a few times. Use the level to level the IFC (inter-fluid cartridge). SPR is a very sensitive technique and these adjustments really matter. In general, the tighter the “finger screws” the better, but be reasonable and **don’t** use anything but your fingers to tighten them.

**Step 1:** allow the sealed sensor chip to **equilibrate at room temperature for 30 minutes** (prevents water vapor on detector side of chip surface)

**Step 2:** Clean the prism surface by with a cotton swab wetted with ethanol (**Do NOT squirt ethanol directly on the prism surface!**). Wipe dry with kimwipe. The prism surface is sapphire and cannot be scratched. This surface **MUST** be extremely clean with no visible streaks prior to mounting a new chip.

**Step 3:** Place **0.75  $\mu\text{L}$**  of Immersion oil type A (Cargille Cat # 16482) on the center of the prism surface (midpoint between the two posts).

**Step 4:** Clean the flowcell body and gasket by wiping with ethanol. Mount the gasket on the flowcell body. An interference fit between the gasket and channel in the flowcell body will hold the gasket securely. The gasket can be pushed into the grooves using a clean tweezers.

If you have some problem with the flowcell body contamination, clean it in a 1 % TritonX-100 solution. If a sonicator is available, sonicate the flowcell and gasket in the 1 % TritonX-100 solution for a few minutes. If a sonicator is not used, soak the flowcell and gasket in the solution for at least 5 min. Then, rinse with ddH<sub>2</sub>O and ethanol.

**Step 5:** To determine which side of the slide should be up, look at the surface of the slide. If it looks smooth like glass, that is the up side. If it looks like it has shiny gold side, that is the bottom of the slide and should be set onto the prism surface.

A chip case has flat and dorm shaped sides. **The dorm shaped side indicates the binding surface side of the chip.**

**Step 6:** **Slowly**, lower the chip in place onto the prism surface being careful not to trap air bubbles between the slide and prism surface. (Use your hands and not tweezers so you don’t scratch it.) Make sure the slide is flush against the back of the prism slot.

**Step 7:** Using the flowcell posts for guides, gently position the flowcell body with gasket over the gold slide. Lower the flowcell into position until the gasket contacts the slide. Tighten the flowcell locking nuts evenly and to the same degree while providing pressure with your finger on the flowcell top.

Visual examination at eye level of the prism-flowcell will allow final leveling of the flowcell. Uneven tightening can result in leaks due to misalignment. Look for an even gap on both sides of the flowcell.

**Step 8:** Flow ddH<sub>2</sub>O for several minutes. Go to “maintenance” on the SPRAutolink window and choose “Detector Scan Data”. When you click “Update”, two minima should be seen of about the same depth at around 250-300 pixels. If you do not see two minima, you may be aspirating air, there may be a slide problem (wrong side is up), or you may have trapped an air bubble under the flowcell. You must correct this problem by checking each connection. You can also increase the flowrate to 300-500  $\mu\text{L}/\text{min}$  to try to push out bubble.

### **Connecting the Flowcell**

Typically, the right side of the flowcell is your sample channel and the left side is the reference channel.

#### **The regular connection**

1. Connect autosampler (or injector) tube to the inlet of flowcell (right front)
2. Connect sample channel (right back) to reference channel (left front)
3. Connect the outlet of flowcell (left back) to waste

#### **When you are immobilize your sample**

1. Disconnect the tubing connecting the reference channel side of the flowcell to the sample side (left back)
2. Disconnect the tubing connecting the outlet to waste (right back)
3. Be sure the flow from the autosampler (or injector) goes directly into the flowcell on the sample channel
4. Connect the waste tubing to the sample side (left back) to go to waste directly on the sample side only
5. Connect the remaining tube to reference side (right back)
6. After your sample binds only to your sample channel, reconnect the tubing in the regular way.

## Priming a Chip (4)

This is designed to make sure you have a clean chip with a re-suspended matrix. Both of which are very important. There are multiple ways to prime a chip (see ref.); this is the method I use. Before installing a new chip to collect data with pass a light flow of Nitrogen gas over it to remove any dust particles that maybe present.

Set flow rate to 100  $\mu\text{L}/\text{min}$

Click on the Run Table “Prime”

1A- **0.1%** SDS (1 min)

1B- 50 mM Glycine-HCl pH 2.2 (10 min)

1C- 2 M NaCl (10 min)

1D- 10 mM NaOH (10 min)

Again, filter all reagents and dilute the 0.5% SDS used in cleaning as it could be too harsh for the chip.

# Modifying a Chip with Streptavidin (5)

Taking advantage of the strong interaction between Biotin and Streptavidin is very useful and reliable. There are of course problems that can be encountered. They most always occur with EDC. **READ how to care for the reagents.**

This is typical SA modification for reference.

1. Do this just after Priming the Chip
2. Use ddH<sub>2</sub>O as back flow rate 50 µL/min and the sample loop *must* be the 1000 µL sample loop
3. Inject ddH<sub>2</sub>O, and the baseline should stay flat (optional)
4. Prepare 12 bottles of 1.5 mL ddH<sub>2</sub>O and place them to the position 2A, 3A, 4A, 2B, 3B, 4B, 2C, 3C, 4C, 2D, 3D, and 4D
5. Prepare 2 mL of 2M NaCl/10 mM NaOH solution and place the sample to the position **1A**
6. Prepare 2 mL of 1 M ethanolamine, pH 9.5 (adjust pH with concentrated HCl) and place it to the position **1D**
7. Prepare 1.5 mL of 30 µg/mL streptavidin/neutralavidin in 20 mM NaOAc, pH 5.2 solution and place it to the position **1C**
8. Prepare *fresh* mixture of 0.22 M EDC/0.20 M NHS (EDC: 1-ethyl-3-(3-dimethylamino propyl) carbodiimide hydrochloride, NHS: N-hydroxysuccinimide) and place it to the position **1B**
9. Choose “Streptavidin Coating” in the run method, and click “run”. It will take ~2 hours.

Summary (Flow rate 50 µL/min)

2A-4A, 2B-4B, 2C-4C, and 2D-4D ddH<sub>2</sub>O

1E HPLS (or your buffer)

1A NaCl/NaOH

1D Ethanolamine

1C Streptavidin

1B EDC/NHS

(The order is based on practicality)

## **How to care for the reagents used in SA modification**

The EDC is by far the most likely reagent to make your modification not work. It must be handled properly or it WILL NOT work.

EDC

It should be room temperature and then open it under the hood. Remove your desired amount as quickly as possible with the vial open for the least amount of time possible. It needs to be stored under N<sub>2</sub> gas and parafilmed. Then dissolve the powder in water and aliquot it for single use. From experience it will not be usable if it is not stored under Nitrogen gas and its effectiveness decays with freeze/thaw cycles.

Also the EDC/NHS mixture must be made just prior to use. I make it last and immediately run the chip modification procedure and it hasn't caused a problem.

#### NHS

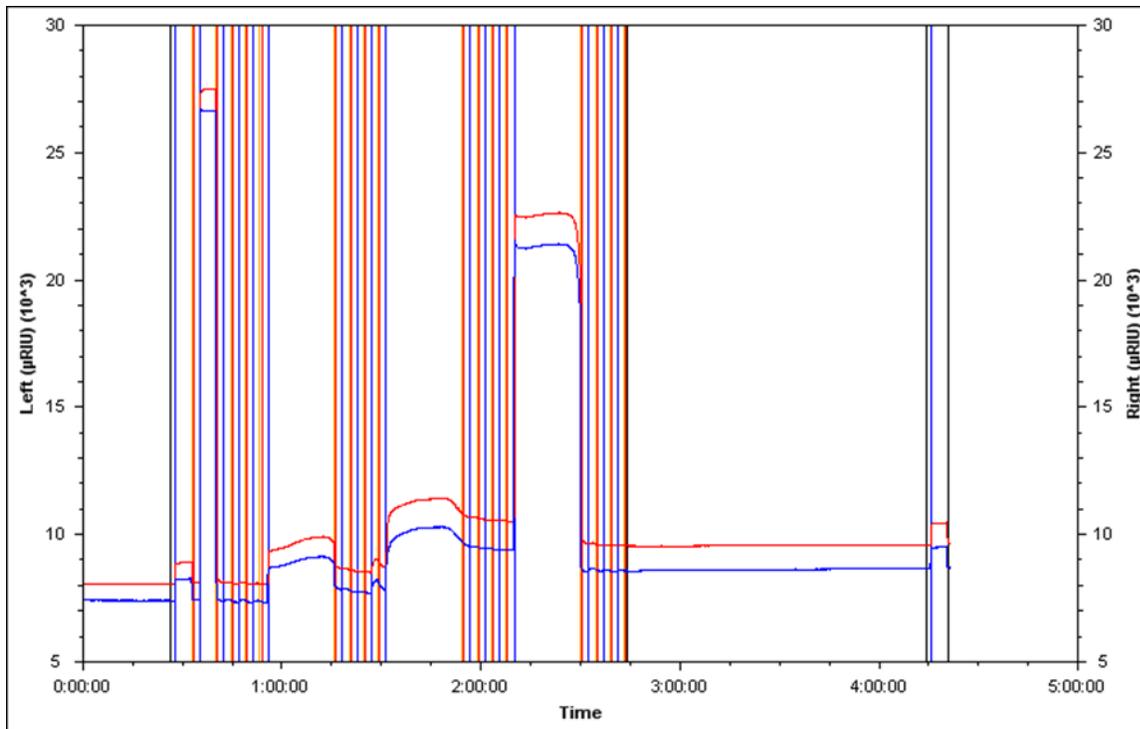
The only problem I've encountered is that if it's not fully dissolved it will cause an SPR error so always double check it after thawing.

#### Streptavidin

The two things to beware of are excessive freeze/thaw cycles and extremely long storage. To counter this I aliquot for single use and store some in the fridge for use in the near future and freeze the other aliquots. Also the pH does matter so it can be useful to check it after preparation because the streptavidin is in buffer and the Sodium Acetate is another buffer and they are combined. I use a stock 40 mM Sodium Acetate pH 5.2 to avoid a drastic dilution.

#### Ethanolamine

Again, excessive freeze/thaw cycles should be avoided. Also it is photosensitive.



**Supplemental Figure 2-1: Typical streptavidin modification.** The increases are due to HLPLS, NaCl/NaOH, EDC/NHS, streptavidin, ethanolamine, and HPLS.

# LoopA/LoopB (6)

## Immobilization of LoopA

1. Change the running buffer to HPLS pH 7.5 and the flow-rate to 100  $\mu\text{L}/\text{min}$
2. Once it appears to be equilibrated (~1 hr.) test the baseline with HPLS.
  - a. It should have a bulk shift <10  $\mu\text{RIU}$  and drift <10  $\mu\text{RIU}$  over 10 min.
3. Next run only through the left channel and change the flow-rate to 10  $\mu\text{L}/\text{min}$
4. Once the baseline is steady place 375  $\mu\text{L}$  of 1  $\mu\text{M}$  of WT LoopA in an insert in position 1A and run "Capture LoopA"
5. Record amount of ligand immobilized
6. Then rinse three times with HPLS
7. Change to having the running buffer (RB) only flowing through the right side
8. Once the baseline is steady place 375  $\mu\text{L}$  of 1  $\mu\text{M}$  of Control LoopA in an insert in position 1A and run "Capture LoopA"
9. Record amount of ligand immobilized
10. Then rinse three times with HPLS
11. Then connect to flow-cell so RB goes through both channels
12. Switch to buffer that will be used for kinetics and leave at 10  $\mu\text{L}/\text{min}$  overnight or at 100  $\mu\text{L}/\text{min}$  until equilibrated

## Kinetics with LoopB

1. Test the baseline with running buffer
2. Then use desired run table

# Analyzing Data (7)

Scrubber 2 is actually a user-friendly program with excellent tutorials. What needs to be covered here is how to get data imported into Scrubber.

1. Click on “Form” then select “Plot-Processing” from the drop-down menu
2. Then click on a desired curve, it will become green
3. Then click “Add”
4. Repeat until you have all your desired curves in the lower window (Remove curves if necessary.)
5. Then click “File” and export as file type .ovr
6. Then open Scrubber, click on “Load” and change file type to all files and then load the desired file
7. When done in Scrubber it is best to save the method so later the data file and method file can both be loaded together

## Ordering Information (8)

<b>Product</b>	<b>Company</b>	<b>Cat. No.</b>	<b>Price</b>
Chip (5/pk)	Reichert	13206061-5	270.94
Glass Vials (100/pk)	VWR	HP-5182-0715	24.39
Screw Cap (100/pk)	VWR	HP-5182-0717 (blue)	24.39
		HP-5182-0718 (green)	24.39
		HP-5182-0719 (red)	24.39
Glass Insert (500/pk)	VWR	HP-5181-3377	74.65
Streptavidin, 1mg	NEB	N7021S	60.00
EDC, 1g	Sigma-Aldrich	E7750-1G	19.80
NHS, 5g	Sigma-Aldrich	130672-5G	13.30
Ethanolamine, 100 mL	Sigma-Aldrich	E9508-100ML	24.40
Sodium Hypochlorite	Sigma-Aldrich	425044-250ML	33.40

**Supplemental Table 2-1: SPR Ordering information.** These are preferred products and reagents for SPR.

# Selected Readings (9)

This book is available through Springer Protocols is very useful and will be referred to throughout the reference section.

Surface Plasmon Resonance: Methods and Protocols

Editor(s): Nico J. Mol<sup>1</sup>, Marcel J. E. Fischer<sup>2</sup>

Affiliation(s): (1)Dept. Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16 Utrecht 3585 CA Netherlands

(2)Dept. Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16 Utrecht 3585 CA Netherlands

Series: Methods in Molecular Biology | Volume No.: 627

Print ISBN: 978-1-60761-669-6

[http://www.springerprotocols.com.proxy2.cl.msu.edu/BookToc/doi/10.1007/978-1-60761-670-2?uri=/Abstract/doi/10.1007/978-1-60761-670-2\\_1](http://www.springerprotocols.com.proxy2.cl.msu.edu/BookToc/doi/10.1007/978-1-60761-670-2?uri=/Abstract/doi/10.1007/978-1-60761-670-2_1)

Link from 11-22-11

## **For Section 1, “Before you attempt SPR”**

Surface Plasmon Resonance: A General Introduction by Nico J. de Mol and Marcel J. E. Fischer  
Chapter 1 of Surface Plasmon Resonance: Methods and Protocols from Springer Protocols- *This is great general overview of SPR*

Chapter 7, Theory of Binding Data Analysis from Invitrogen- *Very good explanations of the technical theory of binding*

<http://tools.invitrogen.com/downloads/FP1.pdf>

Link from 11-22-11

SPR Pages, website (<http://www.sprpages.nl/Index.php>) *Useful website especially when starting out, geared toward Biacore though.*

Surface Plasmon Resonance for Measurements of Biological Interest by Cynthia Bamdad Unit 20.4 in Current Protocols in Molecular Biology (1997) 20.4.1-20.4.12 from Wiley Protocols- *A decent overview*

Laird-Offringa, I. A., Kinetic studies of RNA-protein interactions using surface plasmon resonance, *Academic Press Methods*, 2002, 95-104- *Worth reading if doing RNA-protein interactions some useful information on how-to but mostly overview.*

## **For Section 4, “Priming a Chip”**

Chapter 3 (of Surface Plasmon Resonance: Methods and Protocols) Amine Coupling Through EDC/NHS: A Practical Approach by Marcel J.E. Fischer- *pg. 63 is where you can find multiple chip priming procedures*

# Additional Information (10)

## Experiment Optimization

### 1. Check for nonspecific binding (*J. Mol. Recognit.*, 1999, 12, 280)

Biomolecules often show an interaction with surfaces. It is necessary to check both reagents for nonspecific binding on the chip before running any experiments. The best way to do this is to simply inject each sample over a nonderivatized surface at the highest concentration to be used in the analysis. A low level of background binding may be compensated for using a reference surface. If the nonspecific binding is high, then the experimental conditions may be altered. If the nonspecific binding cannot be eliminated, it is best to couple the sticky protein onto the surface. Be cautious though, as the nonspecific binding may affect the activity of the immobilized ligand.

Myszka, D.G., Improving Biosensor Analysis, *J. Mol. Recognit.*, 1999, 12, 279-284. Quick Instructions for Using Integrated SPRAutolink, Reichert Technologies

Reichert SR7000DC Dual Channel SPR Instrument User Guide, Version 5, Reichert Technologies

## **REFERENCES**

## REFERENCES

1. Draper, D., Grilley, D. and Soto, A. (2005) Ions and RNA folding. *Annual review of biophysics and biomolecular structure*, **34**, 221-243.
2. Leipply, D. and Draper, D.E. (2011) Effects of Mg<sup>2+</sup> on the free energy landscape for folding a purine riboswitch RNA. *Biochemistry*, **50**, 2790-2799.
3. Trachman, R.J. and Draper, D.E. (2013) Comparison of interactions of diamine and Mg<sup>2+</sup> with RNA tertiary structures: similar versus differential effects on the stabilities of diverse RNA folds. *Biochemistry*, **52**, 5911-5919.
4. Draper, D.E. (2004) A guide to ions and RNA structure. *Rna*, **10**, 335-343.
5. Fedor, M. (2000) Structure and function of the hairpin ribozyme. *Journal of molecular biology*, **297**, 269-291.
6. Fedor, M. (2009) Comparative enzymology and structural biology of RNA self-cleavage. *Annual review of biophysics*, **38**, 271-299.
7. Rupert, P. and Ferré-D'Amaré, A. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, **410**, 780-786.
8. Rupert, P., Massey, A., Sigurdsson, S. and Ferré-D'Amaré, A. (2002) Transition state stabilization by a catalytic RNA. *Science (New York, N.Y.)*, **298**, 1421-1424.
9. Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. and Wedekind, J.E. (2006) Water in the active site of an all-RNA hairpin ribozyme and effects of Gua8 base variants on the geometry of phosphoryl transfer. *Biochemistry*, **45**, 686-700.
10. MacElrevey, C., Salter, J.D., Krucinska, J. and Wedekind, J.E. (2008) Structural effects of nucleobase variations at key active site residue Ade38 in the hairpin ribozyme. *RNA (New York, N.Y.)*, **14**, 1600-1616.
11. Sumita, M., White, N.A., Julien, K.R. and Hoogstraten, C.G. (2013) Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis. *RNA Biol*, **10**, 425-435.
12. Hoogstraten, C.G., Sumita, M. and White, N.A. (2014) Unraveling the thermodynamics and kinetics of RNA assembly: surface plasmon resonance, isothermal titration calorimetry, and circular dichroism. *Methods Enzymol*, **549**, 407-432.
13. Cai, Z. and Tinoco, I. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.

14. Butcher, S., Allain, F. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature structural biology*, **6**, 212-216.
15. Ferré-D'amaré, A.R. and Rupert, P.B. (2002) The hairpin ribozyme: from crystal structure to function. *Biochemical Society transactions*, **30**, 1105-1109.
16. Hampel, A. and Cowan, J. (1997) A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage. *Chemistry & biology*, **4**, 513-517.
17. Nesbitt, S., Hegg, L. and Fedor, M. (1997) An unusual pH-independent and metal-ion-independent mechanism for hairpin ribozyme catalysis. *Chemistry & biology*, **4**, 619-630.
18. Young, K., Gill, F. and Grasby, J. (1997) Metal ions play a passive role in the hairpin ribozyme catalysed reaction. *Nucleic acids research*, **25**, 3760-3766.
19. de Mol, N.J. and Fischer, M.J. (2010) Surface plasmon resonance: a general introduction. *Methods in molecular biology (Clifton, N.J.)*, **627**, 1-14.
20. Kuzmin, Y.I., Da Costa, C.P., Cottrell, J.W. and Fedor, M.J. (2005) Role of an active site adenine in hairpin ribozyme catalysis. *Journal of molecular biology*, **349**, 989-1010.
21. Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic acids research*, **15**, 8783-8798.
22. Fischer, M. (2010) Amine coupling through EDC/NHS: a practical approach. *Methods in molecular biology (Clifton, N.J.)*, **627**, 55-73.
23. Myszka, D.G. (1998) Improving biosensor analysis. *Journal of molecular recognition : JMR*, **12**, 279-284.
24. Lebars, I., Legrand, P., Aimé, A., Pinaud, N., Fribourg, S. and Di Primo, C. (2008) Exploring TAR-RNA aptamer loop-loop interaction by X-ray crystallography, UV spectroscopy and surface plasmon resonance. *Nucleic acids research*, **36**, 7146-7156.
25. Nair, T.M., Myszka, D.G. and Davis, D.R. (2000) Surface plasmon resonance kinetic studies of the HIV TAR RNA kissing hairpin complex and its stabilization by 2-thiouridine modification. *Nucleic acids research*, **28**, 1935-1940.
26. Hampel, K.J., Walter, N.G. and Burke, J.M. (1998) The solvent-protected core of the hairpin ribozyme-substrate complex. *Biochemistry*, **37**, 14672-14682.
27. Nahas, M.K., Wilson, T.J., Hohng, S., Jarvie, K., Lilley, D.M. and Ha, T. (2004) Observation of internal cleavage and ligation reactions of a ribozyme. *Nature structural & molecular biology*, **11**, 1107-1113.

28. Liu, S., Bokinsky, G., Walter, N.G. and Zhuang, X. (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.
29. Butcher, S.E., Heckman, J.E. and Burke, J.M. (1995) Reconstitution of hairpin ribozyme activity following separation of functional domains. *The Journal of biological chemistry*, **270**, 29648-29651.
30. Fedor, M.J. (2000) Structure and function of the hairpin ribozyme. *Journal of molecular biology*, **297**, 269-291.
31. Butcher, S.E., Allain, F.H. and Feigon, J. (2000) Determination of metal ion binding sites within the hairpin ribozyme domains by NMR. *Biochemistry*, **39**, 2174-2182.

## **CHAPTER 3**

### **THERMODYNAMICS OF RNA TERTIARY STRUCTURE FORMATION IN THE JUNCTIONLESS HAIRPIN RIBOZYME**

## ABSTRACT

The hairpin ribozyme has two loops that interact in a minor-groove, minor-groove fashion to obtain the catalytically-competent structure. The catalytically-competent structure has the scissile-phosphate incorporated into  $S_N2$  geometry. Also, the catalytically residues of G8, of loop A, and A38, of loop B, are positioned near the cleavage-site. This intricate interaction is one of relatively few RNA-RNA tertiary interactions that occurs in the absence of helix annealing.

The docking process provides an excellent opportunity to understand the underlying characteristics of this interaction and RNA tertiary structure formation in general. We determined the thermodynamic signature for this loop-loop interaction using temperature-dependent surface plasmon resonance (SPR). We found that when the native 2'-OH was present at the cleavage site, the loop-loop interaction, termed docking, was enthalpically favorable and slightly entropically unfavorable. In contrast, with the 2'-*O*-methyl cleavage-site modification present, the docking process is both enthalpically and entropically favorable. The binding affinity was also somewhat less favorable with the 2'-*O*-methyl cleavage-site modification present. This provides evidence that cleavage-site modifications not only prevent cleavage, as intended, but can also significantly disrupt the underlying nature of an interaction.

### 3.1 INTRODUCTION

Three-dimensional RNA structure is now appreciated as able to rival proteins in terms of size and complexity (1-3). However, the understanding of how RNA forms these structures lags significantly behind that of proteins. The focus of this chapter is on studying one of the relatively few examples of RNA tertiary structure formation that occurs without helix annealing.

Primary structure is similar for both RNA and protein. It is a sequence of nucleotides or amino acids. The term secondary structure in RNA, generally refers to simply where base pairing occurs and where it does not (4). For protein, by contrast, it is used to describe the distribution of  $\alpha$ -helices,  $\beta$ -sheets, random coils. Both protein secondary structure and RNA secondary structure describe hydrogen-bonding patterns. However, a small secondary structure unit of a protein is unlikely to be stable outside of the context of the larger protein, whereas a small secondary structure unit of RNA is likely to be stable on its own (5). In contrast to protein folding, RNA folding is hierarchal. The term tertiary structure, for RNA, is typically used to describe one RNA domain interacting with something else, like another RNA domain, as is the case here. More rigorously, it can be defined as the interaction of secondary structure units, where secondary structure is only altered in a minimalistic sense (6).

There are many examples of protein tertiary structure, for water-soluble domains, where there is a hydrophobic core and the periphery is much more hydrophilic. There are far fewer examples of RNA tertiary structure formation. The hairpin ribozyme provides an example of tertiary structure formation. Loop A and loop B have their defined secondary structure and then dock in a minor groove-minor groove fashion. This exemplifies RNA interacting in a hierarchical manner.

In this chapter, the thermodynamic signature for the docking process of the hairpin ribozyme will be presented. It will be put in context with the handful of examples of RNA tertiary structure formation that have been studied.

The structural studies of the hairpin ribozyme have focused on loop A and loop B by themselves and the docked complex of loop A and loop B (7-10). Studies have determined the thermodynamic signature of docking constructs of the hairpin ribozyme with various junctions (see discussion section 3.4.2) (11). There has yet to be a reporting of the thermodynamic signature of docking for the junctionless construct of the hairpin ribozyme, which is necessary for analysis of the properties of the tertiary interface in isolation from other effects.

In structural and biochemical studies a 2'-*O*-methyl modification has often been used at the cleavage site to prevent the reaction from occurring (12). There is biochemical data showing in the two-way junction form of the hairpin ribozyme the methoxy modification does not allow for as stable of a docked complex as the native hydroxyl group (13). There are also simulation studies that suggest the same idea (14). However, there are competing ideas that the cleavage-site modification does not have a significant effect, that are based on in-line geometry observed in crystallographic studies (9).

## **3.2 MATERIALS AND METHODS**

### **3.2.1 RNA preparation**

RNA preparation was performed as described in the previous chapter. Sequences synthesized are schematized in Figure 3-1. As described in the previous chapter the loop A sequences 2'-OH loop A, 2'-*O*-methyl loop A, and the negative control, G+1A loop A all had a 5'-biotin tag for immobilization. The loop B sequences wildtype loop B, A38C loop B, and A38U loop B were all synthesized via *in vitro* transcription, as described in Chapter 2.



### 3.2.2 Temperature-dependent surface plasmon resonance

SPR experiments were performed similarly to those described Chapter 2. Different regeneration conditions were required for the 2'-OH Loop A species due to a slower dissociation rate and that the instrument had to be equilibrated at each desired temperature (see below).

Data was obtained for 2'-OMe loop A and 2'-OH loop A interacting with A38C loop B, using G+1A loop A as the negative control in both instances. The temperatures used were 20° C, 25° C, 30° C, 40° C, and 50° C. It was found that when increasing the temperature it was important to do so by 0.5° C at a time. This helped maintain the integrity of the surface. Temperatures lower than this range were difficult to achieve due to the insufficient cooling of the instrument (see the appendix for this chapter for more details on temperature regulation).

For the interaction between 2'-OMe loop A and A38C loop B, the concentrations of A38C loop B ranged from 150 nM to 750 nM. The parameters of this interaction were a 15 min. association phase, a 10 min. dissociation phase, and the regeneration consisted of a 10 min. contact time for both 1 M NaCl and ddH<sub>2</sub>O. For the 2'-OH loop A and A38C loop B interaction, the A38C loop B concentrations ranged from 30 nM to 750 nM. For this interaction the parameters were an association phase of 15 min., a dissociation phase of 1 hr., and regeneration was achieved with a 10 min. wait period, a 10 min. contact time of 1 M NaCl, and a 10 min. contact time with ddH<sub>2</sub>O.

For all temperatures two replicates (of the entire concentration range) were fit independently and the average was reported. The exception was the interaction of 2'-OH loop A and A38C at 25° C, which was performed five times. Slight variations in dissociation times and regeneration conditions had no discernible impact on the results. In application, rates increase with temperature so less stringent regeneration conditions and/or dissociation times were

necessary at higher temperatures. For instance, at 50° C the dissociation phase was changed to 12 min. because the signal from docking was effectively zero in that duration. Occasionally, an individual concentration would yield useable data and was omitted from data analysis.

### 3.2.3 Data analysis

The SPR data was analyzed in the same manner as describe in the first chapter, using Scrubber 2 with double-referencing, as described in Chapter 2. The van't Hoff and Arrhenius plots we constructed using linear regression with Igor Pro 6 (Wavemetrics). For the van't Hoff plots the reciprocal of temperature ( $K^{-1}$ ) was plotted on the x-axis and  $\ln K_{eq}$  was plotted on the y-axis. The equation for fitting the data is:

$$\ln(K_{eq}) = \frac{-\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R}$$

Error limits on the slope and intercept were as reported by the regression software and were propagated into the thermodynamic parameters using standard formulas. For the Arrhenius plots the temperature ( $K^{-1}$ ) was plotted on the x-axis and the natural logarithms of the rate constants were plotted on the y-axes. The data was fit with the equation:

$$\ln(k_x) = -\frac{E_a}{RT} + \ln(A)$$

where  $k_x$  is  $k_a$  or  $k_d$ ,  $E_a$  is the activation energy for the corresponding step, and  $A$  is the Arrhenius prefactor.

### 3.3 RESULTS

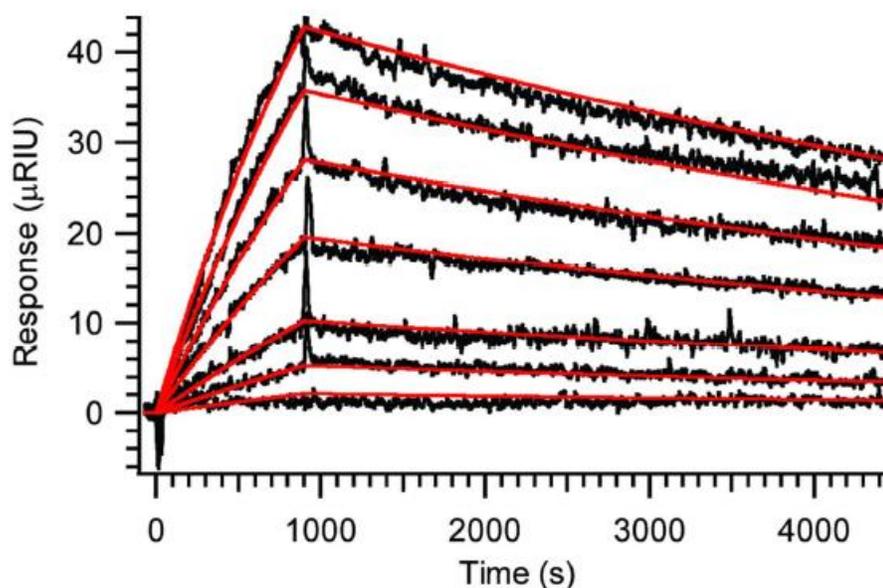
#### 3.3.1 Rate and dissociation constants for the various hairpin ribozyme constructs

In order to study the docking interaction by SPR it is important to have docking-competent but not catalytically-active constructs. The pairings of loop A and loop B studied satisfy those criteria. The results of the interactions can be seen in Table 3-1 and representative data for the interaction between 2'-OH loop A and A38C loop B are displayed in Figure 3-2. For the interactions between 2'-*O*-CH<sub>3</sub> (2'-OMe) loop A and A38C loop B the  $K_d$  was determined to be  $338 \pm 66$  nM. For the interaction between 2'-OH loop A and A38C loop B the  $K_d$  was determined to be  $114 \pm 33$  nM.

We also attempted to utilize another loop B construct, A38U loop B, which was found to yield inconsistent results. In some SPR runs it gave results consistent with A38C loop B and in others was consistent with an anomalously low concentration of loop B (data not shown). By using gel electrophoresis it was clear that A38U loop B has a tendency to dimerize and was not used further (data not shown). Constructs and conditions have previously been implicated as important in whether or not loop B dimerizes (8).

loop A	loop B	$k_a, M^{-1} s^{-1}$	$k_d, s^{-1}$	$K_d, nM$	$\Delta G^\circ_{dock}, kcal mol^{-1}$
2'-O-CH <sub>3</sub> <sup>a</sup>	wild-type	$(1.97 \pm 0.29) \times 10^3$	$(7.1 \pm 1.0) \times 10^{-4}$	$372 \pm 101$	$-8.76 \pm 0.16$
2'-O-CH <sub>3</sub>	A38C	$(1.13 \pm 0.06) \times 10^3$	$(3.79 \pm 0.56) \times 10^{-4}$	$338 \pm 66$	$-8.82 \pm 0.12$
2'-OH	A38C	$(1.24 \pm 0.35) \times 10^3$	$(1.33 \pm 0.23) \times 10^{-4}$	$114 \pm 33$	$-9.47 \pm 0.17$

**Table 3-1: Kinetic and dissociation constants determined by SPR.** The 2'-O-CH<sub>3</sub> (2'-OMe) loop A and A38C loop B are docking-competent and catalytically-competent. The 2'-OH loop A and wild-type loop B can dock as well as cleave. <sup>a</sup>The first line of data has been previously published by our lab (12).



**Figure 3-2: Representative data of 2'-OH loop A and A38C loop B at 25° C.** In this interaction the 2'-OH loop A is immobilized via streptavidin-biotin interaction. The A38C loop B in the mobile phase is at concentrations at 30 nM, 75 nM, 150 nM, 300 nM, 450 nM, 600 nM, and 750 nM. The black trace has been double-referenced with Scrubber 2. The red lines are from the global fit. This figure was prepared with Igor 6.

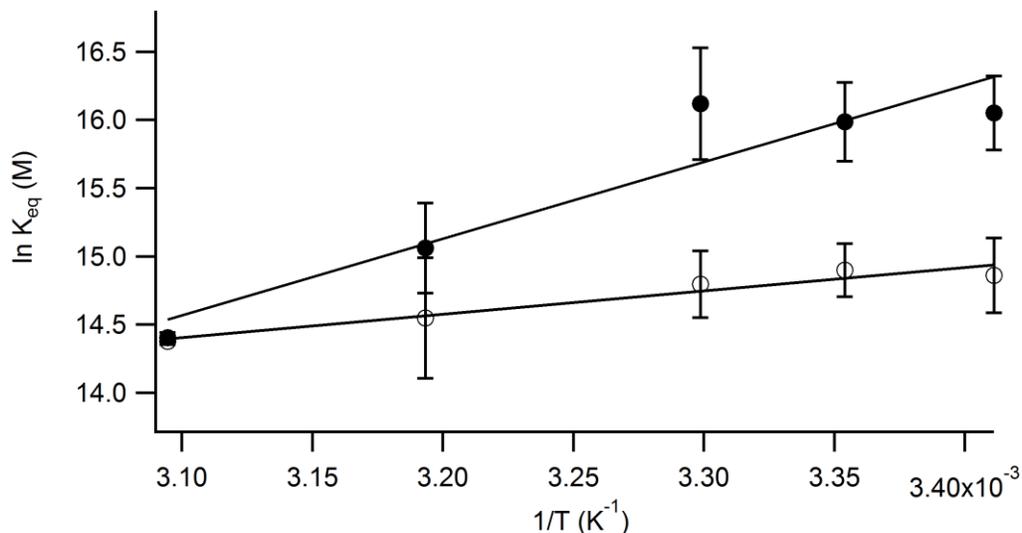
### 3.3.2 Temperature-dependent SPR studies

The interactions between 2'-*O*-CH<sub>3</sub> (2'-OMe) loop A and A38C loop B and 2'-OH loop A and A38C loop B were studied over a range of 20° C to 50° C (Table 3-2). The resulting dissociation rate constants were then fit to van't Hoff plots (Figure 3-3). The thermodynamic signature for the interaction between 2'-OMe loop A and A38C loop B is enthalpically favorable and entropically favorable. The enthalpy term,  $\Delta H^\circ$ , was determined to be  $(-3.40 \pm 0.51)$  kcal mol<sup>-1</sup> and the  $\Delta S^\circ$  term was found to be  $(18.1 \pm 1.7)$  kcal mol<sup>-1</sup> K<sup>-1</sup>. The thermodynamic signature for the interaction with 2'-OH loop A and A38C loop B was found to be more enthalpically favorable and actually entropically unfavorable. The  $\Delta H^\circ$  constant was determined to be  $(-11.2 \pm 2.4)$  kcal mol<sup>-1</sup> and the  $\Delta S^\circ$  constant was  $(-5.7 \pm 4.8)$  kcal mol<sup>-1</sup> K<sup>-1</sup>.

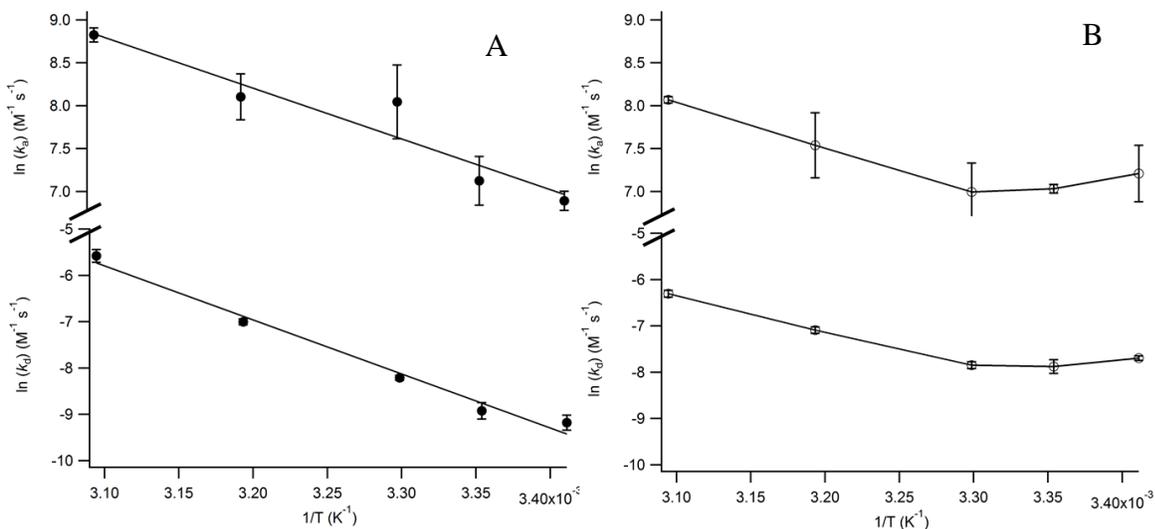
The association and dissociation rates for docking and undocking were plotted using the Arrhenius equation (Figure 3-4). For the 2'-OH loop A and A38C loop B interaction the activation energy for the forward barrier (docking) was determined to be  $(11.7 \pm 2.2)$  kcal mol<sup>-1</sup>. The barrier for the reverse process (undocking) was determined to be  $(23.2 \pm 1.7)$  kcal mol<sup>-1</sup>. For the 2'-OMe loop A and A38C loop the plot yielded non-linear fits. As a result the data was unable to be fit and activation energies were undeterminable for the forward and reverse barriers.

Temperature (°C)	Loop A	$k_a, M^{-1} s^{-1}$	$k_d, s^{-1}$	$K_d, nM$
20° C	2'-O-CH <sub>3</sub>	$(1.35 \pm 0.45) \times 10^3$	$(4.54 \pm 0.23) \times 10^{-4}$	352 ± 96
25° C	2'-O-CH <sub>3</sub>	$(1.13 \pm 0.06) \times 10^3$	$(3.79 \pm 0.56) \times 10^{-4}$	338 ± 66
30° C	2'-O-CH <sub>3</sub>	$(1.09 \pm 0.37) \times 10^3$	$(3.90 \pm 0.28) \times 10^{-4}$	375 ± 92
40° C	2'-O-CH <sub>3</sub>	$(1.88 \pm 0.71) \times 10^3$	$(8.30 \pm 0.57) \times 10^{-4}$	480 ± 212
50° C	2'-O-CH <sub>3</sub>	$(3.19 \pm 0.12) \times 10^3$	$(1.82 \pm 0.13) \times 10^{-3}$	570 ± 14
20° C	2'-OH	$(9.83 \pm 1.10) \times 10^2$	$(1.03 \pm 0.16) \times 10^{-4}$	107 ± 29
25° C	2'-OH	$(1.24 \pm 0.35) \times 10^3$	$(1.33 \pm 0.24) \times 10^{-4}$	114 ± 33
30° C	2'-OH	$(3.12 \pm 0.13) \times 10^3$	$(2.71 \pm 0.15) \times 10^{-4}$	100 ± 41
40° C	2'-OH	$(3.30 \pm 0.88) \times 10^3$	$(9.08 \pm 0.59) \times 10^{-4}$	288 ± 95
50° C	2'-OH	$(6.80 \pm 0.57) \times 10^3$	$(3.77 \pm 0.52) \times 10^{-3}$	555 ± 21

**Table 3-2: Kinetic and dissociation constants determined by temperature-dependent SPR.** The rate and dissociation constants for loop A 2'-O-CH<sub>3</sub> loop A 2'-OH docking with A38C loop B at all temperatures assayed.



**Figure 3-3: van't Hoff plots for docking.** The solid circles are the data from the 2'-OH loop A and A38C loop B interaction at the five temperatures assayed. The open circles are the data from the 2'-OMe loop A and A38C loop B at the same five temperatures. The error bars are the experimental standard deviations for each of the five temperature points for each interaction.



**Figure 3-4: Arrhenius plots for the rates of docking and undocking.** In panel A, for the interaction of 2'-OH loop A and A38C loop B, the association (or on) rate is plotted on the top and the dissociation (off) rate is plotted on the plot against the inverse of temperature. Both traces are fit to the Arrhenius equation. In panel B, for the interaction of 2'-OMe loop A and A38C loop B are plotted. Both the association and dissociation rates could not be fit to the Arrhenius equation, it was non-linear of the range assayed, and the trace is only meant to guide the eye. The error bars are the experimental standard deviations for each of the five temperature points for each interaction.

## 3.4 DISCUSSION

### 3.4.1 The effect of the cleavage-site modification

Transitioning from wild-type loop B to the docking-competent, but catalytically inactive, A38C loop B does not result in a significant change in the dissociation constant, as they are the same within error for the interaction with the 2'-OMe loop A. However, in transitioning from 2'-OMe loop A to 2'-OH loop A with A38C loop B does result in a 3.3-fold increase in affinity. This is due, kinetically, to a 5.3-fold slower dissociation rate. This is indicative of a more stable docked complex being formed with the native 2'-OH at the cleavage site.

Extensive crystallographic studies have shown mutating the catalytic A38 residue does not result in significant structural changes to hairpin ribozyme (9,15). In addition, conclusions drawn from crystallography are that cleavage-site modifications have little effect on the interaction, based on overall structure and in-line geometry. However, this study and single-molecule FRET studies of the two-way junction (hinged) construct from the Walter lab show the cleavage-site modification does have a significant effect. It was found in the hinged construct, that transitioning from the 2'-OMe cleavage site modification to the native 2'-OH resulted in a 20-fold increase in affinity, fundamentally due to a slower dissociation rate (16-19). The values obtained by the Walter lab were by single-molecule FRET, in a unimolecular system and the ones here are from the ensemble technique of SPR for intermolecular docking, so they are not directly comparable. However, increased docking affinity in the transition from the 2'-OMe modification to the native 2'-OH due to slower dissociation is a fundamentally, similar finding in both structural constructs.

In this work, determination of the thermodynamic signature provides more insight on the cleavage-site modification. With the 2'-OMe loop A and A38C loop B the  $\Delta H^\circ$  was determined

to  $(-3.40 \pm 0.51)$  kcal mol<sup>-1</sup> and the  $\Delta S^\circ$  term was determined to be  $(18.1 \pm 1.7)$  kcal mol<sup>-1</sup> K<sup>-1</sup>. In transitioning to the 2'-OH loop A and A38C loop B the  $\Delta H^\circ$  constant was determined to be  $(-11.2 \pm 2.4)$  kcal mol<sup>-1</sup> and the  $\Delta S^\circ$  constant was  $(-5.7 \pm 4.8)$  kcal mol<sup>-1</sup> K<sup>-1</sup>. The cleavage site modification caused a significant difference in underlying thermodynamic terms for the docking process. The modification resulted in making the interactions significantly less enthalpically favorable and actually entropically favorable. These effects suggest that the 2'-OMe modification could be disrupting the network of interactions at this intricate interface. This could be done via the not mutually exclusive possibilities of steric hindrance of the larger methoxy group, the inability to hydrogen bond of the methoxy group, and/or prevention of the optimal ribose conformation of the methoxy group. This could be a fairly novel example of entropy-enthalpy compensation that has been commonly observed with protein binding-sites and small ligands (20).

Also, with the cleavage-site modification present the Arrhenius plots were nonlinear. This is interpreted as a significant disruption to the energy landscape and further evidence that the cleavage site modification does have a significant effect on the docking interaction. The interaction between 2'-OH loop A and A38C loop B yielded an activation energy barrier for the forward process (docking) of  $(11.7 \pm 2.2)$  kcal mol<sup>-1</sup>. The activation barrier for the reverse process (undocking) was determined to be  $(23.2 \pm 1.7)$  kcal<sup>-1</sup>. As mentioned earlier there are significant rearrangements that precede docking and an intricate network of interactions associated with docking. These relatively high barriers are consistent with those known processes.

### **3.4.2 Thermodynamic effects of a junction on docking**

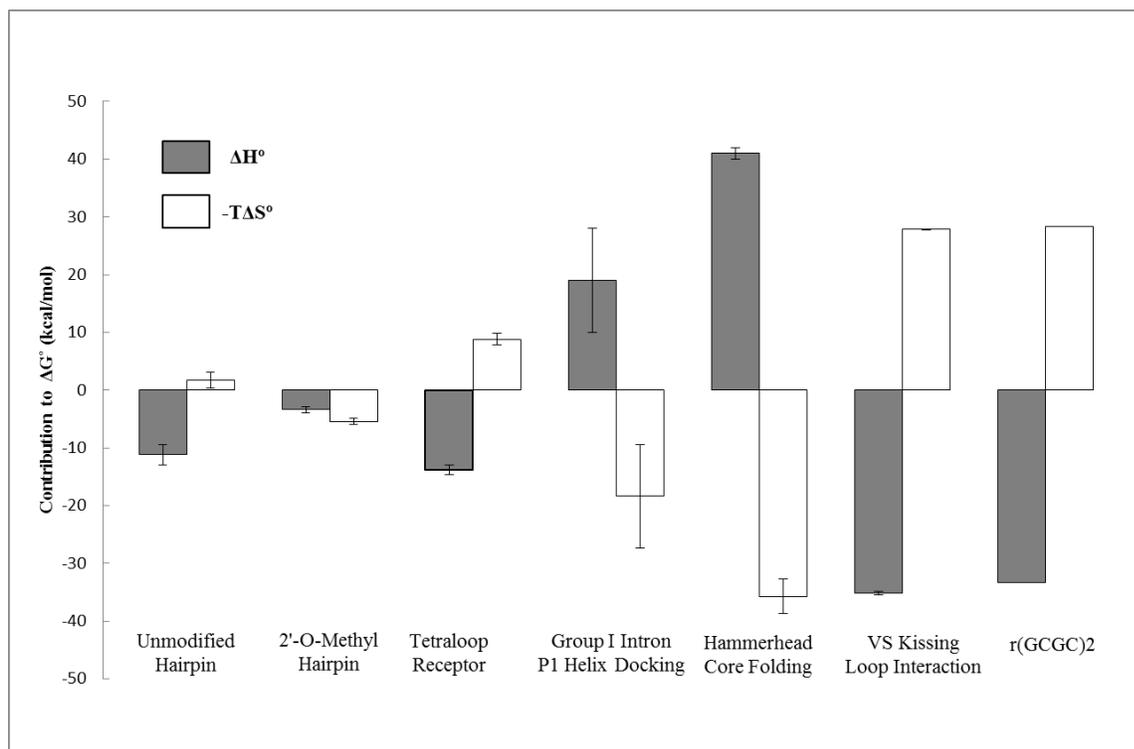
The thermodynamic signatures for the 2WJ, 3WJ, and 4WJ constructs of the hairpin ribozyme have been determined (11). This work presents the first determination, to our knowledge, of the enthalpy and entropy terms for the junctionless construct of the hairpin ribozyme. Comparisons between our construct and their constructs contain several important caveats. The cleavage-site modification utilized in their constructs with a junction forms was a 2'-H and as shown by this study, the cleavage-site modification can have a significant impact. Also, the junction constructs were studied by FRET and all had changes in heat capacity. We found fitting our data with the  $\Delta C_p^0$  term was unjustified based on not resulting in a significant enough improvement in the data fitting.

With those caveats in mind, it is interesting that for the constructs with a junction there is a significant, negative change in heat capacity that hampered docking favorability. This was interpreted, as is common, as burial of hydrophobic surface. With our construct we do not have that term. This suggests that perhaps the presence of junctions is unfavorable in this aspect as it appears to necessitate burial of hydrophobic surface. Also, the 2WJ construct had the most favorable enthalpy term relative to the 3WJ and 4WJ constructs. A possibility is that the 2WJ construct is less restrictive and more favorable contacts are made between loop A and loop B. Our enthalpy term is more favorable for the 2'-OH loop A and A38C loop B than what was determined for the 2WJ construct. Our data in comparison with what the Millar lab determined suggests that thermodynamically the presence of junction does not aid docking.

### **3.4.3 Thermodynamic strategies of RNA tertiary structure formation**

As mentioned earlier there are limited examples of RNA tertiary structure formation, in the absence of helix annealing. In Figure 3-5, this study is put into a broader context of

thermodynamic studies of RNA tertiary structure formation (21-25). With the selected examples there are different systems and different ionic conditions as well as some having changes in heat capacities, so quantitatively, comparisons are complicated. However, with several examples it is clear that RNA tertiary structure formation can utilize a variety of thermodynamic strategies. Also, from our study, it is clear that a seemingly minor difference can greatly affect the thermodynamic signature.



**Figure 3-5: Various thermodynamic signatures of RNA tertiary structure formation.** The hairpin systems are from this work. The rest in succession are from the Butcher, Herschlag, Feig, and Legault labs and the final system is a RNA duplex, included for reference (21-25).

### **3.5 CONCLUSION**

From this study of RNA tertiary structure formation of the junctionless hairpin ribozyme it was found that a cleavage-site modification decreased the stability of the docked complex. This is consistent with findings from the Walter lab. The modification also greatly affected the thermodynamic signature of docking as well as the energy landscape for our junctionless construct. This study shows the importance of complementing RNA structural studies with biochemical or biophysical characterization of the interaction, especially when cleavage site modifications are used.

## **REFERENCES**

## REFERENCES

1. Wedekind, J. (2011) Metal ion binding and function in natural and artificial small RNA enzymes from a structural perspective. *Metal ions in life sciences*, **9**, 299-345.
2. Hoogstraten, C.G. and Sumita, M. (2007) Structure-function relationships in RNA and RNP enzymes: recent advances. *Biopolymers*, **87**, 317-328.
3. Zhang, Q. and Al-Hashimi, H.M. (2009) Domain-elongation NMR spectroscopy yields new insights into RNA dynamics and adaptive recognition. *RNA (New York, N.Y.)*, **15**, 1941-1948.
4. Schroeder, R., Barta, A. and Semrad, K. (2004) Strategies for RNA folding and assembly. *Nat Rev Mol Cell Biol*, **5**, 908-919.
5. Thirumalai, D. and Woodson, S.A. (1996) Kinetics of Folding of Proteins and RNA. *Accounts of Chemical Research*, **29**.
6. Tinoco, I. and Bustamante, C. (1999) How RNA folds. *Journal of molecular biology*, **293**, 271-281.
7. Cai, Z. and Tinoco, I. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.
8. Butcher, S., Allain, F. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature structural biology*, **6**, 212-216.
9. MacElrevey, C., Salter, J.D., Krucinska, J. and Wedekind, J.E. (2008) Structural effects of nucleobase variations at key active site residue Ade38 in the hairpin ribozyme. *RNA (New York, N.Y.)*, **14**, 1600-1616.
10. Klein, D.J. and Ferre-D'Amare, A.R. (2006) Structural basis of glmS ribozyme activation by glucosamine-6-phosphate. *Science*, **313**, 1752-1756.
11. Klostermeier, D. and Millar, D.P. (2000) Helical junctions as determinants for RNA folding: origin of tertiary structure stability of the hairpin ribozyme. *Biochemistry*, **39**, 12970-12978.
12. Sumita, M., White, N.A., Julien, K.R. and Hoogstraten, C.G. (2013) Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis. *RNA Biol*, **10**, 425-435.
13. Liu, S., Bokinsky, G., Walter, N.G. and Zhuang, X. (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. Mlýnský, V., Banás, P., Hollas, D., Réblová, K., Walter, N.G., Sponer, J. and Otyepka, M. (2010) Extensive molecular dynamics simulations showing that canonical G8 and protonated A38H<sup>+</sup> forms are most consistent with crystal structures of hairpin ribozyme. *The journal of physical chemistry. B*, **114**, 6642-6652.
  15. Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. and Wedekind, J.E. (2006) Water in the active site of an all-RNA hairpin ribozyme and effects of Gaa8 base variants on the geometry of phosphoryl transfer. *Biochemistry*, **45**, 686-700.
  16. Hampel, K.J., Walter, N.G. and Burke, J.M. (1998) The solvent-protected core of the hairpin ribozyme-substrate complex. *Biochemistry*, **37**, 14672-14682.
  17. Zhuang, X., Kim, H., Pereira, M.J., Babcock, H.P., Walter, N.G. and Chu, S. (2002) Correlating structural dynamics and function in single ribozyme molecules. *Science (New York, N.Y.)*, **296**, 1473-1476.
  18. Bokinsky, G., Rueda, D., Misra, V.K., Rhodes, M.M., Gordus, A., Babcock, H.P., Walter, N.G. and Zhuang, X. (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.
  19. Rueda, D., Bokinsky, G., Rhodes, M.M., Rust, M.J., Zhuang, X. and Walter, N.G. (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.
  20. Fenley, A.T., Muddana, H.S. and Gilson, M.K. (2012) Entropy-enthalpy transduction caused by conformational shifts can obscure the forces driving protein-ligand binding. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 20006-20011.
  21. Vander Meulen, K.A., Davis, J.H., Foster, T.R., Record, M.T. and Butcher, S.E. (2008) Thermodynamics and folding pathway of tetraloop receptor-mediated RNA helical packing. *Journal of molecular biology*, **384**, 702-717.
  22. Narlikar, G.J. and Herschlag, D. (1996) Isolation of a local tertiary folding transition in the context of a globally folded RNA. *Nature structural biology*, **3**, 701-710.
  23. Mikulecky, P.J., Takach, J.C. and Feig, A.L. (2004) Entropy-driven folding of an RNA helical junction: an isothermal titration calorimetric analysis of the hammerhead ribozyme. *Biochemistry*, **43**, 5870-5881.

24. Lacroix-Labonté, J., Girard, N., Dagenais, P. and Legault, P. (2016) Rational engineering of the *Neurospora* VS ribozyme to allow substrate recognition via different kissing-loop interactions. *Nucleic acids research*, **44**, 6924-6934.
25. Xia, T., SantaLucia, J., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C. and Turner, D.H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry*, **37**, 14719-14735.

## **CHAPTER 4**

### **THE RELATIONSHIP BETWEEN RIBOSE DYNAMICS AND CATALYSIS AT THE CLEAVAGE SITE OF THE LEAD-DEPENDENT RIBOZYME**

## ABSTRACT

In both the NMR solution structure and the X-ray crystallographic structure of the lead-dependent ribozyme or leadzyme, in-line geometry at the cleavage site is not observed. This necessitates conformational rearrangement to be consistent with proposed in-line nucleophilic attack, as with other small-self cleaving ribozymes. This has led us and others to the hypothesis that the inactive, ground state(s) observed in structural studies are present a majority of the time and there is a transition to an active conformation that is present a minority of the time.

To investigate ribose dynamics of the leadzyme we have carried out NMR spin-relaxation studies using a site-specific labelling scheme previously developed by our lab. The labelling scheme makes these measurements possible. We specifically studied the ribose of C6, which contains the nucleophilic 2'-hydroxyl for the self-cleavage reaction, over the ms- $\mu$ s timescale using  $^{13}\text{C}$  CPMG and  $R_{1\rho}$  relaxation dispersion studies. We found that the ribose exhibits conformational exchange lifetimes on the low  $\mu$ s (5-10  $\mu$ s) end of this regime, meaning that the alternate ribose conformation is briefly sampled.

It is important to ascertain if these ribose dynamics are related to catalysis (function). To test that we needed to probe molecular motion as it relates to function, or in other words carry out a dynamics-function study. To accomplish this we used a bicyclo-nucleotide, at the C6 position, which is locked in the ribose conformation found in both structural studies but retains the nucleophilic hydroxyl group. We found that this ribose modification dramatically reduces catalytic activity. This leads to the conclusion that the ribose conformation fluctuation for the observed ground state at C6, adjacent to the scissile phosphate is a critical step in achieving a catalytically competent conformation.

## 4.1 INTRODUCTION

### 4.1.1 The lead-dependent ribozyme

The lead-dependent ribozyme or leadzyme is an artificial ribozyme developed by *in vitro* selection of yeast tRNA<sup>Phe</sup> (1). It performs self-cleavage in the presence of lead (2). Genomic searches were utilized to try to determine the potential importance of this catalytic motif in lead toxicity (3). However, it has been pointed out that at lead concentrations necessary to activate the leadzyme important proteins would have problems that would drastically affect health prior to lead-dependent RNA cleavage (4).

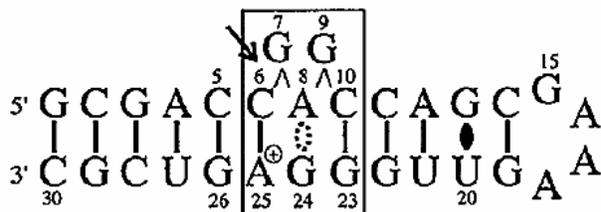
We are studying the leadzyme as a model of how RNA ribose dynamics relate to function. In both the NMR solution structure and X-ray crystallographic structures, the nucleophilic ribose and scissile phosphate are not in the appropriate in-line arrangement for self-cleavage (Figure 4-1) (5-7). This leads to the hypothesis that rearrangement, via conformational dynamics, is obligatory for catalysis.

The secondary structure of the leadzyme is quite simple. There is a six-nucleotide asymmetric internal loop flanked by A-form helices (Figure 1-1, Panel A). The only conserved nucleotides are C6, G9, and G24 of the internal loop (8,9). The leadzyme accelerates cleavage by approximately  $10^3$ -fold, which is less than the minimum of  $10^6$ -fold acceleration from the naturally occurring small, self-cleaving ribozymes (2,10). Another distinction the leadzyme has is not concluding catalytic activity with a 2',3'-cyclic phosphate and 5'-hydroxyl, like the naturally occurring small, self-cleaving ribozymes, but hydrolyzing the 2',3'-cyclic phosphate to a 3'-phosphate terminus also (2).

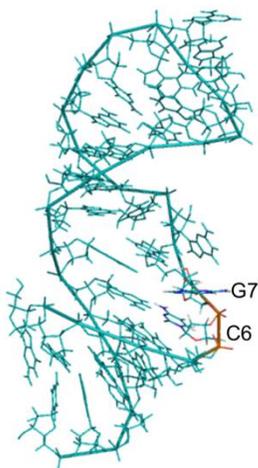
Its relatively small size (30 nucleotides), minimal requirements for cleavage, and previous structural studies that propose dynamics are essential to obtaining an active

conformation, make it an excellent candidate for dynamics studies (11-13). Also, as laid out in chapters 1-3 and 5, structural rearrangements, via conformational changes, are necessary in the hairpin ribozyme obtaining a catalytically-competent state, thus the leadzyme models this pattern. Another feature of the leadzyme (like the naturally occurring hammerhead ribozyme) is that although divalent metal ions ( $\text{Pb}^{2+}$  for the leadzyme and  $\text{Mg}^{2+}$  for the hammerhead ribozyme) are implicated in the active site chemistry for the cleavage reaction, addition of divalent metal ions do not alter the structure of the ground state in a significant manner (12).

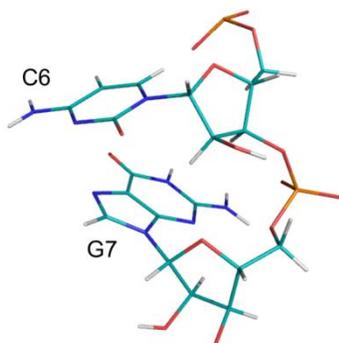
A



B



C



**Figure 4-1: Secondary and tertiary structure of the leadzyme.** Panel A is from ref.(5). It depicts the secondary structure of the leadzyme. Panel B is the NMR solution structure that was presented in ref. (5) (PDB: 2LDZ). Panel C shows the cleavage site not obtaining the in-line geometry between C6 and G7. Panels B and C was rendered using PyMOL (14).

#### 4.1.2 Ribose dynamics and catalysis in the lead-dependent ribozyme

There are two predominant ribose conformations, C3'-endo (“North”), found in A-form helices and C2'-endo (“South”), found in B-form helices (15). The ribose sugars in the NMR solution structure leadzyme, like loop A and loop B of the hairpin ribozyme, are found predominately in the C3'-endo or the C2'-endo conformation or interconverting between the two conformations (11,16,17). In the NMR solution structure of the leadzyme the ribose conformation of C6 was determined to be C3'-endo, known directly from observing J-coupling constants (5).

In general, the primary goal of the work described in this chapter is to investigate the functional relevance of those conformations and transitions between them. This is inherently challenging, for several reasons that will be explained. As an analogy, the well-established technique of alanine scanning mutagenesis allows for insight to be obtained into the relevance of various functional groups of amino acids (18). In our application here, with ribose conformations, we are mutating to remove a motional property (i.e. access to a minor conformation), rather than a functional group.

To this end, our lab has previously used commercially available LNA (locked nucleic acid), which is covalently locked in the C3'-endo conformation (Figure 1-2) (19). In that work, LNA was substituted for the ribose positions of G7, G9, and G24. Interestingly, despite the G9 ribose having the C2'-endo conformation in the both the NMR solution structure and X-ray crystal structure, using LNA to restrict the conformation to C3'-endo resulted in ca. 20-fold increase in catalysis above wildtype (19). That specific result exemplifies the need to go beyond the ground state structure and investigate states that may be only a small fraction of the population. (The other results from this study are summarized in Table 1-1). In general, this

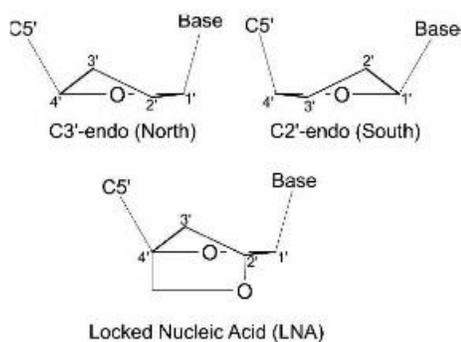
study demonstrates the functional relevance of ribose conformation as it relates to lead-dependent cleavage, i.e. the function of the leadzyme.

With C6 being one of three conserved residues and being at the cleavage site, the ribose dynamics for it are potentially very important in the function of the leadzyme. Due to the fact that LNA does not retain the nucleophile at the 2'-position cleavage would be abolished if it were used, irrespective of any conformational preference. To counter this problem, Dr. Victor Marquez provided our lab with a bicyclo[3.1.0]hexane pseudosugar-nucleotide that is locked in the C3'-endo but retains the nucleophilic hydroxyl.

Dr. Minako Sumita of our lab used this conformationally restricted ribose, with the nucleophile, to test the conformational preference at the cleavage site (Figure 4-3). It was found that restricting the C6 ribose to the C3'-endo conformation resulted in nearly abolishing catalysis. This is a very interesting result, given that the ribose conformation of C6 is C3'-endo in the solved structures (5,6). This leads to the hypothesis that the C6 ribose is undergoing fluctuation from its ground conformation of C3'-endo, to an excited state of C2'-endo, and that is an important part of the active conformation of the leadzyme.

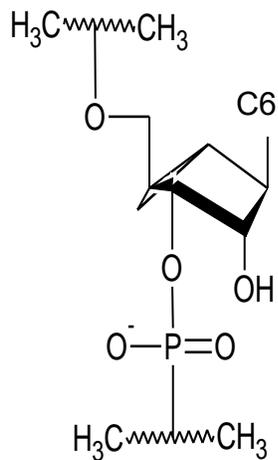
Residue	NMR	X-ray	Conformational Restriction with Probe	Dynamics-Function Probe	Effect on catalysis
C6	C3'-endo	C3'-endo	C3'-endo	Bicyclo-NT	Drastic Decrease
G7	C2'-endo	C3'-endo	C3'-endo	LNA	2-fold Decrease
G9	C2'-endo	C2'-endo	C3'-endo	LNA	22-fold Increase
G24	C3'-endo	C3'-endo	C3'-endo	LNA	17-fold Decrease

**Table 4-1: Effect of dynamics-function probes.** The solution NMR structure is PDB: 2LDZ and the X-ray crystal structure is PDB: 429D (5,6). The conformational probe LNA was used in a previous study performed by our lab (19). The result with the bicyclo-nucleotide is presented here.

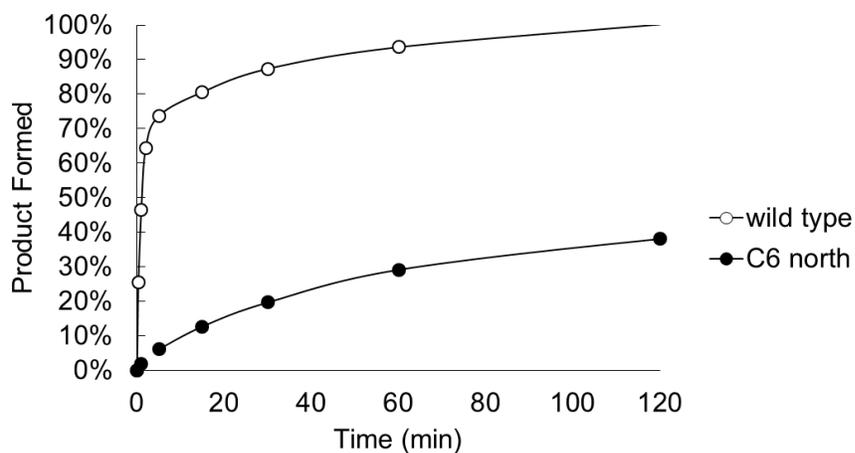


**Figure 4-2: Ribose conformations.** This figure is from reference (19). It shows the two predominant ribose conformations and the covalent structure of the locked nucleic acid (restricted to C3'-endo).

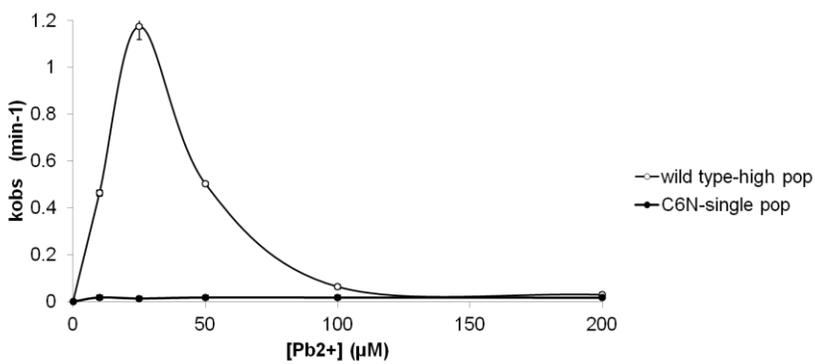
A



B



C



**Figure 4-3: Effect of locking the C6 nucleophilic ribose in the C3'-endo conformation.** This assay, performed by Dr. Minako Sumita of our lab, shows the drastic reduction in cleavage that occurs with this dynamics-function probe.

### 4.1.3 NMR studies of ribose dynamics

Investigating the ribose dynamics at the C6 (cleavage site) position is extremely well-suited to NMR spin-relaxation techniques, which are unmatched in their ability to probe multiple timescales in a site-specific manner (20-24). In order to investigate short-lived, low-population, ribose conformations, it is necessary to probe the ms- $\mu$ s timescale (22). Within this (fast-exchange) regime the techniques of CPMG and power-dependent  $R_{1\rho}$  are directly applicable and they have been used successfully in proteins (see Chapter 1 for a specific example) (22,25). Unfortunately, this technique has been previously greatly hampered in its application to studying ribose dynamics and thus important studies of ribose dynamics lag significantly behind studies of base dynamics and protein, in general (26).

The challenge, for ribose dynamic studies, is that in commercially available, uniformly labelled  $^{13}\text{C}$ -ribose nucleotides are unsuitable for these experiments. In essence, there is magnetic transfer among the carbon nuclei and that leads to uninterpretable data. For example, when measuring  $T_1$  spin-relaxation, the decays are typically multi-exponential, rather than well-described, interpretable single-exponential decays (27,28). This renders the previously mentioned CPMG and power-dependent  $R_{1\rho}$  experiments ineffectual. Hartman-Hahn transfer has been exploited in solid state NMR spectroscopic studies but it is extremely confounding in studying ribose dynamics (29,30).

The innovative approach, developed by our lab, of metabolically-directed isotope, allows for site-specific  $^{13}\text{C}$  labeling of the C2' and C4' carbons of the ribose ring (31). This prevents magnetic transfer within the ring and allows for the CPMG and power-dependent  $R_{1\rho}$  experiments to be utilized to examine the ms- $\mu$ s timescale, typically important to ribose conformational dynamics (26,31). It also allows for analysis of sub-ns disorder via the Model-

Free approach. This allows for a thorough investigation of the fast exchange regime and determination. Specifically, we have applied this approach to examine if the C6 ribose is undergoing conformational fluctuations to a minor C2'-endo state, as functionally implicated by the kinetic data described above.

## **4.2 MATERIALS AND METHODS AND RESULTS**

### **4.2.1 RNA preparation and verification**

The site specific labelling scheme, previously developed by our lab, allows for  $^{13}\text{C}$  2'-4' labelling in the ribose of the nucleotides (see Chapter 2 of Dr. James Johnson's thesis) (31). Briefly, mutant *E. coli* are grown using labelled glycerol, which incorporates the labels in a specific pattern.

The preparation of  $^{13}\text{C}$  2'-4' nucleotides begins with growing *E. coli* on minimal medium with 2- $^{13}\text{C}$  glycerol as the only carbon source. The strain used in this application is the kanamycin resistant *E. coli* strain, JW1841-1. This is an improvement over the K10-15-16 strain, which is not antibiotic resistant and has been used previously (31). The essential trait of the *E. coli* strain (DL323) used here is that they are deficient in glucose 6-phosphate dehydrogenase. A related procedure with different variations in carbon metabolism can also be used to yield favorable labelling in aliphatic and aromatic amino acid side chains (32).

The cells are isolated, lysed, and the nucleic acid is isolated. After isolating the nucleic acid is isolated it is hydrolyzed and DNA and RNA are separated. Incidentally, the DNA is also labeled in the same  $^{13}\text{C}$  2', 4' pattern as the RNA but we have not taken advantage of this material to date. The nucleotides are de-salted using a G-10 column (Sigma Aldrich) prior to NMP separation. Next, the ribonucleotides are separated by their base yielding individual pools of  $^{13}\text{C}$  2', 4' labeled CMP, AMP UMP, and GMP. After separation of the monophosphates they

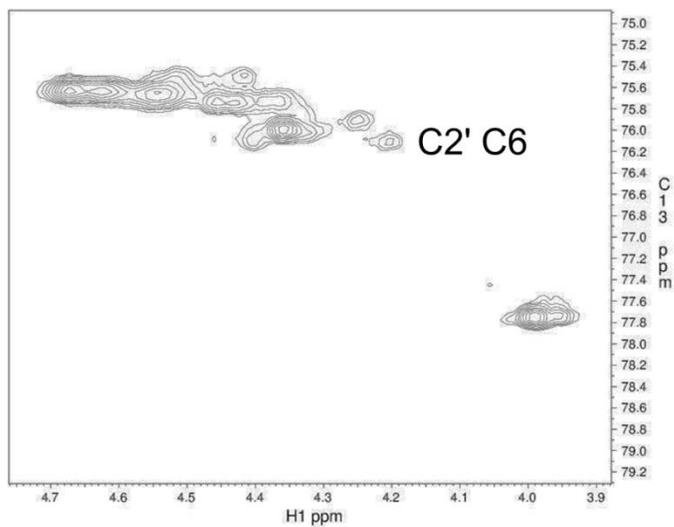
are rephosphorylated and can be used in transcription. Importantly, the only labeled nucleotide used in the transcription reaction for this work was CTP with the ribose labeled  $^{13}\text{C}$  at the 2', 4' positions. It is necessary to optimize the *in vitro* transcription reaction, especially in terms of nucleotide of interest and magnesium concentration. (For more detail refer to the in-house protocol of this procedure.)

Transcription and purification of the RNA was performed as previously described in subsection 2.2.1, with the additional step of repeated lyophilization in 99.8% deuterium oxide ( $\text{D}_2\text{O}$ ). The sample was finally brought up in 255  $\mu\text{L}$  of 99.96%  $\text{D}_2\text{O}$ . The concentration of the leadzyme sample was 0.553 mM and in a buffer of 10 mM sodium phosphate, 100 mM sodium chloride, and 200  $\mu\text{M}$  EDTA at pH 5.5.

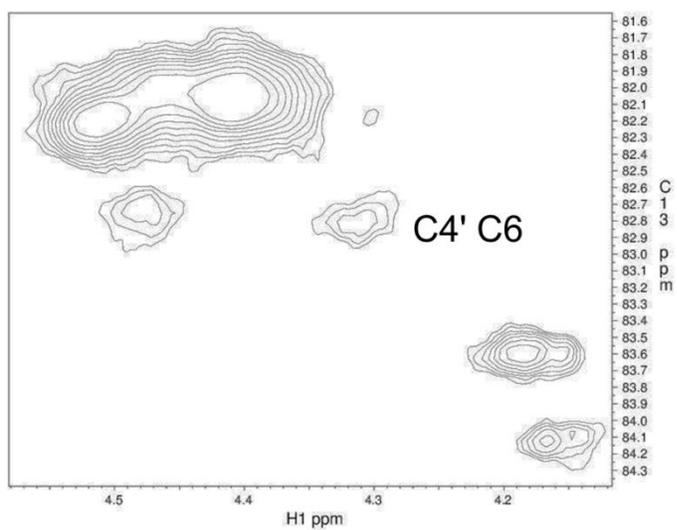
#### **4.2.2 Ribose dynamics (Power-dependent $T_{1\rho}$ relaxation rates and CPMG)**

All the NMR data was acquired with a Varian UNITY INOVA 600 MHz ( $^{13}\text{C}$  at 150 MHz) spectrometer at 25° C. A  $^1\text{H}$ - $^{13}\text{C}$  HSQC (heteronuclear single quantum coherence) spectrum (Figure 4-4) was used to compare to the previous assigned chemical shifts (33). The chemical shifts of the  $^{13}\text{C}$  at the 2'-4' positions of cytosine were unambiguously confirmed via 3D-NOESY (data not shown).

A



B



**Figure 4-4: 2', 4' cytidine leadzyme.** Panel A depicts the C2' region and panel B depicts the C4' region of the spectrum. Both ribose peaks of C6 are well-resolved and suitable for relaxation studies.

The techniques of CPMG and power dependent  $T_{1\rho}$  relaxation have been utilized and well-reviewed by the labs of Palmer and Kay (22,25,34-40). They are used to characterize site-specific motions on the  $\mu$ s-ms timescale. They effectively rely on two conformations, e.g. the ground and excited state of the leadzyme, existing and exchanging between the two conformations. In this chapter, to study ribose dynamics, the relaxation ( $^{13}\text{C}$ ) techniques of CPMG (Carr-Purcell-Meiboom-Gill) and a power dependent  $T_{1\rho}$  series were used as previously described with minor variations (26).

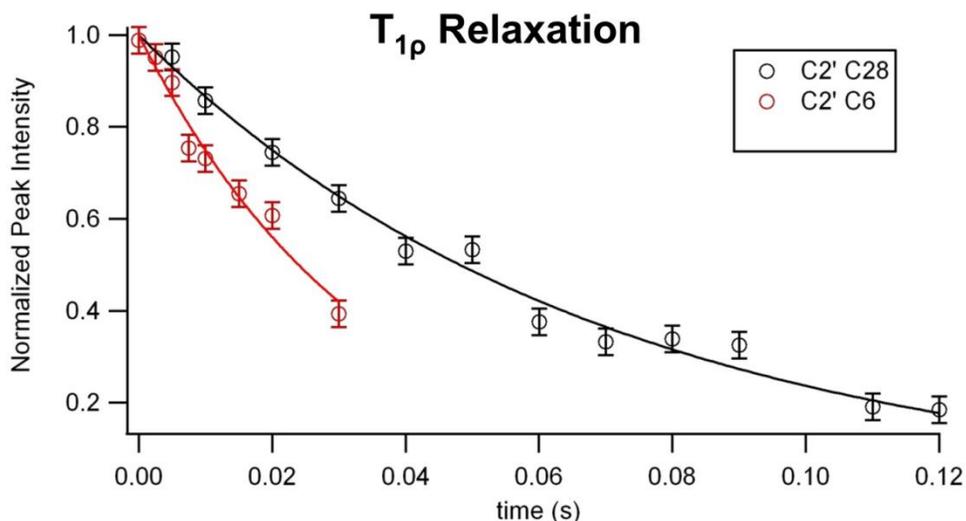
The initial experiment performed was a  $T_{1\rho}$  relaxation series (Figure 4-5 and Figure 4-6). Since the assignments for the nuclear spins of the leadzyme are known we are able to pair the relaxation rate with the appropriate site. The spectra in this series were initially taken with relaxation times ranging from 10-120 ms. For this data 1024 x 80 complex points (corresponding to digital resolution in the  $^1\text{H}$  and  $^{13}\text{C}$  dimensions, respectively) were used with 56 transients. For this data and all subsequent data the proton radio frequency (RF) carrier was centered on the residual HDO and the  $^{13}\text{C}$  carrier was set at 80 ppm. The corresponding spectral widths were 4 ppm and 14 ppm. This range of relaxation times was adequate for all spins except those related to C6, due to its faster relaxation rate. Minor changes to the pulse sequence were made to increase sampling at lower relaxation times. The relaxation series was retaken using that pulse sequence, relaxation times ranging from 0-60 ms, and that yielded an acceptable and comprehensive fit, as shown in Figure 4-4. For this data 256 x 92 complex points were used with 104 transients, to enhance signal to noise.

All NMR spectra were processed using NMRPipe (41). The appendix for this chapter contains the detailed procedure for using this software. We are grateful to Dr. Frank Delaglio for providing specific processing scripts to aid in the determination of relaxation rates. A Lorentz-to-

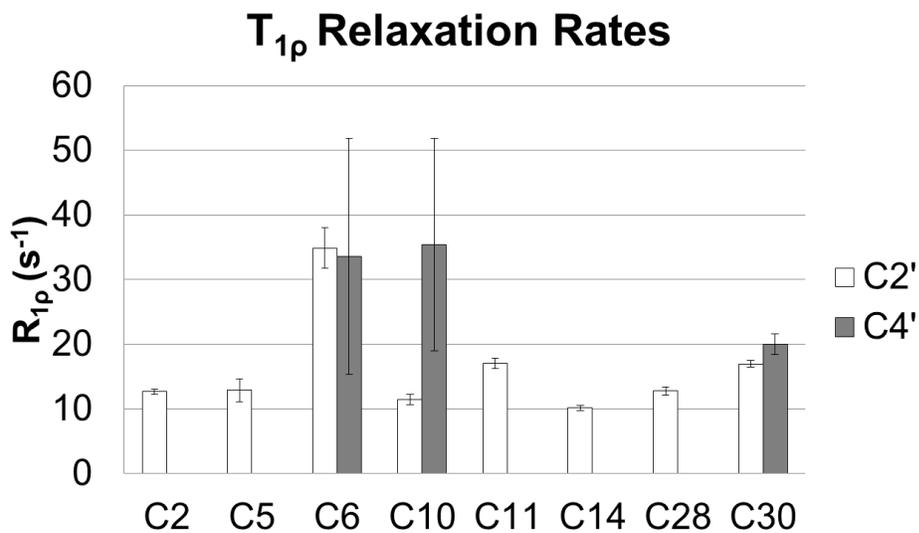
Gauss window function, which is the combination of an inverse exponential function and Gaussian function, was applied in both dimensions, to all spectra used in relaxation. For the proton dimension the Lorentzian line-sharpening was 15 Hz and the Gauss line broadening was 18 Hz. For the carbon dimension the Lorentzian line-sharpening was 6 Hz and the Gauss line broadening was 12 Hz.

The relaxation dispersion series, i.e. power dependent, on-resonance  $T_{1\rho}$  series was taken by varying RF powers,  $\gamma B_1/2\pi$  between 1.5-6.0 kHz (Table 4-2). At each RF power a  $T_{1\rho}$  series was taken with relaxation times from 0-120 ms. By plotting peak intensity as a function of relaxation time a single-exponential decay fit yields the relaxation rate. The parameters for these spectra were 256 x 92 complex points with 64 transients.

CPMG experiments with effective spin-lock powers below the range of the power dependent, on-resonance  $T_{1\rho}$  series were also performed. These experiments fundamentally use high-power-pulses to refocus spins. Improved refocusing as a function of increased frequency ( $\nu_{\text{CPMG}}$ ) of refocusing pulses is evidence of two conformations exchanging on approximately the ms timescale. No variation in refocusing as function of  $\nu_{\text{CPMG}}$  was found (data not shown). The parameters for the CPMG series were the same as those for the relaxation dispersion series.



**Figure 4-5: T<sub>1ρ</sub> relaxation series comparing C6 ribose and a typical helical residue.** In this plot the peak intensity of both sites has been normalized so they can be viewed on the same axis, making the difference distinct.



**Figure 4-6: T<sub>1ρ</sub> relaxation rates for resolved sites in the <sup>13</sup>C, 2'-4' cytidine leadzyme.** Each relaxation rate is the result of the fit for the relaxation curve of each site e.g. Figure 4-5.

	1500	error	2000	error	2500	error	3000	error	3500	error
C2C2'	12.48	0.72	13.32	0.48	12.67	0.41	12.73	0.45	11.42	0.43
C5C2'	12.82	2.60	n.d.		12.90	1.78	12.94	1.43	10.02	1.12
<b>C6C2'</b>	<b>28.92</b>	<b>4.95</b>	<b>27.18</b>	<b>3.38</b>	<b>34.89</b>	<b>6.55</b>	<b>34.38</b>	<b>6.07</b>	<b>29.04</b>	<b>5.49</b>
<b>C6C4'</b>	<b>30.36</b>	<b>11.26</b>	<b>32.02</b>	<b>8.08</b>	<b>33.31</b>	<b>18.52</b>	<b>33.46</b>	<b>17.69</b>	<b>29.08</b>	<b>15.46</b>
C10C2'	9.93	1.44	11.23	1.01	11.46	0.81	11.28	0.93	9.99	0.84
C10C4'	24.00	8.91	26.08	6.46	35.37	16.44	33.63	11.80	23.08	11.57
C11C2'	15.61	1.37	17.46	0.90	17.04	0.83	18.48	0.89	14.97	0.83
C14C2'	9.09	0.73	9.52	0.50	10.15	0.39	10.48	0.44	9.05	0.43
C28C2'	12.17	1.02	13.45	0.73	12.78	0.59	13.36	0.69	11.26	0.62
C30C2'	15.42	0.88	16.42	0.57	16.96	0.52	15.55	0.50	13.96	0.48
C30C4'	17.38	2.32	19.92	1.55	19.99	1.59	22.09	1.72	16.98	1.67

	4000	error	4500	error	5000	error	5500	error	6000	error
C2C2'	12.72	0.48	12.46	0.49	11.54	0.48	11.7	0.53	12.56	0.56
C5C2'	11.81	1.11	15.52	2.91	15.16	4.01	16.25	9.39	17.03	1.98
<b>C6C2'</b>	<b>31.31</b>	<b>6.27</b>	<b>29.52</b>	<b>5.87</b>	<b>25.67</b>	<b>3.77</b>	<b>34.53</b>	<b>5.65</b>	<b>29.67</b>	<b>5.33</b>
<b>C6C4'</b>	<b>32.53</b>	<b>17.13</b>	<b>31.03</b>	<b>12.57</b>	<b>22.05</b>	<b>8.62</b>	<b>28.7</b>	<b>22.62</b>	<b>25.05</b>	<b>13.34</b>
C10C2'	10.31	0.86	10.56	0.97	13.08	3.20	10.20	1.00	11.04	1.09
C10C4'	24.02	11.23	28.78	12.41	22.12	7.42	22.81	19.89	26.78	14.35
C11C2'	18.44	0.90	16.32	0.90	14.91	0.83	16.65	0.99	16.58	1.03
C14C2'	9.88	0.44	9.74	0.45	8.98	0.44	9.29	0.46	9.77	0.54
C28C2'	12.42	0.67	12.19	0.68	11.02	0.68	12.02	0.74	12.15	0.8
C30C2'	16.33	0.57	15.52	0.51	13.25	0.45	16.2	0.57	16.52	0.7
C30C4'	19.97	1.71	19.74	1.57	17.04	1.49	19.07	1.88	20.59	2.06

**Table 4-2: Power-dependent  $T_{1\rho}$  series.** The top row is the power ( $\gamma B_1/2\pi$ ) in Hz that a given experiment was performed at. The observed  $T_{1\rho}$  rates ( $s^{-1}$ ) are listed for each site at a given power along with the error obtained from the fit ( $s^{-1}$ ).

### **4.2.3 Ribose dynamics ( $T_1$ relaxation rates and hNOE experiments)**

For  $T_1$  relaxation rates relaxation delays ranged from 10-1200 ms (Figure 4-7 and Figure 4-8). The spectra were obtained using 1224 x 80 complex points and 40 transients. Saturated and nonsaturated heteronuclear NOE (Nuclear Overhauser Effect) experiments were also performed in an interleaved fashion (Figure 4-8). The parameters for these experiments were 1024 x 96 complex points with 64 transients.

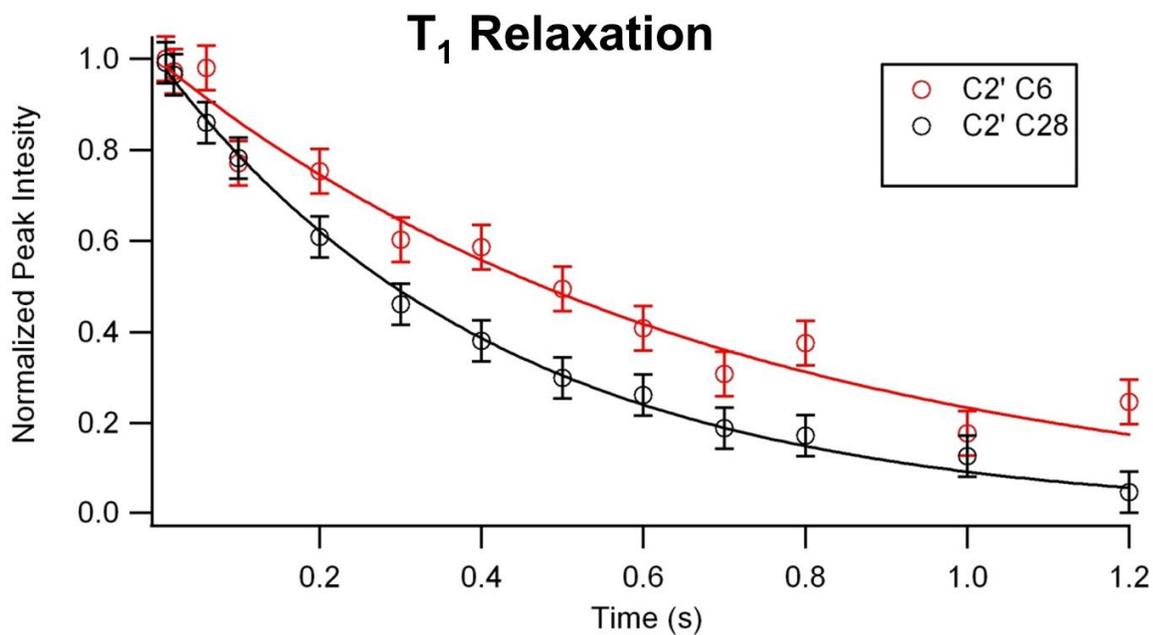


Figure 4-7: T<sub>1</sub> relaxation series comparing C6 ribose and a typical helical residue.

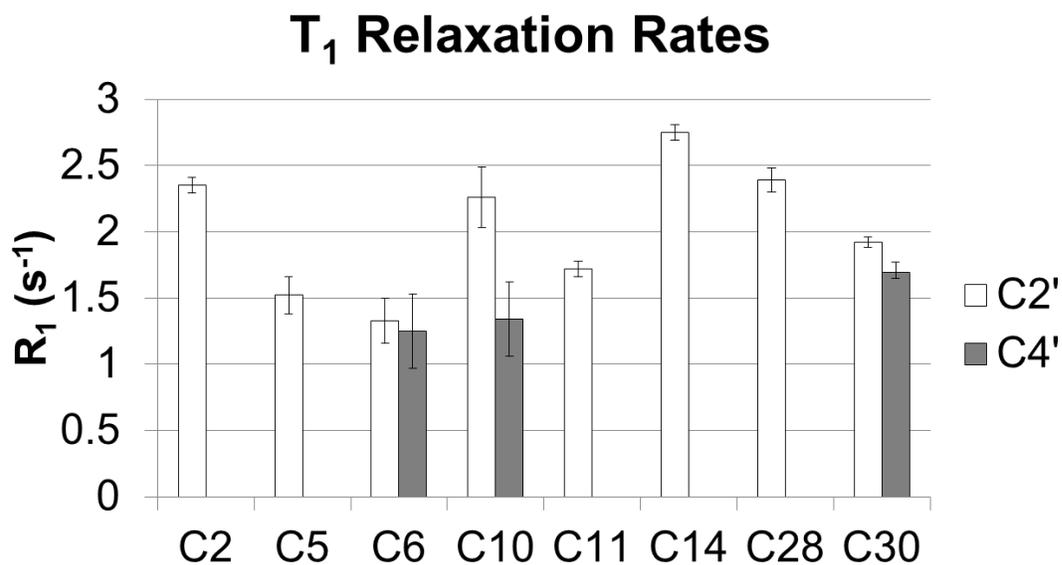


Figure 4-8: T<sub>1</sub> relaxation rates for resolved sites in the <sup>13</sup>C, 2', 4' cytidine leadzyme.

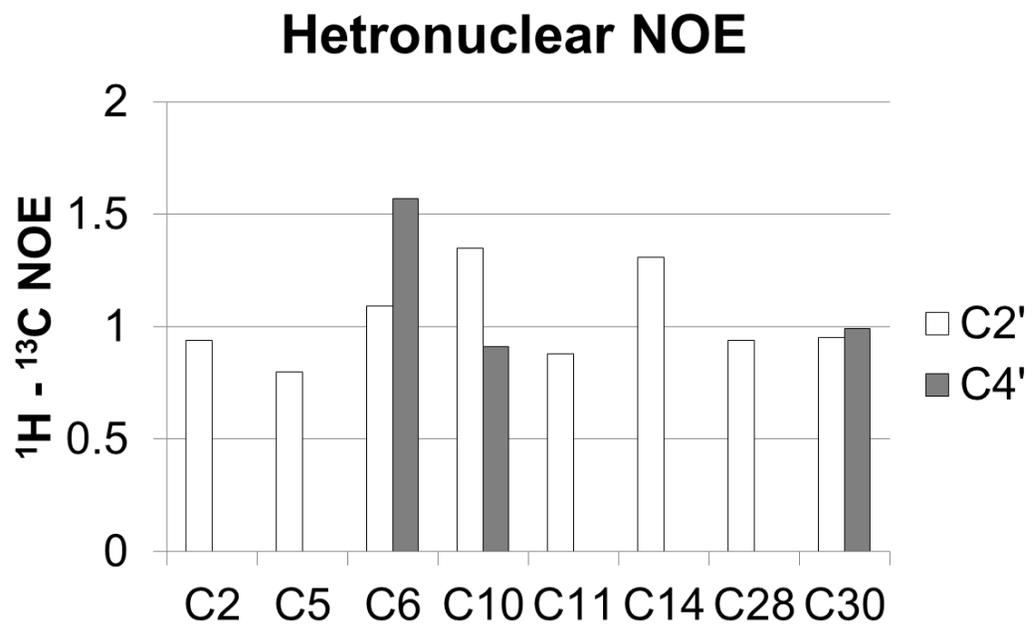


Figure 4-9: hNOE ratios for resolved sites in the  $^{13}\text{C}$ , 2', 4' cytidine leadzyme.

## 4.3 DISCUSSION

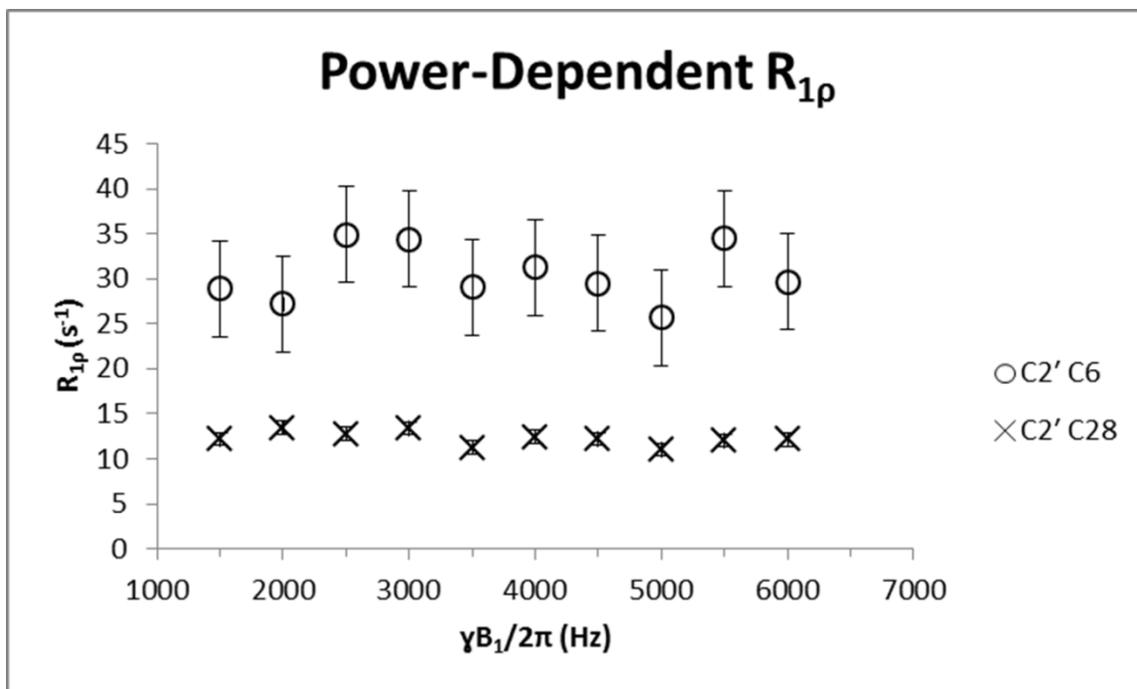
### 4.3.1 Conformational exchange on the $\mu\text{s}$ -ms timescale

The innovative site-specific labelling scheme previously developed by our lab made the relaxation measurements presented here possible (26,31). From the  $T_{1\rho}$  (rotating-frame) relaxation rates (Figure 4-5 and Figure 4-6), it is clear that the C2' and C4' sites of the C6 ribose relax at a faster rate than the helical residues. This is evidence of molecular motion on the  $\mu\text{s}$ -ms timescale, since such motions contribute an extra exchange term to the transverse relaxation. The relaxation rates for helical ribose sites allow a comparison to bond vectors that are rigid with respect to the molecular frame (38). The conformational exchange observed here is most likely the transition from its ground state of C3'-endo to rare excited state of C2'-endo (26,42).

With evidence of the existence of conformational exchange, we sought to study the rate of exchange between the two states by using power dependent  $R_{1\rho}$  (Table 4-2 and Figure 4-10). Figure 4-10 shows that the  $R_{1\rho}$  of the C6 ribose remains elevated and nearly constant over the ranged assayed. Determination of parameters relies on fitting a variation in observed relaxation rate with applied  $B_1$  power. Since that is not observed here, we are only able to put an upper bound on the conformational exchange lifetime of 5-10  $\mu\text{s}$ . These observations are also consistent with the CPMG studies we performed that did not provide evidence of conformational exchange on the ms timescale.

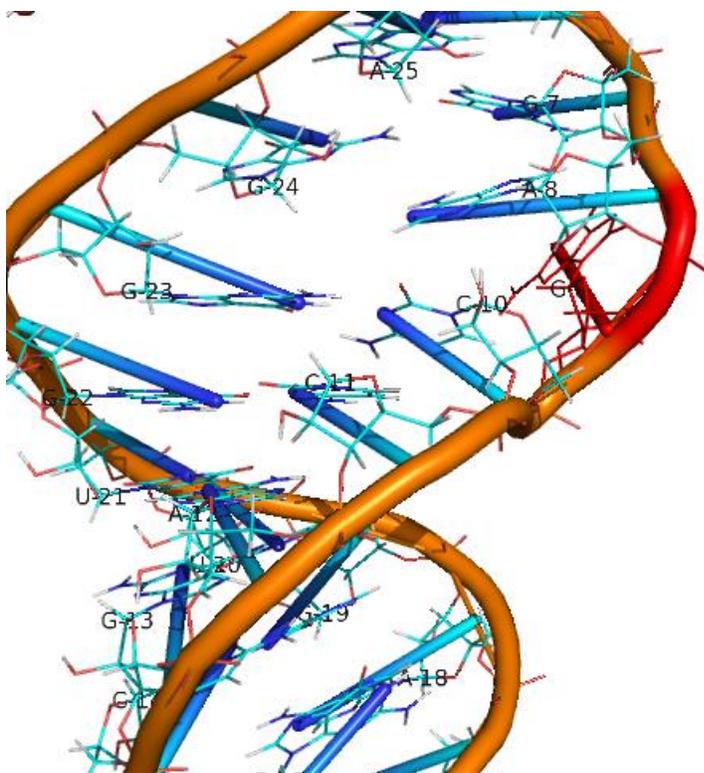
We suspect that utilizing higher power for the power-dependent  $R_{1\rho}$  series will eventually result in decreased observed rates and allow for precise quantification of the rate of conformational exchange. Increasing the power with the NMR spectrometer that was used in these experiments risks damaging the probe on the instrument used. We are currently collaborating with Dr. Alex Hansen of Ohio State University. With the more modern probes

available there, specifically an 800 MHz spectrometer, there is potential to substantially increase the applied  $B_1$  power and thus quantify the rate of conformational exchange.



**Figure 4-10: Comparing the C6 ribose to a typical helical ribose over the power-dependent series.** Data displayed here is from Table 4-2 with the average error in the relaxation rates from each site propagated for each site.

Another site that shows an elevated  $R_{1\rho}$  is C4' of C10 (Table 4-2). In the case of the C10 ribose, however, C2' is not exhibiting an enhanced relaxation rate. Barring a coincidence in chemical shifts for C2' between the two states, it is unlikely that the data are reporting on the same process i.e. ribose conformational exchange. It is more likely that the C4' ribose is reporting on its changing chemical environment (which could yield different chemical shifts). It is possible that the G9 residue, which is unpaired and implicated in conformational transitions to achieve an active state, is affecting the chemical environment of the C4' site of C10 (Figure 4-11)(43,44). This would explain why the C4' site of C10 exhibits an enhanced relaxation rate while the C2' site of C10 that faces an A-form helix does not (Figure 4-11).



**Figure 4-11: Environment of the C10 ribose.** The NMR solution structure was used in this figure (5). Image rendered with PyMOL (14). The G9 residue is shown in red for emphasis.

### 4.3.2 Fast motions

In contrast to the relaxation rates of  $T_{1\rho}$  (for which significant elevation was obvious) the  $T_1$  relaxation rates are not obviously different when comparing the C6 ribose sites to helical ribose sites (Figure 4-6 and Figure 4-8, respectively). A significantly increase in  $R_1$  (shorter  $T_1$ ) would suggest that a particular site is behaving as if it was a smaller molecule and this would be an indication of sub-ns disorder. The  $T_1$  relaxation rates for the various sites (Figure 4-8) and the hNOE ratios (Figure 4-9) do not appear to be significantly varied. In general, these results resemble those for the GCAA RNA tetraloop analyzed by our lab that did not exhibit disorder on the ps-ns timescale (26).  $T_1$ ,  $T_2$  (determined by  $T_{1\rho}$ ), and hNOE ratios feed into model-free analysis. However, without model-free analysis complete, which yields an order parameter at each site, it is important not to draw strong conclusions about the extent of disorder on this timescale, at this particular time (45).

### 4.4 CONCLUSION

The C3'-endo conformation is observed in both the NMR solution structure and the X-ray crystallographic structure (5,6). However, we show that locking the C6 ribose in the C3'-endo conformation drastically reduces catalytic activity. This strongly suggests a need for conformational dynamics at the cleavage-site ribose (C6). The power-dependent  $R_{1\rho}$  experiments demonstrate that the C6 ribose undergoes a functionally relevant conformational transition from a ground state (C3'-endo) to a minor state (C2'-endo). The combination of dynamics and function in this work represents the importance of going beyond static pictures and investigating how dynamics can relate to function. Also, in comparing this work to GCAA RNA tetraloop, previously studied by our lab, it is possible that a general feature of noncanonical RNA structures

may be exhibiting conformational exchange on the  $\mu\text{s}$ -ms timescale and being well-ordered on the ps-ns timescale (26).

## **APPENDIX**

## APPENDIX

### NMR Relaxation Protocol

This picks up after you have already done the NMR relaxation experiments. It is designed to guide you in navigating the software you will use to determine relaxation rates from your data.

#### **Part1: Data Processing**

Varian2Pipe

NMRPipe

NMR Draw

#### **Part2: Data Analysis**

pipe2ucsf

Sparky

Igor

#### **Part3: Data Analysis**

NMRPipe

## Part 1: Data Processing

See: <http://spin.niddk.nih.gov/NMRPipe/doc1/>

### Conversion: Varian → NMRpipe

- 1) Use the command '**varian**' in a Unix shell
  - a) This will invoke a GUI for converting your Varian file to an NMRpipe file.
- 2) Click the green arrow (Spectrometer Input) to select the appropriate fid file.
- 3) Click 'Read Parameters'
  - a) Notice that parameters change. They should be indicative of the experiment you ran.
  - b) If your data is sensitivity-enhanced use the Rance-Kay option for the indirect dimension. If it is not, select Complex.
- 4) Click 'Save Script'
  - a) This will create a file "fid.com" and it will be dumped into the home folder.
    - i) Leave it there for now.
    - ii) Note: You can change the name/destination if you want; in general, I go with default names to make latter processing steps more universal.
  - b) Click 'Execute Script' → a file, "test.fid" will be generated in the folder where your fid is.
    - i) Now, you can move the fid.com into that folder as well. I recommend doing this so that file is save, otherwise the conversion will overwrite it.

### Processing: done with NMRPipe, visualized with NMRDraw.

1. Open the test.fid file with NMRDraw.
  - a. All dimensions should be in time at this point. It is necessary to process the data so it can be visualized in a meaningful way.
  - b. You can do the processing various ways at this point.
    - i. Use "macro edit" (under file menu) in NMRDraw to process via NMRPipe. This is what I started with, using pre-formulated processing schemes to begin processing i.e. Process 2D → Basic 2D. Save and move file (default, nmrproc.com) to folder with previous files. Then, run execute it.
    - ii. Use NMRpipe commands within NMRDraw (processing done one command at a time). I haven't had success with this.
    - iii. Use previous processing file to process data.
2. Processing NOTES
  - a. I highly recommend running the nmrproc.com in a terminal. The alternative is to double click it in the folder. By running it in the terminal you get a readout of the progress as the file is being processed. It's very easy to tell if something might not have work or the entire thing failed from this.
  - b. NMRDraw utilizes virtual memory without writing much to disk. So, be careful that files are actually saved. I think that's why the nmrproc.com file has been more successful then pipe commands in draw.

- c. Phasing must be done by hand. That means changing the phasing values in gedit and then executing the nmrproc.com file again to allow the change to take effect. Phasing is additive so if you change a value to +30 then want to change it -15, input 15 for the value in the file. Also very important, when you have Rance-Kay (SE) or Complex (ZZ) for the data in the indirect dimension (C-13) the phase is p0 = -90, p1 = +180 NOT the typical 90, -180 for complex and p0 = +90, p1 = +180 for Rance-Kay .
  - d. Look at a previous nmrproc.com file and go to <http://www.nmrscience.com/ref/index.html> to learn what the commands are doing.
  - e. We are now going to continue with using a previously obtained nmrproc.com file but know you have other options.
3. Close NMRDraw **'shft+q'**
  4. Copy and Paste nmrproc.com file to folder with previously generated files.
  5. Execute nmrproc.com file
  6. Open ft2.test with NMRDraw
  7. Inspected processed result and phase by hand. (For phasing see Processing Notes.)

## Part 2: Data Analysis

See:

<https://sites.google.com/site/jamiebairdtitus/analysisoft1andt2nmrrelaxationdatausingthesparkyprogram>

<https://www.cgl.ucsf.edu/home/sparky/manual/files.html#ConvertNMRPipe>

<https://www.cgl.ucsf.edu/home/sparky/manual/peaks.html>

<https://www.cgl.ucsf.edu/home/sparky/manual/extensions.html#RelaxFit>

### Conversion: NMRPipe → Sparky

1. Open the folder, in a terminal, with the processed spectrum. Modify the command '**pipe2ucsf noe150.pipe noe150.ucsf**' the first file is your processed spectrum, likely ft2.test, the second is your output file. It's advisable to put the time at the beginning of the output filename because Sparky will be able to automatically detect it later.
2. Create a folder in 'Projects' in Sparky.

### Sparky

1. It essential to assign peaks in one spectrum and copy and paste the assignments through to all spectra before you do the relaxation analysis.
  - a. How-to assign individual peaks is well described by the Sparky Manual.
  - b. Copy and pasting is described well by the other none Sparky link and the Relax Fit link.
2. Relaxation fitting is well described by the Relax Fit link. That uses the **rh** command.

### Part 3: Data Analysis

This is an alternative to using Sparky. It was used for the Leadzyme because it deals with overlapping peaks better than Sparky, in my experience.

There is good general information here and a tutorial:

<http://spin.niddk.nih.gov/NMRPipe/doc2new/#How to fit pseudo 3D spectra>

Specifically the way this works is by marring pseudo 3D spectra to a 2D peak table. The pseudo 3D spectra have the two dimensions of Carbon and Proton and the pseudo third dimension is the relaxation time.

- 1) Copy and paste the files that appear after the experimental folders from a relaxation series that was processed properly. Note: it's probably best to use the ones from Delaglio (Leadzyme) but you could use from the tutorial also. Copy and paste all so you have what you need and can re-write both manually and by executing programs.
- 2) Run **proc.com**. This will output a "ft" folder with processed spectra that can and should be viewed in nmrDraw.
  - i. This transforms the spectra from 2D to pseudo 3D, converts them from Varian to pipe, and processes them.
  - ii. Very important: proc.com may need to be edited prior to use. The file names and "tau" values (relaxation time) need to be correct. Also, refer to Rance-Kay vs. Complex in Part 1 if phasing of spectra is off.
- 3) Run **fit.com**. This will yield folders called "dif" and "sim" and files of axt.tab, nlin.spec.list, and sim.spec.list.
  - i. On the previous link there is a section called "Various Outputs." This will clarify what was generated.
  - ii. Very important: this will use the relax.master.tab file, this your peak table. You can create a new peak table and call it that or use a previously generated one. The same link has a lot of information on peak tables on that webpage.
- 4) Run **model.com**. This will generate plots and txt.
- 5) Use the command **showEvolve.tcl**. This opens a window to view evolutions and you can "fit" to get relaxation rates.

## REFERENCES

## REFERENCES

1. Pan, T. and Uhlenbeck, O.C. (1992) In vitro selection of RNAs that undergo autolytic cleavage with Pb<sup>2+</sup>. *Biochemistry*, **31**, 3887-3895.
2. Pan, T. and Uhlenbeck, O.C. (1992) A small metalloribozyme with a two-step mechanism. *Nature*, **358**, 560-563.
3. Barciszewska, M.Z., Szymanski, M., Wyszko, E., Pas, J., Rychlewski, L. and Barciszewski, J. (2005) Lead toxicity through the leadzyme. *Mutation research*, **589**, 103-110.
4. Hoogstraten, C.G. and Sumita, M. (2013) In Kretsinger, R. H., Uversky, V. N. and Permyakov, E. A. (eds.), *Encyclopedia of Metalloproteins*. Springer New York, New York, NY, pp. 1166-1173.
5. Hoogstraten, C.G., Legault, P. and Pardi, A. (1998) NMR solution structure of the lead-dependent ribozyme: evidence for dynamics in RNA catalysis. *Journal of molecular biology*, **284**, 337-350.
6. Wedekind, J.E. and McKay, D.B. (1999) Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nature structural biology*, **6**, 261-268.
7. Wedekind, J.E. and McKay, D.B. (2003) Crystal structure of the leadzyme at 1.8 Å resolution: metal ion binding and the implications for catalytic mechanism and allo site ion regulation. *Biochemistry*, **42**, 9554-9563.
8. Chartrand, P., Usman, N. and Cedergren, R. (1997) Effect of structural modifications on the activity of the leadzyme. *Biochemistry*, **36**, 3145-3150.
9. Pan, T., Dichtl, B. and Uhlenbeck, O.C. (1994) Properties of an in vitro selected Pb<sup>2+</sup> cleavage motif. *Biochemistry*, **33**, 9561-9565.
10. Wilson, T.J., Liu, Y., Domnick, C., Kath-Schorr, S. and Lilley, D.M. (2016) The Novel Chemical Mechanism of the Twister Ribozyme. *Journal of the American Chemical Society*, **138**, 6151-6162.
11. Hoogstraten, C.G., Wank, J.R. and Pardi, A. (2000) Active site dynamics in the lead-dependent ribozyme. *Biochemistry*, **39**, 9951-9958.
12. Legault, P., Hoogstraten, C.G., Metlitzky, E. and Pardi, A. (1998) Order, dynamics and metal-binding in the lead-dependent ribozyme. *Journal of molecular biology*, **284**, 325-335.

13. Westhof, E. and Hermann, T. (1999) Leadzyme RNA catalysis. *Nature structural biology*, **6**, 208-209.
14. The PyMOL Molecular Graphics System, Version 4.4 Schrödinger, LLC.
15. Saenger, W. (1984) *Principles of nucleic acid structure*. Springer-Verlag, New York.
16. Cai, Z. and Tinoco, I. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.
17. Butcher, S.E., Allain, F.H. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature structural biology*, **6**, 212-216.
18. Cunningham, B.C. and Wells, J.A. (1989) High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*, **244**, 1081-1085.
19. Julien, K.R., Sumita, M., Chen, P.-H.H., Laird-Offringa, I.A. and Hoogstraten, C.G. (2008) Conformationally restricted nucleotides as a probe of structure-function relationships in RNA. *RNA (New York, N.Y.)*, **14**, 1632-1643.
20. Bardaro, M.F. and Varani, G. (2012) Examining the relationship between RNA function and motion using nuclear magnetic resonance. *Wiley interdisciplinary reviews. RNA*, **3**, 122-132.
21. Zhang, Q. and Al-Hashimi, H.M. (2009) Domain-elongation NMR spectroscopy yields new insights into RNA dynamics and adaptive recognition. *RNA (New York, N.Y.)*, **15**, 1941-1948.
22. Baldwin, A.J. and Kay, L.E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nature chemical biology*, **5**, 808-814.
23. Al-Hashimi, H.M. and Walter, N.G. (2008) RNA dynamics: it is about time. *Curr Opin Struct Biol*, **18**, 321-329.
24. Al-Hashimi, H.M. (2013) NMR studies of nucleic acid dynamics. *J Magn Reson*, **237**, 191-204.
25. Palmer, A.G. (2004) NMR characterization of the dynamics of biomacromolecules. *Chemical reviews*, **104**, 3623-3640.
26. Johnson, J.E. and Hoogstraten, C.G. (2008) Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using <sup>13</sup>C NMR spin relaxation and specific isotope labeling. *Journal of the American Chemical Society*, **130**, 16757-16769.

27. Boisbouvier, J., Wu, Z., Ono, A., Kainosho, M. and Bax, A. (2003) Rotational diffusion tensor of nucleic acids from  $^{13}\text{C}$  NMR relaxation. *Journal of biomolecular NMR*, **27**, 133-142.
28. Yamazaki, T., Muhandiram, R. and Kay, L.E. (1994) NMR Experiments for the Measurement of Carbon Relaxation Properties in Highly Enriched, Uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -Labeled Proteins: Application to  $^{13}\text{C}$ .alpha. Carbons. *Journal of the American Chemical Society*, **116**, 8266-8278.
29. Davis, D.G. and Bax, A. (1985) Assignment of Complex H-1-Nmr Spectra Via Two-Dimensional Homonuclear Hartmann-Hahn Spectroscopy. *Journal of the American Chemical Society*, **107**, 2820-2821.
30. Hediger, S., Meier, B.H. and Ernst, R.R. (1995) Adiabatic Passage Hartmann-Hahn Cross-Polarization in Nmr under Magic-Angle Sample-Spinning. *Chem Phys Lett*, **240**, 449-456.
31. Johnson, J.E., Jr., Julien, K.R. and Hoogstraten, C.G. (2006) Alternate-site isotopic labeling of ribonucleotides for NMR studies of ribose conformational dynamics in RNA. *J Biomol NMR*, **35**, 261-274.
32. LeMaster, D.M. (1996) Structural determinants of the catalytic reactivity of the buried cysteine of Escherichia coli thioredoxin. *Biochemistry*, **35**, 14876-14881.
33. Legault, P., Hoogstraten, C.G., Metlitzky, E. and Pardi, A. (1998) Order, dynamics and metal-binding in the lead-dependent ribozyme. *J Mol Biol*, **284**, 325-335.
34. Massi, F., Grey, M.J. and Palmer, A.G. (2005) Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR  $R1\rho$  relaxation experiments. *Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR  $R1\rho$  relaxation experiments*.
35. Loria, J.P., Rance, M. and Palmer, A.G. (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *Journal of the American Chemical ...*
36. Palmer, A.G. (1997) Probing molecular motion by NMR. *Current opinion in structural biology*, **7**, 732-737.
37. Palmer, A.G., Kroenke, C. and Loria, P.J. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods*, **339**, 204-238.
38. Palmer, A.G. and Massi, F. (2006) Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chemical reviews*, **106**, 1700-1719.

39. Kay, L.E. (1998) Protein dynamics from NMR. *Nature structural biology*, **5 Suppl**, 513-517.
40. Korzhnev, D.M. and Kay, L.E. (2008) Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. *Accounts of chemical research*.
41. Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A.D. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. ... *of biomolecular NMR*.
42. Rinnenthal, J., Buck, J., Ferner, J., Wacker, A., Fürtig, B. and Schwalbe, H. (2011) Mapping the landscape of RNA dynamics with NMR spectroscopy. *Accounts of chemical research*, **44**, 1292-1301.
43. Hoogstraten, C.G., Wank, J.R. and Pardi, A. (2000) Active site dynamics in the lead-dependent ribozyme. *Biochemistry*, **39**, 9951-9958.
44. Lemieux, S., Chartrand, P., Cedergren, R. and Major, F. (1998) Modeling active RNA structures using the intersection of conformational space: application to the lead-activated ribozyme. *RNA*, **4**, 739-749.
45. Lipari, G. and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society*.

## **CHAPTER 5**

### **RIBOSE DYNAMICS OF LOOP A OF HAIRPIN RIBOZYME**

## ABSTRACT

Loop A and loop B of the hairpin undergo significant conformational rearrangements in order to come together in a minor-groove, minor-groove fashion, termed docking. We are studying hairpin ribozyme in the junctionless construct, meaning loop A and loop B reside on separate molecules. In this chapter we are studying loop A in isolation. We are using a site-specific  $^{13}\text{C}$  labelling scheme, previously developed by our lab, to study the ribose dynamics of loop A using NMR spin-relaxation techniques.

We are investigating two loop A samples,  $^{13}\text{C}$  2', 4' adenosine loop A and  $^{13}\text{C}$  2', 4' guanosine loop A. These were selected due to the cleavage site being located between what has been denoted A-1 and G+1 and the catalytic residue G8 being located within loop A. There are also other adenosines in the loop region. It is possible that the ribose dynamics of these residues could be involved in the formation of the docked structure.

We observed varied  $T_{1\rho}$  relaxation rates for both the adenosine and guanosine samples. In this initial phase is difficult to know how significant that variation is. It will be necessary to do perform  $R_{1\rho}$  relaxation dispersion studies before making rigorous conclusion about the ribose conformational exchange in loop A. There was one adenosine ribose (C2' peak) that has a significantly elevated  $T_{1\rho}$  relaxation rate. It is possible that these ribose conformational dynamics are important in the formation of a docking-competent conformation. We need to perform NMR experiments to complete the assignment of our loop A construct to be able to definitively determine which peaks are representative of specific ribose sugars.

## 5.1 INTRODUCTION

The hairpin ribozyme must undergo conformational rearrangement for loop A and loop B to come together, in a minor-groove, minor-groove fashion termed docking (1-3). The docked complex of loop A and loop B has an extensive network of interactions. Previously, the rearrangements and extensive interface have been profiled previously by our lab (shown in Figure 2-2) (4). Also, key conformational changes of loop A are illustrated in Figure 1-2 (5).

We are studying the junctionless construct of the hairpin ribozyme; meaning loop A and loop B reside on separate molecules. In this chapter, we are studying loop A in the absence of loop B. This allows for investigation of docked-like conformations in the absence of its binding partner (loop B). Detection of these conformations would be evidence of conformational capture (Figure 1-3).

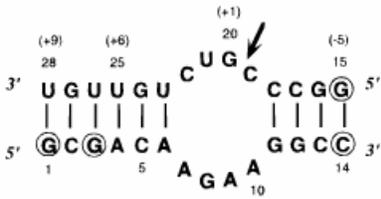
The selection of our loop A construct was a challenge. The NMR solution structure of loop A, previously solved by the Tinoco group, has significant differences from the naturally occurring form and what has been used in extensive crystallographic studies (2,3,6-8). A potentially significant difference is a cytidine base at the A-1 position. As shown in Chapter 3 of this work, a methoxy group compared to a hydroxyl group, at the cleavage site, can have a very significant impact. Thus, changing a base at the cleavage site could have significant effects as well. The Tinoco lab selected their sequence to mimic a cleavage site that would potentially be useful in HIV treatment (2).

However, we are seeking to have the canonical G+1, A-1 cleavage site with the catalytic residue of G8 being present as well. The development of such a stable NMR construct took a significant amount of effort and was done by Dr. Patrick Ochieng of our lab (see Chapter 3 of his

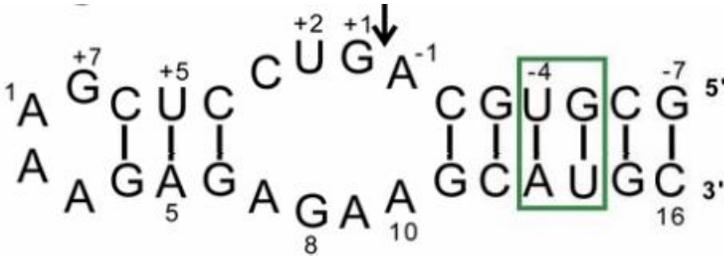
thesis). In Figure 5-1, the sequences and secondary structure for our construct of loop A is presented and compared to the loop A construct utilized by the Tinoco lab.

The site-specific labelling scheme and NMR spin relaxation techniques used here are the same as described in Chapter 4. The goals of using them are to investigate ribose conformational exchange on the  $\mu$ s-ms timescale and characterize sub-ns disorder. This will allow us to hypothesize if the conformational exchange, of specific ribose sugars, is important in obtaining a docking-competent conformation. In this manner, we can go beyond snapshots and bridge the gap between the structures of loop A and loop B in the absence of each other and the structure of the docked complex. Recently, our lab has published a computational study of loop A (in isolation) that suggests loop A does indeed sample docked-like conformations (5). Here we are looking to explore that strong possibility from an experimental approach.

A



B



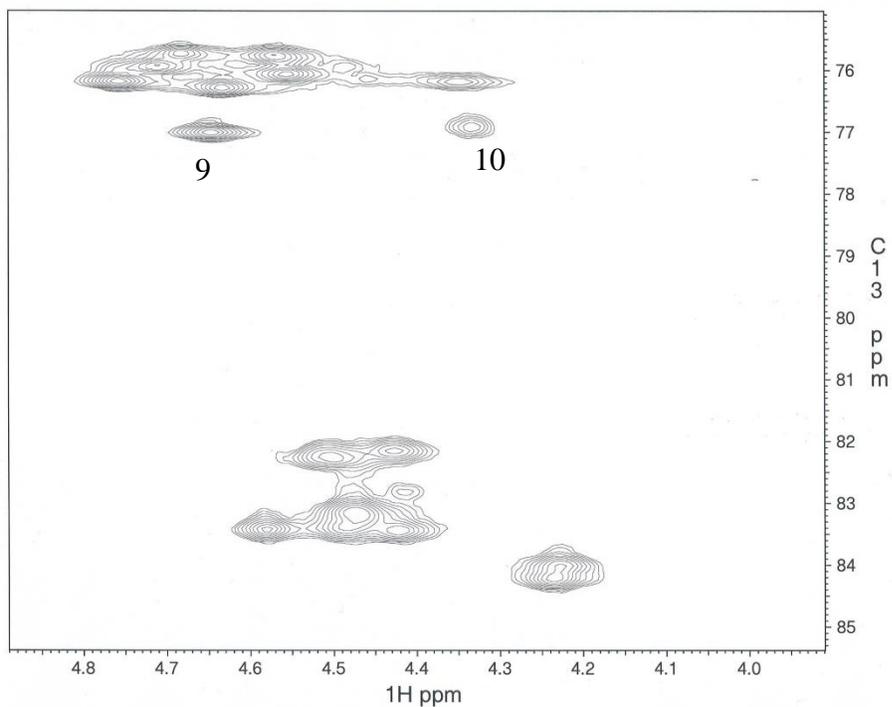
**Figure 5-1: Construct of loop A.** Panel A is the loop A construct that the Tinoco group solved the structure of and that panel is from ref. (2). Panel B is our loop A constructed, termed GAAA extended loop A and that panel is from Figure 3-2 of Dr. Ochieng's thesis.

## 5.2 MATERIALS AND METHODS

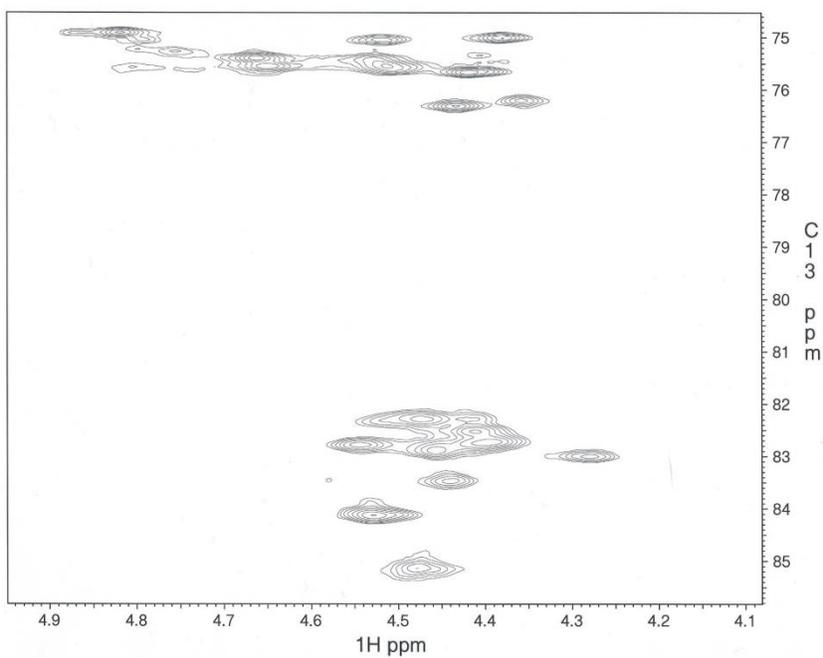
### 5.2.1 Sample preparation and verification

The two samples of loop A,  $^{13}\text{C}$  2', 4' adenosine loop A and  $^{13}\text{C}$  2', 4' guanosine loop A, were prepared in the same manner as the  $^{13}\text{C}$  2', 4' cytosine leadzyme that was previously described in Chapter 4 (9,10).

The  $^{13}\text{C}$  2', 4' adenosine loop A, after repeated lyophilization in 99.8% deuterium oxide ( $\text{D}_2\text{O}$ ) was brought up in 300  $\mu\text{L}$  of 99.96%  $\text{D}_2\text{O}$ . The concentration of the  $^{13}\text{C}$  2', 4' adenosine loop A sample was 0.750 mM and it was in a buffer of 10 mM sodium phosphate, and 20  $\mu\text{M}$  EDTA at pH 5.5. A HSCQ (heteronuclear single quantum coherence) spectrum was used to verify a  $^{13}\text{C}$  2', 4' sample was indeed produced (Figure 4-2). The general parameters for the spectrum, and all following spectra, were similar to those of Chapter 4, with the proton radio frequency (RF) carrier centered on residual HDO and the  $^{13}\text{C}$  was set at 80 ppm, with spectral widths of 4 ppm and 14 ppm, respectively. The specific parameters were 1024 x 64 complex points with 32 transients. The  $^{13}\text{C}$  2', 4' guanosine loop A sample was brought up in 250  $\mu\text{L}$  of 99.96%  $\text{D}_2\text{O}$ , after repeated lyophilization in 99.8%  $\text{D}_2\text{O}$ . The concentration of this sample was 0.501 mM and in a buffer of 10 mM sodium phosphate, and 20  $\mu\text{M}$  EDTA at pH 5.5. To verify this sample a HSQC spectrum was used with the same parameters that were used for the other loop A sample (Figure 4-3).



**Figure 5-2:**  $^{13}\text{C}$  2', 4' adenosine loop A. HSQC spectrum with C2' region (top) and C4' region (bottom) shown. Peak 10 is the peak that relaxes at a faster rate than the rest so it is highlighted. The relaxation rate of Peak 9 may also be somewhat elevated.

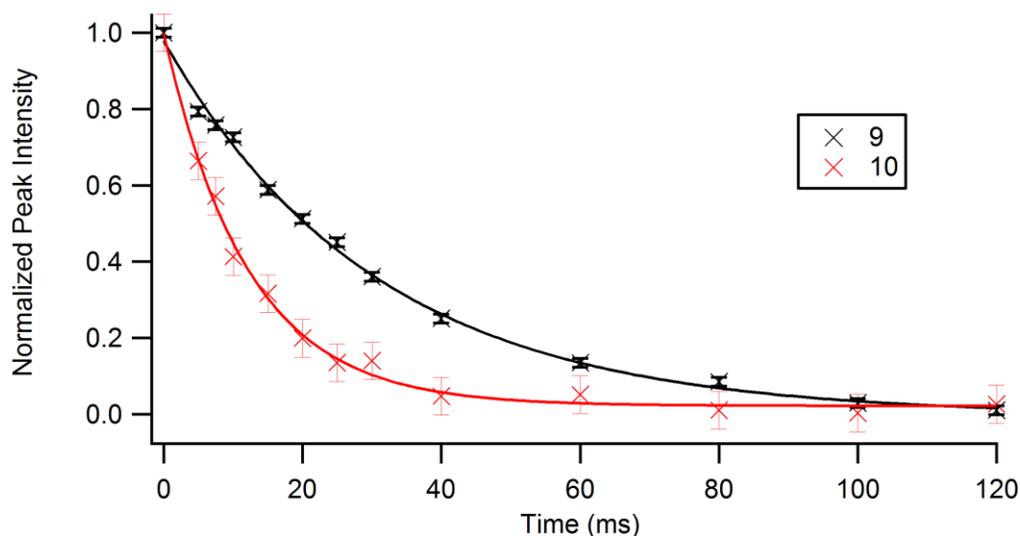


**Figure 5-3:**  $^{13}\text{C}$  2', 4' guanosine loop A. A HSQC spectrum showing the C2' region (top) and C4' region (bottom).

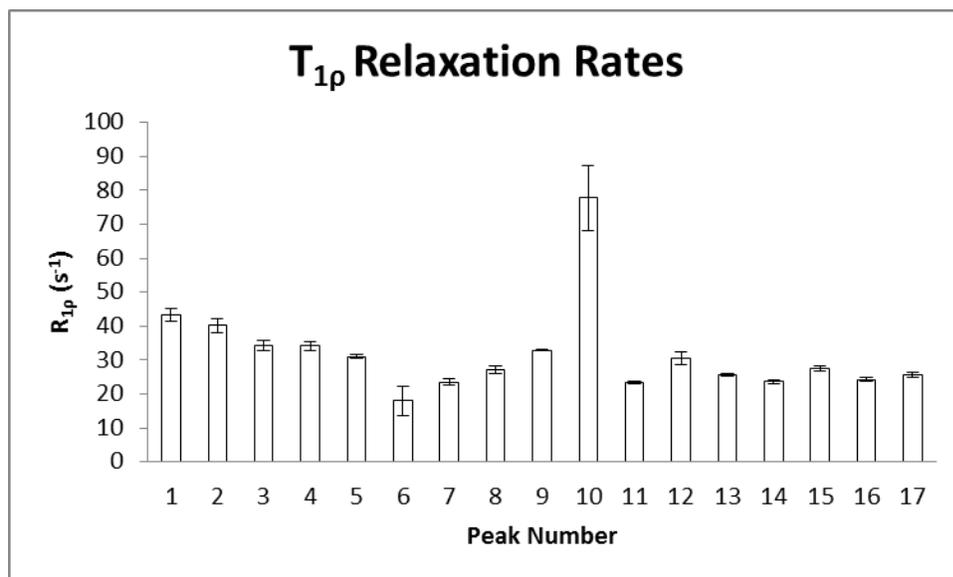
### 5.2.2 $T_{1\rho}$ relaxation series

Preliminary  $T_{1\rho}$  spectra indicated that for both,  $^{13}\text{C}$  2', 4' adenosine loop A and  $^{13}\text{C}$  2', 4' guanosine loop A, some sites (i.e. NMR peaks, demonstrating nuclear spin-relaxation) relaxed faster than others. This is also what was observed with the  $^{13}\text{C}$  2', 4' cytosine leadzyme, in Chapter 4. Due to peaks relaxing at different rates the relaxation time for the  $T_{1\rho}$  relaxation series ranged from 0-120 ms with sufficient sampling between 0-60 ms for peaks that relax faster than others. The general parameters for both ( $^{13}\text{C}$  2', 4' adenosine loop A and  $^{13}\text{C}$  2', 4' guanosine loop A) series were the same as the HSQC spectra. The specific parameters for both loop A samples were 1024 x 96 complex points with 64 transients. As in Chapter 4, the relaxation analysis was done using NMRPipe (11). Again, we are grateful to Dr. Frank Delaglio for scripts that were provided for analysis with the leadzyme because there were useful in analyzing the loop A samples present here.

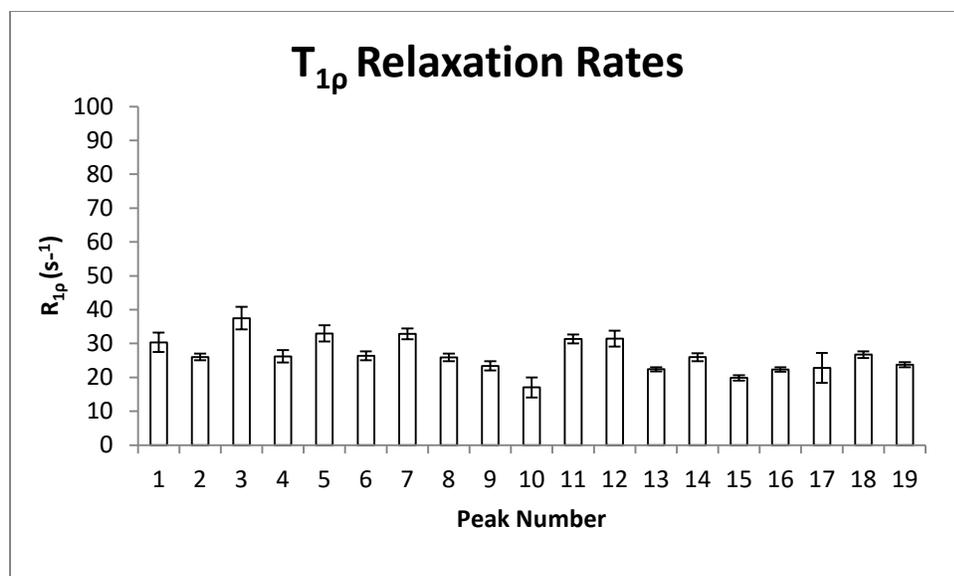
For the  $^{13}\text{C}$  2', 4' adenosine loop A there was one peak that relaxed at a rate significantly faster than all other peaks. It is compared with another peak in Figure 5-4. The  $^{13}\text{C}$  2', 4' adenosine loop A and  $^{13}\text{C}$  2', 4' guanosine loop A  $T_{1\rho}$  relaxation rates for all sites that were measured are reported in Figure 5-5 and Figure 5-6 respectively. There appears to be some variation in relaxation rates at this stage.



**Figure 5-4: Enhanced  $T_{1\rho}$  relaxation rate for peak 10 compared to a more typical site. Both peaks are highlighted in Figure 5-2.**



**Figure 5-5:  $T_{1\rho}$  relaxation rates for  $^{13}C$  2', 4' adenosine loop A. Peaks 1-8 are from the C4' region and peaks 9-17 are in the C2' region. Peak 10 has a significant faster relaxation rate than the other peaks and there appears to be variation among the others.**



**Figure 5-6:**  $T_{1\rho}$  relaxation rates for  $^{13}\text{C}$  2', 4' guanosine loop A. Peaks 1-10 are in the C4' region and peaks 11-19 are located in the C2' region. The scale is the same as Figure 5-5 for comparison. There appears to be some variation in relaxation rates.

### 5.2.3 T<sub>1</sub> relaxation series

The T<sub>1</sub> relaxation series was performed as described in Chapter 4. The relaxation time was varied from 0-1200 ms. The spectra were acquired with 1024 x 96 complex points with 40 transients. At present the T<sub>1</sub> series for <sup>13</sup>C 2', 4' adenosine loop A has been completed. The T<sub>1</sub> series for <sup>13</sup>C 2', 4' guanosine loop A was performed with the same parameters.

The peak numbering is the same for T<sub>1</sub> and T<sub>1ρ</sub> relaxation rates. The two peaks shown in Figure 5-7 are the same peaks that were shown in Figure 5-4. Contrary to the T<sub>1ρ</sub> series, in the T<sub>1</sub> series these two peaks (9 and 10) relax at the same rate within error. There appears to be less variation in general in the T<sub>1</sub> series. All T<sub>1</sub> rates for the peaks studied, associated with <sup>13</sup>C 2', 4' adenosine loop A, are reported in Figure 5-8. All T<sub>1</sub> rates for the peaks studied, associated with <sup>13</sup>C 2', 4' guanosine loop A, are reported in Figure 5-9.

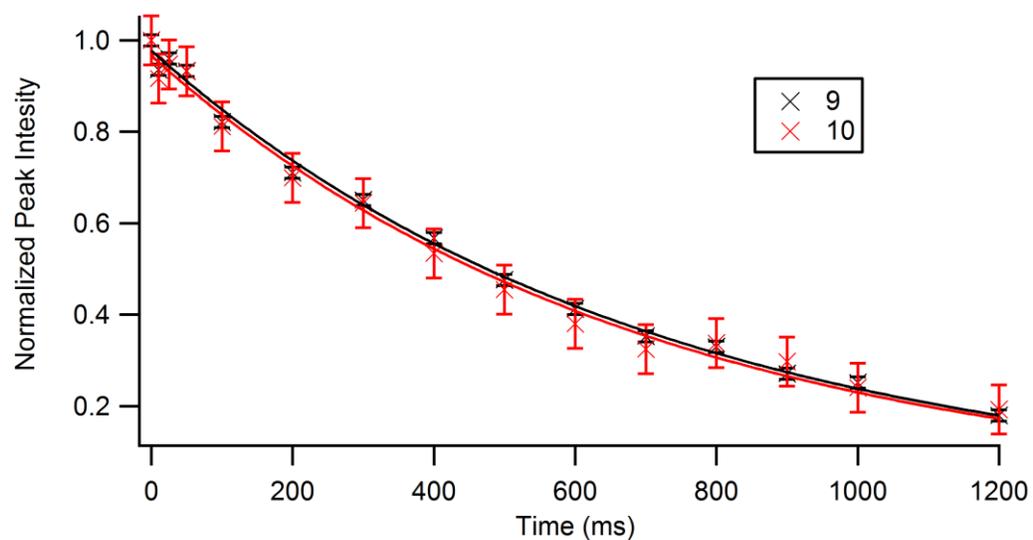


Figure 5-7: T<sub>1</sub> relaxation rates for Peak 9 and 10 <sup>13</sup>C 2', 4' adenosine loop A.

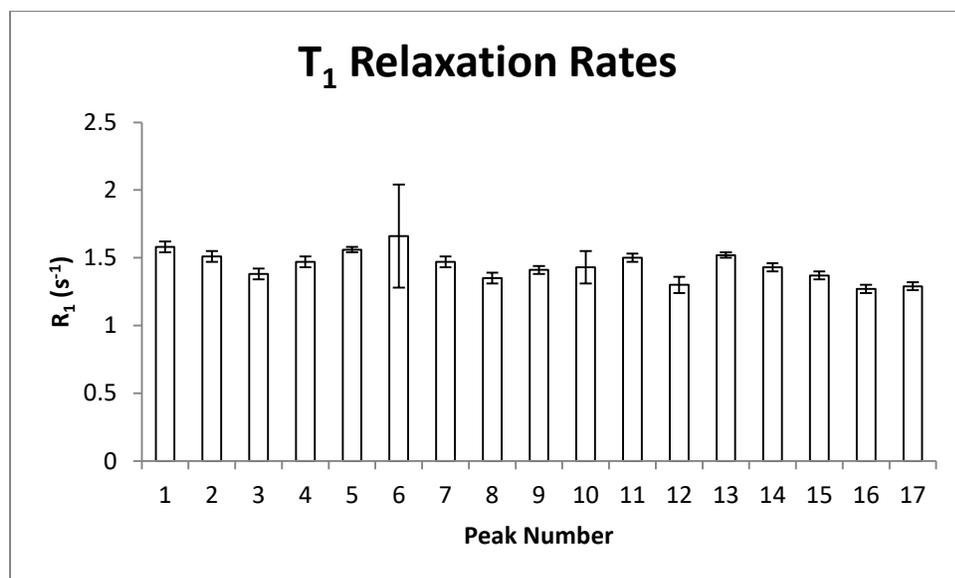
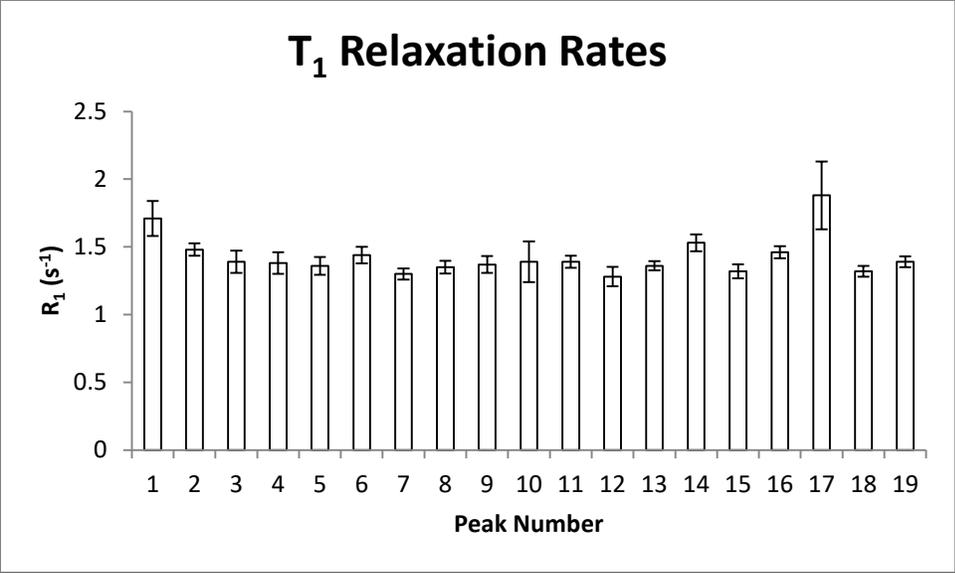


Figure 5-8: T<sub>1</sub> relaxation rates for <sup>13</sup>C 2', 4' adenosine loop A.



**Figure 5-9:** T<sub>1</sub> relaxation rates for <sup>13</sup>C 2', 4' guanosine loop A.

### 5.3 DISCUSSION

The enhanced  $T_{1\rho}$  relaxation rate for Peak 10 in the C2' region of the  $^{13}\text{C}$  2', 4' adenosine loop A seems to be similar to A7 of the GCAA tetraloop, previously studied by our lab (10). However, without assignment it would be premature to draw that conclusion. In the loop region of loop A there are several adenosines. There is A7, A9, and A10 adjacent to the G8 catalytic residue. There is also A-1, which contains the nucleophile. It is possible that variation in relaxation rates is reporting on dynamics within the loop region. It will be necessary to perform  $T_{1\rho}$  relaxation dispersion to examine conformational exchange more rigorously (as done in Chapter 4).

In the structure of loop A, solved by the Tinoco group, the ribose conformation for A7 is interconverting between C2'-endo and C3'-endo and then resides in C3'-endo in the docked structure, from the Wedekind group (2,6). Due to crystallization only one ribose conformation can be observed. A9 and A10 were found to be in the C3'-endo conformation in both isolated loop A and the docked complex.

A-1 (mutated to C-1 in the NMR solution structure) is found in the C3'-endo in isolation, but is in the C2'-endo conformation in the docked complex. However, simulations have suggested the 2'-*O*-methyl group at the cleavage site in the crystal structure disrupts the ribose conformation, as when it was changed to a 2'-OH the C3'-endo conformation was observed (12). While crystallographic work supports the use of the 2'-*O*-methyl group, because of the in-line geometry that is observed (3). In the extensive crystallographic studies by the Wedekind lab transition from the 2'-*O*-methyl group to the 2'-OH does have the same effect as the simulations done by the Walter lab, but when the functional group at the cleavage was changed something else (like a base mutation) was also done so strong conclusions are difficult to draw. It could

potentially be very informative to perform future docking studies with LNA (locked nucleic acid, C3'-endo) at the cleavage site.

The C2' and C4' sites in the  $^{13}\text{C}$  2', 4' guanosine loop A exhibit some variation. The guanosines in loop A are the catalytic residue G8 and the cleavage site residue, with the scissile phosphate, G+1. In comparing the NMR solution structure and the crystal structure G8 transitions from C2'-endo in isolation to C3'-endo in the docked structure (2,6). The G+1 residue is in the C2'-endo conformation in both isolation and the docked complex.

Similar to the Leadzyme (Chapter 4) and GCAA tetraloop, previous studied by our lab, there is a little variation in the  $T_1$  relaxation rates, suggesting that these structured RNA molecules are well-ordered on the ps-ns timescale (10). However, further NMR experiments (hNOE) will need to be performed and Model-Free analysis utilized to draw comprehensive conclusions about the order or disorder on the ps-ns timescale (13).

## 5.4 CONCLUSION

We found a significantly enhanced  $T_{1\rho}$  relaxation rate for a C2' peak with the  $^{13}\text{C}$  2', 4' adenosine loop A. Also we found some variation in  $R_{1\rho}$ , for both the  $^{13}\text{C}$  2', 4' adenosine loop A and  $^{13}\text{C}$  2', 4' guanosine loop A. It is possible that, like the leadzyme described in Chapter 4, there could be conformation exchange associated the nucleophilic ribose at the cleavage site and be that could be important to catalysis (function). Other dynamics could also be important to formation of the docked complex. It will be important to assign the loop A spectra so that identification of the individual peaks can be made.

## **REFERENCES**

## REFERENCES

1. Butcher, S.E., Allain, F.H. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature structural biology*, **6**, 212-216.
2. Cai, Z. and Tinoco, I. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.
3. MacElrevey, C., Salter, J.D., Krucinska, J. and Wedekind, J.E. (2008) Structural effects of nucleobase variations at key active site residue Ade38 in the hairpin ribozyme. *RNA (New York, N.Y.)*, **14**, 1600-1616.
4. Sumita, M., White, N.A., Julien, K.R. and Hoogstraten, C.G. (2013) Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis. *RNA Biol*, **10**, 425-435.
5. Ochieng, P.O., White, N.A., Feig, M. and Hoogstraten, C.G. (2016) Intrinsic Base-Pair Rearrangement in the Hairpin Ribozyme Directs RNA Conformational Sampling and Tertiary Interface Formation. *The journal of physical chemistry. B*.
6. Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. and Wedekind, J.E. (2006) Water in the active site of an all-RNA hairpin ribozyme and effects of Gaa8 base variants on the geometry of phosphoryl transfer. *Biochemistry*, **45**, 686-700.
7. Rupert, P.B. and Ferré-D'Amaré, A.R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, **410**, 780-786.
8. Rupert, P.B., Massey, A.P., Sigurdsson, S.T. and Ferré-D'Amaré, A.R. (2002) Transition state stabilization by a catalytic RNA. *Science (New York, N.Y.)*, **298**, 1421-1424.
9. Johnson, J.E., Jr., Julien, K.R. and Hoogstraten, C.G. (2006) Alternate-site isotopic labeling of ribonucleotides for NMR studies of ribose conformational dynamics in RNA. *J Biomol NMR*, **35**, 261-274.
10. Johnson, J.E., Jr. and Hoogstraten, C.G. (2008) Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using <sup>13</sup>C NMR spin relaxation and specific isotope labeling. *J Am Chem Soc*, **130**, 16757-16769.
11. Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A.D. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. ... *of biomolecular NMR*.
12. Mlýnský, V., Banás, P., Hollas, D., Réblová, K., Walter, N.G., Sponer, J. and Otyepka, M. (2010) Extensive molecular dynamics simulations showing that canonical G8 and

protonated A38H<sup>+</sup> forms are most consistent with crystal structures of hairpin ribozyme. *The journal of physical chemistry. B*, **114**, 6642-6652.

13. Lipari, G. and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society*.

## **CHAPTER 6**

### **RNA BINDING SPECIFICITY OF PPR27**

## ABSTRACT

Pentatricopeptide repeat (PPR) proteins are characterized by having helical repeats that are 35 amino acids long. Here a PPR protein was studied from the parasitic protozoan, *Trypanosoma brucei*, which causes human African trypanosomiasis or sleeping sickness. We selected the smallest known PPR protein from *T. brucei*. Its molecular weight is approximately 27 kDa and will be referred to as PPR27. This was selected to aid future potential structural and dynamics studies by NMR.

*T. brucei* has more PPR proteins than humans and PPR proteins appear to be essential to the viability of the protozoan. This makes PPR proteins a strong candidate to target with therapeutic agents. The PPR proteins in this organism localize to its single mitochondria and their roles are unknown. There is one report that PPR27 associates with the mitochondrial small ribosomal unit via protein-protein interactions. Outside of that its RNA specificity is unknown.

Here, using 11 rounds of *in vitro* selection resulted in enhanced enrichment and the most likely reason for that was the presence of guanosine-rich sequences. Other Hoogstraten lab members have confirmed that PPR 27 preferentially binds G-tract RNA. More specifically, it binds quadruplex polyguanosine and does not disrupt the quadruplex structure.

By searching the mitochondrial genome of *T. brucei*, we found an enrichment of G-tracts in mRNA of maxicircle genes that are significantly edited. This leads to the possibility that PPR27 may function by interacting with pre-mRNA to aid in the processes of processing or translation.

## 6.1 INTRODUCTION

A PPR motif is comprised of antiparallel alpha-helices that are 35 amino acids long (1,2). PPR proteins have repeats of their secondary structure motif. In plants the number of PPR proteins is in the hundreds and they are thought to bind RNA, with their mechanism of binding and role of recruitment unknown (2). They are found plant organelles, chloroplasts and mitochondria (3).

*Trypanosoma brucei* is the parasitic protozoan that causes African sleeping sickness. According to the World Health Organization (last updated Feb. 2016) treatment for African still involves drugs that are derivatives of arsenic. For an example of how poor the treatment is, the drug Melarsoprol, will kill the human being treated 3-10% of the time. This obviously necessitates the need for develop of specific drugs that can target the parasite and have a minimal effect on the human host.

What has made *T. brucei* so challenging to target is that it lives in the tsetse fly and humans. To counter different living condition the parasite's biology contains many eccentricities, such as extensive mitochondrial RNA editing (4-7). As previously stated, plants contain hundreds of PPR proteins; however, non-plant eukaryotes have less than 10, with the exception of *T. brucei* (8). So far, at least 36 PPR proteins have been found in *T. brucei* (9,10). Previously knockdown studies of PPR proteins in *T. brucei* have shown that the PPR proteins studied are important the oxidation pathway, hence located in the mitochondria, and their knockdown lead to adverse growth effects and death of the parasite (10,11).

PPR27 was selected to study from the various PPR proteins of *T. brucei*. This was done because it is the smallest and at the time of selection thought to be a good candidate for studies of structure and dynamics by NMR spectroscopy. This turned out not to be the case as the

solubility of PPR27 is very poor. In this application it was expressed as a fusion protein with a Maltose Binding Protein (MBP) solubility-enhancing tag. This allowed for future studies such as the *in vitro* selection shown here. For a far more comprehensive discussion on this subject please refer to the thesis of Dr. P. F. Kamba of our lab.

In order to determine the RNA sequence specificity of PPR27 we used *in vitro* selection, or what is also called systematic evolution of ligands by exponential enrichment (SELEX) (12,13). The technique works by beginning with a random pool of various RNA sequences. Then round by round the selection process results in the RNA pool transitioning from random to specific and from this a consensus sequence for a RNA-binding protein can be determined.

## **6.2 MATERIALS AND METHODS**

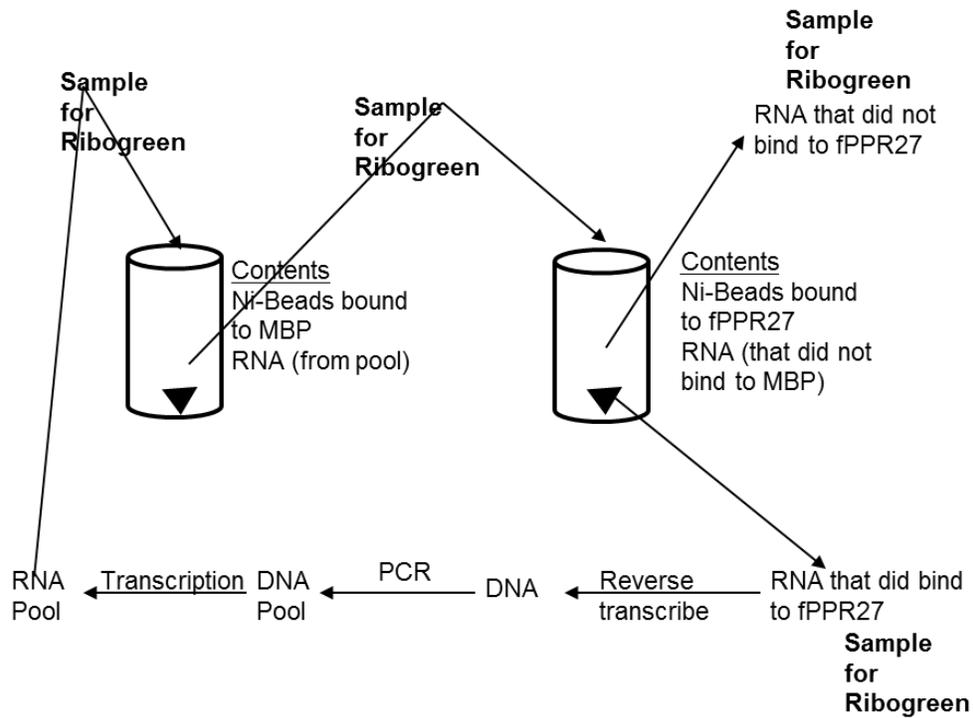
This is the first *in vitro* selection performed by our lab and the protocol for it developed by myself with advice from Dr. Sumita of our lab can be found in the appendix for this chapter. More specifics, i.e. step by step procedures and planning can be found there. In general, we consulted the Berglund lab with specific questions and used one of their publications to aid in the development of our own protocol (14).

The experiment begins with  $2.4 \times 10^{18}$  possible sequences utilizing a DNA template with a 20 nucleotide random region. There is literature to consult to make certain the sequences are represented from a probabilistic standpoint (15). The RNA was then transcribed from this DNA template with a random region using standard methods (16). The DNA was then removed with a DNase assay previously used (14).

For a comprehensive discussion of the expression of this protein review the thesis of Dr. Kamba of our lab. In brief, the protein was expressed as fusion protein with maltose binding protein and a His-tag for solubility and purification purposes. For *in vitro* selection, an amylose

column was used first and then the beads used in binding were Ni-NTA beads. This choice was made because the capacity of amylose beads was lower than Ni-NTA at the time of experimentation.

PPR27 was immobilized to Ni-NTA beads for the interaction with the RNA pool. The interaction occurred at 4° C for 20 minutes in binding buffer (50 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.5, 100 mM KCl, 1.5 mM MgCl<sub>2</sub>, 2.0 mM imidazole). The RNA was in excess of the protein, this is what allows for selection. The unbound RNA is then washed away and the RNA that bound to PPR27 is retained extracted by phenol: chloroform extraction with ethanol precipitation, reverse-transcribed, and amplified by PCR. This is one round and this process was carried out 11 times. There was negative selection step done every other round with His-tagged MBP to remove nonspecific binding sequences. For one round see Figure 6-1. For the concentrations utilized in each round see Table 6-1.



**Figure 6-1: One round of *in vitro* selection.** This details the procedures executed in one round of *in vitro* selection. Not included, is that phenol: chloroform extractions with ethanol precipitations are required after binding, reverse transcriptions, PCR, and transcription.

Round	RNA Conc. ( $\mu\text{M}$ )	MBP Conc. ( $\mu\text{M}$ )	PPR Conc. ( $\mu\text{M}$ )	Fold Excess of RNA to PPR	Fold Excess of RNA to MBP
1	18	-	8.02	2	-
2	30	3.00	2.05	15	10
3	15	-	3.14	5	-
4	30	3.04	2.19	14	10
5	30	-	3.21	9	-
6	30	3.10	0.64	47	10
7	30	-	2.18	14	-
8	30	2.11	2.09	14	14
9	30	-	1.11	27	-
10	5	0.50	0.83	6	10

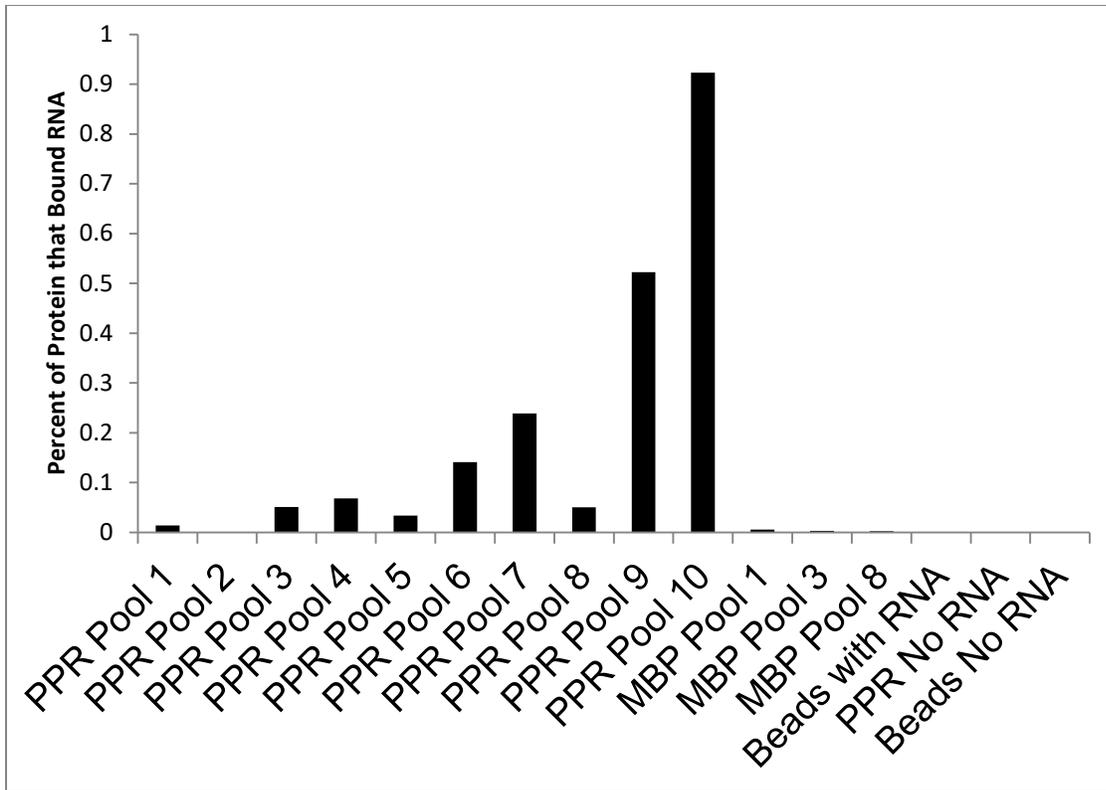
**Table 6-1: SELEX binding scheme.** This shows the effective concentrations used during the round of *in vitro* selection.

Showing enhanced binding can be done several ways but what is ultimately the hallmark that the *in vitro* selection was effective is the determination of a consensus sequence . To show enhanced binding as a result of selection achieved we used a Ribogreen assay (Invitrogen, New York) read with a FLUstar Omega microplate reader (BMG LABTECH, North Carolina). This idea came from seeing other researchers taking a similar approach with DNA (17). This assay was used to determine relative enrichment. In our hands, we found the Ribogreen reagent to not be very stable over a long period of time (months). So, we found it more effective to use the kit that has standards and comes as a 96-well plate.

The cDNA library from round 11 was sequenced using TOPO® TA cloning. This is necessary because the first nucleotides read in sequencing are unreliable so this process places sufficient nucleotides between the region to be read and the first nucleotides sequenced. Sequenced were aligned using MEME and WebLogo to find and generate a representation of conserved elements (18,19).

### **6.3 RESULTS**

The results of the Ribogreen assay show a general trend of increased enrichment of the RNA pool being bound by PPR27 (Figure 6-2). These results show conclusively that his-tagged MBP is not binding in a specific manner relative to PPR27. In order to try to determine the conserved elements MEME and WebLogo were utilized (Figure 6-3).



**Figure 6-2: Ribogreen assay to access relative enrichment.** This assay shows enrichment increasing over the course of many rounds. Also, ample negative controls were ran to avoid spurious results since this was the first time we ran this assay.



**Figure 6-3: WebLogo of RNA sequences in the pool after 11 cycles of selection with PPR27.**

## 6.4 DISCUSSION

The somewhat possible consensus sequence obtained may contained up to 6 Guanosines. However, others in our lab have shown that polyG binds tighter than GGUGGU. There was enrichment over the course of many rounds of SELEX likely due to sufficient G bases being present. Also, using FRET other Hoogstraten lab members have shown PPR27 prefers binds G in its quadruplex form and PPR27 does not disrupt the G-quadruplex. The G-quadruplex is a tertiary interaction were four guanosine bases are in the same plan. It is likely sufficient G bases were present in the pool for G-quadruplex formation and that was the reason for enrichment rather than a typical consensus sequence. Further analysis of mitochondrial genome of *T. brucei* genome by Dr. Kamba of our lab revealed maxicircle genes that are extensively edited have an enrichment of G-tracts.

## 6.5 CONCLUSION

For PPR27 of *T. brucei*, a protein for which neither its function nor RNA binding specificity was known, we now have proposals for both RNA binding specificity and role. An 11 round SELEX showed enrichment that was likely due to G-quadruplex formation or sufficient guanosines being present, rather than a canonical consensus sequence. The role of PPR 27 is likely to bind RNA and mediate RNA editing for maxicircle genes of the mitochondria of *T. brucei*. Also, despite being the smallest of *T. brucei* PPR proteins due to the extensive solubility issues of PPR27 it is not an ideal candidate for structure or dynamics studies using NMR spectroscopy. A SELEX investigation of two other PPR proteins has been undertaken by a mentored undergraduate in the lab. Perhaps they will have more of a conical consensus sequence.

## **APPENDIX**

## APPENDIX

### SELEX for MBP-PPR

Fall 2010

**Prepared by:** Neil Andrew White

**Advised by:** Dr. Mina Sumita

**Principal Investigator:** Dr. Charles Hoogstraten

302C Biochemistry Michigan State University, East Lansing MI, 48824, USA

#### Introduction

The method of SELEX (**S**ystematic **E**volution of **L**igands by **EX**ponential enrichment) is being used with the goal of determining the sequence of RNA that fPPR27 (now called MBP-PPR) binds to. The MBP-PPR protein is obtained by overexpressing recombinant PPR27 as a fusion to a solubility-enhancing tag containing maltose binding protein (MBP), hexahistidine (His6), and a spacer of about 40 amino acid residues. The PPR27 protein one of more than twenty-eight proteins in *Trypanosoma brucei* with the 35-amino-acid pentatricopeptide repeat (PPR) motif. A negative selection step has been inserted for maltose binding protein and the His Tag.

#### Contents

Introduction  
Overall Scheme for SELEX  
Daily Plans  
DNase Assay  
RNA Concentration  
PBinding  
Reverse Transcription  
PCR  
Clean-up of PCR  
DNA Concentration and Annealing  
Transcription  
Miscellaneous

Round	$\mu\text{M}$ RNA	$\mu\text{M}$ MBP-His	$\mu\text{M}$ MBP-His-PPR	$\mu\text{L}$ Amylose Beads	Total Beads
1	18	-	6	400	400
2	20	2	2	150	300
3	10	-	1	75	75
4	5	0.5	0.5	37.5	75
5	2.5	-	0.25	18.75	18.75
6	1.25	0.125	0.125	9.375	18.75
7	0.625	-	0.0625	4.6875	4.6875
8	0.3125	0.03125	0.03125	2.34375	2.34375
9	0.15625	0.015625	0.015625	1.171875	<u>2.34375</u>
					896.875

**Supplemental Table 6-1: Round by round overview.** What was used in each round with and the total amount of magnetic beads.

## **Daily Plans**

This is the plan for doing one round in three days. To prepare for round one it is necessary to start with PCR as there is nothing to reverse transcribe.

### **Day One-Generating DNA Pool**

- 1) Reverse Transcription
- 2) Polymerase Chain Reaction (PCR)

### **Day Two-Generating RNA Pool**

- 1) Clean-up of PCR
- 2) Determine Concentration of DNA Pool and Annealing
- 3) Transcription

### **Day Three-Binding**

- 1) DNase Assay
- 2) Determination Concentration of RNA Pool
- 3) Binding

## Day One

### Reverse Transcription

**Purpose** To use M-MLV RT to generate DNA for PCR from the RNA that bound to fPPR27.

Perform the following steps separately to each tube.

- 1) Add 11  $\mu\text{L}$  MilliQ water to tube with dry RNA. Vortex to dissolve pellet. (Pellet may or may not be able to be seen.)
- 2) Add 1  $\mu\text{L}$  2 $\mu\text{M}$  Primer II.
- 3) Add 1  $\mu\text{L}$  10 mM dNTP Mix.
- 4) Heat at 65° C for 5 minutes and quick chill on ice.
- 5) Collect contents by brief centrifugation.
- 6) Add 4  $\mu\text{L}$  5X First-Strand Buffer.
- 7) Add 2  $\mu\text{L}$  0.1 M DTT.
- 8) Mix contents gently.
- 9) Incubate at 37° for 2 minutes.
- 10) Add 1  $\mu\text{L}$  of M-MLV-RT and mix by pipetting up and down.
- 11) Incubate at 37° C for 50 minutes.
- 12) Inactivate the reaction by heating at 70° C for 15 minutes.

### PCR

**Purpose** To amplify the amount of DNA so a RNA pool can be generated by transcription.

Perform the following steps to three tubes.

- 1) Add 2.5  $\mu\text{L}$  of water to PCR tube.
- 2) Add 2  $\mu\text{L}$  of previous reaction to PCR tube.
- 3) Add 4  $\mu\text{L}$  of Primer I and Primer II (both 10  $\mu\text{M}$ ).
- 4) Add 12.5  $\mu\text{L}$  of the PCR Master Mix.
- 5) Run the program under the guest account for PPR in the Ferguson-Miller Lab. (No “Hot Start” and 25  $\mu\text{L}$  for volume.)
- 6) Phenol:Chloroform extraction and ethanol precipitation.

#### Saved Settings for PCR

Cycle 1: 1x Initial Denaturation at 95.0° C for 2:00

Cycle 2: 30x Amplification

Denature at 95.0° for 0:30

Anneal at 55.0° C for 0:30

Step 3 Extend at 72.0° for 0:30

Cycle 3: Final Extension at 72.0° C for 5:00

Cycle 4: Hold and Cool at 4.0° C for  $\infty$ .

## Day Two

### Clean-Up of PCR

**Purpose** To purify the nucleic acids by removing NTP's and remaining salt.

Do the following to each sample (A and B).

Use a 3,000 MW Filter from Millipore

- 1) Dissolve the pellet in 100  $\mu\text{L}$  of ddH<sub>2</sub>O and transfer to 3 kDa MW centrifugal filtration unit.
- 2) Use another 100  $\mu\text{L}$  of ddH<sub>2</sub>O to attempt to remove any DNA and transfer to filter.
- 3) Centrifuge for 30 minutes at 7,000 RPM.
- 4) Add 100  $\mu\text{L}$  of ddH<sub>2</sub>O, to filter, and centrifuge for 30 minutes at 7,000 RPM.
- 5) Add 100  $\mu\text{L}$  of ddH<sub>2</sub>O, to filter, and centrifuge for 30 minutes at 7,000 RPM.
- 6) Collect in fresh tube.

### Determine Concentration of the DNA Pool and Annealing

**Purpose** To determine the concentration of the DNA pool so the appropriate amount of DNA is used in transcription and anneal them.

Do the following separately for each sample (A and B)

- 1) Measure and record the Absorbance at 260 nm and determine the concentration with the molar extinction coefficient  $\epsilon = 684,750/\text{M}\cdot\text{cm}$ .
- 2) Add 20 pmol of sample to a new micro-centrifuge tube.
- 3) Add 8.39  $\mu\text{L}$  of T7 promoter 2.38  $\mu\text{M}$ .
- 4) Add MilliQ water to bring volume to 33.25  $\mu\text{L}$ .
- 5) To anneal heat at 95° C for 2 minutes.

### Transcription

**Purpose** To generate a RNA pool from the DNA pool to use for binding.

Do the following separately for each sample (A and B)

- 1) Add the following into a micro-centrifuge tube.
  - a. 10  $\mu\text{L}$  400 mM Tris (pH 8.0)
  - b. 4  $\mu\text{L}$  25 mM spermidine
  - c. 10  $\mu\text{L}$  0.1% Triton
  - d. 8  $\mu\text{L}$  ATP, CTP, GTP, UTP 50 mM each
  - e. 3.75  $\mu\text{L}$  0.5 M MgCl<sub>2</sub>
  - f. 2  $\mu\text{L}$  10.65 mg/mL T7 RNA pol
  - g. 5  $\mu\text{L}$  100 mM DTT
- 2) Incubate for 4 hours at 37° C in a water bath.
- 3) Phenol:Chloroform extraction and ethanol precipitation.

## Day Three

### DNase Assay

**Purpose** To remove the DNA from the transcribed RNA.

- 1) Add 16  $\mu\text{L}$  MilliQ water to pellet of RNA.
- 2) Vortex to dissolve as much of the pellet as possible.
- 3) Add 2  $\mu\text{L}$  RQ1 RNase-Free DNase 10X Reaction Buffer.
- 4) Add 2  $\mu\text{L}$  RQ1 RNase-Free DNase.
- 5) Incubate at 37° C for 30 minutes.
- 6) Phenol:Chloroform extraction and ethanol precipitation (Add 90  $\mu\text{L}$  MilliQ water for extraction. The stop solution is not needed since the extraction is used.)

### Clean-Up of Transcript

**Purpose** To purify the nucleic acids by removing NTP's and remaining salt.

Do the following to each sample (A and B).

Use a 3,000 MW Filter from Millipore

- 1) Dissolve the pellet in 100  $\mu\text{L}$  of ddH<sub>2</sub>O and transfer to filter.
- 2) Use another 100  $\mu\text{L}$  of ddH<sub>2</sub>O to attempt to remove any RNA and transfer to filter.
- 3) Centrifuge for 30 minutes at 7,000 RPM.
- 4) Add 100  $\mu\text{L}$  of ddH<sub>2</sub>O, to filter, and centrifuge for 30 minutes at 7,000 RPM.
- 5) Add 100  $\mu\text{L}$  of ddH<sub>2</sub>O, to filter, and centrifuge for 30 minutes at 7,000 RPM.
- 6) Collect content (Invert and centrifuge for 5 minutes at 7,000 RPM).

### Determining Concentration of RNA Pool

**Purpose** Obtain the desired concentration of RNA for the reaction vessel.

- 1) Take and record the absorbance at 260 nm.
- 2) The molar extinction coefficient is 689,450/M\*cm.

## Binding

**Purpose** To obtain only the RNA that binds to fPPR27 and round by round to make the binding more selective.

- 1) Remove MBP-His and MBP-PPR-His from -20° C storage and put on ice for five minutes, then thaw in ice water, and return to ice.
- 2) Dispense \_\_\_Ni-Beads into two different micro-centrifuge tubes labeled as above.
- 3) Wash beads 3 times with Protein SELEX Binding Buffer using volume equal to what beads were suspended in each time.
- 4) Add appropriate amount of protein to each tube.  
MBP-His \_\_\_\_\_ MBP-PPR-His \_\_\_\_\_
- 5) Incubate for 2 hours using end-over-end rotator in the cold room.
- 6) Remove “Supernatant” and Wash all tubes 3 times with 100 µL of Ni-NTA SELEX Buffer (2 mM Imidozal). Take Absorbance of these later.
- 7) Save remaining RNA Pool. \_\_\_\_\_
- 8) To MBP tube add appropriate amount of RNA Pool. \_\_\_\_\_
- 9) Incubate tube for 20 minutes in cold room with end-over-end rotator.
- 10) Take 1.5 µL from MBP-His tube for Ribogreen assay and store in appropriately labeled tube. \_\_\_\_\_
- 11) Transfer the rest of the “supernatant” (NO BEADS) to MBP-PPR-His Tube.
- 12) Incubate tube for 20 minutes in cold room with end-over-end rotator
- 13) Remove the supernatant and record the volume. \_\_\_\_\_ V=\_\_\_\_\_
- 14) Wash with 100 µL of Buffer and save. \_\_\_\_\_
- 15) Phenol:Chloroform extraction and ethanol precipitation.
- 16) Dissolve pellet in 50 µL take 1.5 µL for Ribogreen assay and then dry. \_\_\_\_\_

## Miscellaneous

A note on handling the Ni-Beads.

Be sure to vortex them before withdrawing as the buffer they are in is not visible. Also after spinning the beads down a magnet (stir bar) can be held underneath as you remove the supernatant. The beads generally cluster quite well but the magnet is helpful.

Proteins used.

Thx.

F28-29          64.7  $\mu\text{M}$           Use 50  $\mu\text{L}$  of beads every round and 15  $\mu\text{L}$  of Protein

Stored in 1x Barkum

$\epsilon = 13,940/\text{M}\cdot\text{cm}$

17,044.361 g/mol

fPPR27 (In NAW Selex Box in  $-20^\circ\text{C}$  storage)

F22-37          12.5 $\mu\text{M}$           Purified by FPLC

$\alpha\text{NW}$           65.8  $\mu\text{M}$           Spun Down after Cobalt Column (but no FPLC)

Both stored in 1x Barkum, 10% Glycerol

$\epsilon = 36,130/\text{M}\cdot\text{cm}$

41,610.505 g/mol

A note on concentrations.

They are based on a 55  $\mu\text{L}$  reaction vessel. Originally it was going 100  $\mu\text{L}$ , 50  $\mu\text{L}$  beads, 20  $\mu\text{L}$  RNA, and 30  $\mu\text{L}$  Buffer. But the beads spin down and I removed about 45  $\mu\text{L}$  of Buffer before adding the rest.

I suppose you could use less than 10  $\mu\text{L}$  of beads but the handbook only mentions that small of an amount. Also I think after spinning down it would be very difficult to work with.

## REFERENCES

## REFERENCES

1. Small, I.D. and Peeters, N. (2000) The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci*, **25**, 46-47.
2. Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J. and Mustelin, T. (2004) Protein tyrosine phosphatases in the human genome. *Cell*, **117**, 699-711.
3. Yu, Q.-B.B., Jiang, Y., Chong, K. and Yang, Z.-N.N. (2009) AtECB2, a pentatricopeptide repeat protein, is required for chloroplast transcript accD RNA editing and early chloroplast biogenesis in *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology*, **59**, 1011-1023.
4. Leung, S.S. and Koslowsky, D.J. (2001) Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA (New York, N.Y.)*, **7**, 1803-1816.
5. Leung, S.S. and Koslowsky, D.J. (2001) RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic acids research*, **29**, 703-709.
6. Reifur, L., Yu, L.E., Cruz-Reyes, J., Vanhartervelt, M. and Koslowsky, D.J. (2009) The impact of mRNA structure on guide RNA targeting in kinetoplastid RNA editing. *PloS one*, **5**.
7. Aphasizhev, R. and Aphasizheva, I. (2011) Mitochondrial RNA processing in trypanosomes. *Research in microbiology*, **162**, 655-663.
8. Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B. *et al.* (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089-2103.
9. Aphasizheva, I., Maslov, D., Wang, X., Huang, L. and Aphasizhev, R. (2011) Pentatricopeptide repeat proteins stimulate mRNA adenylation/uridylation to activate mitochondrial translation in trypanosomes. *Molecular cell*, **42**, 106-117.
10. Pusnik, M., Small, I., Read, L.K., Fabbro, T. and Schneider, A. (2007) Pentatricopeptide repeat proteins in *Trypanosoma brucei* function in mitochondrial ribosomes. *Molecular and cellular biology*, **27**, 6876-6888.
11. Mingler, M.K., Hingst, A.M., Clement, S.L., Yu, L.E., Reifur, L. and Koslowsky, D.J. (2006) Identification of pentatricopeptide repeat proteins in *Trypanosoma brucei*. *Mol Biochem Parasitol*, **150**, 37-45.

12. Gold, L., Polisky, B., Uhlenbeck, O. and Yarus, M. (1995) Diversity of oligonucleotide functions. *Annual review of biochemistry*, **64**, 763-797.
13. Ellington, A.D. and Szostak, J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818-822.
14. Garrey, S.M., Voelker, R. and Berglund, J.A. (2006) An extended RNA binding site for the yeast branch point-binding protein and the role of its zinc knuckle domains in RNA binding. *The Journal of biological chemistry*, **281**, 27443-27453.
15. Ciesiolka, J., Illangasekare, M., Majerfeld, I., Nickles, T., Welch, M., Yarus, M. and Zinnen, S. (1995) Affinity selection-amplification from randomized ribooligonucleotide pools. *Methods in enzymology*, **267**, 315-335.
16. Milligan, J.F., Groebe, D.R., Witherell, G.W. and Uhlenbeck, O.C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic acids research*, **15**, 8783-8798.
17. Wochner, A. and Glökler, J. (2007) Nonradioactive fluorescence microtiter plate assay monitoring aptamer selections. *BioTechniques*, **42**.
18. Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. and Noble, W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*, **37**, W202-208.
19. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res*, **14**, 1188-1190.

## **CHAPTER 7**

### **CONCLUSIONS AND FUTURE DIRECTIONS**

## 7.1 CONCLUSIONS AND FUTURE DIRECTIONS

### 7.1.1 Exploring the kinetics of the hairpin ribozyme

Kinetic parameters for trans-docking of the hairpin ribozyme were presented in Chapters 2 and 3 (1,2). The association rate was very slow,  $(1.97 \pm 0.29) \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ . This is likely representative of the significant conformational rearrangements that are necessary for docking (3-6). The dissociation rate,  $(7.0 \pm 1.0) \times 10^{-4}$ , is indicative a stable complex being formed, which can be attributed to the intricate interface in this loop-loop interaction.

There are several environmental factors that could be varied to provide insight into docking. The metal ion concentration could be decreased, which will result in lower binding affinity (1). Whether the association or dissociation rate is more affected or they are affected evenly could provide insight into the structural role of metal ions. Metal ions could potentially be needed to form the docked complex (stabilizing docking-competent conformations) or they could play a greater role in stabilization of the docked complex.

An osmolyte or molecular crowder could also have a significant effect on docking. Molecular crowders have resulted in a human hepatitis delta virus-like ribozyme favoring a more compact form (7). Osmolytes have been shown to stabilize tertiary structure and destabilize tertiary structure (8,9). Whether an osmolyte increases or decreases binding affinity, that could provide insight into the nature of the docking interaction.

### 7.1.2 The effect of cleavage site modifications on the thermodynamic signature

A 2'-*O*-methyl modification at the cleavage site had some effect on the binding affinity of docking, compared to the native 2'-OH ( $338 \pm 66 \text{ nM}$  and  $114 \pm 33 \text{ nM}$  respectively). The cleavage-site modification significantly impacted the thermodynamic signature compared to the native function group. The thermodynamic signature with 2'-*O*-methyl was  $\Delta H^\circ$  ( $-3.40 \pm 0.51$ )

kcal mol<sup>-1</sup> and  $\Delta S^\circ$  ( $18.1 \pm 1.7$ ) kcal mol<sup>-1</sup> K<sup>-1</sup>. In contrast the thermodynamic signature with the native, 2'-OH was  $\Delta H^\circ$  ( $-11.2 \pm 2.4$ ) kcal mol<sup>-1</sup> and  $\Delta S^\circ$  ( $-5.7 \pm 4.8$ ) kcal mol<sup>-1</sup> K<sup>-1</sup>. Non-mutually exclusive explanations for this phenomenon include the modification disrupting the network of interactions, steric hindrance, or disruption of the optimal ribose conformation. The effects of other cleavage site modifications on the thermodynamic signature are also of future interest. For example the effect a 2'-H (deoxy) at the cleavage site would be interesting. It could reduce potential steric hindrance, relative to the 2'-*O*-methyl, but would not have the capacity to hydrogen-bond, like the 2'-OH.

### **7.1.3 Ribose dynamics at the cleavage site of the lead-dependent ribozyme**

We proposed that the C6 ribose is undergoing conformational exchange from its inactive, ground C3'-endo conformation to a functional, excited C2'-endo conformation. The conclusion that it is exhibiting conformational exchange lifetimes, on the low  $\mu$ s (5-10  $\mu$ s) end of the  $\mu$ s-ms regime, is based on elevated  $T_{1\rho}$  relaxation rates over the range of the relaxation dispersion experiments. The conclusion that the C3'-endo conformation is not optimal was derived from a bicyclo-nucleotide, inserted at the C6 position, which was locked in that conformation and nearly abolished catalysis. The remaining inquiry for the cleavage-site ribose is whether, with a higher applied  $B_1$  power, the rate of exchange can be quantified. Also, model-free analysis will yield an order parameter on sub-ns disorder and that will be performed in the near future.

### **7.1.4 Ribose dynamics of the hairpin ribozyme**

A C2' adenosine ribose peak in loop A of the hairpin ribozyme has a significantly elevated  $T_{1\rho}$  relaxation rate. Beyond that, in general, there was some variation in the  $T_{1\rho}$  relaxation rates for both the adenosine and guanosine labeled loop A samples. There is not enough data to conclude about conformational exchange on the  $\mu$ s-ms regime. It will be

necessary to perform relaxation dispersion experiments, as done with the leadzyme, to make rigorous conclusions about ribose conformational exchange. It will also be important to assign the peaks in the NMR spectra to the specific sites in loop A, so we can determine for which ribose sites conformational exchange is or is not occurring. Efforts to accomplish this are ongoing. Also, as with the leadzyme, preliminary inquiry does not seem to indicate fast motions on the ps-ns timescale. However, further experimentation and model free analysis will also be needed to make a strong conclusion.

As shown by our group previously, and in Chapter 4 of this work, the specific ribose conformation for various residues of the leadzyme can have a significant impact on catalysis. This could be the case for the hairpin ribozyme, as well. Surface plasmon resonance docking studies with locked nucleic acid (LNA), which is covalently locked in the C3'-endo conformation, could provide insight into the ribose conformational preference of docking. These docking studies would provide an excellent complement to the ongoing dynamics studies.

## **REFERENCES**

## REFERENCES

1. Sumita, M., White, N.A., Julien, K.R. and Hoogstraten, C.G. (2013) Intermolecular domain docking in the hairpin ribozyme: metal dependence, binding kinetics and catalysis. *RNA Biol*, **10**, 425-435.
2. Hoogstraten, C.G., Sumita, M. and White, N.A. (2014) Unraveling the thermodynamics and kinetics of RNA assembly: surface plasmon resonance, isothermal titration calorimetry, and circular dichroism. *Methods Enzymol*, **549**, 407-432.
3. Cai, Z. and Tinoco, I. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite. *Biochemistry*, **35**, 6026-6036.
4. Butcher, S.E., Allain, F.H. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. *Nature structural biology*, **6**, 212-216.
5. Rupert, P.B. and Ferré-D'Amaré, A.R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. *Nature*, **410**, 780-786.
6. Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. and Wedekind, J.E. (2006) Water in the active site of an all-RNA hairpin ribozyme and effects of Gua8 base variants on the geometry of phosphoryl transfer. *Biochemistry*, **45**, 686-700.
7. Strulson, C.A., Yennawar, N.H., Rambo, R.P. and Bevilacqua, P.C. (2013) Molecular crowding favors reactivity of a human ribozyme under physiological ionic conditions. *Biochemistry*, **52**, 8187-8197.
8. Strulson, C.A., Boyer, J.A., Whitman, E.E. and Bevilacqua, P.C. (2014) Molecular crowders and cosolutes promote folding cooperativity of RNA under physiological ionic conditions. *RNA*, **20**, 331-347.
9. Lambert, D. and Draper, D.E. (2007) Effects of osmolytes on RNA secondary and tertiary structure stabilities and RNA-Mg<sup>2+</sup> interactions. *J Mol Biol*, **370**, 993-1005.