

**MOLECULAR DYNAMICS IN THE HAIRPIN RIBOZYME:
CALCULATIONAL AND EXPERIMENTAL ASPECTS**

By

Patrick Omondi Ochieng

A DISSERTATION

Submitted to
Michigan State University
in partial fulfilment of the requirements
for the degree of

Biochemistry and Molecular Biology - Doctor of Philosophy

2015

ABSTRACT

MOLECULAR DYNAMICS IN THE HAIRPIN RIBOZYME: CALCULATIONAL AND EXPERIMENTAL ASPECTS

By

Patrick Omondi Ochieng

The increasing role of RNA therapy in targeting diseases has inspired several RNA studies and especially structural RNA. Of interest to many scientists is how such RNA can perform their work with limited functional groups available to RNA. The structural versatility of RNA seems to underscore the importance of dynamics in performing several functions. Ribozymes are a good example of structured RNA involved in RNA backbone cleavage with a range of strategies. Hairpin ribozyme invokes domain-domain docking to activate the cleavage process. The major loop rearrangements observed upon docking, as well as the kinetically unfavorable docking process, both argue for conformational selection by pre-organization of the catalytically-competent active site of the hairpin ribozyme. In this thesis, we sought to study the behavior of loop A in sampling the docked-like conformation as evidence for conformational selection. We addressed three major aims which involved (i) understanding the dynamics in loop A using molecular dynamics simulation as a tool for assessing conformational sampling (ii) determining the right loop A construct for NMR studies and resonance assignments for structure determination and (iii) elucidation of fast and slow dynamics in loop A using NMR relaxation techniques.

In aim 1 (Chapter 2), molecular dynamics simulation was used to determine conformational heterogeneity in RNA based on alternate base-pair formation within a subset of residues in the loop region of domain A of the hairpin ribozyme. Three main conformers and several minor conformations were observed in our simulations as analyzed by the Markov State model

analysis. RNA base residues and backbone dynamics played a major role in the conformational heterogeneity and ensemble in loop A. Of the conformations that were sampled, the most populated conformer, AA/CA, closely sampled conformational properties similar to the activated (docked) loop A conformation, suggesting the activating role induced by conformational heterogeneity.

In aim 2 (Chapter 3), we determined the suitable loop A construct for NMR studies by NMR secondary structure analysis on various constructs of loop A. Using exchangeable and non-exchangeable NMR experiments we assigned certain specific proton, nitrogen and carbon resonances of loop A for structural determination and relaxation measurements.

In aim 3 (Chapter 4), we assessed loop A dynamics using the ^{13}C -NMR relaxation measurements. Fast internal motions in the order of ps - ns timescale were analyzed by Model-free approach using data from ^{13}C R_1 , $R_{1\rho}$, and heteronuclear NOE of loop A. Loop A was generally found to be a rigid molecule on this timescale with internal generalized order parameters, S^2 , of at least 0.9 in the helical regions. Several residues reported correlation times indicative of fast motions on the ps to ns timescale, while a few reported slow exchange in the μs -ms timescale.

Our data underscores the importance of fast and slow dynamics in the formation of conformational states with structures similar to the activated form. These conformational variability and structural transitions seem to activate RNA thereby facilitating RNA-RNA interaction via the conformational selection mechanism.

Copyright by
PATRICK OMONDI OCHIENG
2015

I dedicate this thesis to my late parents Philemon and Phoebe Ochieng who despite their sacrifice for me to attend school, never lived to see me begin my graduate school. My late grandpa (Ayub Oricho), who also died a few months before my thesis defense, inspired me to the end.

ACKNOWLEDGEMENTS

After several years of graduate school at Michigan State University, I wish to thank the Department of Biochemistry and Molecular Biology for according me the opportunity to pursue my dream. My application to Michigan State University was inspired by my former mentor and master's thesis advisor at Western Michigan University, Dr. David Huffman, who not only provided me with an opportunity to gamble my research skills in his lab, but also mentored me methodically to be a future scientist and scholar. While on my visit tour at Michigan State University, I met Dr. Charles Hoogstraten who had been assigned to meet all the visiting prospective graduate students for breakfast. Dr. Hoogstraten was excited about his research as he shared with us, and I remember him particularly emphasizing the use of Nuclear Magnetic Resonance (NMR) as a tool for understanding molecular dynamics, not just structure. This was a new concept to me, since I had all along understood NMR as a tool for structure determination both in solution and solid state. When I finally joined Michigan State University, I was definitely interested in pursuing my graduate studies in the NMR field. Dr. Hoogstraten accepted me in his lab and has been a pillar, a mentor, an advisor who I will forever cherish. His insights and brilliance in biophysics, and NMR in particular, has been a useful resource in my pursuit for this project. In the course of my research, I also developed an interest in pursuing computational biochemistry alongside my experimental work. Dr. Michael Feig readily offered to mentor me and quickly provided me a space and computer in his lab. As an amateur in Linux commands and computational biophysics, Dr. Feig encouraged me to take the necessary courses that prepared me for research in this complex field. He has since mentored me and offered me high quality training by always asking the hard questions that spur my thinking. I count myself fortunate to have gone through Dr. Feig's mentorship.

I want to express my gratitude to all the former and current members of the Hoogstraten and Feig labs. Dr. Minako Sumita and Dr. James Johnson Jr. inducted me in the Hoogstraten lab, offering to train me in synthesizing and studying RNA. Dr. Shayantani Mukherjee and Dr. Sean Law trained me in Molecular Dynamics Simulation for which I am grateful. Dr. Monika Sharma, Dr. Beibei Wang, Dr. Parimal Kar and Dr. Vahid Mirjalili provided me with useful suggestions and technical support in the Molecular dynamics project. Dr. Fadhiru Kamba was a colleague in Dr. Hoogstraten's lab whose friendship and insights I will live to cherish. Together, we encouraged one another during good and difficult times. I also wish to thank Neil White for his support and friendship.

I want to thank my graduate guidance committee for volunteering to mentor me as well as their suggestions that improved the quality of my research and presentations. Special thanks to Dr. Bill Henry, Dr. Leslie Kuhn and Dr. John McCracken for serving in this voluntary role. I wish to thank my friends Dr. Joel Lwande and Dr. David Achila for inspiring me through graduate school right from Western Michigan University to Michigan State University.

I also thank Dr. Dan Holmes and Kermit Johnson for always providing me with NMR time, for offering solutions to my NMR challenges and for providing me with technical support.

Last but not least, I want to thank my lovely wife Lilly Ochieng, and son Phillip Ochieng for providing me with the desired support during my time in graduate school and for always being there for me.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
KEY TO ABBREVIATIONS.....	xiv
CHAPTER 1	1
INTRODUCTION TO RIBOZYMES AND THE APPROACHES TO ASSESSING RNA DYNAMICS-FUNCTION RELATIONSHIP	1
1.1 BACKGROUND.....	2
1.1.1 Thesis outline and research goals	2
1.1.2. RNA conformational transitions.....	4
1.1.3. RNA tertiary structure	8
1.1.4. RNA and ribozymes	10
1.1.5. Mechanism of catalysis in ribozymes.....	11
1.1.6. Therapeutic role of ribozymes.....	13
1.1.7. Hairpin ribozyme	16
1.1.8. Structure and active site architecture of the hairpin ribozyme	17
1.1.9. Computer molecular dynamics (MD) simulations	19
1.2.0. Experimental assessment of dynamics	23
1.2.1. Functional probing of RNA dynamics.....	27
REFERENCES	29
CHAPTER 2	40
PROBING ALTERNATE BASE-PAIR REARRANGEMENTS AND CONFORMATIONAL SAMPLING IN A RIBOZYME ACTIVE-SITE INTERNAL LOOP USING MOLECULAR DYNAMICS SIMULATIONS	40
ABSTRACT	41
2.1 INTRODUCTION.....	42
2.2 MATERIALS AND METHODS	47
2.2.1 Molecular dynamics (MD) simulations using NAMD and the CHARMM force field	47
2.2.2 MD simulations using sander and the Amber force field	50
2.2.3 Targeted molecular dynamics (TMD) simulation	50
2.2.4 Analysis	52
2.3 RESULTS.....	52
2.3.1 Force field validation.....	52
2.3.2 Conformational heterogeneity of wildtype loop A.....	53
2.3.3 Conformational transitions between major states.....	65
2.4 DISCUSSION	70
2.4.1 Sampling of docked loop A structure	71
2.4.2 Functional relevance.....	76
2.5 CONCLUSION	77
REFERENCES	78

CHAPTER 3	86
CONSTRUCT OPTIMIZATION, DOCKING STUDIES AND NMR RESONANCE ASSIGNMENT OF THE LOOP A DOMAIN OF THE HAIRPIN RIBOZYME	86
3.1 INTRODUCTION.....	86
3.2 MATERIALS AND METHODS	92
3.2.1 RNA preparation.....	92
3.2.2 Docking transitions between mutant loop A and wildtype loop B.....	93
3.2.3 NMR studies	94
3.3 RESULTS.....	96
3.3.1 Docking studies	96
3.3.2 Construct optimization for NMR spectroscopy.....	99
3.3.3 NMR assignments	105
3.3.4 Assignment of exchangeable proton resonances	108
3.3.5 Non exchangeable protons.....	111
3.4 DISCUSSION	124
ACKNOWLEDGEMENT	127
REFERENCES	128
 CHAPTER 4	 132
INSIGHTS INTO THE INTERNAL DYNAMICS OF LOOP A USING NMR SPIN RELAXATION.....	132
ABSTRACT	133
4.2 MATERIALS AND METHODS	138
4.2.1 RNA preparation.....	138
4.2.2 NMR studies	140
4.2.3 Data analysis.....	141
4.2.4 Model-free analysis	142
4.3 RESULTS AND DISCUSSION	142
4.3.1 ¹³ C Relaxation measurements	142
4.3.2 Relaxation dispersion	144
ACKNOWLEDGMENT	155
REFERENCES.....	156
 CHAPTER 5	 161
SUMMARY, DISCUSSION AND FUTURE WORK.....	161
5.1 SUMMARY AND DISCUSSION	162
5.2 FUTURE DIRECTION.....	165
5.2.1 Base dynamics for the hairpin ribozyme	166
5.2.2 Functional read-out of dynamics	167
REFERENCES	170

LIST OF TABLES

Table 2-1: Summary of molecular dynamics simulations of loop A constructs	49
Table 2-2: Comparison of NOE distance constraints with simulation averages using Amber and CHARMM	55
Table 2-3: Representative backbone and base-pair parameters of average structures derived from CHARMM36 and Amberff10 simulations respectively	56
Table 2-4: Characterization of the kinetic states observed during loop A simulation.....	61
Table 2-5: Residue dynamics for selected loop residues during hpA simulations	62
Table 2-6: Conformational energies of loop A conformers calculated with MMPB/SA	67
Table 2-7: Percentage population that sample specific interactions observed in the docked form of loop A during isolated loop A simulation	76
Table 3-1: Resonance assignments for loop A (GAAA LpAext) using NMR experiments.....	110
Table 3-2: Non exchangeable assignments of H6/H8 aromatic protons in loop A	118
Table 3-3: Non exchangeable assignments of adenine H2 aromatic protons in loop A	119
Table 3-4: Assignments of H1'/H2' and C1'/C2' in loop A.....	120
Table 3-5: Aromatic resonance assignments of H1'/H2' and C1'/C2' in loop A.....	122
Table 4-1: ¹³ C R ₁ , R _{1ρ} and heteronuclear NOE measurements.....	148
Table 4-2: Model-free analysis of loop A relaxation measurements.....	149
Table 4-3: Chemical exchange parameters from T _{1ρ} dispersion analysis of loop A.....	154

LIST OF FIGURES

Figure 1-1: Docked structure of the RNA hairpin ribozyme loop A (cyan) and loop B (green) domains.....	5
Figure 1-2: Illustration of conformational and chemical transition in a functional macromolecule along a reaction co-ordinate.....	6
Figure 1-3: The crystal structure of self-splicing group II intron.....	9
Figure 1-4: Secondary structure representation of the trans-acting hairpin ribozyme in the undocked form.....	15
Figure 1-5: Timescales and dynamics sampled by chemical and biochemical processes.....	22
Figure 2-1: Trans-acting hairpin ribozyme secondary structure representation of metal-dependent loop-loop docking.....	45
Figure 2-2: Schematic representation of free loops A and B as well as docked loop A and B.....	46
Figure 2-3: Secondary structure representation of loop A with their respective heavy atom RMSD profiles.....	54
Figure 2-4: Population distribution between non-methylated loop A (green) and methylated loop A (red).....	57
Figure 2-5: Schematic representations (A) and structural renderings (B) of loop A simulation structures.....	60
Figure 2-6: Schematic representation of loop A conformational classification based on loop base-pairing and stacking.....	61
Figure 2-7: Conformational transition of loop A RNA capturing top 18 macrostates derived from Markov state modeling.....	64
Figure 2-8: Targeted Molecular Dynamics (TMD) simulation of conformer transitions from AA/UA to GA/UA and AA/CA.....	68
Figure 2-9: Free energy diagrams depicting the energy profile of TMD-driven conformations..	69
Figure 2-10: Loop A sampling docked conformation.....	73
Figure 3-1: Cartoon representation of hairpin ribozyme active site region.....	89
Figure 3-2: Hairpin ribozyme trans-constructs used in our studies.....	91

Figure 3-3: Comparison between the circular dichroism (CD) spectra of docking between wild type loop A and loop A(U+2C/C+3U)	97
Figure 3-4: $\text{Co}(\text{NH}_3)_6^{3+}$ titration studies to assay for metal-dependent interaction between loop A(U+2C/C+3U) and loop B.....	98
Figure 3-5: 1D imino NMR peaks of various loop A constructs.....	101
Figure 3-6: RNA base-pairing between A-U and G-C highlighting the involvement of imino protons of guanine and uracil.....	102
Figure 3-7: 1D imino spectra of GUAA loop A at different annealing conditions.....	103
Figure 3-8: Temperature-dependent 1D imino spectrum of extended GAAA loop A.....	104
Figure 3-9: 2D imino NOESY spectrum of GAAA extended loop A.....	106
Figure 3-10: 2D amino NOESY spectrum of GAAA extended loop A.....	107
Figure 3-11: 2D imino ^1H , ^{15}N HSQC spectrum of GAAA extended loop A.....	109
Figure 3-12: 2D ^1H , ^{13}C -HSQC contour spectrum of the ribose region in loop A.....	114
Figure 3-13: 2D ^1H , ^{13}C -HSQC spectrum of aromatic region in loop A.....	115
Figure 3-14: 2D ^1H , ^{13}C CT-HSQC spectrum of aromatic region in loop A.....	116
Figure 3-15: Representative 2D planes of 3D HCCH-COSY spectrum of the ribose region in loop A.....	117
Figure 3-16: Representative 2D planes of ^{13}C -edited NOESY-HSQC spectrum of loop A.....	121
Figure 3-17: Schematic representation of NOE connectivities that correlate intra- and inter-residue H8/H8 to H1' and H2'	122
Figure 4-1: Timescale of macromolecular internal motions relevant in biological processes...	137
Figure 4-2: The hairpin ribozyme loop A construct used in NMR relaxation studies.....	139
Figure 4-3: Representative ^{13}C R_1 and $R_{1\text{pobs}}$ curves for A2 and A9	146
Figure 4-4: Histograms of the measured relaxation rates in loop A.....	147
Figure 4-5: Histogram representation of order parameter (S^2) and exchange contribution (R_{ex}) derived from relaxation measurements	150
Figure 4-6: Simulation structure of loop A highlighting base stacking in the GAAA tetraloop involving A2 stacked between A1 and A3	152

Figure 4-7: Representative relaxation dispersion curves of selected C8 resonances.....	153
--	-----

KEY TO ABBREVIATIONS

1D	one dimensional
2D	two dimensional
2WJ	two-way junction
3D	three dimensional
4WJ	four-way junction
ATP	adenosine triphosphate
CD	circular dichroism
COSY	correlated spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CSA	chemical shift anisotropy
CT	constant time
CTP	cytosine triphosphate
DC	decoupler
DNA	Deoxyribonucleic acid
DSS	sodium 2,2 dimethyl-2-silapentane-5-sulfonate
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
E-S	enzyme-substrate
FID	free induction decay
FRET	fluorescence resonance energy transfer
GlcN6P	glucosamine-6-phosphate
GTP	guanosine triphosphate
HDV	hepatitis delta virus

His	histidine
HIV	human immunodeficiency virus
hNOE	heteronuclear nuclear Overhauser enhancement
HSQC	heteronuclear single quantum coherence
LNA	locked nucleic acid
MALDI	matrix-assisted laser desorption ionization
MD	molecular dynamics
MMPB/SA	molecular mechanics Poisson-Boltzmann surface area
mRNA	messenger RNA
MSM	Markov state model
MWCO	molecular weight cutoff
NADP	nicotinamide adenine dinucleotide phosphate
NAMD	nanoscale molecular dynamics
NMP	nucleotide monophosphate
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser enhancement
NOESY	nuclear Overhauser enhancement spectroscopy
NTP	nucleotide triphosphate
PDB	protein data bank
Phe	phenylalanine
PME	particle-mesh Ewald
R_1	longitudinal relaxation rate
$R_{1\rho}$	rotating-frame transverse relaxation rate
R_2	transverse relaxation rate
RF	radio frequency

RMSD	root mean square deviation
RNA	ribonucleic acid
RNAse A	ribonuclease A
RNAP	ribonucleic acid polymerase
RZ	ribozyme
SBS	substrate binding sequence
T ₁	longitudinal relaxation time
T ₂	transverse relaxation time
T _{1ρ}	rotating-frame transverse relaxation time
TMD	targeted molecular dynamics
TOCSY	total correlated spectroscopy
TOF	time of flight
tRNA	transfer ribonucleic acid
UTP	uridine triphosphate
UTR	untranslated region

CHAPTER 1

INTRODUCTION TO RIBOZYMES AND THE APPROACHES TO ASSESSING RNA DYNAMICS-FUNCTION RELATIONSHIP

1.1 BACKGROUND

1.1.1 Thesis outline and research goals

Although the factors underlying secondary structure in DNA and RNA are relatively well understood, the understanding of the driving forces and mechanisms of the formation of complex tertiary structures in nucleic acids has not kept pace with the great recent advances in RNA structural biology. Due to the importance of catalytic RNAs (ribozymes) as potential therapeutics, their tertiary structure interactions in the context of binding and catalysis remain an active subject of investigation. Specifically, domain-domain interactions are an inherent feature in the active site assembly of many ribozymes, including the hairpin ribozyme. These interactions, although rate limiting, prepare the required residues to engage in catalysis through phosphodiester backbone cleavage. The loop-loop binding transition in the hairpin ribozyme therefore presents us with a suitable model for understanding the mechanism behind RNA-RNA interactions with possible insights on RNA secondary interactions, tertiary structure interactions, dynamics and function. The biological timescale for such transitions and interactions usually occur in the regime of picoseconds to milliseconds. The RNA-RNA interaction in hairpin ribozyme implicates two independent RNA domains (loop A and loop B) with each domain containing two stems interrupted by loops A and B respectively. The interaction between these domains, often referred to as docking, is required for hairpin ribozyme activation and subsequent cleavage of the RNA backbone. The major loop rearrangements observed upon docking, as well as the kinetically unfavorable docking process, both argue for a tight and mechanistic role for conformational dynamics in the pre-organization of the catalytically-competent active site of the hairpin ribozyme. However, the details such dynamics have not been explored in the context of

dynamics-function relationship. Even so, conformational transitions and their role in activating the ground-state docking-competent forms of domains A and B of the hairpin ribozyme are still not well understood. Here, we have integrated Molecular Dynamics (MD) simulation and NMR to assess the possible transitions taking place in the hairpin ribozyme's loop A domain as an initial step towards understanding hairpin ribozyme docking mechanism. Due to the slow association rates reported for interaction between loop A and B (1) and the extensive structural rearrangements observed upon docking (**Figure 1-1**) (2-4), we hypothesize that the interaction between loop A and loop B is driven by a double conformational capture mechanism, in which both loops undergo dynamic fluctuations to conformations resembling their docked forms and that only collisions between loops that have sampled the docked forms are capable of recognition and subsequent docking. We have probed the conformational sampling and dynamics of loop A domain of the hairpin ribozyme as an initial step to understanding the active site dynamics in RNA. A deeper understanding of dynamics-function relationship is important in tuning the rate and mechanism of RNA-RNA interaction. Thus, this feature may be a useful tool in the design of RNA aptamers that can be used in effective gene therapy treatments which target RNA cleavage. Investigating RNA dynamics also presents us with a unique opportunity to study the physicochemical factors underlying RNA tertiary structure in general, and has the potential to provide generalizable insights in RNA. We have used molecular dynamics (MD) simulations to map out the conformational ensemble visited by loop A. In addition, we have pursued loop A NMR studies to probe dynamics and conformational sampling in this cleavage domain. Our studies present new insights in conformational dynamics in hairpin ribozyme loop A *vis a vis* its mechanistic and functional role.

1.1.2. RNA conformational transitions

A critical challenge in understanding biological processes at the molecular level is the elucidation of conformational variations of macromolecules towards the formation of active conformers. Indeed, these conformational changes are observed on various timescales. Of critical importance are changes that occur in timescales necessary for biological function. Many biological processes rely on transduction of information through conformational changes in proteins and nucleic acids. Examples of such dynamics are observed in how RNA recognizes their interacting partners including proteins and small ligands (5) and in the utility of ribosome (6). Conformational dynamics may facilitate pre-organization of an inactive conformation of a given macromolecule (E-S) into an activated form often referred to as the pre-catalytic form (**Figure 1-2**). Both fast and slow exchange processes incorporate proteins and RNA into biological networks usually effected by internal molecular effectors in conjunction with external cues. Thus, RNA conformational transitions become part of biochemical pathways which overall influence biological events. RNA with various conformational features can bind to protein which may stabilize one or multiple conformations within a subset of conformations. This feature has been observed in tyrosyl-tRNA synthetase derived from *Neurospora crassa* which binds to group I introns (7). Group 1 introns are large self-splicing ribozymes that catalyze their own cleavage from mRNA, tRNA and rRNA precursors. In some cases, protein binding may selectively lower the surrounding energy barrier (8, 9) which activates the dynamic transitions in the RNA. An example of this behavior is typical in the U1A protein which binds to its cognate RNA and induces dynamics in parts of the RNA especially in direct contact with the protein.

Alternate secondary structural changes also occur in RNA although they are limited by energy barriers associated with base-pair opening and formation of new base-pairs. In this case, RNA

gets kinetically trapped in local conformations which limit their ability to sample multiple conformations. To effect these transitions *in vivo*, cells have a variety of proteins capable of modulating secondary structural changes in RNA by overcoming large energetic barriers (10) as exemplified in hammerhead ribozyme. In our work, the mechanisms by which RNA explores wells involving only variations in tertiary structure, with limited changes in Watson-Crick base paired regions, are investigated.



Figure 1-1: Docked structure of the RNA hairpin ribozyme loop A (cyan) and loop B (green) domains. Structural rearrangement of loop residues in loop A and B characterize the docked form with a G+1 (red residue) extra-helically base-paired to C25 (magenta) of loop B. Figure adapted and modified from ref 3.

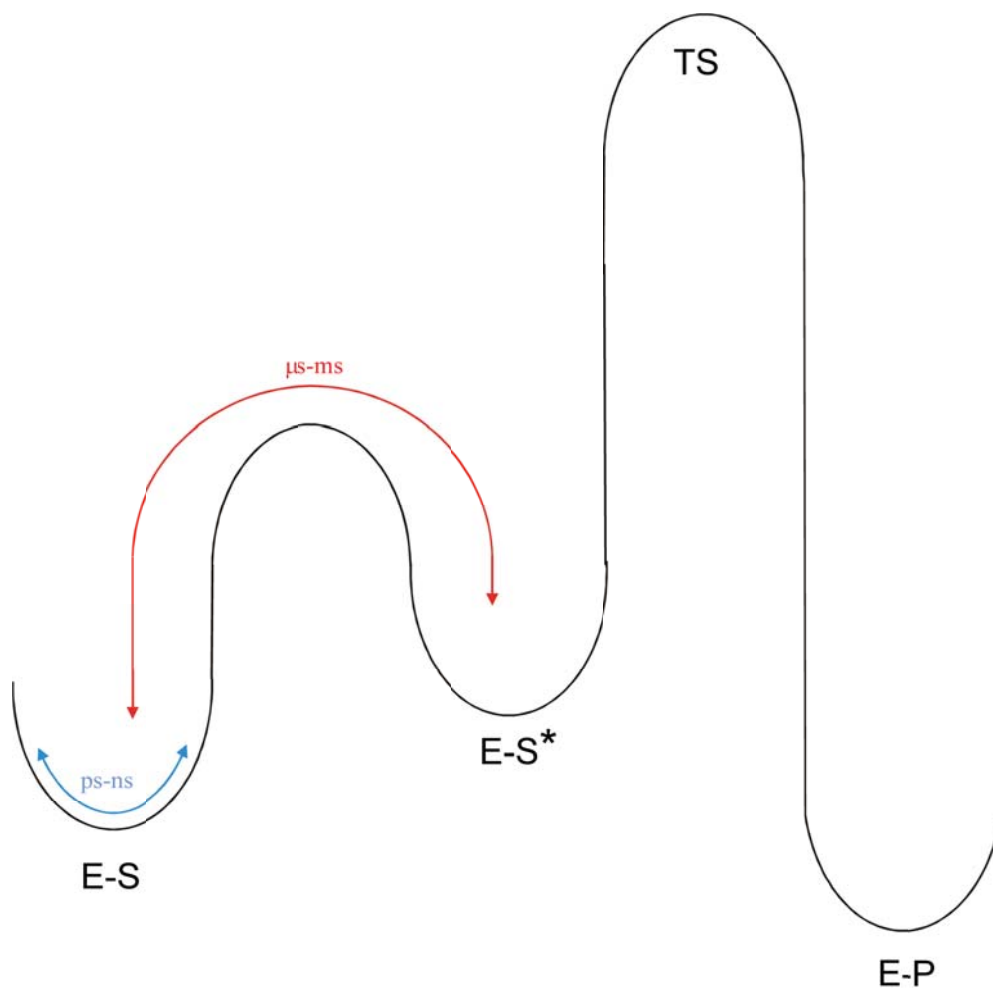


Figure 1-2: Illustration of conformational and chemical transition in a functional macromolecule along a reaction co-ordinate. E-S is enzyme-substrate complex at ground state, E-S* is the activated state, TS represents the transition state, and E-P is the enzyme product complex. RNA ribozymes undergo both local transitions and large conformational exchange at different timescales within ps-ns and μs-ms timescales respectively (Figure reprinted from James Johnson Jr. doctoral thesis).

In vivo, conformational transitions occur when signals reaching the cell modulate the formation of multiple conformations and consequent stabilization of such conformations. A good example of this is in riboswitch where metabolite binding stabilizes the riboswitch by inducing a conformational redistribution to a state that isolates RNA element into the aptamer domain (11). Riboswitches are a segment of RNA genetic elements usually inserted in the 5' untranslated region (5' UTR) of bacterial genes. These elements regulate the expression of bacterial metabolic genes with changes in cellular metabolite concentration (12). In such processes RNA samples various conformations within the energy landscape spanning through a series of deep local energy minima interrupted by shallow energy wells. These deep energy minimas are usually characterized by conformations that result from equilibrium motions which are stabilized by cellular signals to effect conformational transitions (13, 14). Tertiary conformational transitions are mostly seen in small ribozymes such as the hairpin and hepatitis delta virus (HDV). These transitions involve large conformational changes in the position of helical arms, usually after domain interaction to another RNA domain promoting the conformational changes necessary for RNA catalysis. Following catalysis, another set of undocking changes promote the release of the cleavage product (15). The importance of these transitions is underscored by the robust impact they have on the overall catalytic rate constant. The *Tetrahymena* group I ribozyme has been shown to transition between tertiary conformations while reporting different substrate binding affinities but with similar enzyme activities (16). The inter-conformational conversion rates were slower than the catalytic rate, implying the existence of multiple conformations.

1.1.3. RNA tertiary structure

Most of the current knowledge about RNA tertiary structures comes from X-ray crystallography and NMR. These techniques reveal precise snapshots of a dynamic reality, but provide little information about the dynamics at atomic resolution over time within the three-dimensional fold. Additional biophysical techniques such as single-molecule optical traps, time-resolved fluorescent resonance energy transfer (FRET) and hydroxyl radical footprinting may also provide some dynamics information. With these tools, it is possible to monitor conformational changes occurring within RNA molecules during their folding and assembly. Three-dimensional structures reveal that RNAs have an ordered organization where the primary sequence informs the architecture of secondary structure which also dictates the extent of tertiary folding. The secondary structural elements (helices, loops, bulges and junctions) are networked to each other within a complex tertiary structure (17). RNA helices are usually A-form Watson-Crick duplexes while the loops, bulges and junctions are mostly non-Watson-Crick regions terminated by one or more helices (18). Although the structural integrity of helices are maintained by Watson-Crick base-pairs and monovalent ions, the tertiary contacts are usually stabilized by non-canonical base-pairs and generally require the presence of divalent ions, especially magnesium ions. RNA architecture is weaved by a series of base-pairing and base stacking interactions leading to co-axial stacks of helical domains packed parallel or orthogonal to one another, as displayed by a recent structure of a group II self-splicing intron (**Figure 1-3**) (19).

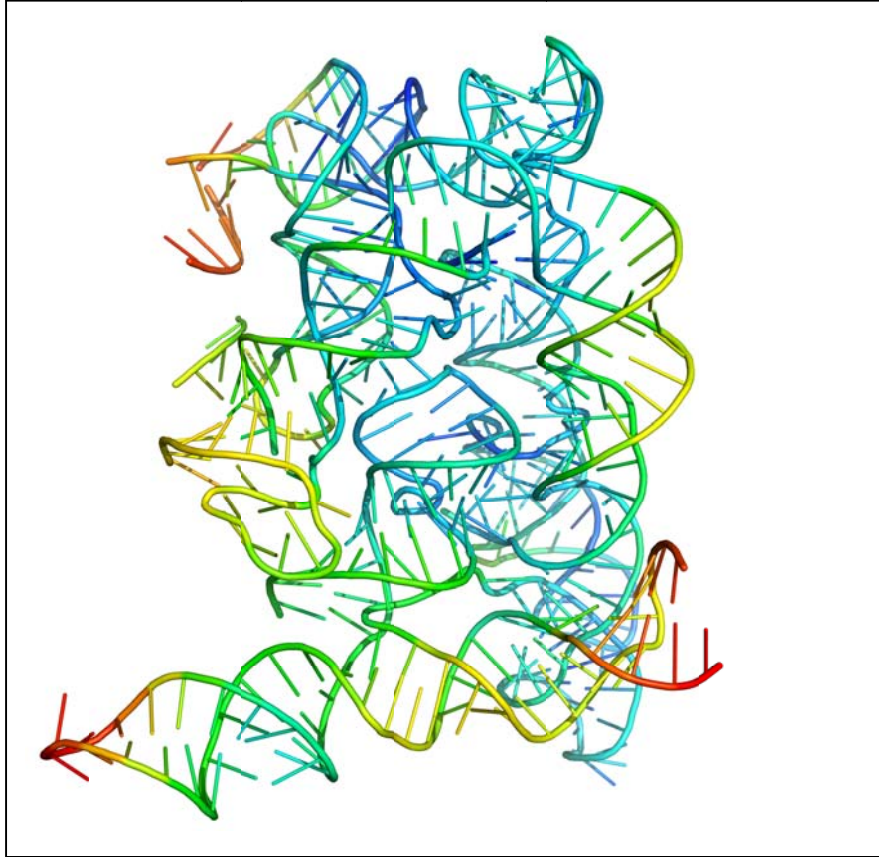


Figure 1-3:The crystal structure of self-splicing group II intron. Co-axial stacking and base-stacking interactions support the tertiary. Figure adapted from ref 3.

Junctions promote conformational diversity that can be modulated by either RNA-RNA or RNA-ligand interactions. The hammerhead ribozyme provides a good example of this phenomenon. The hammerhead active site displays a three-way helical junction containing a central core with fifteen highly conserved nucleotides essential for catalytic activity. Its activity, however, is dependent on loop-loop tertiary interactions in non-conserved regions far away from the active site (20). These long-range tertiary interactions stabilize the active conformation of the junction positioning the relevant nucleotides in the right positions for catalysis. Understanding the formation of RNA tertiary structure is pertinent towards unravelling the mechanisms of elaborate cellular machines such as the eukaryotic spliceosome, in which numerous RNA structures and RNA protein interactions form and are disrupted at precise points along the splicing pathway (21, 22). We propose that the docking transition of in the hairpin ribozyme can serve a unique role as a small, tractable RNA-RNA interaction formed entirely by tertiary structure that is accompanied by a direct functional readout in the form of self-cleavage (23, 24).

1.1.4. RNA and ribozymes

The importance of RNA was suggested as early as the late 1950s when it was established that living cells harbor much more RNA than DNA. The discovery of the details of protein synthesis (25) revealed that RNA molecules are implicated in a variety of processes within the cells. Orgel (26) and Crick (27) independently proposed that RNA acted both as catalyst and as carrier of information. The idea of RNA as catalysts received unquestionable proof after the discovery of natural RNA enzymes (ribozymes), by the independent groups of Sydney Altman and Thomas Cech (28, 29). Ribozymes are RNA molecules that catalyze reactions on themselves or other molecules. Over the years, six naturally occurring self-cleaving ribozymes have been discovered. These are the hairpin, hammerhead, hepatitis delta virus (HDV), glmS, Varkud

satellite (VS) and twister ribozymes (23, 30-34). These small ribozymes catalyze self-cleavage of the RNA backbone using a chemical strategy similar to RNase A (35).

1.1.5. Mechanism of catalysis in ribozymes

As a protein enzyme that functionally mimics ribozymes, RNase A deploys the general acid-base catalytic mechanism with unique strategies to promote RNA backbone cleavage reaction. The following strategies are adopted by specific critical amino acids to facilitate this process. These involve (1) alignment of the hydroxyl nucleophile, electrophilic phosphate, and the leaving group by reorienting the adjacent nucleotides to activate the nucleophile, (2) transition-state stabilization, and (3) protonation of the leaving group. In RNase A, the 5' of pyrimidine nucleotide forms hydrogen bonds with the main chain and backbone oxygen atoms of Threonine 45. Also, base-stacking interactions between the nucleotide 5' of the scissile phosphate with Phe120 helps in facilitating stability (36). His12 is the general base responsible for deprotonating the 2'-OH nucleophile while the negatively charged transition state is stabilized by the positively charged Lys41. Studies introducing a lysine to arginine mutation at position 41 does not significantly reduce RNase A activity indicating the requirement for a positively charged residue in that position (37).

An intriguing question in understanding the catalytic mechanism of ribozymes is the ability of RNA to catalyze diverse chemical processes with limited set of functional groups unlike the diversity in protein enzymes. The lack of neutral functional groups in RNA with pK_a close to neutrality or positively charged functional groups present a dilemma on how the chemical processes requiring these functional groups are achieved. However, a number of factors are thought to be essential for catalysis. Specific nucleotides, metals and water may be involved

in facilitating the chemistry of different sets of ribozymes. The clear deficiencies in what was thought to be critical functional groups informed an incorrect assumption that most small ribozymes were metalloenzymes (38). The participation of divalent metal ions, nucleobases and RNA backbone functional groups could provide an explanation for the general acid-base chemistry involving a general acid, base and electrostatic stabilization (39). RNAs mostly require positively charged ions to enhance a folded state. These ions are usually divalent. However, higher (molar) concentrations of monovalent cations or inert polyvalent ions are still sufficient for most ribozymes to achieve full activity (40, 41).

Despite lacking the chemical diversity exhibited by amino acid residues in RNase A, also a protein enzyme that cleaves RNA backbone, ribozymes are still capable of catalyzing RNA cleavage reactions albeit 15-500 fold slower than their RNase A counterparts (23, 42, 43). Using a divalent metal, the HDV ribozyme employs a specifically bound magnesium ion and specific nucleotides to promote phosphate backbone cleavage to a rate of ~150-fold slower than the more efficient RNase A (44), while glmS ribozyme uses a glucosamine-derivative cofactor (GlcN6P) and nucleotides to attain cleavage rates ~100-fold slower than RNase A (45). Despite the similar general catalytic mechanism, there is no general consensus on the origin of the catalytic proficiency showed by various ribozymes. The use of different strategies in promoting chemistry in ribozymes has sparked curiosity in the RNA world and remains an active field of study both computationally and experimentally.

1.1.6. Therapeutic role of ribozymes

The attempted use of RNA molecules as therapeutic agents is relatively novel, but has received growing interest over the few decades. This curiosity has been inspired by various scientific discoveries that feature the role of RNA molecules in relaying genetic instructions in living systems as well as their natural abundance and adaptability. Interestingly, RNA molecules can assume a diverse set of conformations and perform a variety of cellular functions. For example some RNAs can fold into catalytic centers, while others may adopt conformations that facilitate specific RNA–protein, RNA–DNA or RNA–RNA interactions.

The discovery of ribozymes (28, 29) has informed the development of a new class of trans-cleaving therapeutic RNAs with great potential in human medicine. This strategy is designed around inactivating viral RNA genome by cleavage. Therapeutic trans-ribozymes are designed to specifically bind RNA substrate via base-pairing interactions. This base-pairing directs cleavage of the target RNA and subsequent release the cleavage products which may go through a recycling process to facilitate a multi-step catalytic process. The realization that ribozymes can be directed to a specific RNA target to cleave pathogenic viral transcripts *in vitro* has cultivated promise on their potential therapeutic value *in vivo* (46, 47). Significant progress towards assessing the potential use of trans-cleaving ribozymes has been made, with the hairpin and hammerhead ribozymes receiving more attention in this translational effort (48). Ribozymes can potentially cleave a range of RNA transcripts if well designed to efficiently bind and cleave these transcripts. Such transcripts may include oncogene transcripts, viral genomes and many other pathogenic transcripts. The success of a synthetic ribozyme to deactivate these transcripts via cleavage will depend on the efficiency of delivering the trans-forms of ribozymes *in vivo*, the concentration and duration of inhibition required to change the pathophysiology of the disease.

For example, the hammerhead ribozyme inhibited the replication of a murine retrovirus about 90% by co-localizing this ribozyme with its viral target *in vivo* (49); however, significant increases in the levels of inhibition were not achieved even in excess concentrations of the ribozyme. This suggests that the efficacy of trans-cleaving therapeutic ribozymes could be limited to cellular conditions and therefore a combination of therapies (e.g. chemotherapy) and other therapeutics can augment the role of ribozyme. Another major focus of research on ribozyme therapy has been on inhibiting the replication of RNA viruses and retroviruses, specifically HIV-1 (50, 51), hepatitis C virus (52) and hepatitis B virus (53). These are appropriate targets since their genomes are composed of RNA, and thus replication can be directly targeted for inhibition.

1.1.7. Hairpin ribozyme

The hairpin ribozyme is a small ribozyme that catalyzes reversible, site-specific phosphodiester bond cleavage giving a 5'-OH and a 2'- 3'-cyclic phosphate termini products (54). Derived from the negative strand of tobacco ringspot virus satellite RNA, the hairpin ribozyme is involved in processing the product of rolling circle replication through backbone cleavage and ligation respectively (55, 56). The cognate substrate is recognized by canonical base-pairing to a single-stranded region of the ribozyme referred to as the substrate binding sequence (SBS) (**Figure 1-4**). The substrate and the substrate binding sequence form the loop A domain while the rest of the ribozyme forms the loop B domain. Loop A and loop B together with their flanking helices constitute the catalytic core of the hairpin ribozyme (57). The most critical nucleotides for catalysis are found in the two loop regions (58). The interaction between loop A and loop B is a necessary step that precedes catalysis (59) and is referred to as docking. The details of this process are not clearly understood. However, structural evidence infers possible conformational changes through which domain-domain interaction and subsequent catalysis is achieved. The chemical steps of the reaction proceed without the necessity of multivalent metal ions (59, 60) but the rate acceleration is still realized, a process similar to metal-dependent ribozymes (61). There origin of catalytic proficiency in hairpin ribozymes has elicited debate with a major focus on the role of active site nucleobases with respect to general acid and base catalysis. Due to the limited diversity of the four native RNA residues, the role of these residues in nucleophilic activation, stabilization of the charged transition-state and activation of the leaving group via protonation has been extensively investigated albeit with different conclusions. Critical geometric restraints facilitate bond making and breaking in the trans-phosphorylation reaction promoted by nucleolytic ribozymes (62). For example the nucleophilic S_N2 reaction aligns the

torsion angle between the 2'-OH nucleophile, the scissile phosphate, and the 5'-O leaving group should be $\sim 180^\circ$ to facilitate an inline nucleophilic attack on the scissile phosphate. Hairpin ribozymes provides a unique set of stabilizing interactions to achieve this unfavorable conformation. This inline architecture was supported by base-pairing interactions between residues in loop A and B as well as base-stacking interactions. Interestingly, the same trans-phosphorylation reaction is catalyzed by the protein enzyme ribonuclease A (RNase A), with two histidine residues used as general acid and base respectively (63). Small ribozymes therefore may employ many similar RNase A tactics in promoting chemistry. However, the hairpin ribozyme can cleave in the presence of cobalt hexamine, $[\text{Co}(\text{NH}_3)_6]^{3+}$, a non-hydrolyzable and exchange-inert mimic of magnesium hexahydrate, $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ (64, 65). This strengthens the argument that metals may serve unique roles by modulating hairpin ribozyme tertiary interactions rather than via direct involvement in catalysis.

1.1.8. Structure and active site architecture of the hairpin ribozyme

The X-ray crystal structures of the hairpin ribozyme provide valuable insights into the active site conformation of the docked state (3, 66), transition state, and product state structures (67, 68) observed during ribozyme catalysis. In each structure, the active sites had generally similar conformations. The G+1 nucleotide is completely flipped out of loop A and inserted in loop B pocket where it stacks between A26 and A38 residues (3). The flipped-out G+1 forms a tertiary Watson-Crick base-pair with C25 of loop B, and is stabilized by additional hydrogen-bonding interactions with A38 and G36 residues. The specific contacts among G+1 and these loop B residues seem to highlight the specificity of G at that +1 position (3). Mutation of G+1 to other bases results in the loss of catalytic activity which is partially restored by compensatory mutations at C25 in loop B (69). Important structural information is available in the docked

structures of loop A and loop B, but additional structural information have also been obtained from the solution structures of the isolated loop A and loop B which have been separately determined (4, 70). The separate RNA pieces can be mixed together *in vitro* to reconstitute a functional ribozyme (71). The nuclear magnetic resonance (NMR) structures of loops A and B probably represent the conformation of these two loops before docking. In the isolated loop A structure, an anti-conformation G+1 is accommodated in the helical stack and forms a sheared pair with A9 (4). This residue is involved in a G+1:C25 tertiary base pairing with domain B in the docked structure (3), indicating a dramatic conformational reshuffle before or after docking. Also, the minor groove width of the isolated loop A conformation observed in NMR is considerably narrower than the docked state. The conformation of isolated loop B (70) also differs greatly from the conformation observed in the crystal structure (docked form). Only two of the seven non-canonical base pairs observed by NMR in isolated loop B are retained in the docked conformation. Whereas U41 is extruded into an S-turn in the undocked loop B, it forms base pair with A22 in the docked structure (3, 70). A detailed comparison between the docked crystal structure and the isolated loop structures (loop A and loop B) shows large conformational rearrangements taking place in both RNA loops in consistent with the active site assembly. The rearrangements involve dramatic changes in base-pairing schemes and significant alteration of backbone geometry concomitant with the formation of tertiary interactions. These observations support the idea that conformational dynamics could at play in the process of loop-loop interaction, substrate orientation and subsequent catalysis. Our thesis is that static structures are important but insufficient to understand the mechanism of ribozyme action. Previously, an NMR structural study of the lead-dependent ribozyme (leadzyme) observed that the highly populated conformation in solution is not well-oriented to sample its activated state, suggesting that

conformational rearrangement was necessary for to achieve catalytic transitions (72). Similarly, we also hypothesize that independent loops of the hairpin ribozyme may be sampling other higher energy (activated) states via specific dynamic modes and that these modes may be coupled to formation of active structures in facilitating hairpin ribozyme function. Our investigation has harnessed contemporary Molecular dynamics (MD) simulation and nuclear magnetic resonance (NMR) techniques which have in the past provided the richest information on dynamics of macromolecules. Previous work has shown that “trans-docking” constructs, loops A and B, still retain catalytic activity (73) and that the two pieces can bind to each other with the K_D in the low μM range (74). In this thesis, we have mainly probed loop A sampling the high energy docked form in the process of conformational fluctuation. *We hypothesize* that the relatively slow docking rates may be due to the isolated loop A sampling other conformers including the docked (activated) form. In other words, only when structurally favorable conformers appear is loop A competent for docking. This phenomenon has precedents in several carefully-studied protein enzymes. Dihydrofolate reductase-NADPH complex, for example, samples a higher energy substate in which the empty substrate-/product-binding pocket adopts a conformation similar to that of the ligand-bound state (75). A similar observation has been reported for ribonuclease A (76).

1.1.9. Computer molecular dynamics (MD) simulations

Molecular dynamics (MD) simulation uses the computational approach to interrogate the atomic transitions and molecular motions using classical empirical potential force fields. It describes the time evolution in a set of discrete particles by solving the Newton’s equation of classical mechanics according to equation 1-1.

$$F_i = m_i a_i \quad 1-1$$

where m_i and a_i are the mass and acceleration of atom i respectively while F_i is the force acting on the mass calculated from the gradient of potential energy function, U .

$$F_i = -\nabla_i U \quad 1-2$$

MD simulation method calculates the time-dependent behavior of a molecular system within the molecular force field that defines the level of interactions among particles in their environment. MD simulations give a variety of information ranging from fast fluctuations to slower conformational changes of proteins, nucleic acids and even lipids. Computational methods are currently attractive for use in investigating the chemical transitions and structural dynamics of many biological molecules as well as biological complexes. The robustness of this technique in assessing dynamics coupled with its cost-effective sampling has provoked several studies in this field. Some of the most commonly used force fields are CHARMM (77-79), AMBER (80), GROMOS (81), and OPLS (82, 83). While the specifics of the potential energy functions used in various force fields are different, they generally use the general mathematical bonded and non-bonded interaction terms. The results of a simulation are determined by the quality of the force field used, the sampled time (timescale) and the accuracy of the initial structures. Therefore, MD simulation is a tool primarily suitable to assess conformational dynamics within a wide range of time scales (**Figure 1-5**) using both the all-atom simulation method (84) as well as enhanced sampling techniques used to explore the conformational space of a given molecule in the relevant timescale. Modern MD simulations define the potential energy function, U , as constituting the bonding terms and the non-bonding terms as a function of the Cartesian co-ordinates (85-87) as shown in equation 1-3.

$$\begin{aligned}
E_{MM} = & \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 \\
& + \sum_{dihedrals} K_\phi [1 + \cos(n\phi - \delta)] + \sum_{impropers} K_\omega (\omega - \omega_0) \\
& + \sum_{i < j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i < j} \frac{q_i q_j}{\epsilon r_{ij}}
\end{aligned} \tag{1-3}$$

The bonding terms are defined by the bond lengths (b), bond angles (θ), dihedrals (ϕ), and improper dihedrals (ω) (Eqn. 1-3). Equilibrium values are denoted by subscript 0 whereas K_b , K_θ , K_ϕ and K_ω are the force constants for the bonding terms of bond lengths, bond angles, dihedrals and improper dihedrals respectively. The dihedral angle term is modeled along the sinusoidal function with n and δ representing periodicity and phase shift respectively. The non-bonding terms are van der Waals and electrostatic interactions modeled by distances between atoms i and j respectively. Also, q_i and q_j are the point charges between which electrostatic interactions are calculated using Coulombic potential. The van der Waals component of the equation models the repulsion and attraction forces.

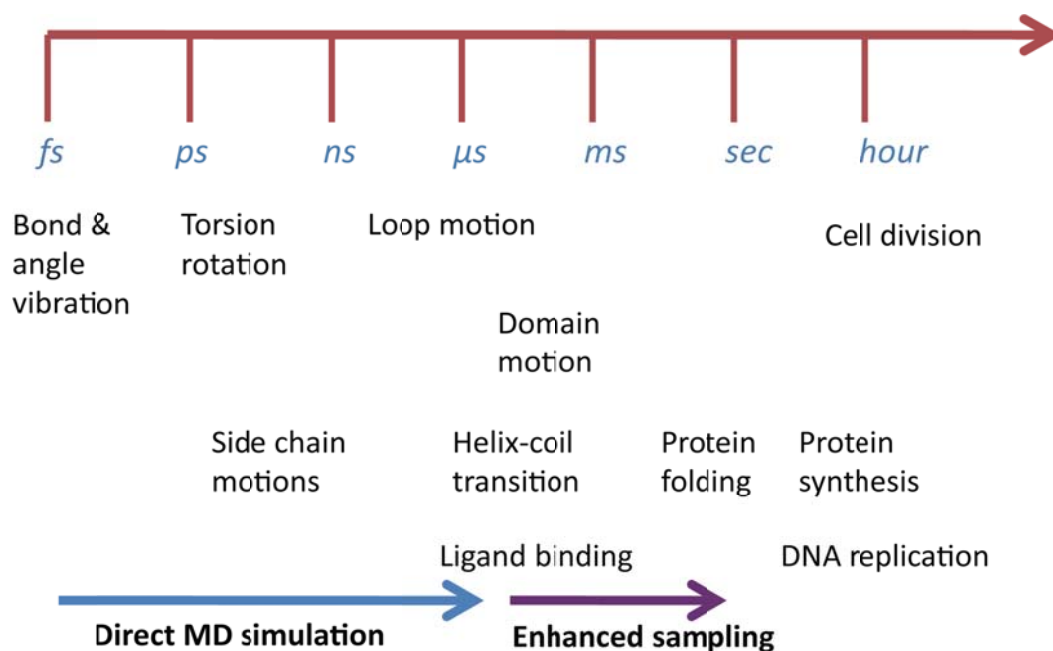


Figure 1-5: Timescales and dynamics sampled by chemical and biochemical processes. Molecular dynamics simulation can sample a range of dynamic processes. Important biochemical processes involving local transitions and large conformational exchange occur at different timescales within picoseconds to hour timescales.

All-atom simulations have unique capabilities to map out the motions available to the molecule in three-dimensional detail. The use of state-of-the-art molecular dynamics calculations to build and complement structural and functional data is a key ingredient of the synergistic approach to dynamics-function relationship. The atomic-level dynamics that cannot be observed directly are easily deciphered and used to complement experimental techniques like NMR and X-ray crystallography. The hairpin ribozyme is a uniquely suitable model system for the pursuit of a deeper understanding of conformational sampling and dynamics, and the MD approach maximizes the possibilities of successfully obtaining this understanding in a rigorous and comprehensive fashion. Previous investigators have focused on simulation studies on the structure and catalysis of the hairpin ribozyme (88-93), but the great majority of these have mostly explored the properties of the post-docking, catalytically competent state. Our exploration of ground-state loop A and the docking transition presents a mechanistic detail for the activation of RNA domains as the initial step towards the formation of the docked tertiary structure that is a prerequisite for catalysis. This phenomenon has been comprehensively studied and reported in chapters two and three of this thesis.

1.2.0. Experimental assessment of dynamics

The first protein X-ray structure of myoglobin (94) opened up the field of structural biology to ask certain fundamental questions on how ligands could access the deeply buried heme iron center. This observation stimulated decades of research into the dynamics of proteins and today, it is reasonably conventional that protein structures are in continuous motion and that such fluctuations are important for function. Early X-ray crystal structures of RNA also predicted the role of conformational dynamics especially due to the huge rearrangements observed in the

tRNA's helical arms upon binding tRNA synthetase (95). Similarly the conformational change theory in ribozymes well explains the formation of active states (96-98).

To infer dynamics from X-ray structures, several crystal structures capturing the molecules at different stages of its function can be considered as part of the structural ensemble through which structural transitions can be deciphered. Also, the flexible regions based on B-factors can be assumed to be dynamic. However, more definitive answers can be obtained by using NMR solution studies. NMR spectroscopy can be used to interrogate macromolecules motions of over a range of timescales using relaxation measurements. Relaxation in NMR is a phenomenon that is induced by field fluctuations that occur during molecular motion, and subsequently imparting variations in the local fields within the molecular movement. This property facilitates the NMR relaxation experiments to capture fluctuations that are derived from such molecular motions. Relaxation in macromolecules probes ^{13}C and ^{15}N nuclei in solution which arises from dipolar interaction with covalently bound protons. The NMR techniques that are actively in use to study these motions are longitudinal (T_1), transverse (T_2) relaxation and heteronuclear nuclear Overhauser enhancement (hNOE) measurements. In the early days, natural abundance ^{13}C -NMR T_1 measurements were used to study motions within ribonuclease A (99). This technique was successful since it annotated multiple dynamics for β - and ϵ -carbons. A year later, the same research group analyzed detailed segmental motions of ribonuclease (100) to assess the robustness of T_1 . Over the next several years, researchers continued to obtain relaxation measurements using ^{13}C natural abundance in DNA for ribose and base carbons. Due to the low sensitivity of the ^{13}C natural abundance, minimal ^{13}C enrichment was eventually used to study dynamics in DNA (101) and also RNA (102). The first attempt to quantify relaxation in uniformly ^{13}C labeled RNA (103) with relaxation measurements involving longitudinal

relaxation, transverse relaxation and NOE parameters was interpreted using the model-free analysis despite the challenges of dipolar couplings in the non-isolated spin systems. The model free analysis formalism (104-106) was developed by Lipari and Szabo to extract the fast ps-ns dynamics parameters from the measured relaxation rates. In the “model-free” approach, fast internal motion is described by a generalized order parameter, S^2 , which assesses the atomic spatial restriction, and an effective correlation time, τ_e , which provides information on the rate of the motion. This formalism removes the ambiguity of internal dynamics interpretation. ^{13}C relaxation can be complicated by crowded spin-systems in RNA or DNA, and this informs the use of C8/C6 and C2 resonances in simplifying relaxation data. Also, magnetic interactions between neighboring ^{13}C nuclei can significantly enhance relaxation measurements, and this may limit the accuracy of data collected especially for nucleic acids. Fortunately, various approaches can successfully suppress such unwanted cross-correlation effects. The false effects of spin echo-modulation during the ^{13}C relaxation delay usually caused by the ^{13}C - ^{13}C dipolar coupling can be suppressed by measuring $T_{1\rho}$ experiment instead of T_2 (107) which is measured as a time constant during the loss of spin magnetization under the magnetic spin-lock conditions. T_1 , T_2 , and heteronuclear NOE are Fourier transformations of the autocorrelation function for molecular motion. Autocorrelated (T_1 , T_2 , $T_{1\rho}$, NOE) and cross-correlated relaxation measurements (108) are useful in monitoring fast internal motions in the ps–ns timescales.

The ability of NMR to detect chemical exchange between various conformational species usually occurs in the μs – ms timescale. The most frequently used NMR methods for probing conformational exchange the μs – ms timescale are relaxation dispersion measurements. Conformational exchange in the μs – ms timescale results in broad NMR signals. The broadening, reflected by conformational exchange, R_{ex} , contributes to the measured transverse relaxation rate

($R_{2\text{eff}}$). For a two-site dynamic process, three physical parameters are obtained. The parameters are; interconversion rates (k_{ex}), the population fractions of the exchanging species (pA and pB) and the chemical shifts between pA and pB ($\Delta\omega$). The value of the effective transverse relaxation rates ($R_{2(\text{eff})}$ or $R_{1\rho(\text{eff})}$) derive contributions from R_2^* (R_2 relaxation from dipolar and chemical shift anisotropy relaxation) or $R_{1\rho\infty}$ (relaxation rate at infinite spin-lock power) and R_{ex} , is the relaxation attributed to conformational exchange. Equations 1-4 and 1-5 summarize the expressions.

$$R_{2(\text{eff})} = R_2^* + R_{\text{ex}} \quad 1-4$$

$$R_{1\rho(\text{eff})} = R_{1\rho\infty} + R_{\text{ex}} \quad 1-5$$

There are several studies particularly of protein dynamics based on relaxation dispersion and off-resonance $R_{1\rho}$ (109-113). However, there are only a few studies focusing on conformational exchange in RNA, and even so in the of the power dependence of $T_{1\rho}$ (112, 114-116). An example of studies exploring $T_{1\rho}$ relaxation in the leadzyme indicates that the thermal fluctuation of the reactive groups are involved in transitioning highly populated inactive state to the low populated active state (115). Also, NMR work done by Johnson *et al* (112), derived the transverse relaxation rates of RNA ribose during conformational exchange which was determined to depend on a range of factors including the nuclei exchange lifetime, relative populations of exchanging species, pulse rate, and the chemical shift difference between the conformational states. In this thesis, we have used hairpin ribozyme loop A as a model to further characterize similar dynamic modes in the loop region of RNA using NMR. The findings of this investigation have been presented in detail in Chapter four of this thesis.

1.2.1. Functional probing of RNA dynamics

The conformational dynamics of RNA can originate from the phosphate backbone, base, or ribose. However, it remains a challenge to directly probe the functional significance of a specific RNA conformational change or dynamic region. Our lab introduced a specific isotope-labeling scheme that addresses the challenges of ribose ^{13}C NMR analysis of dynamics in RNA. Previous attempts at probing RNA dynamics by taking advantage of conformational preferences of nucleotides with substitutions at the 2'-carbon have been thwarted by the relatively modest magnitude of such preferences, representing a maximum of ~ 0.8 kcal/mol at 298 K for the most commonly used nucleotides. Our lab has also demonstrated the use of conformationally restricted locked nucleic acid (LNA) to perturb ribose dynamics in a lead-dependent ribozyme and other systems (117). An LNA nucleotide features a methylene linkage between 2'-oxygen and 4'-carbon which conformationally restricts the nucleotide to a C3'-endo pucker. Here, Julien *et al* demonstrated the role of modulating dynamics with a 20-fold increase in catalytic rate in the leadzyme upon LNA substitution at G9. This observation provides a precise tool for mapping conformational change at a specific ribose site that appears to be a key part of the generation of the active state of the ribozyme. LNAs have also previously been incorporated into helical regions and recognition arms of 10-23 catalytic DNA and the hammerhead ribozymes in an effort to improve targeting of cellular and RNA sequences resulting in dramatic effects upon structural stability and target affinity (118, 119). This strategy can be applied to a variety of ribozymes that exhibit C3'/C2' sugar pucker transitions. Most notably, critical dynamic residues in loop A of the hairpin ribozyme including G8, U+2, C+3 and A-1 whose sugar puckers are either C2'-endo or mixed C2'/C3' (3, 4) can provide good probe sites for comprehensively assessing the effects of dynamics on the function of hairpin ribozyme.

In catalytic RNA the role of dynamics is even more prominent than most protein enzymes. It appears that the conformational fluctuations from a stable ground state to an activated conformer are an inherent part of most ribozymes. One notable example includes the conformational rearrangements observed between crystal structures of precursor and product forms of the hepatitis delta virus (HDV) with catalytic or inhibitory metal ions binding only to the precursor form (120). Conformational rearrangements have also been observed between isolated and docked hairpin ribozyme loop A structures (3, 4). These features show that conformational dynamics is not only vital to catalytic activity but could also be involved in directing chemistry. In this thesis, we have taken a multi-pronged approach of using Molecular dynamics (MD) simulations and NMR spectroscopy to characterize conformational dynamics and sampling in loop A of the hairpin ribozyme. Both MD simulation and NMR spectroscopy are robust techniques which present us with a unique ability to investigate dynamic properties of molecules over a range of different timescales with atomic resolution, and integrate both results in understanding the mechanism behind RNA-RNA tertiary interactions. To achieve this goal we have done extensive MD simulations and herein presented our detailed results in chapter two. We have also isotopically labeled the hairpin ribozyme loop A using ^{15}N and ^{13}C , assigned the resonances and carried our NMR relaxation studies also presented in chapter four of this thesis.

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. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat Struct Biol* 6, 212-216.
3. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis, *Nature* 410, 780-786.
4. Cai, Z., and Tinoco, I., Jr. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite, *Biochemistry* 35, 6026-6036.
5. Frankel, A. D. (1999) If the loop fits, *Nat Struct Biol* 6, 1081-1083.
6. Green, R., and Noller, H. F. (1997) Ribosomes and translation, *Annu Rev Biochem* 66, 679-716.
7. Paukstelis, P. J., Chen, J. H., Chase, E., Lambowitz, A. M., and Golden, B. L. (2008) Structure of a tyrosyl-tRNA synthetase splicing factor bound to a group I intron RNA, *Nature* 451, 94-97.
8. Shajani, Z., Drobny, G., and Varani, G. (2007) Binding of U1A protein changes RNA dynamics as observed by ¹³C NMR relaxation studies, *Biochemistry* 46, 5875-5883.
9. Bokinsky, G., Nivon, L. G., Liu, S., Chai, G., Hong, M., Weeks, K. M., and Zhuang, X. (2006) Two distinct binding modes of a protein cofactor with its target RNA, *J Mol Biol* 361, 771-784.
10. Herschlag, D., Khosla, M., Tsuchihashi, Z., and Karpel, R. L. (1994) An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis, *Embo J* 13, 2913-2924.
11. Greenleaf, W. J., Frieda, K. L., Foster, D. A. N., Woodside, M. T., and Block, S. M. (2008) Direct observation of hierarchical folding in single riboswitch aptamers, *Science* 319, 630-633.
12. Winkler, W., Nahvi, A., and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression, *Nature* 419, 952-956.

13. Bajor, M. H., Mustoe, A. M., Brooks, C. L., 3rd, and Al-Hashimi, H. M. (2011) Topological constraints: using RNA secondary structure to model 3D conformation, folding pathways, and dynamic adaptation, *Curr Opin Struct Biol* 21, 296-305.
14. Schultes, E. A., Spasic, A., Mohanty, U., and Bartel, D. P. (2005) Compact and ordered collapse of randomly generated RNA sequences, *Nat Struct Mol Biol* 12, 1130-1136.
15. Wilson, T. J., Nahas, M., Ha, T., and Lilley, D. M. (2005) Folding and catalysis of the hairpin ribozyme, *Biochem Soc Trans* 33, 461-465.
16. Solomatin, S. V., Greenfield, M., Chu, S., and Herschlag, D. (2010) Multiple native states reveal persistent ruggedness of an RNA folding landscape, *Nature* 463, 681-U117.
17. Cruz, J. A., and Westhof, E. (2009) The dynamic landscapes of RNA architecture, *Cell* 136, 604-609.
18. Tinoco, I., and Bustamante, C. (1999) How RNA folds, *J Mol Biol* 293, 271-281.
19. Toor, N., Keating, K. S., Taylor, S. D., and Pyle, A. M. (2008) Crystal structure of a self-spliced group II intron, *Science* 320, 77-82.
20. Martick, M., and Scott, W. G. (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis, *Cell* 126, 309-320.
21. Madhani, H. D., and Guthrie, C. (1994) Dynamic RNA-RNA Interactions in the Spliceosome, *Annu Rev Genet* 28, 1-26.
22. Wahl, M. C., Will, C. L., and Luhrmann, R. (2009) The Spliceosome: Design Principles of a Dynamic RNP Machine, *Cell* 136, 701-718.
23. Fedor, M. J. (2000) Structure and function of the hairpin ribozyme, *J Mol Biol* 297, 269-291.
24. Fedor, M. J. (2009) Comparative enzymology and structural biology of RNA self-cleavage, *Annu Rev Biophys* 38, 271-299.
25. Crick, F. H. (1958) On protein synthesis, *Symp Soc Exp Biol* 12, 138-163.
26. Orgel, L. E. (1968) Evolution of the genetic apparatus, *J Mol Biol* 38, 381-393.
27. Crick, F. H. (1968) The origin of the genetic code, *J Mol Biol* 38, 367-379.
28. Guerriertakada, 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.

29. 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.
30. Been, M. D., and Wickham, G. S. (1997) Self-cleaving ribozymes of hepatitis delta virus RNA, *Eur J Biochem* 247, 741-753.
31. Lilley, D. M. (2004) The Varkud satellite ribozyme, *RNA* 10, 151-158.
32. Lim, J., Grove, B. C., Roth, A., and Breaker, R. R. (2006) Characteristics of ligand recognition by a glmS self-cleaving ribozyme, *Angew Chem Int Ed Engl* 45, 6689-6693.
33. McKay, D. B. (1996) Structure and function of the hammerhead ribozyme: an unfinished story, *RNA* 2, 395-403.
34. Roth, A., Weinberg, Z., Chen, A. G. Y., 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.
35. Raines, R. T. (1998) Ribonuclease A, *Chem Rev* 98, 1045-1066.
36. Wlodawer, A., Miller, M., and Sjolín, L. (1983) Active site of RNase: neutron diffraction study of a complex with uridine vanadate, a transition-state analog, *P Natl Acad Sci USA* 80, 3628-3631.
37. Trautwein, K., Holliger, P., Stackhouse, J., and Benner, S. A. (1991) Site-directed mutagenesis of bovine pancreatic ribonuclease: lysine-41 and aspartate-121, *FEBS Lett* 281, 275-277.
38. Pyle, A. M. (1993) Ribozymes: a distinct class of metalloenzymes, *Science* 261, 709-714.
39. Sigel, R. K., and Pyle, A. M. (2007) Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry, *Chem Rev* 107, 97-113.
40. Murray, J. B., Seyhan, A. A., Walter, N. G., Burke, J. M., and Scott, W. G. (1998) The hammerhead, hairpin and VS ribozymes are catalytically proficient in monovalent cations alone, *Chem Biol* 5, 587-595.
41. Roth, A., Nahvi, A., Lee, M., Jona, I., and Breaker, R. R. (2006) Characteristics of the glmS ribozyme suggest only structural roles for divalent metal ions, *RNA* 12, 607-619.
42. Canny, M. D., Jucker, F. M., Kellogg, E., Khvorova, A., Jayasena, S. D., and Pardi, A. (2004) Fast cleavage kinetics of a natural hammerhead ribozyme, *J Am Chem Soc* 126, 10848-10849.

43. Zamel, R., Poon, A., Jaikaran, D., Andersen, A., Olive, J., De Abreu, D., and Collins, R. A. (2004) Exceptionally fast self-cleavage by a *Neurospora* Varkud satellite ribozyme, *P Natl Acad Sci USA* 101, 1467-1472.
44. Brown, T. S., Chadalavada, D. M., and Bevilacqua, P. C. (2004) Design of a highly reactive HDV ribozyme sequence uncovers facilitation of RNA folding by alternative pairings and physiological ionic strength, *J Mol Biol* 341, 695-712.
45. Wilkinson, S. R., and Been, M. D. (2005) A pseudoknot in the 3' non-core region of the glmS ribozyme enhances self-cleavage activity, *RNA* 11, 1788-1794.
46. Cech, T. R. (1988) Ribozymes and Their Medical Implications, *J Am Med Assoc* 260, 3030-3034.
47. Usman, N., and Blatt, L. M. (2000) Nuclease-resistant synthetic ribozymes: developing a new class of therapeutics, *J Clin Invest* 106, 1197-1202.
48. Symons, R. H. (1992) Small Catalytic RNAs, *Annu Rev Biochem* 61, 641-671.
49. Sullenger, B. A., and Cech, T. R. (1993) Tethering Ribozymes to a Retroviral Packaging Signal for Destruction of Viral-RNA, *Science* 262, 1566-1569.
50. Ojwang, J. O., Hampel, A., Looney, D. J., Wongstaal, F., and Rappaport, J. (1992) Inhibition of Human-Immunodeficiency-Virus Type-1 Expression by a Hairpin Ribozyme, *P Natl Acad Sci USA* 89, 10802-10806.
51. Rossi, J. J. (2000) Ribozyme therapy for HIV infection, *Adv Drug Deliver Rev* 44, 71-78.
52. Lieber, A., He, C. Y., Polyak, S. J., Gretch, D. R., Barr, D., and Kay, M. A. (1996) Elimination of hepatitis C virus RNA in infected human hepatocytes by adenovirus-mediated expression of ribozymes, *J Virol* 70, 8782-8791.
53. Pulitz, J. Z., Yu, Q., Burke, J. M., and Wands, J. R. (1999) Combinatorial screening and intracellular antiviral activity of hairpin ribozymes directed against hepatitis B virus, *J Virol* 73, 5381-5387.
54. Buzayan, J. M., Gerlach, W. L., Bruening, G., Keese, P., and Gould, A. R. (1986) Nucleotide sequence of satellite tobacco ringspot virus RNA and its relationship to multimeric forms, *Virology* 151, 186-199.
55. 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, *P Natl Acad Sci USA* 83, 8859-8862.

56. Walter, N. G., and Burke, J. M. (1998) The hairpin ribozyme: structure, assembly and catalysis, *Curr Opin Chem Biol* 2, 303.
57. Hampel, A., and Tritz, R. (1989) RNA catalytic properties of the minimum (-)sTRSV sequence, *Biochemistry* 28, 4929-4933.
58. Anderson, P., Monforte, J., Tritz, R., Nesbitt, S., Hearst, J., and Hampel, A. (1994) Mutagenesis of the hairpin ribozyme, *Nucleic Acids Res* 22, 1096-1100.
59. Walter, N. G., and Burke, J. M. (1998) The hairpin ribozyme: structure, assembly and catalysis, *Curr Opin Chem Biol* 2, 24-30.
60. Lilley, D. M. (1999) Structure, folding and catalysis of the small nucleolytic ribozymes, *Curr Opin Struct Biol* 9, 330-338.
61. Fedor, M. J. (2000) Structure and function of the hairpin ribozyme, *J Mol Biol* 297, 269-291.
62. Cochrane, J. C., and Strobel, S. A. (2008) Catalytic strategies of self-cleaving ribozymes, *Acc Chem Res* 41, 1027-1035.
63. Raines, R. T. (1998) Ribonuclease A, *Chem Rev* 98, 1045-1066.
64. Kuimelis, R. G., and McLaughlin, L. W. (1998) Mechanisms of Ribozyme-Mediated RNA Cleavage, *Chem Rev* 98, 1027-1044.
65. Lilley, D. M. (2003) The origins of RNA catalysis in ribozymes, *Trends Biochem Sci* 28, 495-501.
66. 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.
67. Rupert, P. B., Massey, A. P., Sigurdsson, S. T., and Ferre-D'Amare, A. R. (2002) Transition state stabilization by a catalytic RNA, *Science* 298, 1421-1424.
68. Torelli, A. T., Krucinska, J., and Wedekind, J. E. (2007) A comparison of vanadate to a 2'-5' linkage at the active site of a small ribozyme suggests a role for water in transition-state stabilization, *RNA* 13, 1052-1070.
69. Pinard, R., Lambert, D., Walter, N. G., Heckman, J. E., Major, F., and Burke, J. M. (1999) Structural basis for the guanosine requirement of the hairpin ribozyme, *Biochemistry* 38, 16035-16039.

70. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat Struct Biol* 6, 212-216.
71. Butcher, S. E., Heckman, J. E., and Burke, J. M. (1995) Reconstitution of hairpin ribozyme activity following separation of functional domains, *J Biol Chem* 270, 29648-29651.
72. Hoogstraten, C. G., Legault, P., and Pardi, A. (1998) NMR solution structure of the lead-dependent ribozyme: evidence for dynamics in RNA catalysis, *J Mol Biol* 284, 337-350.
73. Butcher, S. E., Heckman, J. E., and Burke, J. M. (1995) Reconstitution of hairpin ribozyme activity following separation of functional domains, *J Biol Chem* 270, 29648-29651.
74. 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.
75. Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis, *Science* 313, 1638-1642.
76. Beach, H., Cole, R., Gill, M. L., and Loria, J. P. (2005) Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state, *J Am Chem Soc* 127, 9167-9176.
77. Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., and Mackerell, A. D., Jr. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles, *J Chem Theory Comput* 8, 3257-3273.
78. Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., and MacKerell, A. D. (2010) CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields, *J Comput Chem* 31, 671-690.
79. Pastor, R. W., and MacKerell, A. D. (2011) Development of the CHARMM Force Field for Lipids, *J Phys Chem Lett* 2, 1526-1532.
80. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1996) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules (vol 117, pg 5179, 1995), *J Am Chem Soc* 118, 2309-2309.
81. Oostenbrink, C., Villa, A., Mark, A. E., and Van Gunsteren, W. F. (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6, *J Comput Chem* 25, 1656-1676.

82. Jorgensen, W. L., Maxwell, D. S., and TiradoRives, J. (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, *J Am Chem Soc* 118, 11225-11236.
83. Jorgensen, W. L., and Tiradorives, J. (1988) The Opls Potential Functions for Proteins - Energy Minimizations for Crystals of Cyclic-Peptides and Crambin, *J Am Chem Soc* 110, 1657-1666.
84. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) An All Atom Force-Field for Simulations of Proteins and Nucleic-Acids, *J Comput Chem* 7, 230-252.
85. Karplus, M., and Petsko, G. A. (1990) Molecular-Dynamics Simulations in Biology, *Nature* 347, 631-639.
86. Mackerell, A. D. (2004) Empirical force fields for biological macromolecules: Overview and issues, *J Comput Chem* 25, 1584-1604.
87. Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The Biomolecular Simulation Program, *J Comput Chem* 30, 1545-1614.
88. Rhodes, M. M., Reblova, K., Sponer, J., and Walter, N. G. (2006) Trapped water molecules are essential to structural dynamics and function of a ribozyme, *P Natl Acad Sci USA* 103, 13380-13385.
89. McDowell, S. E., Spackova, N., Sponer, J., and Walter, N. G. (2007) Molecular dynamics simulations of RNA: An in silico single molecule approach, *Biopolymers* 85, 169-184.
90. Ditzler, M. A., Sponer, J., and Walter, N. G. (2009) Molecular dynamics suggest multifunctionality of an adenine imino group in acid-base catalysis of the hairpin ribozyme, *RNA-A Publication of the RNA Society* 15, 560-575.
91. Ditzler, M. A., Otyepka, M., Sponer, J., and Walter, N. G. (2010) Molecular Dynamics and Quantum Mechanics of RNA: Conformational and Chemical Change We Can Believe In, *Accounts Chem Res* 43, 40-47.
92. Mlynsky, V., Banas, P., Hollas, D., Reblova, K., Walter, N. G., Sponer, J., and Otyepka, M. (2010) Extensive Molecular Dynamics Simulations Showing That Canonical G8 and Protonated A38H(+) Forms Are Most Consistent with Crystal Structures of Hairpin Ribozyme, *J Phys Chem B* 114, 6642-6652.

93. Mlynsky, V., Banas, P., Walter, N. G., Sponer, J., and Otyepka, M. (2011) QM/MM Studies of Hairpin Ribozyme Self-Cleavage Suggest the Feasibility of Multiple Competing Reaction Mechanisms, *J Phys Chem B* 115, 13911-13924.
94. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., and Phillips, D. C. (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis, *Nature* 181, 662-666.
95. Rould, M. A., Perona, J. J., Soll, D., and Steitz, T. A. (1989) Structure of E. coli glutamyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 Å resolution, *Science* 246, 1135-1142.
96. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) Three-dimensional structure of a hammerhead ribozyme, *Nature* 372, 68-74.
97. Scott, W. G., Finch, J. T., and Klug, A. (1995) The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage, *Cell* 81, 991-1002.
98. Wang, S., Karbstein, K., Peracchi, A., Beigelman, L., and Herschlag, D. (1999) Identification of the hammerhead ribozyme metal ion binding site responsible for rescue of the deleterious effect of a cleavage site phosphorothioate, *Biochemistry* 38, 14363-14378.
99. Allerhan, A., Doddrell, D., Glushko, U., Cochran, D. W., Wenkert, E., Lawson, P. J., and Gurd, F. R. N. (1971) Carbon-13 Fourier Transform Nuclear Magnetic Resonance .3. Conformation and Segmental Motion of Native and Denatured Ribonuclease-A in Solution - Application of Natural-Abundance Carbon-13 Partially Relaxed Fourier Transform Nuclear Magnetic Resonance, *J Am Chem Soc* 93, 544-546.
100. Glushko, V., Lawson, P. J., and Gurd, F. R. N. (1972) Conformational States of Bovine Pancreatic Ribonuclease a Observed by Normal and Partially Relaxed Carbon 13 Nuclear Magnetic-Resonance, *Journal of Biological Chemistry* 247, 3176-3185.
101. Kojima, C., Ono, A., Kainosho, M., and James, T. L. (1998) DNA duplex dynamics: NMR relaxation studies of a decamer with uniformly ¹³C-labeled purine nucleotides, *J Magn Reson* 135, 310-333.
102. Brutscher, B., Boisbouvier, J., Pardi, A., Marion, D., and Simorre, J. P. (1998) Improved sensitivity and resolution in ¹H-¹³C NMR experiments of RNA, *J Am Chem Soc* 120, 11845-11851.
103. King, G. C., Harper, J. W., and Xi, Z. (1995) Isotope labeling for ¹³C relaxation measurements on RNA, *Methods Enzymol* 261, 436-450.

104. Lipari, G., and Szabo, A. (1981) Nuclear magnetic resonance relaxation in nucleic acid fragments: models for internal motion, *Biochemistry* 20, 6250-6256.
105. 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, *J Am Chem Soc* 104, 4546-4559.
106. Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .2. Analysis of Experimental Results, *J Am Chem Soc* 104, 4559-4570.
107. Yamazaki, T., Muhandiram, R., and Kay, L. E. (1994) Nmr Experiments for the Measurement of Carbon Relaxation Properties in Highly Enriched, Uniformly C-13,N-15-Labeled Proteins - Application to C-13(Alpha) Carbons, *J Am Chem Soc* 116, 8266-8278.
108. Palmer, A. G., Williams, J., and McDermott, A. (1996) Nuclear magnetic resonance studies of biopolymer dynamics, *J Phys Chem* 100, 13293-13310.
109. Mulder, F. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Measurement of slow (micros-ms) time scale dynamics in protein side chains by (15)N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme, *J Am Chem Soc* 123, 967-975.
110. Traaseth, N. J., and Veglia, G. (2010) Probing excited states and activation energy for the integral membrane protein phospholamban by NMR CPMG relaxation dispersion experiments, *Biochim Biophys Acta* 1798, 77-81.
111. Loria, J. P., Berlow, R. B., and Watt, E. D. (2008) Characterization of enzyme motions by solution NMR relaxation dispersion, *Accounts Chem Res* 41, 214-221.
112. Johnson, J. E., Jr., and Hoogstraten, C. G. (2008) Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using 13C NMR spin relaxation and specific isotope labeling, *J Am Chem Soc* 130, 16757-16769.
113. Palmer, A. G., 3rd, Kroenke, C. D., and Loria, J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules, *Methods Enzymol* 339, 204-238.
114. Shajani, Z., and Varani, G. (2005) 13C NMR relaxation studies of RNA base and ribose nuclei reveal a complex pattern of motions in the RNA binding site for human U1A protein, *J Mol Biol* 349, 699-715.
115. Hoogstraten, C. G., Wank, J. R., and Pardi, A. (2000) Active site dynamics in the lead-dependent ribozyme, *Biochemistry* 39, 9951-9958.

116. Blad, H., Reiter, N. J., Abildgaard, F., Markley, J. L., and Butcher, S. E. (2005) Dynamics and metal ion binding in the U6 RNA intramolecular stem-loop as analyzed by NMR, *J Mol Biol* 353, 540-555.
117. Julien, K. R., Sumita, M., Chen, P. H., Laird-Offringa, I. A., and Hoogstraten, C. G. (2008) Conformationally restricted nucleotides as a probe of structure-function relationships in RNA, *RNA* 14, 1632-1643.
118. Christiansen, J. K., Lobedanz, S., Arar, K., Wengel, J., and Vester, B. (2007) LNA nucleotides improve cleavage efficiency of singular and binary hammerhead ribozymes, *Bioorg Med Chem* 15, 6135-6143.
119. Donini, S., Clerici, M., Wengel, J., Vester, B., and Peracchi, A. (2007) The advantages of being locked. Assessing the cleavage of short and long RNAs by locked nucleic acid-containing 8-17 deoxyribozymes, *J Biol Chem* 282, 35510-35518.
120. Ke, A., Zhou, K., Ding, F., Cate, J. H., and Doudna, J. A. (2004) A conformational switch controls hepatitis delta virus ribozyme catalysis, *Nature* 429, 201-205.

CHAPTER 2

PROBING ALTERNATE BASE-PAIR REARRANGEMENTS AND CONFORMATIONAL SAMPLING IN A RIBOZYME ACTIVE-SITE INTERNAL LOOP USING MOLECULAR DYNAMICS SIMULATIONS

ABSTRACT

Dynamic fluctuations in RNA structure steer changes that correlate with a broad spectrum of functions. These conformational fluctuations yield various populations based on the free energy and the transition rates between various conformers along the free energy landscape. Conformational redistribution in non-coding RNA is driven by cellular signals that affect the rugged RNA free energy landscape. To assess conformational sampling in the pre-docking form of the hairpin ribozyme, we have run several explicit solvent Molecular Dynamics (MD) simulations of the hairpin ribozyme loop A domain totaling 2.4 μ s. We observed one dominant conformer and other minor states identified using hydrogen bonding and base stacking interactions in the active site loop region of loop A. Targeted Molecular Dynamics (TMD) was used to model inter-conformer transitions. Using the Markov State Model (MSM) analysis we constructed the kinetic pathways within each conformer and the transition network among a set of conformers. The Molecular Mechanics Poisson-Boltzmann Surface Area (MMPB/SA) approximation was applied to predict conformational energies and energy barriers between conformations. The barrier height between two of the observed conformations (named AA/UA and AA/CA) is accessible based on modest energies reported for this transition. However, transitions to a third conformer (GA/UA) appear to be energetically inaccessible. This study indicates that RNA undergoes structural versatility through several kinetic states and transitions, a phenomenon that underscores the formation of a rugged tertiary energy landscape with multiple accessible states for RNA, consistent with a role for conformational sampling in the pre-organization of loop A RNA prior to RNA-RNA interactions.

2.1 INTRODUCTION

The critical role of conformational dynamics in macromolecular function, allostery, catalysis, and ligand recognition is well appreciated (1-4). The dynamics-function relationship provides the missing link in more traditional structure-function paradigms where the motional variables have been frozen out using mutation or modification (5, 6). While the structure-dynamics-function relationship has been explored extensively in proteins, less is known about how RNA structure and dynamics relates to its biological roles (7, 8). In the context of ribozymes, a detailed understanding of the how RNA structure and dynamics limit catalytic rates involves dynamic aspects of RNA-RNA interaction processes. Two frameworks are often discussed to describe the formation of macromolecular interactions: induced fit and conformational capture. In the induced fit mechanism, binding of the most populated state is followed by conformational change within the bound state. In contrast, the conformational capture mechanism invokes ground-state fluctuations of the binding partners that create a state resembling the conformation of the bound state. Binding then occurs when the partner(s) visit the respective binding conformations. When the energy landscape of the binding partners is complex because one or both partners in isolation sample multiple conformations, the presence of a defined state or measurable population corresponding to a conformation resembling the bound state would be taken as evidence for conformational capture.

Molecular Dynamics (MD) simulation and Nuclear Magnetic Resonance (NMR) have showcased their power in assessing macromolecular motions as well as conformational sampling (9, 10) and are thus useful for interrogating timescales relevant for the conformational capture process. In the case of RNA, the energy landscape is typically more rugged than for proteins (11-13). This means that multiple alternate states, separated by kinetic barriers, may be populated to

a significant extent at physiological temperatures. This has hindered past efforts to fully characterize the conformational dynamics of RNA and in particular to explore to what extent conformational capture or induced fit processes are at play in RNA-RNA interactions.

Here, we have used the hairpin ribozyme as a model to study RNA dynamics in ensemble formation and inter-conformer transitions. The hairpin ribozyme catalyzes reversible, site-specific phosphodiester bond cleavage to yield 5'-OH and a 2'- 3'-cyclic phosphate terminus (14, 15). The ribozyme consists of two RNA loops A and B that interact (dock) to form a multi-domain structure that acts to cleave the phosphodiester bond within loop A (**Figure 2-1**). Loops A and B and their flanking helices constitute the catalytic core of the hairpin ribozyme derived from tobacco ringspot virus satellite RNA (16). This catalytic core reveals a network of stacking and hydrogen-bonding interactions within the active site that orient the reactive phosphate in the in-line orientation for an S_N2 -type nucleophilic attack mechanism. In addition, nucleotide base functional groups are oriented near the reactive phosphate to facilitate catalytic chemistry (17, 18). The catalytic function of the hairpin ribozyme represents a system that can serve as a potential therapeutic agent. Previously, the hairpin ribozyme has been engineered to specifically cleave HIV-1 RNA (19, 20). NMR structures of the isolated loop A and B have previously been determined as have been the crystal structures of the numerous docked forms (17, 21-28). Significant structural rearrangements of both loops are observed upon formation of the docking complex (17, 21, 22, 29-31). For example, U+2 in the free form of loop A is extrahelical, while it forms a non-canonical G8-U+2 base-pair in the docked form. Similarly, U41 in the free loop B is extrahelical, but it forms the A22-U41 canonical base-pair in the docked form (**Figure 2-2**). A critical residue G+1 in loop A interacts with C25 in loop B to form a G+1-C25 inter-domain

base-pair by interrupting the G+1-A9 and C25-U37 non-canonical base-pairs in the individual loops A and B, respectively.

Recently, loop A and B interaction kinetic rates have been reported by the Hoogstraten group (32). The reported docking association rate reveals a very slow but tight docking process between loop A and loop B (on-rate five orders of magnitude below the diffusion limit). This slow association rate as well as structural versatility between the free and docked forms, especially for the key residues involved in the interaction, has led us to the hypothesis that isolated loops A and B may both sample the docking-competent states in a possibly double conformational capture docking mechanism. The major loop rearrangements observed upon docking, as well as the kinetically unfavorable docking process, both argue for a tight and mechanistic role for conformational dynamics in the pre-organization of the catalytically-competent active site of the hairpin ribozyme. An alternative induced fit mechanism would be expected to be kinetically accelerated. Here, we focus on probing the conformational sampling of just the loop A domain of the hairpin ribozyme as an initial step to understanding the active site dynamics in this RNA system. As the main tool, we have used MD simulations to map out the conformational ensemble visited by loop A and we describe here the energetics and the pathway for transitions between alternate conformers identified from the simulations via transitions through kinetic states.

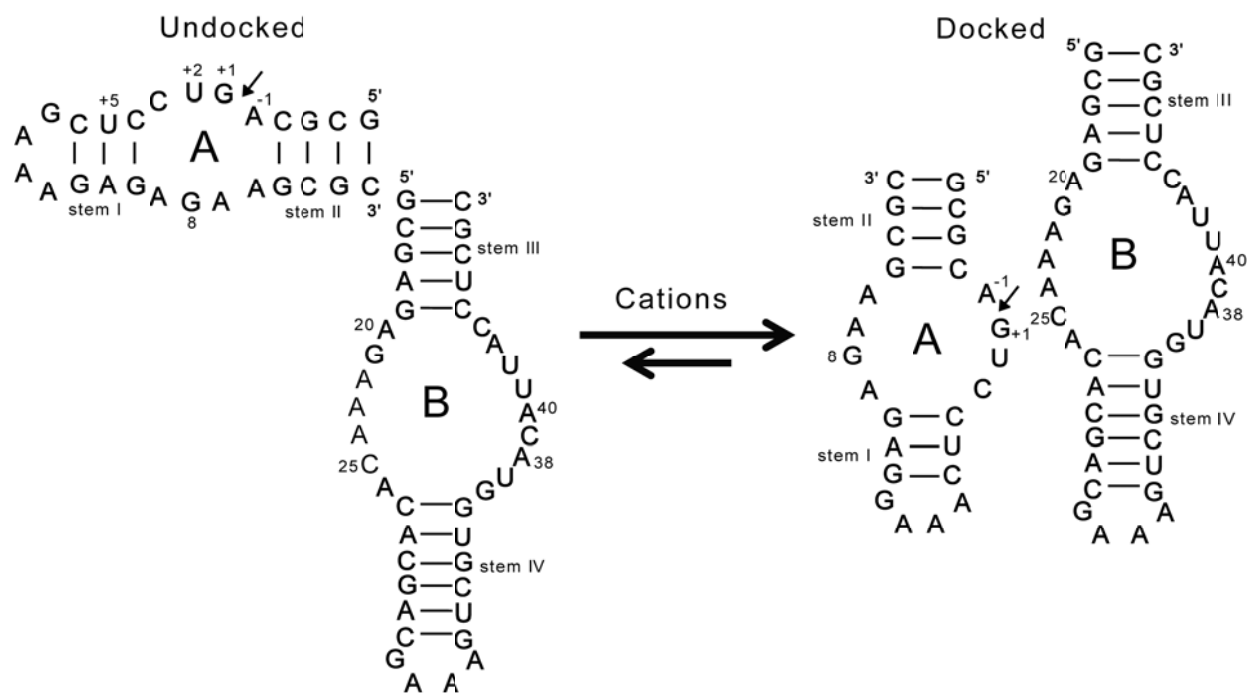


Figure 2-1: Trans-acting hairpin ribozyme secondary structure representation of metal-dependent loop-loop docking. A and B denote loops A and B respectively. Docking of loops A and B precedes cleavage between A-1 and G+1 residues in the loop A region. Cleavage and ligation site are shown by small arrows in loop A.

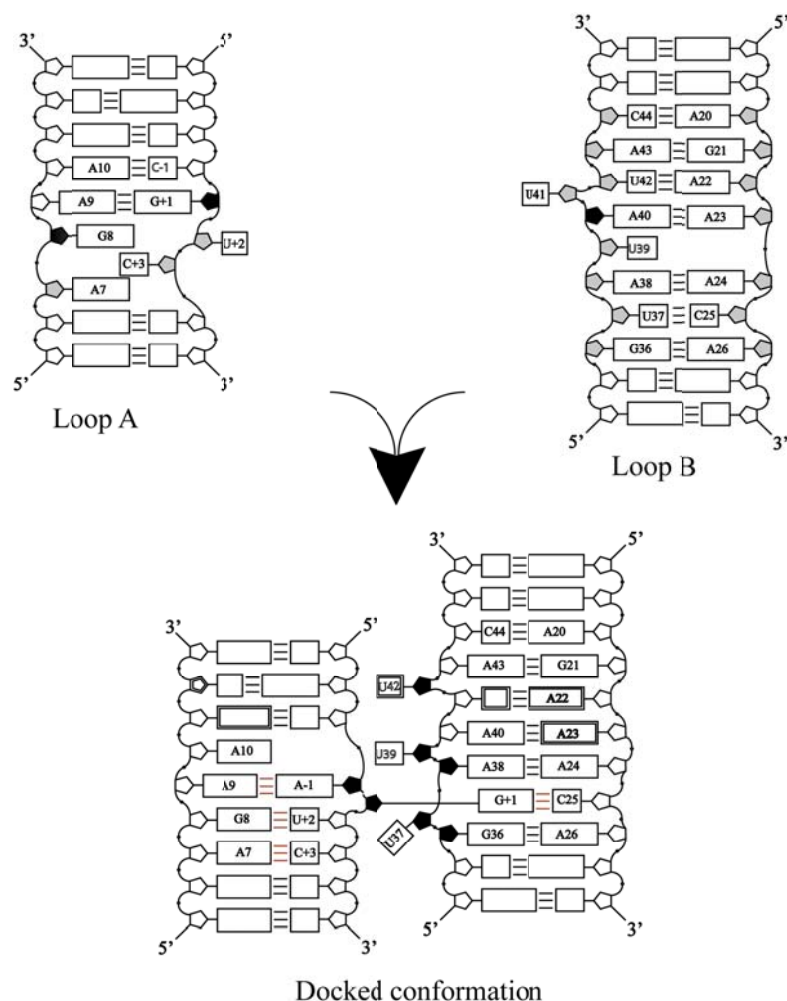


Figure 2-2: Schematic representation of free loops A and B as well as docked loop A and B based on the NMR structures (free loop A and B) and crystal structure (docked loop A and B). Pentagons represent ribose sugar, while boxes represent bases. Ribose puckers are indicated as follows: C3'-endo (open), C2'-endo (solid), or equilibrium (grey). The double-boxed residues form structured binding pocket for U42. Red lines represent new set of base-pairs in docked structure.

2.2 MATERIALS AND METHODS

2.2.1 Molecular dynamics (MD) simulations using NAMD and the CHARMM force field

Explicit solvent Molecular Dynamics (MD) simulations were carried out initially with loop A coordinates kindly provided by Ignacio Tinoco Jr. based on the loop A structure determined by NMR (22) to validate the force field. In additional simulations, stem 1 of Tinoco's loop A (denoted as LpA) was modeled to incorporate a GAAA tetraloop from an RNA structure in complex with theophylline (PDB code 1O15) (33). Residue C-1 was mutated to A-1 using the MMTSB Tool Set (34) to restore the binding sequence of loop A to its native sequence. One G-C base pair was added at the end of stem II to enhance its stability during the simulation time. This construct is denoted hpA. The hpA 2'-OH of the A-1 residue was also methylated in additional simulations to explore experimental constructs. The 2'-O-methylated hpA is denoted hpAome. Experimental data from our group (unpublished) had previously shown significantly lower docking affinities for 2'-O-methylated loop A (hpAome) with loop B compared to non-methylated loop A (hpA). Constructs are illustrated in Figure 3 of the Results section.

The loop A RNA structures were solvated in a cubical box with TIP3P water molecules (35, 36) using a minimal distance of 10 Å from any RNA atom to the edge of the simulation box. The systems were neutralized with 18 Na⁺ ions for Tinoco's loop A (LpA) and 26 Na⁺ counterions for both hpA and hpAome constructs. The cubical box dimensions were (56.8 Å)³, (62.3 Å)³, and (63.4 Å)³ for Tinoco's loop A (LpA), hpA, and hpAome, respectively. Periodic boundaries were applied in all simulations. The particle-mesh Ewald (PME) method was used to model the long-range electrostatics interactions (37, 38). The solute-solvent systems were minimized with an initial harmonic restraint potential applied to heavy atoms with a force constant of 5 kcal/(mol Å²). The same restraints were also applied during the initial heating phases. Finally, the systems

were allowed to relax in several equilibration runs with a decreasing harmonic potential applied to the heavy atoms at 10 K temperature increments until a temperature of 200 K was attained. Additional equilibration without restraints was then completed to allow for full relaxation of the system while slowly heating to 298 K over 100 ps. All bonds involving hydrogens were constrained holonomically by the SETTLE algorithm, which allows us to use a time-step of 2 fs. The CHARMM36 force field was used (39-43) in NAMD (44) to run Tinoco's loop A (LpA) construct for 100 ns. CHARMM36 was also used to simulate the loop A constructs, hpA and hpAome. Six independent simulations were run for the hpA and hpAome systems using different random seeds. Individual simulations are labeled hpA1-6 and hpAome1-6 respectively. All simulations are summarized in **Table 2-1**.

Table 2-1: Summary of molecular dynamics simulations of loop A constructs

Simulation	LpA ^a	LpA ^a	hpA ^b	hpAome ^c
Length of simulation (ns)	100	100	200	200
Force field	CHARMM36	Amber ff10	CHARMM36	CHARMM36
Number of simulation	1	1	6	6

^aLpA denotes explicit solvent simulation on Tinoco's loop A (**Figure 3A**) using CHARMM36 and Amber ff10 force fields respectively.

^bhpA represents simulation of our loop A construct (**Figure 3B**) without any methylation. The simulation replicates (six) were designated hpA1, hpA2, hpA3, hpA4, hpA5 and hpA6.

^chpAome is our simulation construct with 2'-O-methylation at residue A-1. This simulation has also six replicates named hpAome1, hpAome2, hpAome3, hpAome4, hpAome5, hpAome6.

2.2.2 MD simulations using sander and the Amber force field

To compare force field performance, we also carried out simulations with the Amber force field. Again, the starting structure was obtained from Tinoco's loop A NMR structure (LpA). The RNA was neutralized with 18 Na^+ counter ions using xleap at positions of high negative electric potential and solvated in a rectangular water box measuring 66.4 Å x 69.4 Å x 80.4 Å with a minimum 10 Å thick layer of TIP3P water molecules around the loop A RNA. Long distance electrostatic interactions were accounted for again by using the particle-mesh Ewald summation (38). Minimization of the solute-solvent system was done in two steps. First, the RNA was restrained with a harmonic potential force of 500 kcal/(mol Å²) in 500 steps of steep descend minimization followed by 500 steps of conjugate gradient minimization. Second, the whole system (RNA, solvent and ions) was minimized together without restraints during 2,500 steps at a constant volume with periodic boundaries. After the dual-step minimization, the system was gradually equilibrated by heating from 0 K to 300 K over 100 ps using the Langevin thermostat (45) with 10 kcal/(mol Å²) positional restraint using NVT followed by a 100 ps equilibration at 300 K under NPT conditions without restraints. The SHAKE algorithm (46) was used to constrain covalent hydrogen bonds during the simulation and allow a 2 fs time step. The sander module of Amber 11 with the amberff10 force field was used for this simulation (47, 48) followed by a production phase of 100 ns.

2.2.3 Targeted molecular dynamics (TMD) simulation

Targeted Molecular Dynamics (TMD) (49) simulations were applied to sample the transition between two sets of conformers identified in the unbiased simulations. In particular, TMD was used to drive one of the observed simulation structures (denoted AA/UA, see results)

to two different structures (denoted as GA/UA and AA/CA) in separate TMD simulations. In TMD, the following biasing potential was applied to drive one conformation to the other:

$$U_{TMD} = \frac{1}{2} \frac{k}{N} (RMSD(t) - RMSD_0(t))^2$$

where $RMSD(t)$ is the root mean square deviation of heavy atom coordinates with respect to the target coordinates, $RMSD_0(t)$ is the target RMSD value at a given time, k is the force constant and N denotes the number of atoms. Initial values of $RMSD(t)$ were set to 4.01 Å for the AA/UA-GA/UA TMD simulation and to 3.79 Å for the AA/UA-AA/CA TMD simulation. We initially tried several different force constants to optimize the transition and ability to reach a final RMSD difference of less than 1 Å between the starting and target structures. After several trial runs, we decided on a force constant of 300 kcal/(mol Å²) that was applied to all heavy atoms. TMD was carried out using explicit solvent with a periodic boundary box of 62 x 62 x 62 Å³ to solvate RNA neutralized by Na⁺ counterions with the same parameters as in the unbiased CHARMM simulations described above. After initial minimization and equilibration, TMD simulations were run at 298K for 2 ns. Both sets of TMD simulations were repeated ten times each time with a different random seed. The simulation trajectories were combined and clustered using a k-means clustering algorithm based on RMSD to obtain conformational transition pathways between the initial and target conformations. The observed conformations along the TMD pathway obtained via clustering were then subjected to Molecular Mechanics Poisson-Boltzmann Surface Area (MMPB/SA) analysis (50-59) to estimate the conformational free energies in each ensemble cluster along the TMD transition pathways.

2.2.4 Analysis

The analysis of the simulations was carried out using the MMTSB (Multiscale Modeling Tools for Structural Biology) Tool Set (34) in conjunction with CHARMM. Markov state model (60) analysis was also used to identify and model conformational states and the kinetic transitions between those states using the MSMBuilder 2.0 software package (61). RMSD values were calculated for all heavy atoms excluding the terminal bases due to fraying. Previously reported NMR nuclear Overhauser effect (NOE) distances were measured for average CHARMM36 and Amber ff10 structures using PyMOL (62) and VMD (63) visualization software.

2.3 RESULTS

2.3.1 Force field validation

Molecular dynamics simulations were carried out on the loop A domain of the hairpin ribozyme to study the conformational sampling of RNA, especially in the context of encounter complexes. Since stability of force fields is an ongoing problem in the RNA simulation field, , we initially carried out simulations of the loop A structure obtained from Cai and Tinoco Jr. (22) using the two recent force fields from the Amber and CHARMM families (Amber ff10 and CHARMM36) to evaluate which force field results in better agreement with the experimental data. **Figure 2-3A** shows the heavy atom root mean square deviation (RMSD) with respect to Tinoco's loop A starting structure during 100 ns of MD simulation with either the CHARMM36 or Amber ff10 force fields. Both simulations were stable as evidenced by RMSD values averaging about 3 Å from the NMR starting structure during the simulation. Furthermore, the agreement with NMR-derived Nuclear Overhauser Effect (NOE) distance constraints (22) was evaluated. The distances between specific protons that registered strong and/or medium NOEs

for the loop A average simulation structures were analyzed from the simulations. The results, summarized in **Table 2-2**, focus on five characteristic distances that are not present in regular A-form structures for which NOE constraints are available. As can be seen, in the simulation with the Amber force field, the average simulation structure violates all but one NOE distance constraint, while only two NOEs are missed with CHARMM36 (A9H1'-C+3H6 and A9H1'-C+3H1') and the A9H1'-C+3H6 distance in the CHARMM36 simulation is closer to the NOE range than with the Amber force field. We note, however, that in both of these cases, the reported solution structure has NOE distances at the upper limit of the NOE constraint range, indicating that these distance constraints may be problematic. Finally, we analyzed ribose sugar puckering and base geometries (shear, buckle etc.) and again found better agreement with the experimental data for the average structure generated by the CHARMM36 force field (**Table 2-3**) compared to the Amber ff10 average structure. Based on these initial simulation results, we chose the CHARMM36 force field for carrying out subsequent simulations of the wild-type loop A (**Figure 2-3B**) construct.

2.3.2 Conformational heterogeneity of wildtype loop A

The wild-type loop A construct was modeled based on the mutant loop A NMR structure but with C-1 mutated to A-1. While we were initially interested in comparing the methylated and non-methylated constructs to explain differences in experimental association rates, analysis of the simulation results does not suggest significant differences between the two forms in terms of the conformational sampling of loop A alone (**Figure 2-4**). Using the, statistical F-test, we obtained an F-value of 1.1164 between the two population variances which suggests that the conformational distribution between methylated and non-methylated loop A was similar in all our simulations.

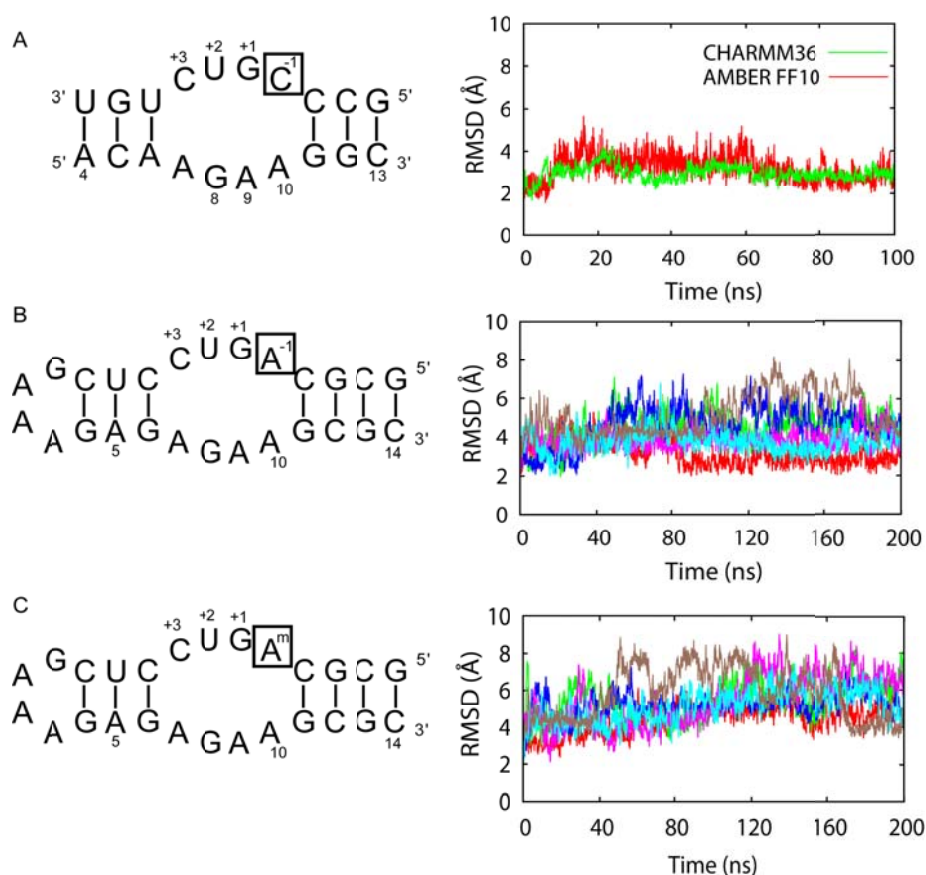


Figure 2-3: Secondary structure representation of loop A with their respective heavy atom RMSD profiles from Tinoco's loop A (A), non-methylated (hpA) loop A (B) and 2'-O-methylated (hpAome) loop A (C). The boxes highlight the major sequence variations and modifications of the wildtype A-1 residue. A^m represents 2'-O-methylated adenosine. In A, Tinoco's loop A was simulated using Amber ff10 and CHARMM36 force fields respectively. Red profile represents Amberff10 while green represents CHARMM36 simulations respectively. In B and C, the non-methylated and methylated loop A variants were simulated using CHARMM36 force field. The RMSD profiles in B and C represent independent simulation runs with different random seed for each variant. The profile colors represent different runs.

Table 2-2: Comparison of NOE distance constraints with simulation averages using Amber and CHARMM^a

Atom Pair	NOE Distance constraint (Å)	Tinoco's structure (Å)	CHARMM36 MD average (Å)	Amber ff10 MD average (Å)
U+2H1'- C+3H5	1.8-3.0	3.0	2.5	5.9
A9H1'- C+3H6	2.5-4.0	3.9	5.4	8.5
A9H1'- C+3H5	3.0-6.5	4.4	6.3	9.5
A9H1- C+3H1'	3.0-6.5	3.2	4.5	5.4
A9H4'- C+3H1'	1.8-4.0	4.0	7.6	7.7

^a NOE distances as reported in Tinoco's loop A structure (22) in comparison to distances in the average structures derived from simulation with CHARMM36 and Amber ff10 force fields respectively. Atom distances were obtained from 100 ns CHARMM36 and Amber ff10 simulation average structures.

Table 2-3: Representative backbone and base-pair parameters of average structures derived from CHARMM36 and Amberff10 simulations respectively

The numbers in bold indicate a significant conformational deviation of residue/base-pair with respect to Tinoco's structure.

Pseudorotation phase angle (degrees)				χ torsion angle (degrees)		
Residue	Tinoco's structure	CHARMM36 MD average	Amberff10 MD average	Tinoco's structure	CHARMM36 MD average	Amberff10 MD average
A4	7.2	14.8	11.5	-165.6	-161.9	-79.1
C5	6.5	13.8	19.3	-156.8	-156.8	-147.3
A6	6.7	9.0	11.9	-162.9	-152.7	-158.2
A7	59.5	10.1	14.9	-169.5	-153.6	-148.3
G8	181.3	173.7	17.3	-113.4	-107.8	-173.3
A9	19.9	8.2	12.7	-162.9	-163.8	175.6
A10	9.9	10.2	8.9	-171.6	-165.1	-169.7
G11	0.4	13.3	7.2	-175.6	-168.5	-154.7
G12	4.8	10.3	10.8	-170.4	-165.7	-166.9
C13	7.1	37.5	55.1	-161.1	-152.9	-147.5
G16	4.8	15.9	11.7	-179.2	-165.3	67.6
C17	2.6	11.8	17.5	-162.9	-160.3	-173.5
C18	16.3	14.4	17.6	-154.6	-160.9	-156.5
C19	5.2	13.4	15.1	-172.6	-156.7	-143.3
G20	186.2	175.7	169.4	-95.5	-116.8	-121.5
U21	16.8	152.6	173.3	-173.4	-109.4	-144.9
C22	40.1	11.9	24.7	-140.9	-158.8	-165.6
U23	15.42	22.5	25.4	-153.9	-152.3	-160.0
G24	2.07	13.6	11.9	-171.8	-149.2	-162.4
U25	8.8	ND	13.1	ND	ND	-154.8

Shear (degrees)				Buckle (degrees)		
Base-pair	Tinoco's structure	CHARMM36 MD average	Amberff10 MD average	Tinoco's structure	CHARMM36 MD average	Amberff10 MD average
A4-U25	0.2	-1.1	7.7	6.8	30.6	35.2
C5-G24	1.2	-0.4	-0.1	12.6	4.8	4.2
A6-U23	-0.4	-0.5	0.5	29.8	17.7	3.41
A7-C22	4.1	5.4	6.9	-14.1	-7.9	-11.4
G8-U21	-0.3	1.9	-4.8	163.8	38.8	-2.5
A9-G20	-4.9	-5.5	-2.6	-19.3	-19.2	-166.9
A10-C19	-2.9	-2.8	-2.6	-14.6	0.6	-167.1
G11-C18	-0.6	0.1	-2.1	-2.4	0.5	172.5
G12-C17	-0.5	-0.4	-2.1	1.9	-3.9	168.8
C13-G16	0.2	0.5	4.5	6.8	0.6	175.2

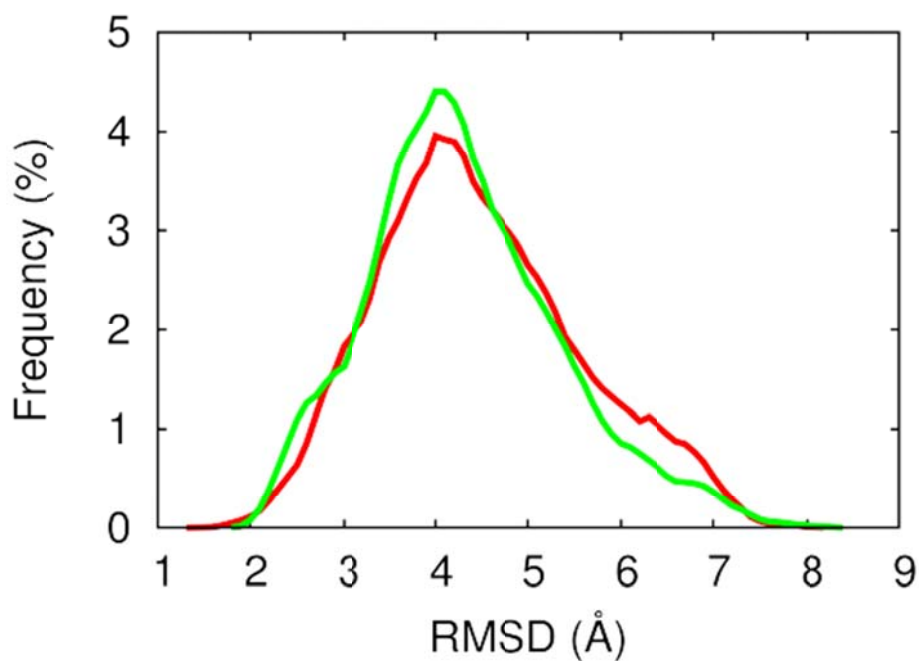


Figure 2-4: Population distribution between non-methylated loop A (green) and methylated loop A (red) based on RMSD from explicit solvent simulation of loop A. The distribution plot captures RMSD data between 40-200 ns of the 12 x 200 ns simulation (six simulations each). The statistical F-test between the two distributions is similar within 95% confidence level

Therefore, we combined the sampling from both sets of simulations to yield a total of 2.4 μ s explicit solvent MD simulations to describe the overall conformational ensemble of loop A. As can be seen in **Figure 2-3B**, there were significant RMSD variations in some of the replicas indicating conformational heterogeneity. Representative simulation structures were subsequently obtained via k-means clustering based on mutual RMSD. The resulting conformations have been shown in **Figure 2-5**. The top and bottom three base-pairs in stems I and II maintained similar Watson-Crick base-pairing, as expected, but there were significant variations in the loop region (A-1 to C+3 and A7 to A10) involving alternate base-pairing and extra-helical bases. At the bottom of stem I, A-1 and G+1 residues were in base-pair competition with A10. In cases where A-1 was base-paired to A10, G+1 was either base-paired with A9 or stacked between other residues. G8 was always unpaired or stacked between other residues. At the top of stem II, U+2 and C+3 were in base-pair competition with A7. Meanwhile the other unpaired bases in the loop were either stacked or flipped out of the helix.

Further inspection suggested grouping into three major populations, namely AA/CA, AA/UA and GA/UA reflecting the combination of two major conformations each in stem I (AA and GA) and stem II (UA and CA). AA and GA were distinguished by A-1:A10 (AA) vs. G+1:A10 (GA) base-pairing, whereas CA and UA were distinguished by C+3:A7 (CA) vs. U+2:A7 (UA) base-pairs. The schematic representations in **Figure 2-6** highlight key variations in structural features of these populations. AA/CA (46%) was the dominant conformer while GA/UA (22%), AA/UA (24%) and GA/CA (4%) were minor conformers (**Figure 2-6**). We further characterized additional variations of these conformers within the conformers outlined above. Using Markov state model analysis (60, 61) we developed a network of kinetic states (macrostates) connected by transitions between and within other conformers (see **Figure 2-7**). Six kinetically different

macrostates (a-f) were determined from the AA/UA ensemble simulation trajectory. The relative populations and energies of the kinetic states (macrostates) with respect to their primary conformations are detailed in **Table 2-4**. The dominant AA/UA macrostate (a, 77%) had a flipped out C+3 and G+1:A9 base-pair which undergoes kinetic transitions involving various orientations of C+3 flipping, tilting of U+2:A7 base-pair, movement of C+3 into the helix and breaking of G+1:A9 base-pair towards the formation of kinetically different macrostates. Within the AA/CA conformer, seven macrostates were identified. The dominant kinetic state within AA/CA (53%) featured flipped out U+2 and A9 residues. Other macrostates had A9 either flipped out or base-paired to G+1, different orientations of U+2 flipping, stacked U+2 and C+3 flipping. Various C+3 flipping orientations and base stacking transitions were observed within the GA/UA conformer in the observed macrostates. Finally, the dominant GA/UA macrostate (64%) exhibited a flipped out C+3. A unique G+1:A9 base-pair was observed in one macrostate albeit in low (1%) population. Structural transitions among kinetic states (macrostates) and conformers were facilitated by dynamics of various key residues in the active site of loop A. Most notably U+2, C+3 and A9 adopted different conformations inside or outside the helix involving base-pairing, base-flipping and stacking.

The kinetic transitions between the various conformations were facilitated by torsional dynamics involving particular transitions in the backbone ζ torsion angle accompanied by transitions in the sugar pucker phase angles of U+2, C+3 and A9. Base-flipping seems to be correlated with a C2'/C3'-endo mixed sugar pucker transition in these dynamic residues. Quantitative correlation analysis between base and sugar pucker dynamics in selected loop A residues has been shown (**Table 2-3**). Residues A9, U+2 and C+3 reported a 1:1 C2'/C3'-endo mixed sugar pucker to base-flipping ratio while G+1 reported a 2:1 ratio respectively.

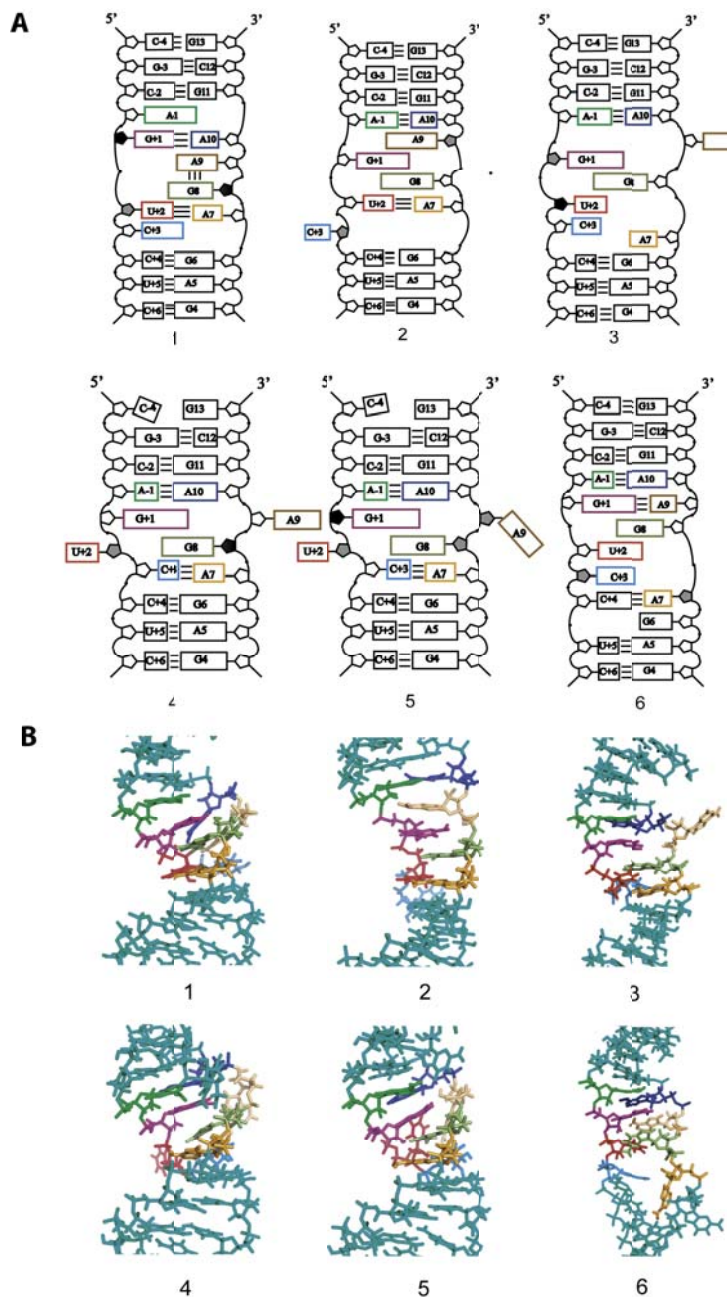


Figure 2-5: Schematic representations (A) and structural renderings (B) of loop A simulation structures. The numbering in A and B correlate schematic structural representations with 3D renderings. The structures shown are representative structures for the six independent loop A simulations. The schematic representation adopts the nomenclature used in **Figure 2-2**.

Table 2-4: Characterization of the kinetic states observed during loop A simulation

The kinetic transitions have been highlighted with AA/UA, GA/UA and AA/CA conformers using Markov state model analysis. The structures used in calculating RMSD are representative structures within each macrostate (a-u) population.

Conformer	Macrostate	% Macrostate population within conformer	RMSD (Å) relative to highest population	$\Delta\Delta G$ (kcal/mol)
AA/UA	a	77.2	0.0	0.0
	b	4.6	3.2	1.7
	c	12.7	3.2	1.1
	d	2.3	3.2	2.1
	e	2.0	4.0	2.2
	f	1.0	2.8	2.5
GA/UA	g	64.3	0.0	0.0
	h	3.5	1.4	1.7
	i	3.8	1.9	1.7
	j	17.9	1.6	0.8
	k	0.5	2.8	2.9
	l	7.7	1.9	1.3
	m	1.3	2.1	2.3
	n	1.0	3.0	2.5
AA/CA	o	52.9	0.0	0.0
	p	1.4	1.9	2.2
	q	40.7	0.7	0.2
	r	2.8	0.9	1.7
	s	1.0	2.2	2.6
	t	1.0	2.6	2.3
	u	0.5	2.6	2.8

Table 2-5: Residue dynamics for selected loop residues during hpA simulations

Residues	% C2'/C3'-endo sugar pucker	% Base flipped out
A7	16.7	0.0
G8	16.7	0.0
A9	33.3	40.0
G+1	16.7	8.33
U+2	50.0	40.0
C+3	33.3	30.0

^aSugar pucker and base flipped-out percentages for each residue were calculated as a fraction of observed mixed (C2'/C3') pucker and complete flip-out with respect to all the six hpA unbiased simulations. A cutoff percentage $\geq 25\%$ was considered as evidence of dynamics within the residue.

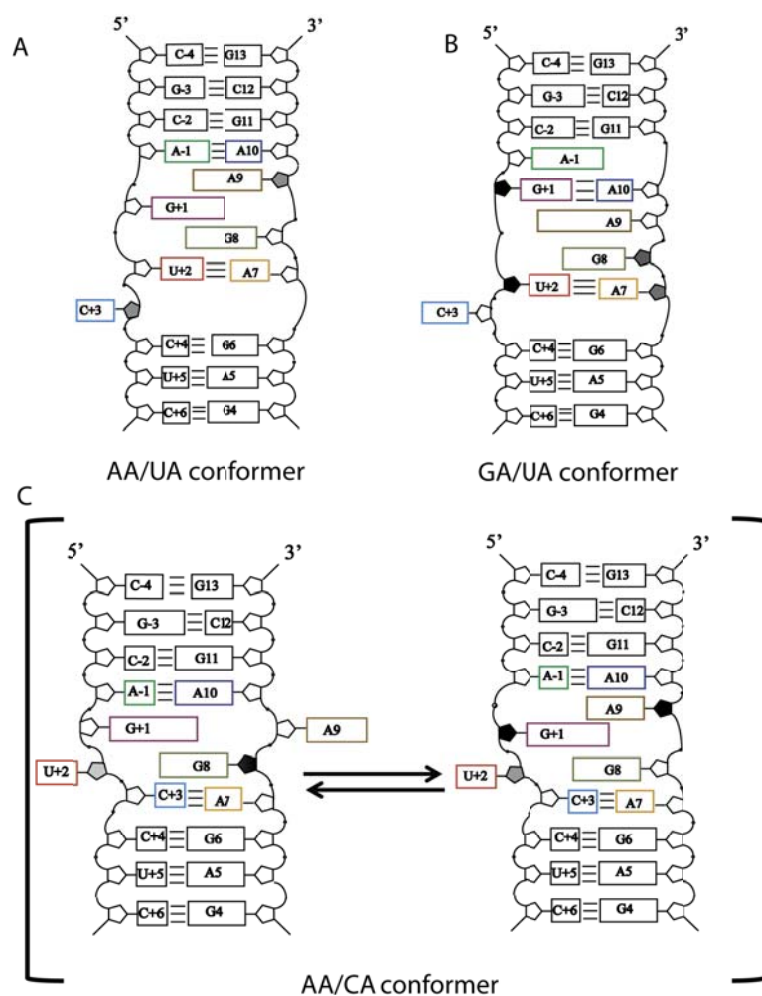


Figure 2-6: Schematic representation of loop A conformational classification based on loop base-pairing and stacking. These structures were obtained after analysis of all simulations by k-means clustering based on heavy-atom RMSD. A and B represent observed conformations AA/UA and GA/UA while C represents the major conformation which depicts equilibrium between flipped-out A9 base and helically-stacked A9 in the major conformer AA/CA.

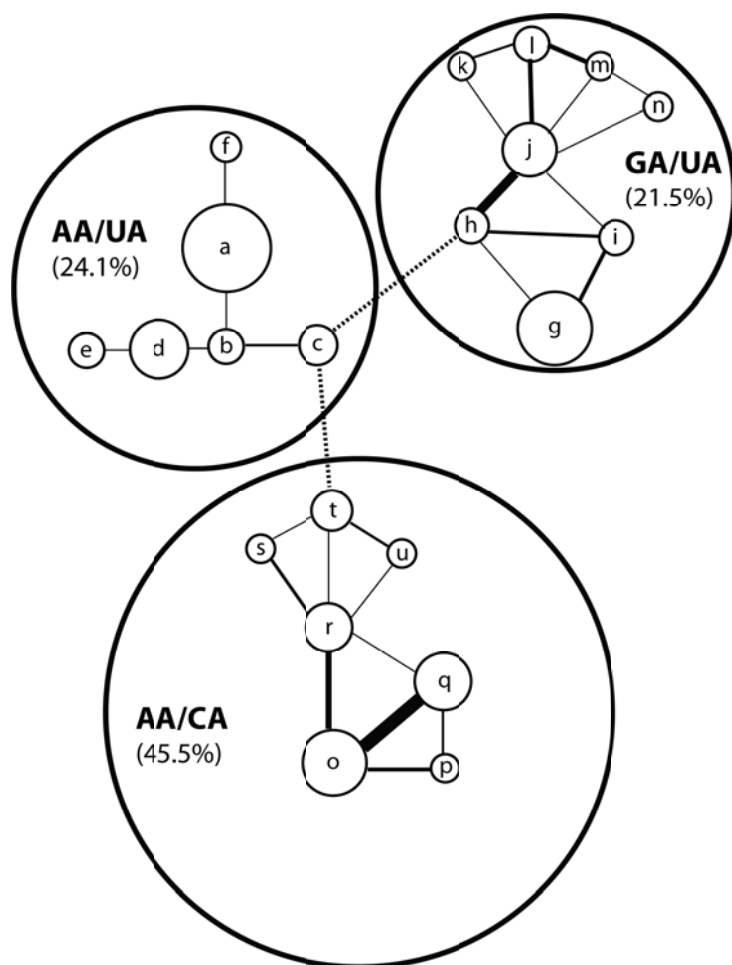


Figure 2-7: Conformational transition of loop A RNA capturing top 18 macrostates (labeled a-u) derived from Markov state modeling. The large circles represent the observed major and minor conformers AA/CA, AA/UA and GA/UA respectively, while inset circles represent kinetic macrostates sized by relative population within each conformer. The flux between individual macrostates is represented by solid line, with the intensity of each solid line depicting the transition frequency. The dotted line represents transition points between conformational pairs.

The base-flipping with mixed pucker correlation has been observed in DNA in the methylation of target cytosine by the bacterial DNA cytosine methyltransferase M.Hha (64). A9 adopts a broad range of conformations including base-pairing to G+1 to intrahelical base-stacking whereas C+3 samples various extrahelical orientations and U+2 assumes various extrahelical orientations. C+3 is mostly base-paired but can transition to extrahelical base-flipping with the base adopting various extrahelical orientations. (**Figure 2-8, Table 2-4**).

2.3.3 Conformational transitions between major states

The MD simulations, when started from different random seeds, reached different conformational states as detailed above and illustrated in the Markov state model (**Figure 2-7**). While minor transitions were observed in some of the simulations such as G+1:A10 to G+1:A9 in stem I, transitions between the major states were not observed, presumably because of significant kinetic barriers. In order to sample the major transitions we performed targeted MD simulation (TMD) from AA/UA to GA/UA and AA/CA conformers respectively (**Figure 2-8**) in two separate TMD simulations. In one TMD simulation (AA/UA to GA/UA), there was a transition of the A-1:A10 base-pair to G+1:A10 due to the reorientation of A10 residue. This rearrangement led to transitional stacking of A9 between G+1 and G10. A9 moved out of position giving way for G+1 to base-pair with A10. The G+1:A10 base-pair formed is subsequently stabilized in the last 0.2 ns of the simulation (**Figure 2-8**). Similarly, in another TMD simulation transition (AA/UA to AA/CA), there was a loss of U+2:A7 base-pair due to extra-helical movement of U+2. Meanwhile, the previously extrahelical C+3 moved into the helix to form a non-canonical base-pair with A7 that was previously base-paired to U+2. These U+2 and C+3 trans-directional dynamics propagated the formation of a new non-canonical C+3:A7 base pair observed in the AA/CA conformer (**Figure 2-8**). This conformational

transition was completed within 60% of the TMD simulation with the new C+3:A7 base-pair formed and stabilized in the remaining 40% of our simulations. Similarly to the kinetic transitions, A9, U+2 and C+3 residues displayed dynamic behavior that drove the conformational transitions and subsequently conformational sampling in loop A.

We calculated the energies for each conformational ensemble and the barrier height between the two sets of conformational transitions in TMD simulations via Molecular Mechanics Poisson-Boltzmann Surface Area (MMPB/SA). The calculated conformational energies, for AA/UA and AA/CA were -2007 and -2003 kcal/mol, respectively (see **Table 2-6**) with the difference within the 2-4 kcal/mol error of the MMPB/SA method. The U+2:A10 to C+3:A10 base-pair transition was estimated to have an energy barrier of ~13 kcal/mol (**Figure 2-8A**). The GA/UA conformer, however, reported a higher energy (-1990 kcal/mol) compared to AA/UA and AA/CA conformers consistent with the low relative population (21.5%) observed during explicit solvent simulation. The A-1:A10 to G+1:A10 base-pair barrier height was estimated as ~30 kcal/mol (**Figure 2-8B**) indicating that this transition is limited by a large kinetic barrier between AA/UA and GA/UA. This high energy barrier for the formation of GA/UA conformer supports the dominance of A-1:A10 over G+1:A10 base-pair in stem I effectively favoring the lower energy AA/UA and AA/CA conformations. The stem II base-pair transition between U+2:A7 to C+3:A7 base-pair was readily achieved by U+2 and C+3 dynamics, especially by their ability to flip in and out of the helix to facilitate base-pair formation. These dynamics are likely responsible for the lower energetic barrier between the AA/UA and AA/CA conformations.

Table 2-6: Conformational energies of loop A conformers calculated with MMPB/SA

Loop A conformer	Signature base pair(s)	Energy (kcal/mol)
AA/UA	A-1:A10 and U+2:A7	-2007.1 $\pm$ 0.7
GA/UA	G+1:A10 and U+2:A7	-1990.3 $\pm$ 0.6
AA/CA	A-1:A10 and C+3:A7	-2002.6 $\pm$ 0.6

^aConformational energy values for loop A conformers calculated by the Molecular Mechanics Poisson-Boltzmann Surface Area (MMPB/SA) for AA/UA, GA/UA and AA/CA conformers. The reported energies are averages for the cluster ensembles representing each conformer. The statistical error was estimated based on the standard error between block averages.

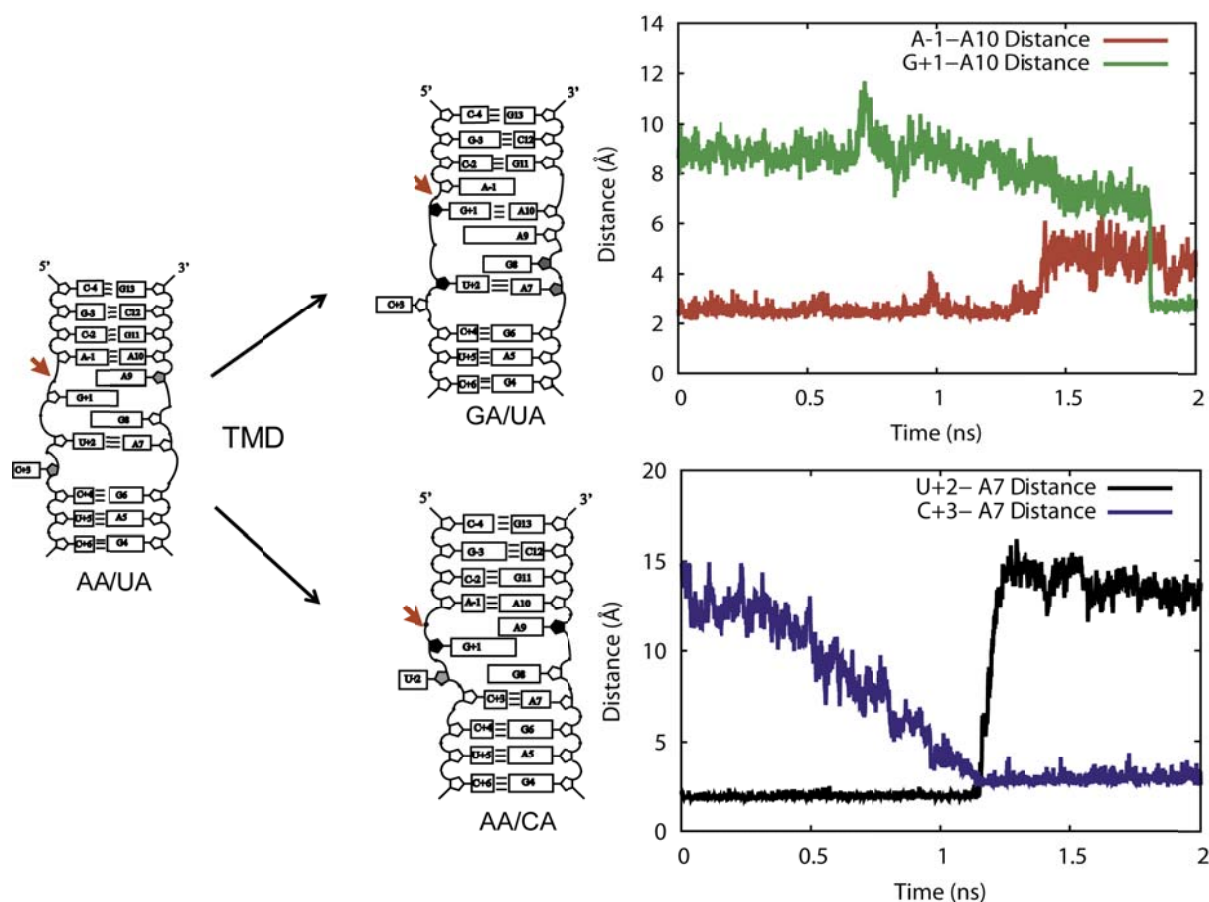


Figure 2-8: Targeted Molecular Dynamics (TMD) simulation of conformer transitions from AA/UA to GA/UA and AA/CA. The target conformations represent major and minor conformations observed in simulations of loop A. Top and bottom figures show the progress of base-pair loss and formation.

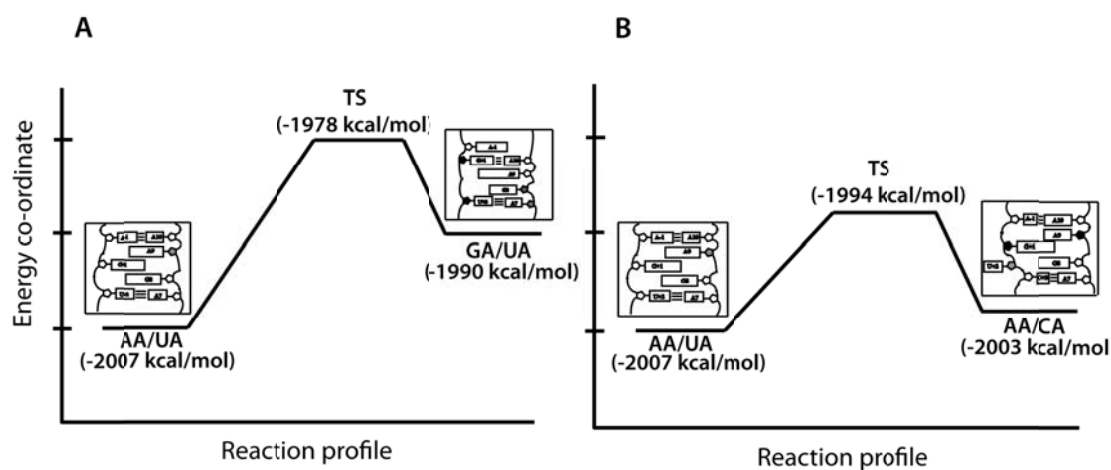


Figure 2-9: Free energy diagrams depicting the energy profile of TMD-driven conformations AA/UA to GA/UA (A) and AA/UA to AA/CA (B). The conformational free energies were estimated with the MMPB/SA method based on conformational ensembles derived from TMD simulations. The structural transition from AA/UA to GA/UA (A) is characterized by a higher energy barrier (~ 30 kcal/mol) compared to the energy barrier between AA/UA to AA/CA (~ 13 kcal/mol) conformations (B).

2.4 DISCUSSION

The conformational dynamics of RNA are key for understanding its functional properties. Specifically, RNA internal loops are common motifs that constitute the active site of many functional RNAs including ribozymes. The understanding of RNA internal loop base-pair and stacking interactions and their conformational variations are therefore key features in predicting RNA interaction and function. In this work, we have characterized the conformational variations of the hairpin ribozyme loop A domain containing an internal loop that is essential in forming RNA-RNA tertiary interactions with subsequent backbone cleavage. Using unbiased explicit solvent simulation we identified a surprisingly complex conformational energy landscape with three major conformations (conformers), distinguished by alternate base-pairing and stacking interactions within the internal loop of loop A, as well as numerous minor kinetic states within each of the major conformers. Transitions between different states primarily involved the dynamics of conserved bases U+2, C+3 and A9 through backbone torsional dynamics. We find significant kinetic barriers, especially between the three major conformers that reinforce the idea of an RNA energy landscape that consists of many deep local minima separated by kinetic barriers and complex transition pathway networks (65, 66). For example, the high kinetic barrier between AA/UA and GA/UA conformations (**Figure 2-8A**) can be overcome by sampling several states that ultimately permit population of the minor GA/UA conformation (**Figure 2-6**), among other conformers. Similar mechanism may be at play in the transition between AA/UA to AA/CA. The kinetic states therefore provide a pathway for accessing loop A shallow and deep conformational wells that stabilize the observed conformations. The conformational transitions implicated several active site residues including U+2 and C+3. The dynamics of U+2 has previously been reported by Cai and Tinoco's observation of broad U+2 NMR cross-peaks in

NOESY and COSY experiments performed in loop A (22) implying that this residue indeed undergoes conformational exchange. Our results indicate that this exchange is predominantly base-flipping facilitated by pucker exchange. Similar NMR studies have shown that both U+2 and C+3 sugar puckers reported mixed C2'/C3'-endo within the NMR structure of isolated loop A further supporting the conformational exchange of these two residues.

2.4.1 Sampling of docked loop A structure

A key question is how the conformational sampling of loop A in isolation relates to the conformation that is adopted when docked to loop B. Multi-conformational sampling of loop A results in a conformation (AA/CA) with structural similarity to the docked loop A, capable of docking with an activated loop B. By assessing the specific conformational changes inherent in the docked loop A crystal structure, we tabulated some of these conformational similarities with the docked structure with their relative population (**Table 2-7**). G+1 sampled extrahelical conformers approximately 8% of the time, mostly in our 2'-O-methylated loop A simulation, suggesting that the A-1 2'-O-methylation (also in the docked form crystal structure) may slightly stabilize the G+1 extrahelical stacking. Within the structures of isolated loop A and docked loop A, the base-pair and base stacking interactions reported significant variations. For example, G+1 was base-paired with A9 in the isolated loop A, while it was extra-helically base-paired with C25 in loop B (**Figure 2-2**) in the docked form. Although C+3 stacked between A7 and G8 in isolated loop A, the same residue was base-paired with A7 in the docked conformation. The dynamics of residues U+2 and C+3 in our simulations play a significant role in the formation of C+3:A7 base-pair in the AA/CA conformer, also observed in the docked structure (17). RNA dynamics and C+3:A7 base-pair formation suggests pre-organization from an inactive to active loop A RNA during our simulations, in which some features of the docked form are sampled. By

structurally comparing loop A's active site between our observed conformers with docked form of loop A (**Figure 2-10**), we determined the relative structural similarity of the AA/CA conformer and the crystallographically-observed docked conformation of loop A with an RMSD of 2.6 Å. This conformational sampling was facilitated by reorientation of specific active site residues and RNA backbone towards the structure of docked loop A besides the formation of a unique C+3:A7 base-pair interaction also observed in the docked conformation but not in isolated loop A. This is consistent with the free energy landscape of loop A RNA allowing sampling of an activated conformer resembling the docked structure. This may suggest that conformational selection at least plays a role in activating loop A. However, other conformers (AA/UA and GA/UA) are not structurally similar with docked loop with the heavy atom loop RMSDs of 4.2 Å and 3.2 Å respectively as referenced to the docked structure.

Although the AA/CA conformer is closer to the docked form than are the other major conformers, not all base-pair and stacking configurations match the docked crystal structure. It is possible that the docked structure is visited during longer (millisecond) timescales which are not reached in our simulations. However, it is more likely that the presence of loop B is necessary to fully form the docked structure. This is supported by the formation of several tertiary interactions between loop A and loop B, mostly stabilized by loop B. In the crystal structure of docked loop A, G+1 is extra-helical while base-pairing with C25 in loop B (17). In our simulations we see dynamics of G+1 with a minor G+1:A10 base-pair. This suggests a hybrid interaction mechanism for the docking of loop A to loop B where initial conformational selection is followed by an induced fit mechanism. A similar mechanism has recently been observed in U1A-RNA molecular recognition where Helix-C is structurally reoriented to allow RNA access in the absence of RNA (67).

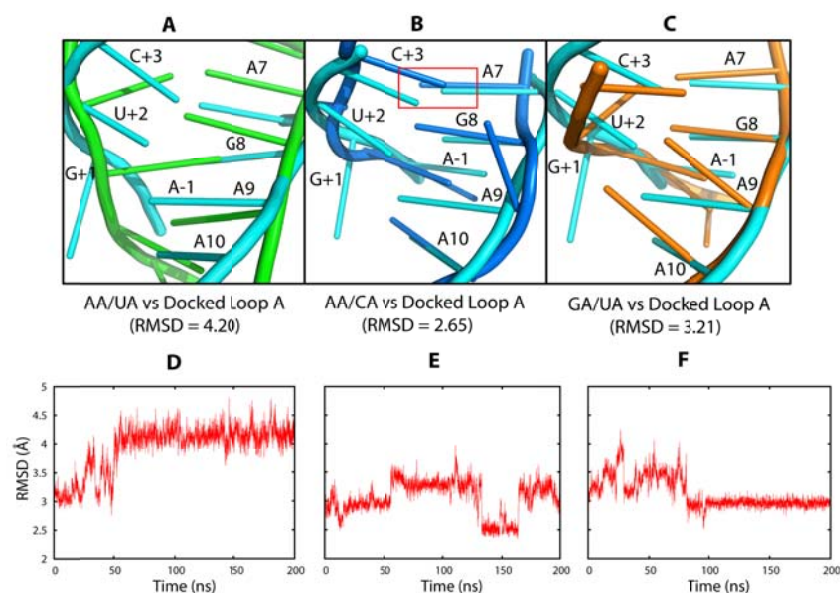


Figure 2-10: Loop A sampling docked conformation provides initial evidence of conformational capture mechanism. Structural comparison between docked loop A crystal structure (cyan) with simulation conformers AA/UA (green), AA/CA (blue) and GA/UA (orange) respectively within the active site (A, B, and C). The critical C+3:A7 base-pair (boxed red) observed in the docked loop A crystal structure is similarly seen in AA/CA but not AA/UA and GA/UA. The reported RMSD values indicate closer AA/CA structural sampling for docked loop A. Corresponding heavy atom RMSD profiles of active site residues for AA/UA (D), AA/CA (E) and GA/UA (F) referenced to docked loop A in three representative simulations.

Most docking-competent interactions were sampled during our simulations albeit in various populations (**Table 2-7**). However, we did not observe conformational similarity of residue A-1 between our simulation structures and the docked loop A structure. This active site residue constitutes the cleavage site which may undergo rearrangements similar to the docked form in only the presence of loop B or at longer timescales. Intra-domain and inter-domain tertiary interactions jointly prepare loop A for docking to its partner loop B.

The role of the GNRA tetraloop in stabilizing RNA provided a good control for assessing dynamics in loop A. It is worth noting that the GAAA tetraloop incorporated in our RNA remained quite stable over the simulation time consistent with the NMR structure (68) and other simulations done at similar temperatures (69, 70). This suggests that our observations were not biased by force field artifacts but that they provide a realistic behavior of this independently folded RNA. Previous studies of RNA indicate the formation of AA and GA sheared-type non-canonical base-pairing within the RNA loop (71-74) suggesting a possible alternate base-pair transition. This observation is consistent with alternate A-1:A10 and G+1:A10 observed in our conformers. An earlier simulation of the docked loop A (75) showed limited dynamics and structural conservation with the crystal structure, suggesting that flexibility of loop A in our simulation is necessary for its pre-organization to the docked form. The observation of multiple conformations and internal loop dynamics lays down a framework for understanding RNA dynamics and especially in pre-organizing RNA for activation within a complicated rugged energy landscape.

Table 2-7: Percentage population that sample specific interactions observed in the docked form of loop A during isolated loop A simulation. The cutoff for base-pair formation was set at 3.5Å

	Residue(s)	Conformation	% Population
Base-pair	A-1:A9	Base-pair	0.0
	U+2:G8	Base-pair	2.3
	C+3:A7	Base-pair	51.2
Stacking	G+1	Extra-helical	8.3
Ribose pucker	A-1	C2'-endo	0.0
	U+2	C3'-endo	41.7
	C+3	C3'-endo	75.0
	A7	C3'-endo	83.2
	G8	C3'-endo	41.7

2.4.2 Functional relevance

In RNA, dynamics is important to facilitate the diverse structural rearrangements associated with transitioning kinetic barriers to functionally competent structures. To accomplish these transitions, thermal and conformational activation provides the energy to drive RNA transitions to their dynamic functional roles. For example, the internal motions leading to the melting of base-pairs near the internal loop of HIV-1 stem loop 1 have been observed using NMR similar to a secondary structural transition that occurs during viral maturation (76). Moreover, cross-linking experiments have shown the formation of multiple functional folds of the hairpin ribozyme under different metal ion conditions implicating U+2 and C+3 residues (77). This suggests that under physiological conditions, solution conditions and other factors facilitate functionally relevant RNA transitions. The assessment of loop A provides the initial step to understanding tertiary structure formation and RNA-RNA interaction mechanism and dynamics in RNA. An NMR study of loop B by the Feigon group reported multiple mutually-inconsistent interproton NOEs, which is also consistent with the existence multiple conformations (21). Some of the NOEs were eventually suggested to be consistent with the docked state. These experimental observations also provide support for a possible conformational sampling in loop B, the docking partner of loop A. Taken together, the dynamics in loop A and subsequent formation of multiple kinetic substates and transitions present initial evidence for conformational sampling as a way of exploring the rugged free energy landscape. This process results in activated molecules capable of RNA-RNA interactions to effect functional output. Loop A domain of the hairpin ribozyme provides us with an excellent model for understanding the detailed mechanism for this process.

2.5 CONCLUSION

Unbiased MD simulation was used to determine conformational heterogeneity in RNA based on alternate base-pairs within a subset of residues in the loop region of domain A of the hairpin ribozyme. The conformers were determined using non-canonical base-pair combinations in stems I and II. The Markov state model analysis of our simulations determined several macrostates that undergo kinetic transitions within different local minima based on subtle structural changes and dynamics within a set of residues. Base and backbone dynamics also play an important role in alternate base-pair formation and subsequently conformational sampling in loop A. This suggests that conformational sampling and transition in loop A RNA is a key strategy to avoiding the kinetic traps that localize RNA in non-functional conformations. Kinetic transitions are therefore achieved through sampling a series of kinetic states which subsequently prepare the RNA to overcome the kinetic energy barriers to form functional (active) RNA. The most populated conformer, AA/CA, closely sampled conformational properties similar to the activated (docked) loop A conformation. This is significant because it highlights the role of conformational sampling in activating loop A (conformational selection) for the critical tertiary RNA-RNA interaction. This unique intrinsic base-pair rearrangement within a subset of kinetic states and transitions to conformations along the potential energy surface support the rugged but accessible free energy landscape of RNA with mechanistic properties that direct pre-organization to the activated state. The multiple conformers and inter-conformer transitions observed lay a foundation for understanding tertiary RNA interactions in the hairpin ribozyme, as well as in functional RNA systems more generally.

REFERENCES

REFERENCES

1. Al-Hashimi, H. M. (2005) Dynamics-based amplification of RNA function and its characterization by using NMR spectroscopy, *ChemBiochem* 6, 1506-1519.
2. Kern, D., and Zuiderweg, E. R. (2003) The role of dynamics in allosteric regulation, *Curr Opin Struct Biol* 13, 748-757.
3. Perez-Canadillas, J. M., and Varani, G. (2001) Recent advances in RNA-protein recognition, *Curr Opin Struct Biol* 11, 53-58.
4. Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., and Kern, D. (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair, *Nat Struct Mol Biol* 11, 945-949.
5. Julien, K. R., Sumita, M., Chen, P. H., Laird-Offringa, I. A., and Hoogstraten, C. G. (2008) Conformationally restricted nucleotides as a probe of structure-function relationships in RNA, *RNA* 14, 1632-1643.
6. Zhuang, X. W., Kim, H., Pereira, M. J. B., Babcock, H. P., Walter, N. G., and Chu, S. (2002) Correlating structural dynamics and function in single ribozyme molecules, *Science* 296, 1473-1476.
7. Fedor, M. J. (2000) Structure and function of the hairpin ribozyme, *J Mol Biol* 297, 269-291.
8. Ferre-D'amare A, R., and Rupert, P. B. (2002) The hairpin ribozyme: from crystal structure to function, *Biochem Soc Trans* 30, 1105-1109.
9. Dethoff, E. A., Hansen, A. L., Musselman, C., Watt, E. D., Andricioaei, I., and Al-Hashimi, H. M. (2008) Characterizing complex dynamics in the transactivation response element apical loop and motional correlations with the bulge by NMR, molecular dynamics, and mutagenesis, *Biophys J* 95, 3906-3915.
10. Frank, A. T., Stelzer, A. C., Al-Hashimi, H. M., and Andricioaei, I. (2009) Constructing RNA dynamical ensembles by combining MD and motionally decoupled NMR RDCs: new insights into RNA dynamics and adaptive ligand recognition, *Nucleic Acids Res* 37, 3670-3679.
11. Sosnick, T. R., and Pan, T. (2003) RNA folding: models and perspectives, *Curr Opin Struct Biol* 13, 309-316.

12. Pljevaljcic, 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.
13. Russell, R., Das, R., Suh, H., Traver, K. J., Laederach, A., Engelhardt, M. A., and Herschlag, D. (2006) The paradoxical behavior of a highly structured misfolded intermediate in RNA folding, *J Mol Biol* 363, 531-544.
14. Buzayan, J. M., Gerlach, W. L., Bruening, G., Keese, P., and Gould, A. R. (1986) Nucleotide sequence of satellite tobacco ringspot virus RNA and its relationship to multimeric forms, *Virology* 151, 186-199.
15. Walter, N. G., and Burke, J. M. (1998) The hairpin ribozyme: structure, assembly and catalysis, *Curr Opin Chem Biol* 2, 24-30.
16. Hampel, A., and Tritz, R. (1989) RNA catalytic properties of the minimum (-)sTRSV sequence, *Biochemistry* 28, 4929-4933.
17. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis, *Nature* 410, 780-786.
18. Pinard, R., Lambert, D., Walter, N. G., Heckman, J. E., Major, F., and Burke, J. M. (1999) Structural basis for the guanosine requirement of the hairpin ribozyme, *Biochemistry* 38, 16035-16039.
19. Ojwang, J. O., Hampel, A., Looney, D. J., Wong-Staal, F., and Rappaport, J. (1992) Inhibition of human immunodeficiency virus type 1 expression by a hairpin ribozyme, *P Natl Acad Sci USA* 89, 10802-10806.
20. Yu, M., Poeschla, E., Yamada, O., Degrandis, P., Leavitt, M. C., Heusch, M., Yees, J. K., Wong-Staal, F., and Hampel, A. (1995) In vitro and in vivo characterization of a second functional hairpin ribozyme against HIV-1, *Virology* 206, 381-386.
21. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat Struct Biol* 6, 212-216.
22. Cai, Z., and Tinoco, I., Jr. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite, *Biochemistry* 35, 6026-6036.
23. Rupert, P. B., Massey, A. P., Sigurdsson, S. T., and Ferre-D'Amare, A. R. (2002) Transition state stabilization by a catalytic RNA, *Science* 298, 1421-1424.
24. 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.

25. 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.
26. Torelli, A. T., Spitale, R. C., Krucinska, J., and Wedekind, J. E. (2008) Shared traits on the reaction coordinates of ribonuclease and an RNA enzyme, *Biochem Bioph Res Co* 371, 154-158.
27. Torelli, A. T., Krucinska, J., and Wedekind, J. E. (2007) A comparison of vanadate to a 2'-5' linkage at the active site of a small ribozyme suggests a role for water in transition-state stabilization, *RNA-A Publication of the RNA Society* 13, 1052-1070.
28. Liberman, J. A., Guo, M., Jenkins, J. L., Krucinska, J., Chen, Y. Y., Carey, P. R., and Wedekind, J. E. (2012) A Transition-State Interaction Shifts Nucleobase Ionization toward Neutrality To Facilitate Small Ribozyme Catalysis, *J Am Chem Soc* 134, 16933-16936.
29. Hampel, K. J., and Burke, J. M. (2001) A conformational change in the "loop E-like" motif of the hairpin ribozyme is coincidental with domain docking and is essential for catalysis, *Biochemistry* 40, 3723-3729.
30. Komatsu, Y., Kumagai, I., and Ohtsuka, E. (1999) Investigation of the recognition of an important uridine in an internal loop of a hairpin ribozyme prepared using post-synthetically modified oligonucleotides, *Nucleic Acids Res* 27, 4314-4323.
31. Walter, N. G., Chan, P. A., Hampel, K. J., Millar, D. P., and Burke, J. M. (2001) A base change in the catalytic core of the hairpin ribozyme perturbs function but not domain docking, *Biochemistry* 40, 2580-2587.
32. 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.
33. Clore, G. M., and Kuszewski, J. (2003) Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation, *J Am Chem Soc* 125, 1518-1525.
34. Feig, M., Karanicolas, J., and Brooks, C. L. (2004) MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology, *J Mol Graph Model* 22, 377-395.
35. Jorgensen, W. L. (1981) Quantum and Statistical Mechanical Studies of Liquids .10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers - Application to Liquid Water, *J Am Chem Soc* 103, 335-340.

36. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water, *J Chem Phys* 79, 926-935.
37. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A Smooth Particle Mesh Ewald Method, *J Chem Phys* 103, 8577-8593.
38. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems, *J Chem Phys* 98, 10089-10092.
39. Foloppe, N., and MacKerell, A. D. (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data, *J Comput Chem* 21, 86-104.
40. MacKerell, A. D., and Banavali, N. K. (2000) All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution, *J Comput Chem* 21, 105-120.
41. Best, R. B., Zhu, X., Shim, J., Lopes, P. E., Mittal, J., Feig, M., and Mackerell, A. D., Jr. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles, *J Chem Theory Comput* 8, 3257-3273.
42. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J Phys Chem B* 102, 3586-3616.
43. Huang, J., and MacKerell, A. D. (2013) CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data, *J Comput Chem* 34, 2135-2145.
44. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J Comput Chem* 26, 1781-1802.
45. Adelman, S. A., and Doll, J. D. (1976) Generalized Langevin equation approach for atom/solid-surface scattering: General formulation for classical scattering off harmonic solids, *J Chem Phys* 64, 2375-2388.
46. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J Comput Phys* 23, 327-341.

47. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1996) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules (vol 117, pg 5179, 1995), *J Am Chem Soc* 118, 2309-2309.
48. Zgarbová, M., Otyepka, M., Šponer, J., Mládek, A., Banáš, P., Cheatham, T. E., and Jurečka, P. (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles, *J Chem Theory Comput* 7, 2886–2902.
49. Schlitter, J., Engels, M., Kruger, P., Jacoby, E., and Wollmer, A. (1993) Targeted Molecular-Dynamics Simulation of Conformational Change - Application to the T $\leftrightarrow$ R Transition in Insulin, *Mol Simulat* 10, 291-308.
50. Floris, F., and Tomasi, J. (1989) Evaluation of the Dispersion Contribution to the Solvation Energy - a Simple Computational Model in the Continuum Approximation, *J Comput Chem* 10, 616-627.
51. Gallicchio, E., Kubo, M. M., and Levy, R. M. (2000) Enthalpy-entropy and cavity decomposition of alkane hydration free energies: Numerical results and implications for theories of hydrophobic solvation, *J Phys Chem B* 104, 6271-6285.
52. Honig, B., and Nicholls, A. (1995) Classical Electrostatics in Biology and Chemistry, *Science* 268, 1144-1149.
53. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S. H., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models, *Accounts Chem Res* 33, 889-897.
54. Lu, Q., and Luo, R. (2003) A Poisson-Boltzmann dynamics method with nonperiodic boundary condition, *J Chem Phys* 119, 11035-11047.
55. Luo, R., David, L., and Gilson, M. K. (2002) Accelerated Poisson-Boltzmann calculations for static and dynamic systems, *J Comput Chem* 23, 1244-1253.
56. Kar, P., and Knecht, V. (2012) Mutation-Induced Loop Opening and Energetics for Binding of Tamiflu to Influenza N8 Neuraminidase, *J Phys Chem B* 116, 6137-6149.
57. Kar, P., and Knecht, V. (2012) Origin of Decrease in Potency of Darunavir and Two Related Antiviral Inhibitors against HIV-2 Compared to HIV-1 Protease, *J Phys Chem B* 116, 2605-2614.
58. Kar, P., and Knecht, V. (2012) Energetic basis for drug resistance of HIV-1 protease mutants against amprenavir, *Journal of computer-aided molecular design* 26, 215-232.

59. Kar, P., Lipowsky, R., and Knecht, V. (2013) Importance of Polar Solvation and Configurational Entropy for Design of Antiretroviral Drugs Targeting HIV-1 Protease, *J Phys Chem B* 117, 5793-5805.
60. Bowman, G. R., Huang, X. H., and Pande, V. S. (2009) Using generalized ensemble simulations and Markov state models to identify conformational states, *Methods* 49, 197-201.
61. Beauchamp, K. A., Bowman, G. R., Lane, T. J., Maibaum, L., Haque, I. S., and Pande, V. S. (2011) MSMBuilder2: Modeling Conformational Dynamics on the Picosecond to Millisecond Scale, *J Chem Theory Comput* 7, 3412-3419.
62. Schrödinger, L., (Ed.) (2010) *The PyMOL Molecular Graphics System, Version 1.3*. Schrödinger, LLC, New York, New York.
63. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics, *J Mol Graph Model* 14, 33-38.
64. Shieh, F. K., Youngblood, B., and Reich, N. O. (2006) The role of Arg165 towards base flipping, base stabilization and catalysis in M.HhaI, *J Mol Biol* 362, 516-527.
65. Bailor, M. H., Mustoe, A. M., Brooks, C. L., 3rd, and Al-Hashimi, H. M. (2011) Topological constraints: using RNA secondary structure to model 3D conformation, folding pathways, and dynamic adaptation, *Curr Opin Struct Biol* 21, 296-305.
66. Schultes, E. A., Spasic, A., Mohanty, U., and Bartel, D. P. (2005) Compact and ordered collapse of randomly generated RNA sequences, *Nat Struct Mol Biol* 12, 1130-1136.
67. Kurisaki, I., Takayanagi, M., and Nagaoka, M. (2014) Combined mechanism of conformational selection and induced fit in U1A-RNA molecular recognition, *Biochemistry* 53, 3646-3657.
68. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops, *J Mol Biol* 264, 968-980.
69. Sorin, E. J., Engelhardt, M. A., Herschlag, D., and Pande, V. S. (2002) RNA simulations: Probing hairpin unfolding and the dynamics of a GNRA tetraloop, *J Mol Biol* 317, 493-506.
70. Zichi, D. A. (1995) Molecular-Dynamics of RNA with the Opls Force-Field - Aqueous Simulation of a Hairpin Containing a Tetranucleotide Loop, *J Am Chem Soc* 117, 2957-2969.
71. Chen, G., Znosko, B. M., Jiao, X. Q., and Turner, D. H. (2004) Factors affecting thermodynamic stabilities of RNA 3x3 internal loops, *Biochemistry* 43, 12865-12876.

- 72. Santalucia, J., Kierzek, R., and Turner, D. H. (1991) Functional-Group Substitutions as Probes of Hydrogen-Bonding between GA Mismatches in RNA Internal Loops, *J Am Chem Soc* 113, 4313-4322.
- 73. Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared A(anti)center dot A(anti) base pairs in a destabilizing 2x2 internal loop: The NMR structure of 5 '(rGGCAAGCCU)(2), *Biochemistry* 41, 14969-14977.
- 74. Schroeder, S. J., and Turner, D. H. (2000) Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA, *Biochemistry* 39, 9257-9274.
- 75. Mlynsky, V., Banas, P., Hollas, D., Reblova, K., Walter, N. G., Sponer, J., and Otyepka, M. (2010) Extensive Molecular Dynamics Simulations Showing That Canonical G8 and Protonated A38H(+) Forms Are Most Consistent with Crystal Structures of Hairpin Ribozyme, *J Phys Chem B* 114, 6642-6652.
- 76. Sun, X. Y., Zhang, Q., and Al-Hashimi, H. M. (2007) Resolving fast and slow motions in the internal loop containing stem-loop 1 of HIV-1 that are modulated by Mg²⁺ binding: role in the kissing-duplex structural transition, *Nucleic Acids Res* 35, 1698-1713.
- 77. Lambert, D., Heckman, J. E., and Burke, J. M. (2006) Cation-specific structural accommodation within a catalytic RNA, *Biochemistry* 45, 829-838.

CHAPTER 3

CONSTRUCT OPTIMIZATION, DOCKING STUDIES AND NMR RESONANCE ASSIGNMENT OF THE LOOP A DOMAIN OF THE HAIRPIN RIBOZYME

3.1 INTRODUCTION

Since globular RNAs lack a central hydrophobic core, the driving force for the interaction of multiple RNA helices and loop regions into a highly ordered tertiary structure is still unclear in comparison to proteins. A clear understanding of the formation of RNA tertiary structure is thus vital to unravel the mechanisms behind cellular machines such as the eukaryotic spliceosome, in which several RNA structures and RNA-protein interactions form and are subsequently disrupted at specific points along the splicing pathway (1, 2). As discussed in Chapter 1, the hairpin ribozyme provides a model system for understanding RNA tertiary structure interactions correlated to function. The hairpin ribozyme's active site constitutes two independently folding domains (loop A and loop B) (3). Loop A has a symmetrical internal loop interspersed between two helical stems. Docking between loop A and B pre-organizes the active site to effect catalysis (phosphodiester bond cleavage). In nature, these two domains interact as a four-way helical junction (4WJ) (4) but two way junction (2WJ) and trans-constructs are also actively under investigation (**Figure 3-1**). Previous *in vitro* reconstitution studies have shown that a two-way helical junction (2WJ) can form a functional minimal hairpin ribozyme (3) while independent trans-docking domains also support full activity (5, 6). The cleavage occurs between G+1 and A-1 in loop A (7). Chemical cleavage by various self-cleaving RNA structural motifs is essential to the replication of several plant viroids (8), the human hepatitis δ virus (9), the mitochondrial Varkud satellite from *Neurospora* (10) and glmS riboswitch found in gram-positive bacteria (11). In the specific case of the hairpin ribozyme, the cleavage reaction is reversible, and the ligation reaction is ~ 10 times faster than that of cleavage (12, 13). These small ribozymes are tractable systems to use in the study of the critical principles of RNA catalysis using molecular structure and functional approaches (14, 15). Mutational studies have

identified residues that are critical for driving catalysis in the hairpin ribozyme, especially in the active site (16-19). Burke's group investigated binding and cleavage with single base substitutions at positions A-1 and U+2 (19). U+2 was suggested to be involved in tertiary interactions, possibly to pre-organize the active site. The direct role of U+2 in catalysis, however, was not determined. Structural studies of loop A established a conformational C2'/C3'-endo mixed sugar pucker in U+2 with a flipped out (extra-helical) base (20). Using time-resolved fluorescence resonance energy transfer (FRET) experiments to probe how U+2 mutations affect docking, cleavage rates were shown to significantly decrease despite the fact that U+2 nucleotide substitutions did not considerably affect docking (21, 22), A uracil to cytosine mutation at position +2 (U+2C) could introduce a new Watson-Crick base pair with G8 and limit dynamics at in this position. Such mutations may quench certain critical dynamics required to facilitate faster docking and catalysis. This hypothesis has been motivated by sequence comparisons with several residue mutations. An interesting double mutant (revertant) U+2C/C+3U introduced in loop A (**Figure 3-2**) completely abolishes docking under non-equilibrium, gel shift assay (22). U+2C/C+3U sequence stayed inactive with the native G8 in loop A, but this sequence was partially rescued by a G8U mutation. Unlike the U+2C mutant, the double mutant abrogates docking, implying a role for U+2 in docking. The X-ray crystal structures of the docked ribozyme confirmed the contribution of U+2 in the active site formation, since it is involved in a non-canonical base-pair interaction with G8 (23-25). The loop A NMR solution structure solved by Cai and Tinoco (20), however, contained a non-native A-1 to C mutation previously designed to cleave HIV-1 RNA (26). This structure therefore may not represent the conformation of the pre-catalytic form of loop A.

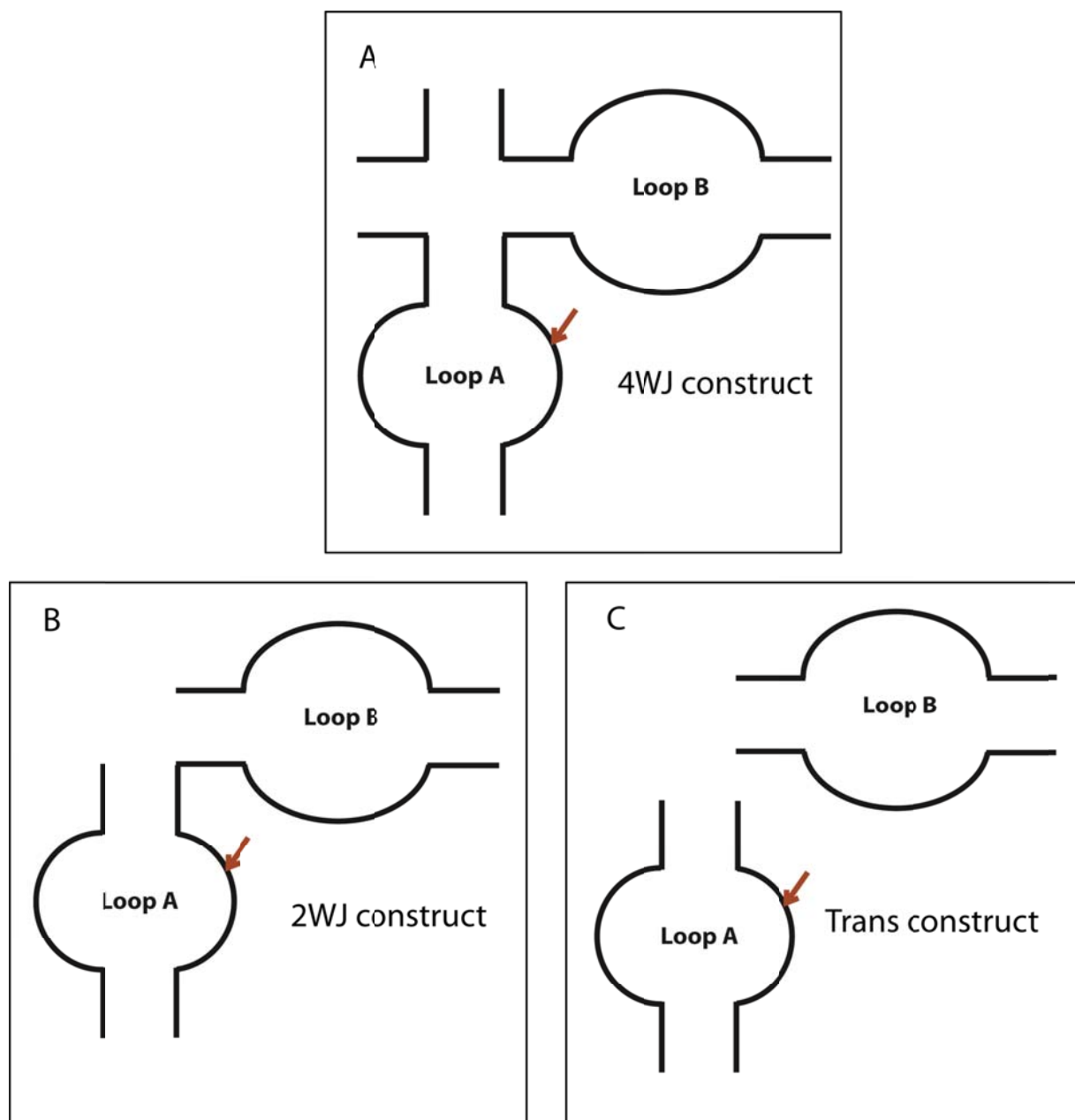


Figure 3-1: Cartoon representation of hairpin ribozyme active site region. (A) Four-way junction, B) two-way junction construct, and (c) isolated (trans-) constructs

To address this gap, we designed a trans-construct of native loop A sequence for our biophysical studies (**Figure 3-2**). Studies of the trans-docking domains of hairpin ribozyme is important because of its potential to provide new insights into the docking transition using direct measurements which assay for bimolecular equilibrium and binding kinetics using various spectroscopic techniques (27). Here, we employed spectroscopic circular dichroism (CD) equilibrium docking assay to monitor interaction between loop A or double mutant U+2C/C+3U loop A and loop B. This assay was developed by Dr. Minako Sumita, then a post-doctoral associate in our lab. All our trans-constructs, including loop A, loop A(U+2C/C+3U) and loop B, were designed to incorporate a GNRA tetraloop to facilitate stem I stability and NMR assignments. Taken together, these results implicate U+2 in aligning G8 in the active site of hairpin ribozyme to facilitate efficient catalysis. In the U+2C/C+3U double mutant sequence, abrogated internal loop dynamics (by loop constriction) may limit certain key tertiary interactions between loop A and loop B in assembling the active site. Our CD assay tested the docking transitions between native or double mutant loop A with loop B. NMR was used to assess the RNA folding and conformation. For the process of NMR assignments of loop A, we evaluated several loop A constructs to determine a well-behaved loop A construct. We used *in vitro* transcription to incorporate the isotopically labeled ^{15}N and ^{13}C (using ^{15}N and ^{13}C nucleotide triphosphates) for NMR resonance assignments. We also used NMR to study loop A by first assigning the critical ^1H , ^{13}C and ^{15}N resonances, an important step in probing loop A dynamics using the NMR relaxation measurements described in Chapter 4 of this thesis. Also, we performed several NMR experiments (1D, 2D and 3D) to obtain resonance assignment data for loop A. The resonance assignment data obtained thus far are presented here in detail.

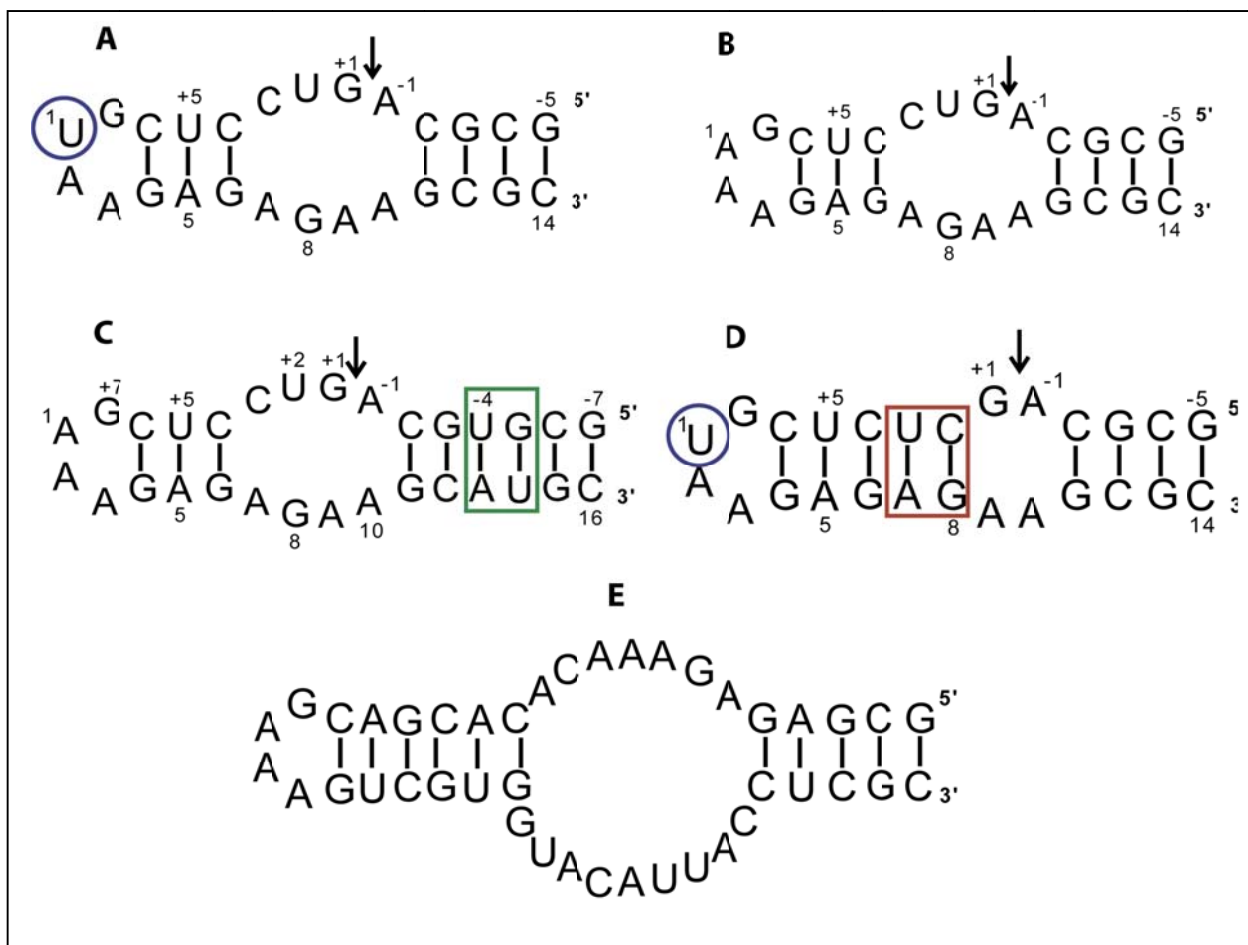


Figure 3-2: Hairpin ribozyme trans-constructs used in our studies: (A) GUAA loop A, (B) GAAA loop A, (C) GAAA loop A extended, (D) loop A(U+2C/C+3U) and (E) GAAA loop B used in our studies. Blue circle highlights a U1 mutation in constructs A and D, green box highlights residues used for stem II extension in C, while the red box highlights U+2C/C+3U double mutation which introduces two additional base-pairs in the loop region.

3.2 MATERIALS AND METHODS

3.2.1 RNA preparation

RNA was synthesized by incorporating nucleotides of the desired RNA constructs (see **Figure 3-2**) via *in vitro* transcription (28, 29) using recombinant phage T7 RNA polymerase enzyme in the presence of a synthetic DNA template. The synthetic template was annealed to a short complementary T7 promoter sequence. The transcription was initiated by recombinant T7 RNA polymerase following addition of DNA template, unlabeled or uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled nucleotide triphosphates (ATP, UTP, CTP, and GTP) and Mg^{2+} . Transcription was continued for 3 hours at 37 °C then stopped by the addition of EDTA. To analyze the success of *in vitro* transcription, samples from the reaction cocktail were analyzed by 15% denaturing gel electrophoresis after staining using Stains All solution.

Transcribed RNA was recovered by ethanol precipitation from the transcription cocktail. The ethanol precipitate was then reconstituted in H_2O , and subjected to purification through the HiLoad 26/60 Superdex⁷⁵ column (30) monitored at A_{260} with an isocratic flow rate of 3 mL/min sodium phosphate buffer, pH 6.40, while collecting fractions. Fractions corresponding to loop A, were pooled together and systematically exchanged with H_2O using Amicon Ultra-4 centrifugal filter units with 3000 MWCO. RNA purity was assessed by denaturing gel and MALDI-TOF mass spectrometry. Purified RNA was then dried under vacuum and reconstituted in appropriate buffers for Circular Dichroism (CD) and NMR studies.

3.2.2 Docking transitions between mutant loop A and wildtype loop B

Docking studies of the double mutant were performed using the difference CD experiment previously developed in our lab (27). For docking studies, ten 5 μM samples of loop A(U+2C/C+3U) and loop B were prepared from each stock of loop A(U+2C/C+3U) and loop B docking in increasing $[\text{Co}(\text{NH}_3)_6]^{3+}$ concentration. The samples were dialyzed with different concentrations of $[\text{Co}(\text{NH}_3)_6]^{3+}$ and then left to equilibrate on the bench for at least 1 hour before taking the CD scans. The $[\text{Co}(\text{NH}_3)_6]^{3+}$ concentrations used were 0 μM , 15 μM , 30 μM , 60 μM , 90 μM , 120 μM , 150 μM , 200 μM , 250 μM and 500 μM . Three sets of CD scans were acquired between 220-400 nm for the individual RNA constructs as well as loop A(C+3U/U+3C) and loop B at increasing concentrations of $[\text{Co}(\text{NH}_3)_6]^{3+}$ as outlined above. The scans were averaged and normalized with buffer conditions. Docking was monitored by circular dichroism (CD) where the CD spectra of individual loop A(U+2C/C+3U) and loop B were acquired separately and mathematically added against the CD spectrum of both loop A(U+2C/C+3U) and loop B together in the same tube. The signal corresponding to docking was then calculated by the formula shown below:

$$\text{CD docking} = (\text{CD U+2C/C+3U loop A} + \text{loop B}) - (\text{CD U+2C/C+3U loop A} + \text{CD loop B}) \quad 3-1$$

The above docking experiment was repeated (by Dr. Mina Sumita) using native loop A instead of the double mutant loop A(C+3U/U+3C). All the docking experiments were done in triplicate and the data obtained exported and analyzed in MS excel. RNA was quantified by extinction coefficients at 260 nm according to Beer-Lambert's law ($A = \epsilon \times b \times c$). The molar extinction coefficients (ϵ) were 214,000 $\text{M}^{-1} \text{cm}^{-1}$ for wild-type loop A, 267,400 $\text{M}^{-1} \text{cm}^{-1}$ for loop A(U+2C/C+3U) and 354,000 $\text{M}^{-1} \text{cm}^{-1}$ for loop B (27).

3.2.3 NMR studies

To analyze samples by NMR, purified loop A samples were resuspended in 10 mM phosphate buffer, 100 μ M EDTA pH 5.5 or pH 6.5 and loaded in a reduced-volume NMR tube (Shigemi) for NMR data acquisition and optimization of pH. For exchangeable proton NMR, loop A was constituted in a total volume of 250 μ L in 90% H₂O and 10% D₂O respectively. Sample concentrations were between 0.8-1.2 mM. Exchangeable 1D proton spectra were acquired with a 1-1 echo sequence (31) at various temperatures ranging from 5, 10, 15, 20 and 25 °C for GUAA loop A, GUAA loop A(U+2C/C+3C), GAAA loop A, and GAAA loop A extended constructs. NMR experiments were done on Varian Unity INOVA 600 MHz or Bruker 900 MHz. During acquisition, the proton radiofrequency (RF) carrier was placed on the HDO residual peak. All NMR experiments were acquired using RNA biopack pulse sequences (Varian/Agilent), unless otherwise stated, with slight modifications after optimization. One-dimensional *imino*-optimized spectroscopy was acquired using 1-1 echo pulse sequence (31) with 13106 complex points in the t_1 dimension, a 3 s recycle delay, and 96 steady-state scans. The first excitation maximum was targeted between imino and aromatic region for the second excitation maximum to fall at the center of the imino region. The sweep width was set at 12,000 Hz. ¹H-¹H homonuclear NOE of exchangeable protons was acquired in Bruker 900 MHz with cryo probe using 1024 x 256 complex points in t_2 and t_1 dimensions, 1.5 s recycle delay, 300 ms mixing time with 22,500 Hz spectral width on both dimensions. ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) spectra was acquired using 1024 $\times$ 256 complex points in the t_2 and t_1 dimensions, a 1.5 s recycle delay and 128 steady-state scans with corresponding spectral widths of 12000 and 6078.76 Hz. The ¹⁵N carrier frequency was set at 118 ppm as determined from indirect referencing to internal reference standard 1 mM sodium 2,2 dimethyl-2-

silapentane-5-sulfonate (DSS) (32). The non-exchangeable protons were acquired in 100% D₂O at 25 °C. Similarly, the proton RF carrier was positioned at the HDO residual peak while the ¹³C carrier at 110 ppm based on indirect referencing to internal reference standard DSS. Both ribose and aromatic ¹H-¹³C HSQC and constant time (CT)-HSQC spectra were acquired with 2048 × 256 complex points in the t₂ and t₁ dimensions, respectively, with corresponding spectral widths of 12000 and 6033 Hz, and a 2 s recycle delay. ¹³C carrier frequency was set at 85 ppm for ribose ¹H-¹³C HSQC and 110 ppm for aromatic ¹H, ¹³C-HSQC, as determined from indirect referencing to internal reference standard DSS. In the CT-HSQC experiment, the constant time delay was set at 0.028 s while the number of points in the ¹³C dimension was set at 52. ¹H-¹H homonuclear ¹³C-edited nuclear Overhauser spectroscopy (NOESY-HSQC) sub-spectra were acquired at 900 MHz NMR with 1024 × 256 complex points in the t₂ and t₁ dimensions, respectively, with corresponding spectral widths of 8992, 12589, and 7914 Hz, a 2 s recycle delay, 256 steady-state scans and mixing time of 100 ms. A 3D ¹³C NOESY-HSQC experiment was done at 900 MHz NMR in 100% D₂O. The ¹H carrier was set to 4.7 ppm, while the ¹³C carrier was set at 145 ppm. The sweep widths were 12590 (ω₁), 7914 (ω₂), and 8992.8 Hz (ω₃). Real data points acquired for the three dimensions were 144 in ω₁, 64 in ω₂, and 2048 in ω₃. The NOESY mixing time was 150 ms. 3D HCCH COSY was acquired at 600 MHz with 1024 × 256 complex points in the t₂ and t₁ dimensions, respectively, with corresponding spectral widths of 6000 × 6000 × 9000 Hz, a 2 s recycle delay, and 256 steady-state scans. 3D HCCH TOCSY were acquired with 512 × 256 complex points with spectral widths of 2999.18, 2999.18 and 8295.31 Hz, a 1 s recycle delay, and 64 steady-state scans. 3D Adenine HCCH-TOCSY was acquired through sweep widths of 5997.9 × 1679.5 × 4524.8. The t₃ dimension had 256 complex points with 30 and 32 data points in t₂ and t₁ dimensions respectively. All NMR data were processed using FELIX 2002

(Accelrys). Before Fourier transformation in the t_2 dimension, data was zero filled, a 20% DC offset was applied, and a 5 Hz exponential line broadening function was used. A cosine-squared apodization function was also applied in the indirect dimensions before spectral transformation

3.3 RESULTS

3.3.1 Docking studies

Docking is a required step that precedes the chemical steps of catalysis in the hairpin ribozyme reaction mechanism. In these docking studies, we hypothesized that certain stabilizing modifications in loop A limit docking by stabilizing ground-state interactions or alternate structures and consequently eliminating functionally important interactions which drive docking and catalysis. To test our hypothesis, we started by analyzing the possible docking transitions in mutant U+2C/C+3U loop A together with loop B and then comparing these results with the docking studies reported for wild type loop A and loop B (27). Docking studies were monitored by circular dichroism (CD) over a wavelength range of 220-400 nm. Our data demonstrates the abolition of docking in the double mutant loop A(U+2C/C+3U) at 250 μ M $[\text{Co}(\text{NH}_3)_6]^{3+}$. Similar non-docking observations were made in the absence of $[\text{Co}(\text{NH}_3)_6]^{3+}$ (data not shown). This suggests that loop A (U+2C/C+3U) does not undergo structural transitions necessary for docking in the presence of loop B. This observation is also consistent with what Burke's group has reported using their gel shift assay, albeit under slightly different metal ion conditions (22). The wild type loop A actively docked with loop B under similar conditions with docking transitions observed both at 230 nm and 270 nm respectively (27) (**Figure 3-3**). Analysis of docking as a function of the increasing concentration of $[\text{Co}(\text{NH}_3)_6]^{3+}$ between loop A (U+2C/C+3U) and loop B has been presented (**Figure 3-4**). This data shows the lack of a specific interaction between loop A(U+2C/C+3U) and loop B between 0 - 250 μ M $[\text{Co}(\text{NH}_3)_6]^{3+}$.

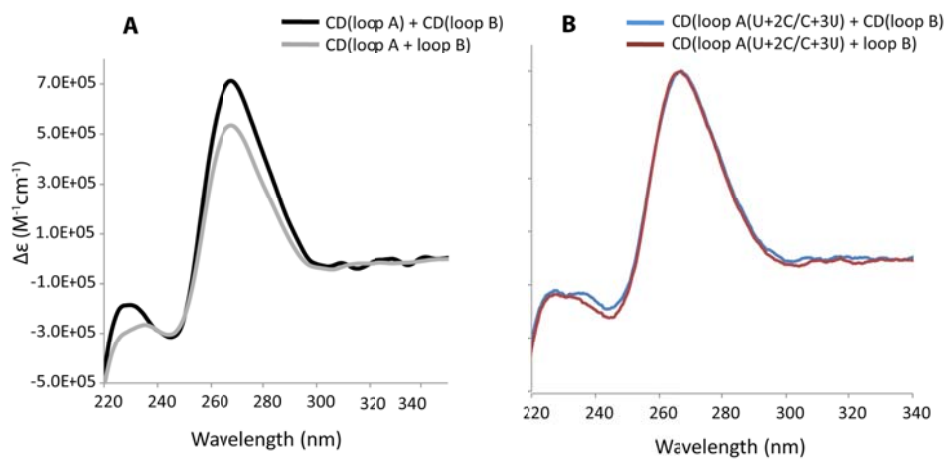


Figure 3-3: Comparison between the circular dichroism (CD) spectra of docking between wild type loop A (A) and loop A(U+2C/C+3U) (B). Docking was monitored in the presence of loop B. The spectra were obtained in the following buffer conditions: 20 mM HEPES, 20 μ M EDTA, pH7.5, 250 μ M $[Co(NH_3)_6]^{3+}$.

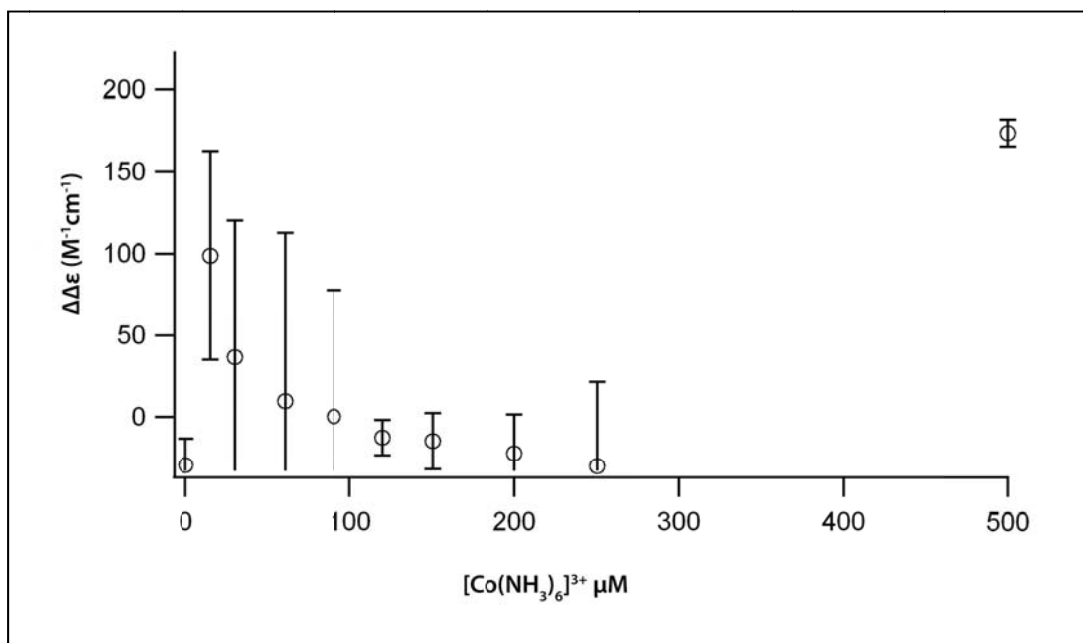


Figure 3-4: $\text{Co}(\text{NH}_3)_6^{3+}$ titration studies to assay for metal-dependent interaction between loop A(U+2C/C+3U) and loop B. The observed scatter indicates non-specific RNA-RNA interaction with increasing $[\text{Co}(\text{NH}_3)_6]^{3+}$.

Since the error bars are large within this range, we interpret this as random interaction between 2 RNA molecules. Conversely, metal-dependent docking transition between wild type loop A and loop B was demonstrated (27) in cobalt hexamine with an observed $(\text{Co})_{1/2}^{\text{dock}}$ of $\sim 50 \mu\text{M}$. These results not only indicate that our trans-domain constructs were active, but also demonstrate the manifestation of structural changes in real time indicated as docking, while the lack of docking in loop A(U+2C/C+3U) is likely due to the over-stabilization of loop A's active site by an additional two Watson-Crick base pairs. This extra stabilization abolishes detectable docking transitions at accessible concentrations of $[\text{Co}(\text{NH}_3)_6]^{3+}$ that support docking in the wild-type system. However, we cannot rule out structural transitions in much higher metal ion concentrations. Detailed analysis of this possibility was not possible due to the tendency of RNA to aggregate in the presence of elevated $[\text{Co}(\text{NH}_3)_6]^{3+}$.

3.3.2 Construct optimization for NMR spectroscopy

A critical challenge to the success of RNA structure determination is the behavior of the RNA construct to be studied. Various constructs of loop A (both unlabeled and $^{15}\text{N}/^{13}\text{C}$ isotopically labeled) were generated by *in vitro* transcription using T7 RNA polymerase and the synthesized RNA purified at high concentrations for NMR studies. The purity of the RNAs was assessed by gel electrophoresis and mass spectrometry (MALDI-TOF) respectively. The constructs were subjected to NMR experiments under various annealing conditions, buffer conditions and temperatures to optimize the secondary structure folding of loop A RNA. We determined pH and buffer conditions within which loop A easily folded into a hairpin structure and these conditions were replicated in all our NMR experiments.

Using a combination of 1D NMR, 2D ^1H , ^{15}N heteronuclear single quantum coherence (HSQC) and 2D nuclear Overhauser effect spectroscopy (NOESY), the behavior of each construct was assessed based on its *imino*-region resonance (**Figure 3-5**). The *imino* protons of G and U nucleotides in RNA resonate between 10-14 ppm especially when protected from exchange (**Figure 3-6**). **Figure 3-5** depicts the results of a 1D *imino* spectrum for all loop A constructs tested. Overall, the number of peaks observed correspond to the base-pairs expected except in most constructs except for the GUAA tetraloop construct which showed extra uridine imino peaks as well as lack of a prominent tetraloop G-A shear base-pair peak at ~ 10 ppm (**Figure 3-5B**). The extra uridine imino peak in GUAA tetraloop and the lack of G-A base-pair imino peak at ~ 10 ppm are indicative of the formation of a duplex structure in which the presumed tetraloop uridine and adenine are base-paired, , explaining the additional uridine imino peak. Various annealing protocols that favor hairpin formation were not successful in folding GUAA loop A into a hairpin (**Figure 3-7**) as shown by the lack of the tetraloop G imino peak at ~ 10.5 ppm. The GUAA loop A(U+2C/C+3U) construct (**Figure 3-5A**), however, clearly recorded the tetraloop G-A sheared imino peak at ~ 10.4 ppm, suggesting that the stabilization of base-pairing in loop A favors formation of the hairpin as opposed to the duplex observed in the previous GUAA loop A. To address the duplex-formation issue, we built a comparable construct with a GAAA tetraloop variant (**Figure 3-5C**), which we hypothesized to be less vulnerable to duplex formation. In this construct, we observed two peaks at ~ 10 ppm (10.6 and 10.8 ppm), one of which was assigned to the G-A sheared base-pair in the GAAA tetraloop region. The observation of the G-A sheared base-pair peak indicated tetraloop formation and the desired hairpin structure. Finally, since the initial GAAA construct showed some evidence of helical fraying at moderate temperatures (data not shown), we designed another construct with the

GAAA tetraloop introducing two additional base-pairs in stem II (GAAA loop A extended). The extended structure showed a 1D imino spectrum (**Figure 3-5D**) consistent with the number of expected Watson-Crick base-pairs with stable hydrogen bonding up to 25°C (**Figure 3-8**). These data indicate appropriate thermodynamic stability of helical regions for use in NMR structural and dynamic studies.

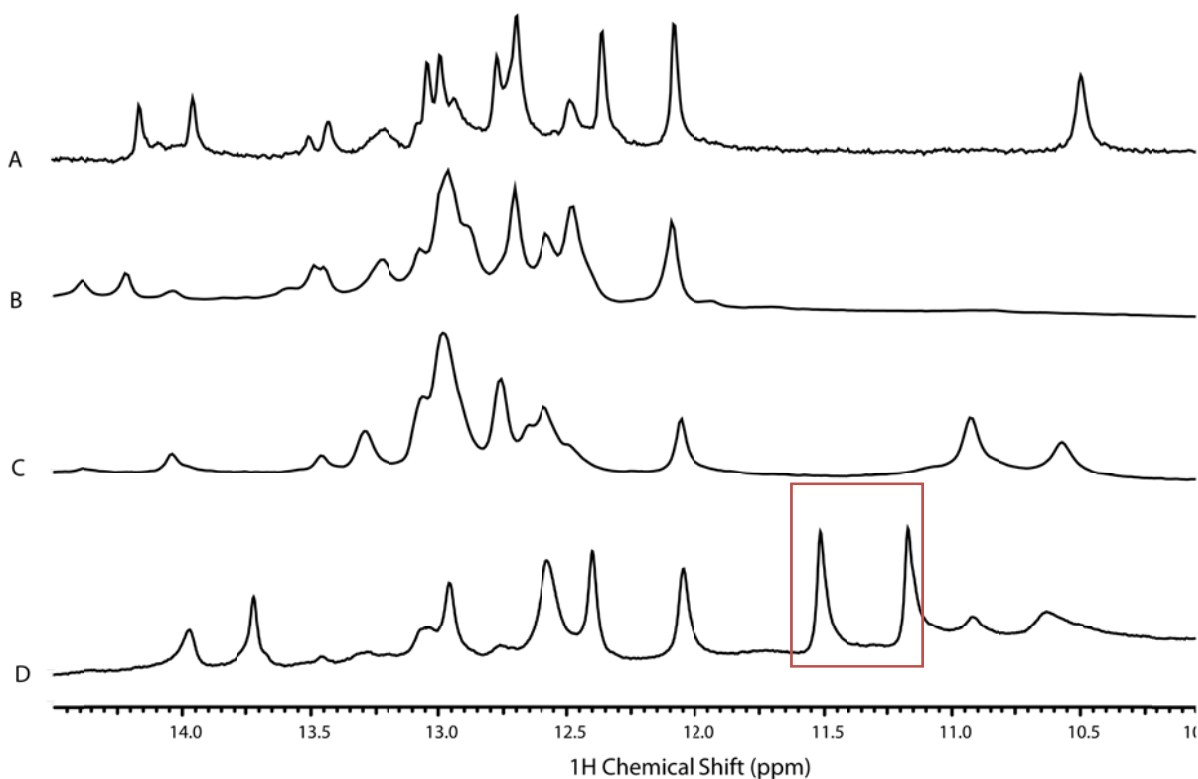


Figure 3-5: 1D imino NMR peaks of various loop A constructs. GUAA tetraloop with U+2C/C+3U mutant (A), GUAA tetraloop (B), GAAA tetraloop (C) and GAAA tetraloop with extended stem I highlighted in red (D). Peaks highlighted in red box are the extended imino residues G and U

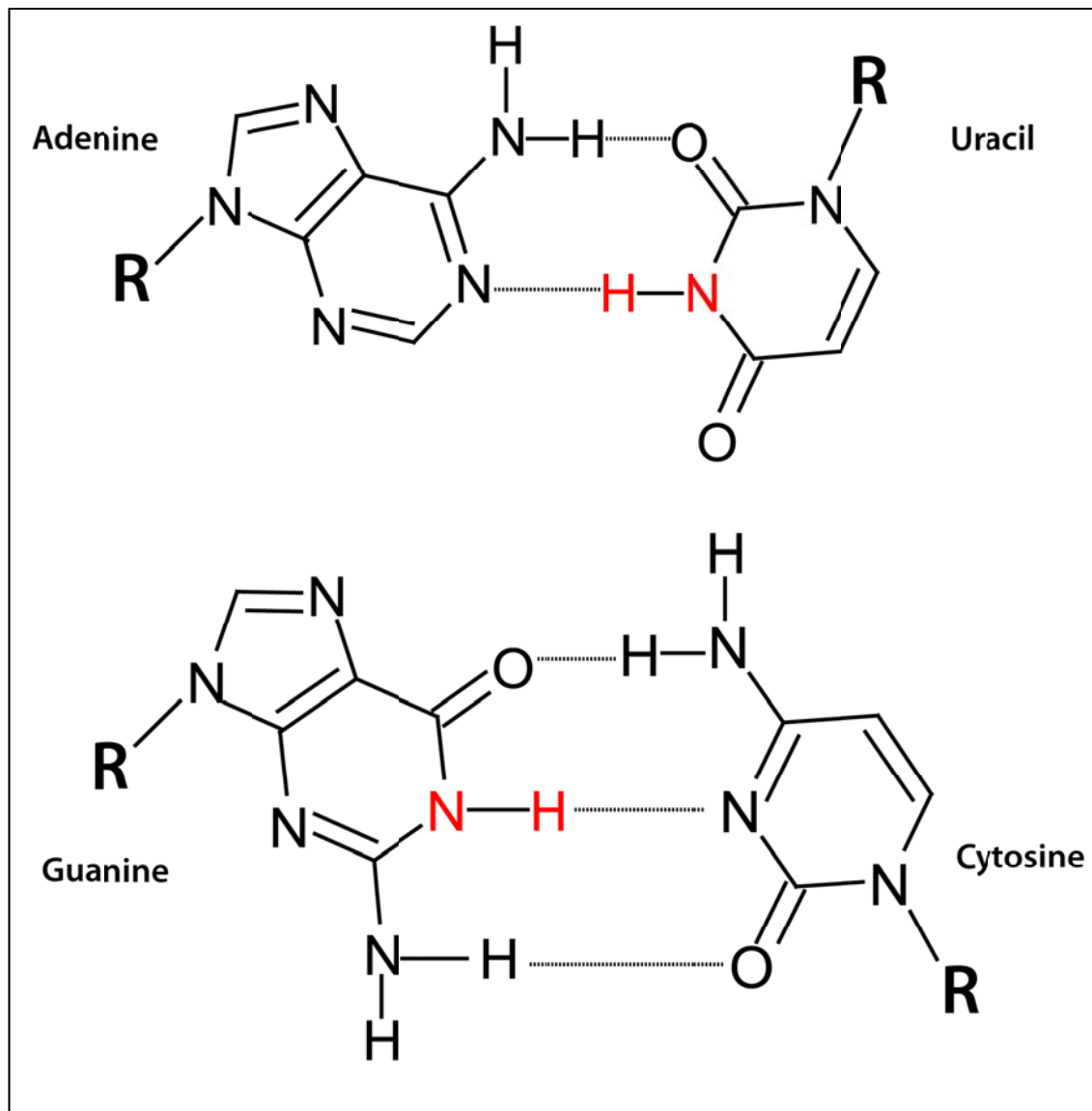


Figure 3-6: RNA base-pairing between A-U and G-C highlighting the involvement of imino (highlighted red) protons of guanine and uracil; and amino protons in hydrogen bond formation.

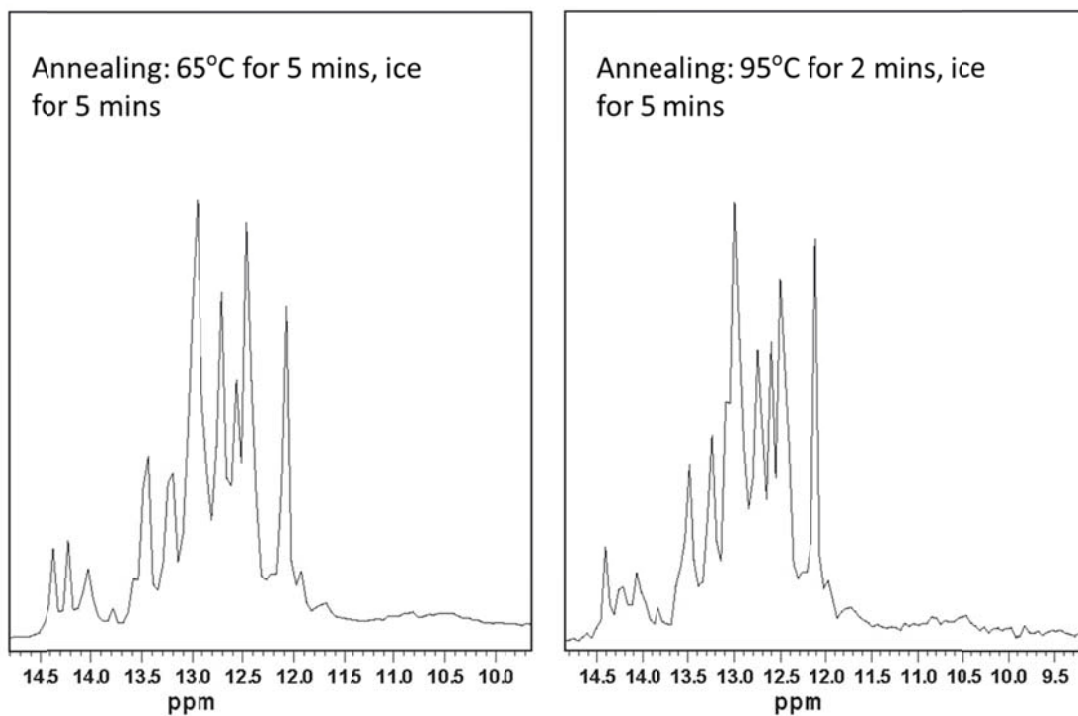


Figure 3-7: 1D imino spectra of GUAA loop A at different annealing conditions. Sample condition: 10 mM phosphate, 150 mM NaCl, 100 μ M EDTA, pH 6.5. Spectra were acquired at 15°C under two annealing conditions that favor hairpin formation.

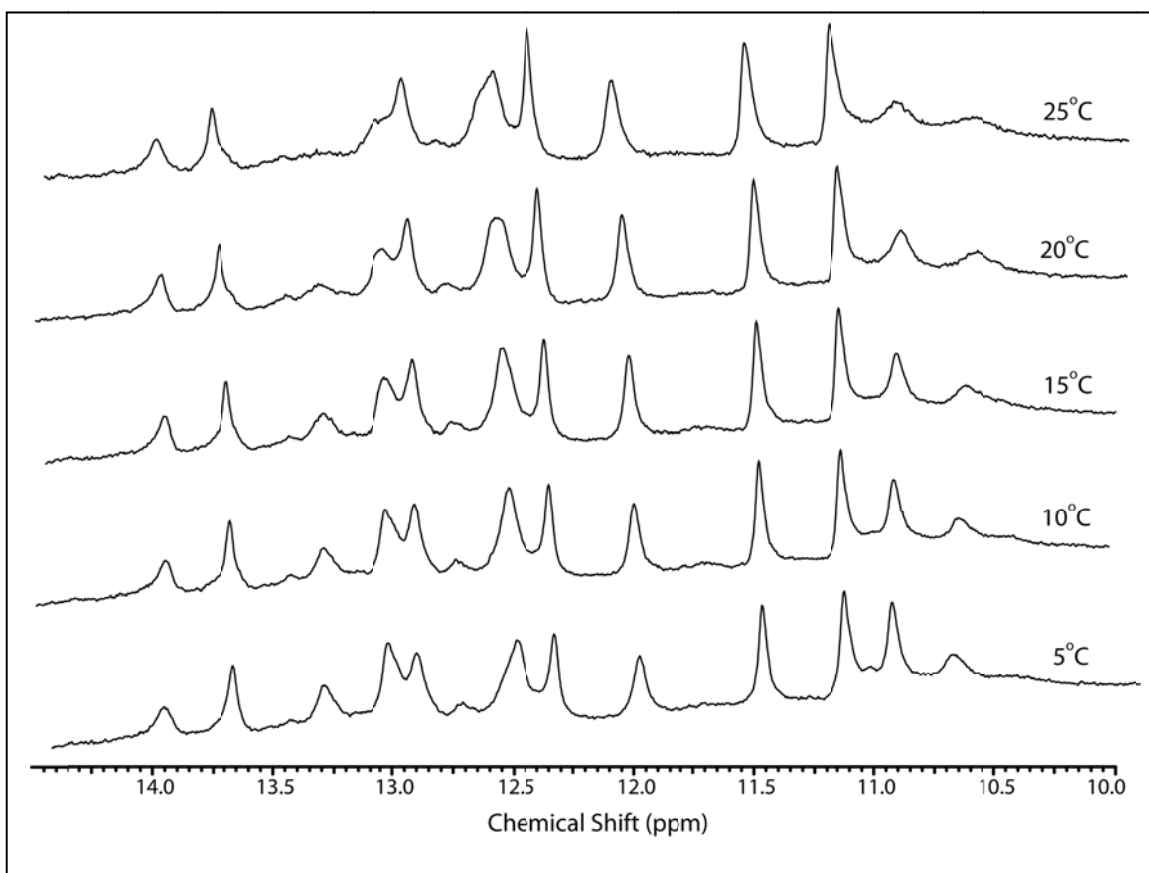


Figure 3-8: Temperature-dependent 1D imino spectrum of extended GAAA loop A. The spectrum was acquired using 600 MHz NMR in 10 mM phosphate, 100 μ M EDTA, pH 5.5 buffer conditions.

3.3.3 NMR assignments

To enable further spectroscopic analysis of the system, we pursued ^1H , ^{15}N , and ^{13}C NMR assignments in the extended-GAAA construct of the hairpin ribozyme loop A. NMR data were primarily acquired in phosphate buffer without added NaCl to limit duplex formation. Our results show that, as expected, hairpin structures were preferred under low salt conditions as indicated by the observation of imino peak at ~ 10.5 ppm (**Figure 3-8**). Starting from the imino resonance in the G-A base-pair of the GAAA tetraloop (with a characteristic chemical shift of ~ 10.5 ppm), the imino-imino region of a two-dimensional NOESY spectrum displays the sequential connectivity for the Watson–Crick base-pairs in stem I (**Figure 3-9**) of loop A. Regions of A-form helical RNA were assigned via a sequential walk along the backbone, since neighboring imino protons are ~ 4 Å apart and give rise to a moderate NOE cross-peak at longer mixing times. The G-U wobble base-pair was easily identified as a strong cross-peak between the resonances of G-H1 and U-H3 imino protons at chemical shifts of 10 to 12 ppm. A 2D NOESY experiment correlating exchangeable protons was acquired at 900 MHz for GAAA loop A and the GAAA extended loop A. For the non-extended GAAA loop A construct, imino-imino cross-peaks were mostly visible in stem I but not stem II. In the GAAA extended loop A construct, designed to stabilize stem II, there were cross-peaks in both helices, with very intense G-U cross-peaks observed at 11.17 and 11.5 ppm, resulting in the assignment of G-5 and U14 respectively. Imino NOE cross-peak pattern from the NOESY spectrum showed a base-pairing scheme consistent with the expected secondary structure of (**Figure 3-9**). The GAAA extended loop A was the best-behaved of all the constructs we tested and was thus selected as the primary sequence for all our NMR experiments to facilitate complete resonance assignment for loop A reported herein.

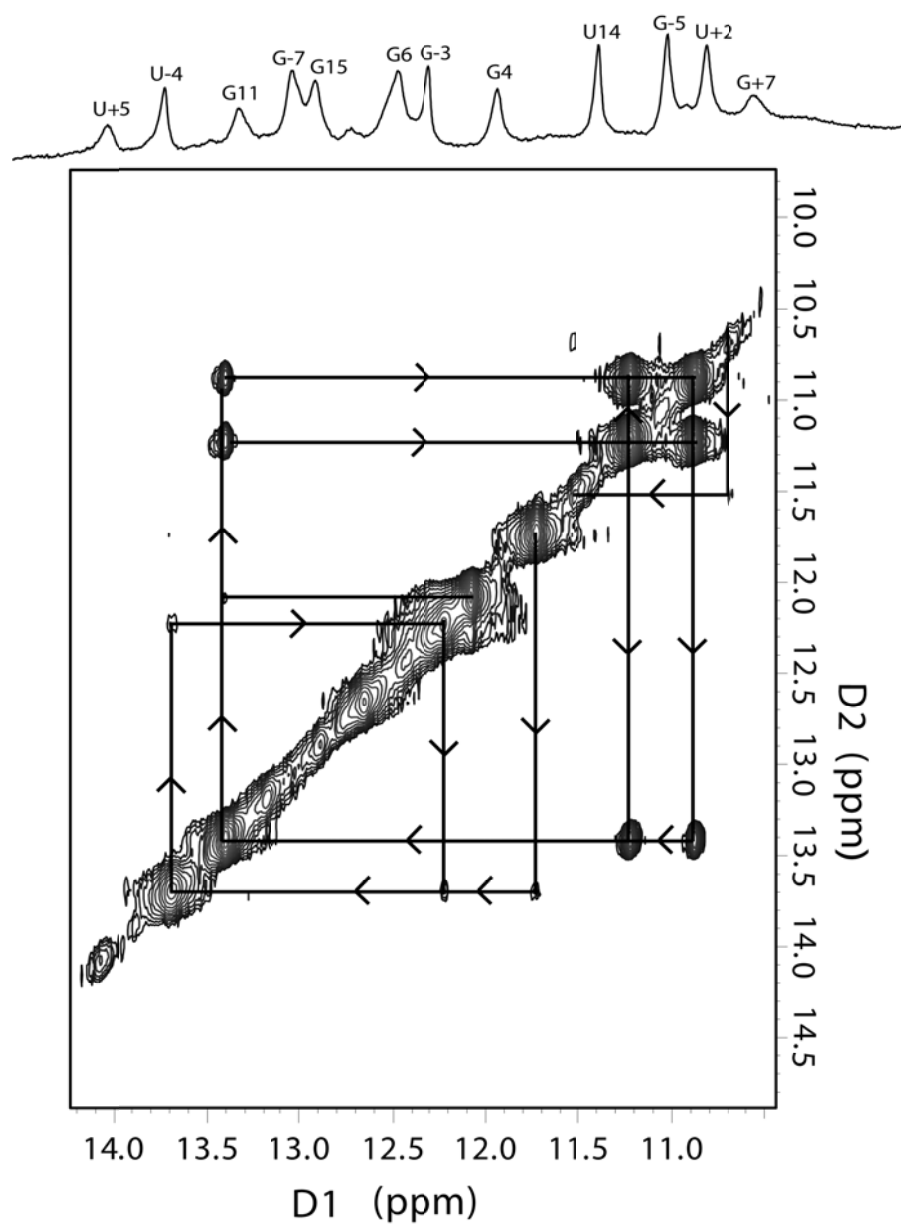


Figure 3-9: 2D imino NOESY spectrum of GAAA extended loop A. D1 and D2 axis are ¹H chemical shift at direct (D1) and indirect (D2) dimensions respectively. The spectrum was acquired by 900 MHz at a mixing time of 0.3 s. Buffer conditions: 10 mM phosphate, 100 μ M EDTA, pH 5.5.

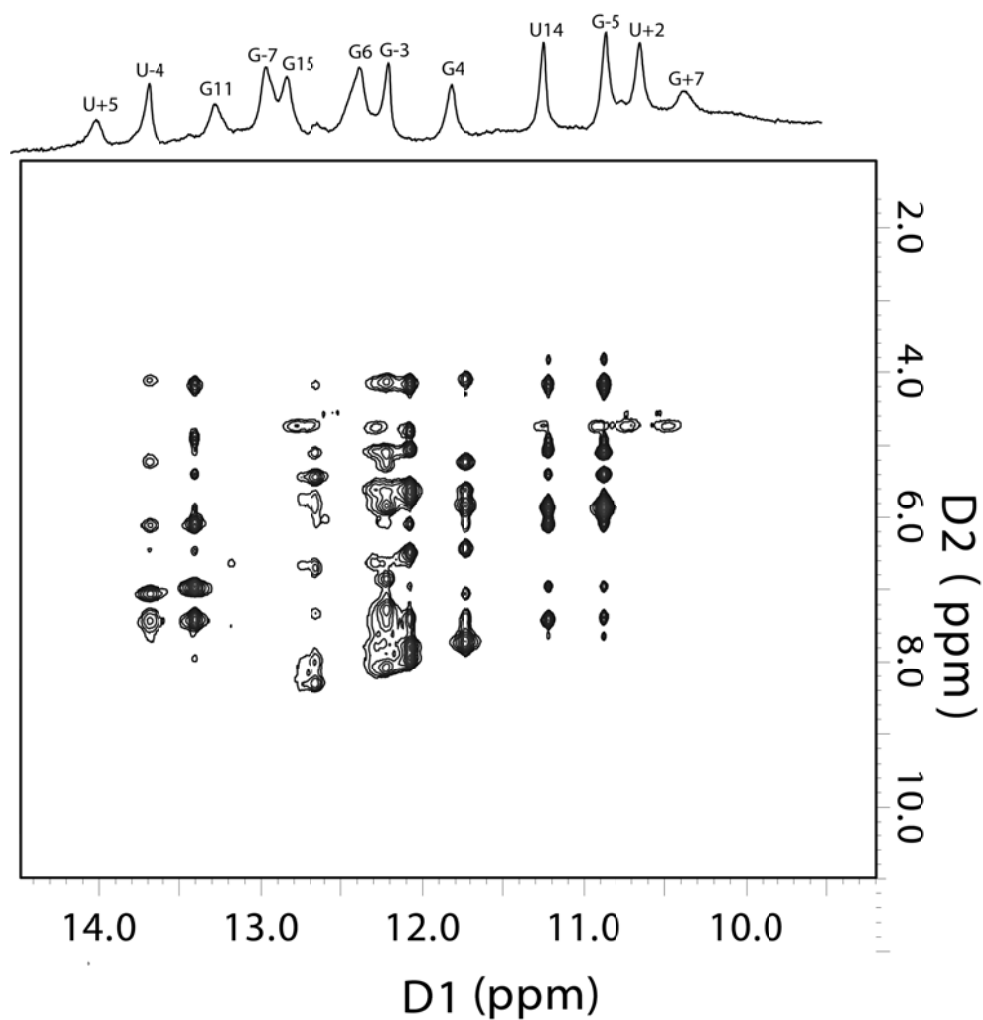


Figure 3-10: 2D amino NOESY spectrum of GAAA extended loop A. These spectra were acquired using 900 MHz at 300 ms mixing time in 10 mM phosphate, 100 μ M EDTA, pH 5.5, 10^oC.

3.3.4 Assignment of exchangeable proton resonances

Our assignment process was initiated by 1D imino spectrum. The peak at 10.63 ppm was easily assigned to the tetraloop G (G+7) because G imino of a G-A sheared base pair resonates at around 10.5-10.8 ppm. The G-U wobble mismatch was also easily assigned based on intense NOE cross-peaks (**Figure 3-9**). The small number of uridines and their downfield ^1H (13.5-14.5 ppm) ^{15}N resonance frequencies (153-159 ppm) facilitated an easier assignment of imino uridines in loop A. Consequently, U-4, U+5 and U14 were assigned as shown in **Table 3-1**. G-5 was assigned from its intense NOE connectivity to U14, which led to the assignment of all the remaining guanine imino peaks. Both aminos and imino protons were detected by 2D exchangeable NOESY (**Figure 3-9**, **Figure 3-10**), and 2D ^1H , ^{15}N HSQC (**Figure 3-11**). The proton resonance peak frequencies between 1D ^1H and 2D ^1H , ^{15}N -HSQC imino peaks were very similar in the proton dimension validating the frequencies obtained for each peak. The corresponding ^{15}N resonances were also determined and tabulated alongside their proton resonances (**Table 3-1**).

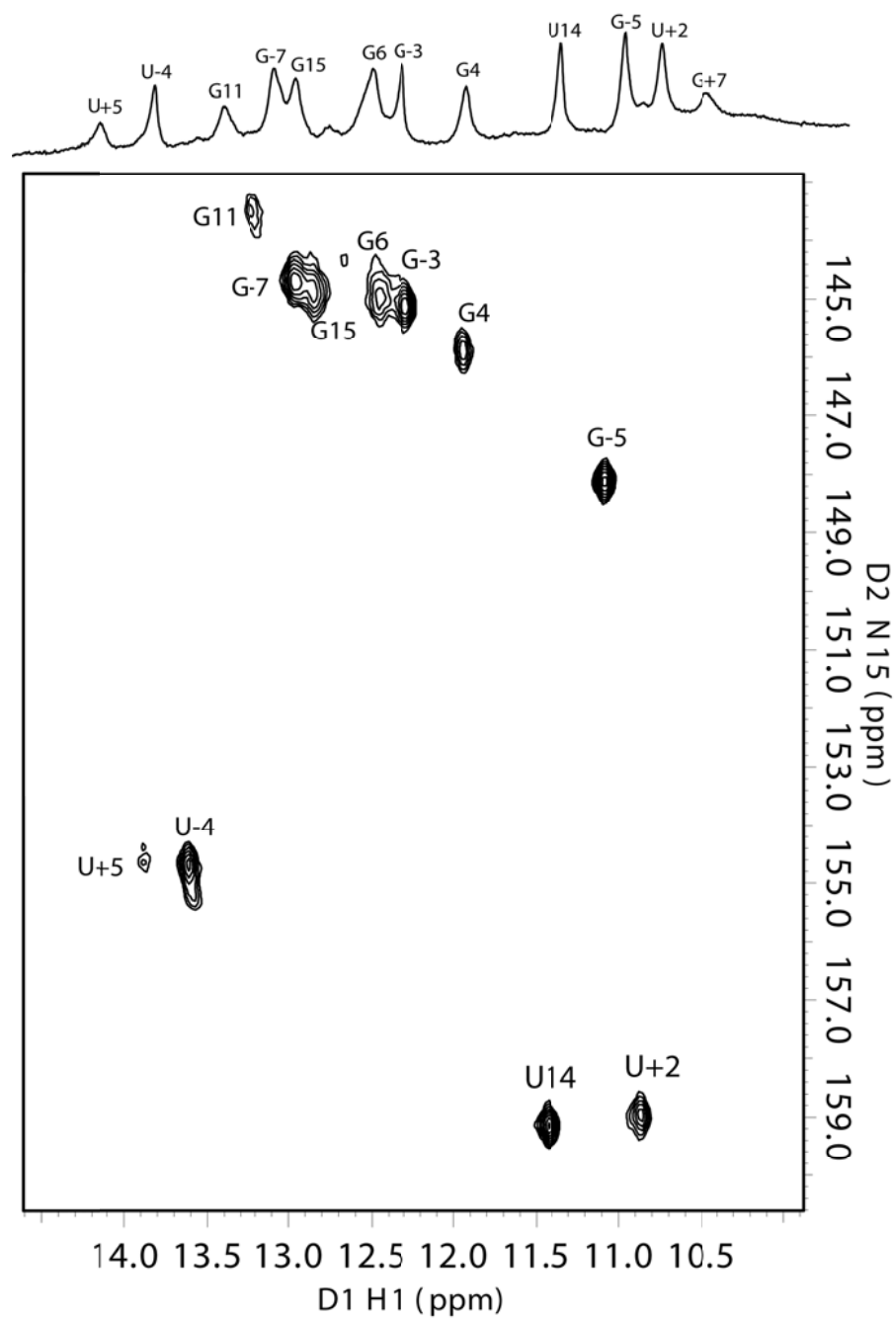


Figure 3-11: 2D imino ^1H , ^{15}N HSQC spectrum of GAAA extended loop A.
 Data was acquired at 600 MHz in 10 mM phosphate, 100 μM EDTA, pH 5.5.

Table 3-1: Resonance assignments for loop A (GAAA LpAext) using NMR experiments

¹ H ppm	¹⁵ N ppm	Assignment
10.63	-	G+7
10.92	159.04	U+2
11.17	148.08	G-5
11.51	159.20	U14
12.05	145.78	G4
12.40	145.06	G-3
12.58	144.73	G6
12.96	144.8	G15
13.04	144.63	G-7
13.29	143.48	G11
13.72	154.72	U-4
13.97	154.65	U+5

3.3.5 Non exchangeable protons

The base and sugar non-exchangeable protons were detected using ribose-optimized 2D ^1H , ^{13}C -HSQC, aromatic-optimized 2D ^1H , ^{13}C -HSQC, 3D HCCH-COSY, 3D HCCH-TOCSY and 3D ^{13}C -edited NOESY-HSQC spectra. ^1H - ^{13}C resonances were observed from ^1H , ^{13}C -HSQC experiments acquired at 600 MHz (**Figure 3-12**, **Figure 3-13**) for loop A. For the ribose resonances, heavy overlap was observed for H1'/C1', H2'/C2', H3'/C3', H4'/C4', and H5'/C5 cross-peaks (**Figure 3-12**), as is typical of RNA molecules within this size. Extensive overlap observed in the ribose resonances can be an impediment for accurate chemical shift assignments in RNA. However, further use of three-dimensional RNA correlation experiments can reduce overlap and resonance ambiguity by the detection of cross-peaks from multiple resonances within a spin system. For example, a 3D HCCH-COSY experiment was used to correlate covalently bound ^1H - ^1H - ^{13}C spins usually present within the ribose ring. The HCCH-COSY pulse sequence transfers magnetization to adjacent protons through one-bond ^{13}C - ^{13}C coupling within the H-C-C-H spin system. The ^1H - ^1H HCCH-COSY sub-spectrum correlating H1' and H2' proton chemical shifts in loop A has been presented (**Figure 3-14**). Twenty two well-resolved H1'-H2' cross-peaks (in both upper and lower quadrants) were observed and their chemical shift resonances have also been tabulated (**Table 3-4**). To provide comprehensive chemical shift assignments for ribose protons, a 3D HCCH-TOCSY data was also acquired (data not shown). HCCH-TOCSY experiments are similar to HCCH-COSY experiments except that the HCCH-TOCSY pulse sequence uses a mixing period where magnetization is transferable through several ^{13}C - ^{13}C bonds that connect all ^1H - ^{13}C pairs within the ribose. This provides additional H3', H4', and H5' correlations to H1' which facilitates assignment of the entire ribose. The aromatic optimized ^1H , ^{13}C -HSQC data (**Figure 3-13**) provided a fingerprint spectrum for

loop A, especially with the peaks within the 140-150 ppm range in the ^{13}C dimension. This region contained approximately 34 observable peaks, some of which were the paired upfield and downfield components of ^{13}C - ^{13}C doublets for pyrimidine residues. These peaks were interpreted as C6/C8 peaks reflecting the number of nucleotides in our sequence (30 nucleotides). The interpretation of this spectrum was confirmed by a ^1H , ^{13}C CT-HSQC spectrum of loop A (**Figure 3-14**), which eliminated the C5-C6 dipolar splitting in pyrimidines and effectively reduced the number of peaks observed to twenty-five. In RNA of this size, some resonances will not give observable peaks due to overlap, end fraying, or other effects, whereas end heterogeneity will sometimes give rise to spurious resonances; thus, this spectrum is roughly consistent with expectations for a 30mer RNA. It should be noted that the sensitivity of ^1H , ^{13}C CT-HSQC is reduced during constant time acquisition, and this spectrum thus had low signal to noise and digital resolution. The peaks between 155 – 160 ppm were assigned as adenine C2 peaks indicative of the number of adenines in the construct (**Table 3-3**). However, we observed seven peaks in the adenine C2 spectral range, instead of the expected nine peaks, due to resonance overlap in some residues. Using the 3D HCCH-COSY data (that provided us with H1'-H2' intra-residue resonances, **Figure 3-15**) and the aromatic ^1H , ^{13}C -HSQC (which provided the H6/H8 resonances), we also obtained a ^{13}C -edited NOESY-HSQC spectrum (**Figure 3-16**) to make the intra and inter-residue NOE sequential walk to correlate and assign base H6/H8 protons to the ribose H1' and H2' protons respectively (**Figure 3-17**). The detailed ^{13}C peak assignments using the three-dimensional NOESY spectrum is still in progress. As at the time of writing this chapter, we had made good progress in loop A assignments to the extent of correlating the ^{13}C -NOESY-HSQC connectivity with 3D-HCCH COSY and ^{13}C -HSQC data to finish the aromatic residue assignments especially of the ^1H , ^{13}C -HSQC resonances. We used ^1H , ^{13}C -CT-HSQC

(**Figure 3-14**), to distinguish between the resonances of purines (positively phase) and pyrimidines (negatively phased) as we continue with the NOE sequential walk (**Table 3-5**). Pyrimidines were mostly clustered in one region of the spectrum while the purines were well dispersed within the ^1H , ^{13}C -CT-HSQC spectrum. Several pyrimidine doublet peaks were collapsed into single peaks during ^1H , ^{13}C -CT-HSQC experiment thus facilitating their easy assignment. To isolate guanine from adenine among the purines, we used a 3D-Adenine HCCH-TOCSY to record adenine resonances within the aromatic ^1H , ^{13}C -HSQC region, also presented in **Table 3-5**. Other additional assignments for the GAAA tetraloop and loop residues were accomplished by resonance comparison with A730 loop of the *Neurospora* Varkud satellite (VS) ribozyme (33) and the partial assignments reported for the mutant form of hairpin ribozyme loop A (20). Tables 3-2 through 3-5 summarize the current state of assignments for the extended-GAAA loop A construct. Given the good behavior of the construct and the quality of data obtained so far, we look forward to completing assignments for aromatic and ribose resonances for further dynamics studies.

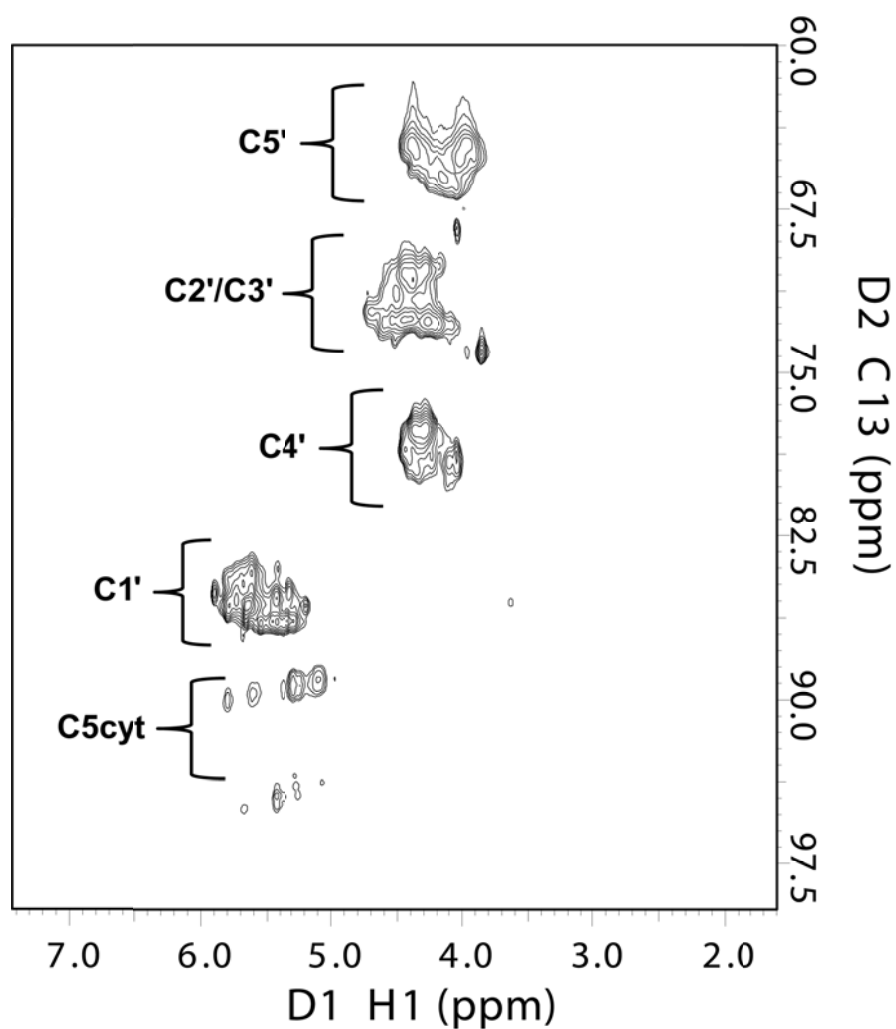


Figure 3-12: 2D ^1H , ^{13}C -HSQC contour spectrum of the ribose region in loop A. Data was acquired in 100% D_2O at 600 MHz NMR in the following buffer conditions; 10 mM phosphate, 100 μM EDTA, pH 5.5.

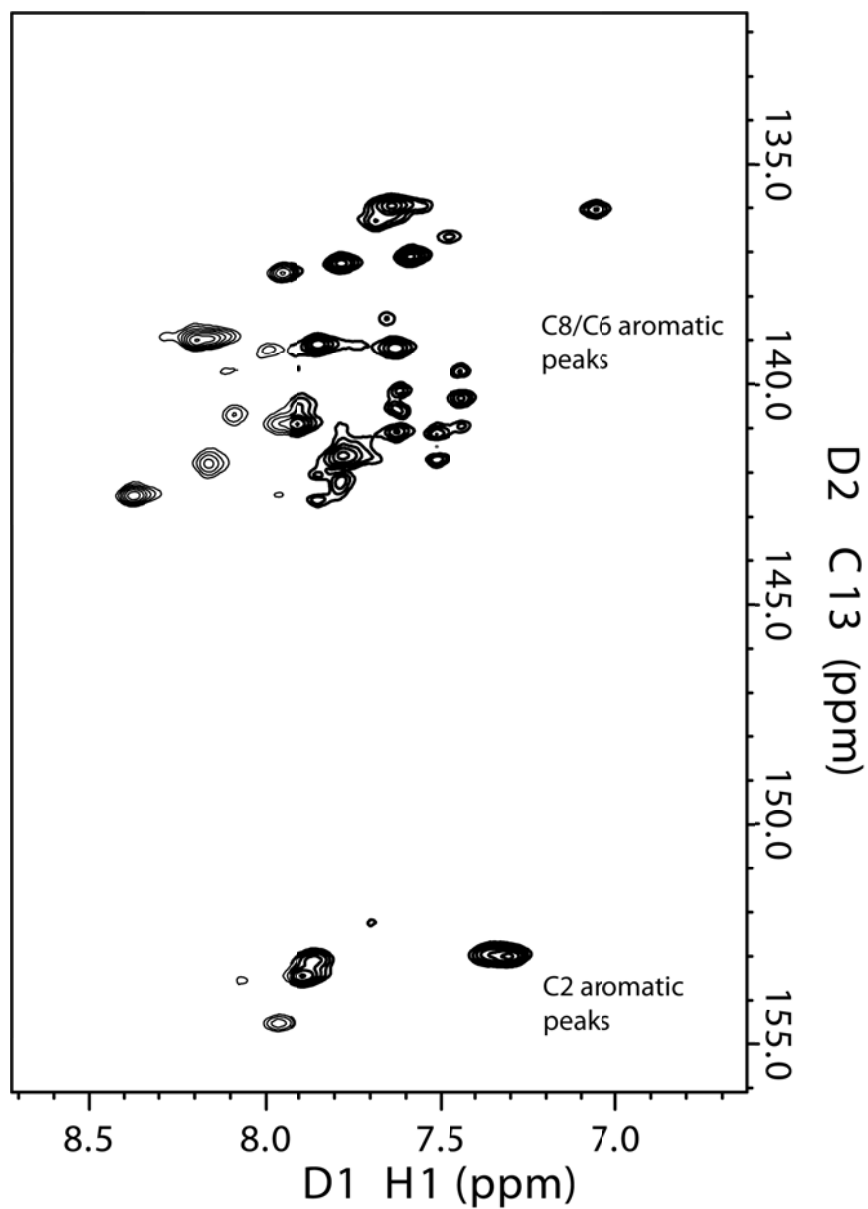


Figure 3-13: 2D ^1H , ^{13}C -HSQC spectrum of aromatic region in loop A. Data was acquired in 100% D_2O at 600 MHz NMR in the following buffer conditions; 10 mM phosphate, 100 μM EDTA, pH 5.5.

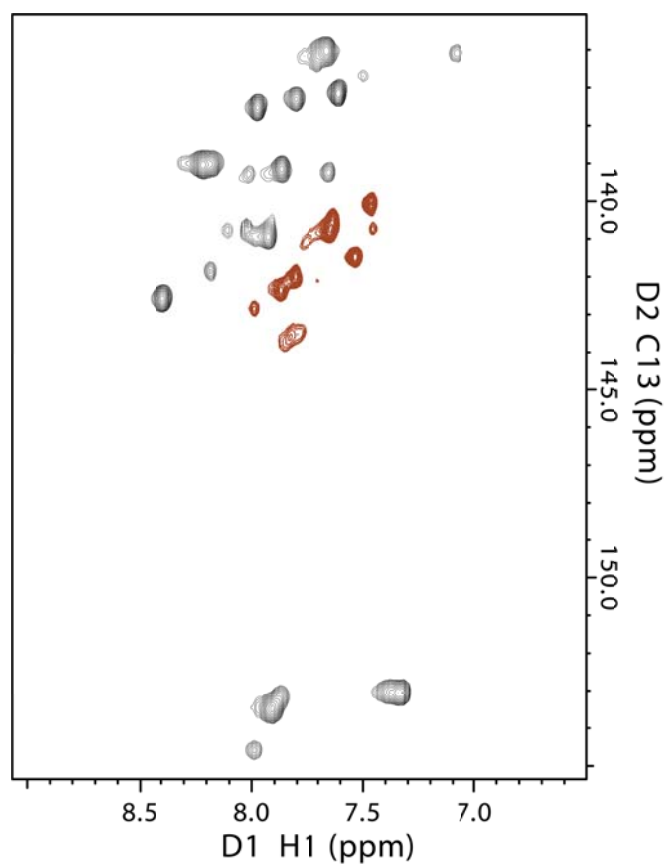


Figure 3-14: 2D ^1H , ^{13}C CT-HSQC spectrum of aromatic region in loop A acquired at 600 MHz NMR. The red peaks are negatively phased pyrimidine peaks, while the black peaks represent the positively-phased purine residues

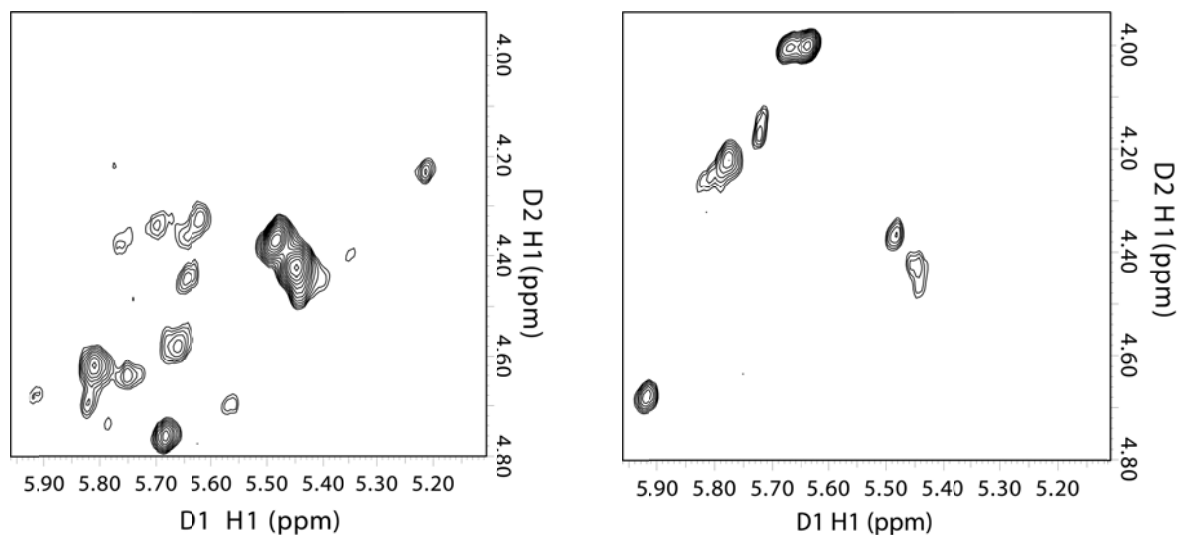


Figure 3-15: Representative 2D planes of 3D HCCH-COSY spectrum of the ribose region in loop A. Data was acquired in 100% D₂O at 600 MHz NMR in buffer conditions as follows; 10 mM phosphate, 100 μ M EDTA, pH 5.5.

Table 3-2: Non exchangeable assignments of H6/H8 aromatic protons in loop A

H8/H6 (ppm)	C6/C8 (ppm)	Assignment
7.06	136.01	G11
7.42	136.57	G+7
7.48	139.78	Peak 14
7.48	140.28	Peak 24
7.50	140.77	Peak 23
7.53	141.07	Peak 22
7.53	141.62	Peak 33
7.56	135.89	Peak 4
7.58	137.05	Peak 5
7.61	136.43	Peak 3
7.62	139.11	Peak 13
7.63	135.95	Peak 3
7.65	140.95	Peak 21
7.65	140.22	Peak 19
7.66	138.26	G8
7.73	141.68	Peak 27
7.74	142.23	Peak 29
7.76	137.26	Peak 6
7.79	140.92	Peak 17
7.81	140.92	Peak 18
7.81	141.54	Peak 27
7.82	142.15	Peak 28
7.87	143.47	Peak 32
7.89	139.13	A2
7.91	140.45	A10
7.92	142.62	Peak 30
7.98	142.77	Peak 31
7.99	137.52	G4
8.03	139.22	Peak 9
8.13	139.75	A3
8.18	141.57	A9
8.19	139.02	Peak 8
8.20	140.48	Peak 16
8.38	142.24	Peak 26

Table 3-3: Non exchangeable assignments of adenine H2 aromatic protons in loop A

H2 (ppm)	C2 (ppm)	Assignment
7.35	153.08	Peak 41
7.42	153.01	Peak 38
7.86	153.12	Peak 37
7.89	153.44	Peak 36
7.85	154.31	Peak 37
7.96	154.51	Peak 34
8.07	153.55	A1

Table 3-4: Assignments of H1'/H2' and C1'/C2' in loop A

H1' (ppm)	H2' (ppm)	C1' (ppm)	C2' (ppm)
5.35	4.22	93.20	76.10
5.49	4.41	94.00	75.60
5.49	4.53	94.00	75.60
5.49	4.59	94.00	75.70
5.60	4.42	92.80	76.10
5.61	4.36	94.00	76.00
5.71	4.69	94.40	75.60
5.78	4.44	93.00	75.70
5.78	4.84	91.00	75.20
5.79	4.55	93.20	75.70
5.81	4.57	93.20	75.70
5.82	4.01	93.20	77.80
5.82	4.69	94.00	75.60
5.82	4.75	93.80	76.10
5.85	4.13	92.40	77.40
5.89	4.63	92.20	76.50
5.89	4.87	93.00	75.20
5.92	4.21	92.40	77.30
5.95	4.62	93.20	76.10
5.96	4.69	93.60	76.00
6.05	4.11	92.00	78.20
6.06	4.67	92.60	77.10

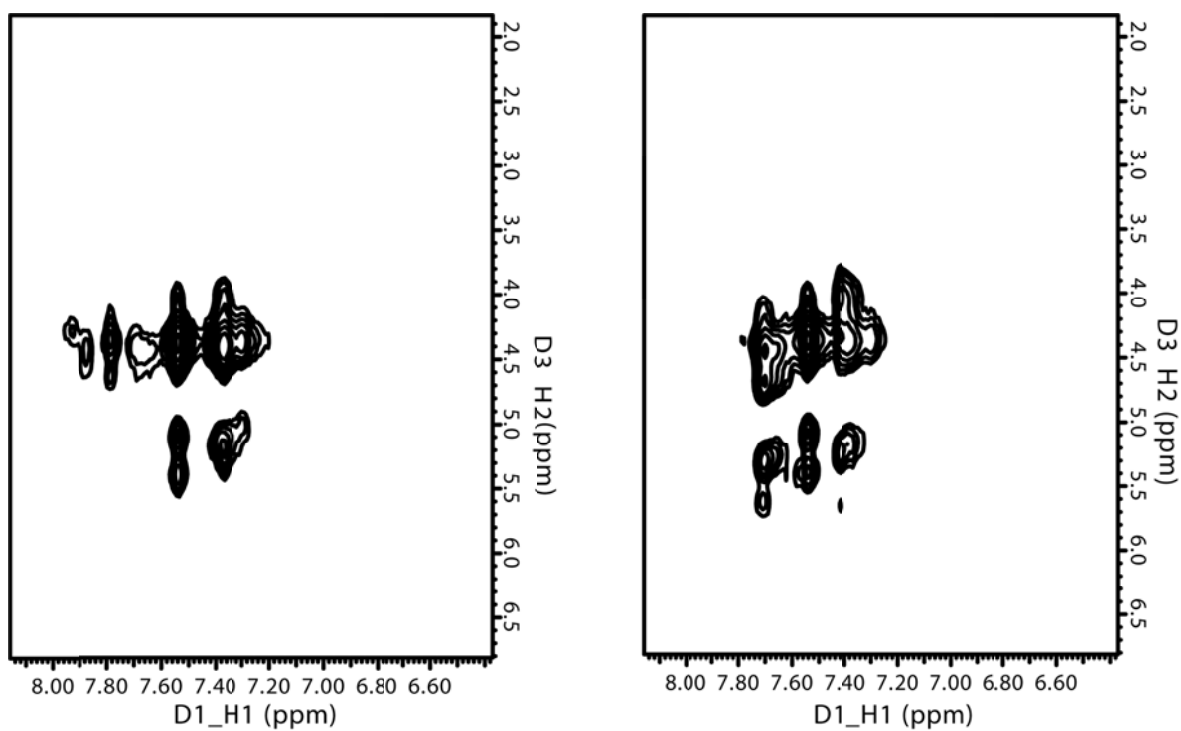


Figure 3-16: Representative 2D planes of ^{13}C -edited NOESY-HSQC spectrum of loop A. Data was acquired at 900 MHz in the following buffer condition; 10 mM phosphate, 100 μM EDTA, pH 5.5.

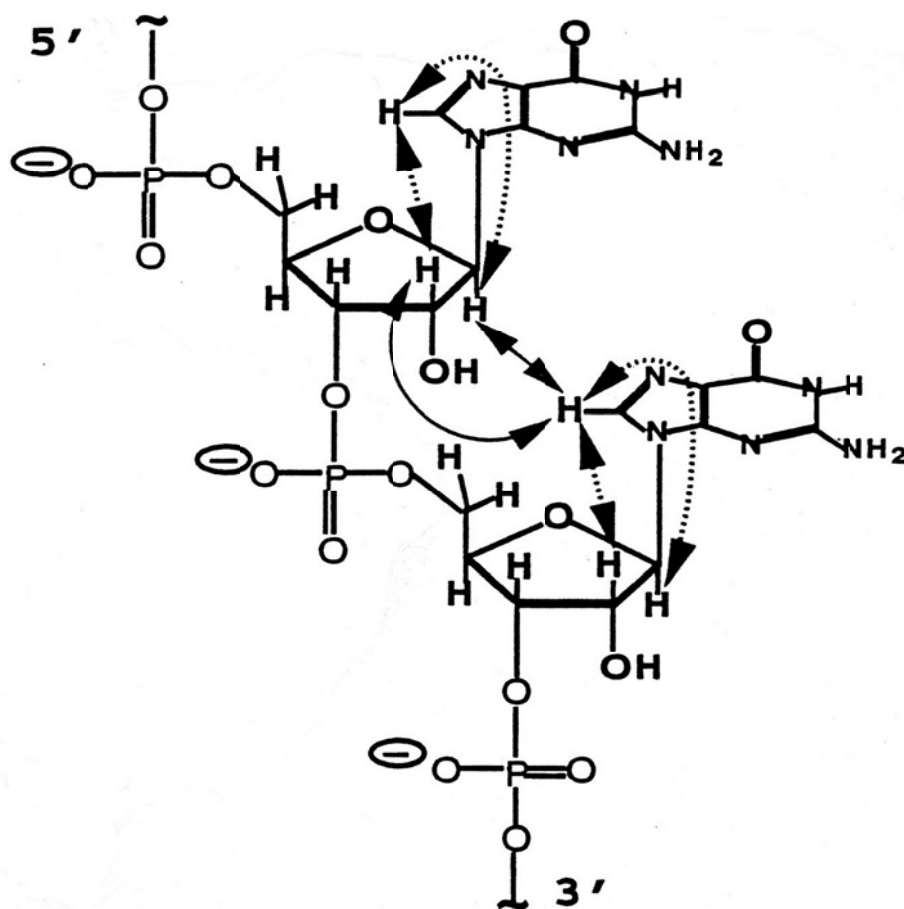


Figure 3-17: Schematic representation of NOE connectivities that correlate intra- and inter-residue H8/H8 to H1' and H2'. Intra-residue NOEs are represented by dashed lines while inter-residue NOEs are represented by solid lines with arrows. Reprinted with permission from Pascale Legault doctoral thesis

Table 3-5: Aromatic resonance assignments of H1'/H2' and C1'/C2' in loop A

H8/H6 (ppm)	C6/C8 (ppm)	Assignment
7.06	136.01	G11
7.42	136.57	G+7
7.48	139.78	purine
7.48	140.28	pyrimidine
7.50	140.77	pyrimidine
7.53	141.07	pyrimidine
7.53	141.62	purine
7.56	135.89	purine
7.58	137.05	purine
7.61	136.43	purine
7.62	139.11	purine
7.63	135.95	purine
7.65	140.95	pyrimidine
7.65	140.22	pyrimidine
7.66	138.26	G8
7.73	141.68	purine
7.74	142.23	pyrimidine
7.76	137.26	purine
7.79	140.92	purine
7.81	140.92	purine
7.81	141.54	purine
7.82	142.15	purine
7.87	143.47	pyrimidine
7.89	139.13	A2
7.91	140.45	A10
7.92	142.62	pyrimidine
7.98	142.77	purine
7.99	137.52	G4
8.03	139.22	purine
8.13	139.75	A3
8.18	141.57	A9
8.19	139.02	purine
8.20	140.48	purine
8.38	142.24	purine

3.4 DISCUSSION

Our docking studies of loop A and loop A(U+2C/C+3U) demonstrated that trans-constructs with wild type sequence remained active under docking condition, and that this activity was metal-driven. However, the double mutant loop A(U+2C/C+3U) lacked activity, as expected, not because of the removal of the important functional groups or disruption of the structure but because of the stabilization of the ground-state conformer that hinder transitions to the docked state. The lack of docking transitions imparted by additional base-pairing featuring U+3:A7 and C+2:G8 base-pairs, emphasizes the importance of conformational fluctuations in the formation of the docked state. In the mutant sequence, the catalytically necessary G8 was stabilized by formation of G8-C+2 base pair. This base-pair locks loop A into its ground-state conformation leading to a stabilized form of undocked loop A. The conformational stabilization of loop A(U+2C/C+3U) is supported by NMR spectra (**Figure 3-5**) which mostly showed similar chemical shift resonance patterns for exchangeable protons in both mutant loop A(U+2C/C+3U) and wild type loop A implying similar global structure adopted by the two constructs. The use of Circular dichroism (CD) in measuring structural rearrangements that characterize docking at equilibrium conditions is a robust technique with the ability to monitor such secondary and tertiary structural changes. The docking data obtained on U+2C/C+3U loop A suggests docking perturbation based on impaired structural transitions visible by CD. Since, the native loop A recorded we still acknowledge the challenges that may come with interpreting the data obtained from this double mutant sequence, loop A(U+2C/C+3U). It is possible that the two additional base-pairs may have virtually rendered the sequence non-functional. Our argument is that a single G8U mutation of this inactive sequence partially rescues the activity of this loop to dock with its docking partner, loop B (22).

The core of the optimization of the best-behaved construct for NMR studies was in the design of the two stems I and II while maintaining active wild-type sequence within the internal loop active site. The stabilization of both stems was useful in limiting our observation of unwanted duplex structures. Better results were achieved when stem I was capped with a GAAA instead of GUAA tetraloop. GUAA may destabilize hairpin formation by formation of an intra-sequence A-U base-pair which is not available for a GAAA tetraloop. Also, the stabilization of stem II with additional base-pairs generally stabilized loop A for further NMR studies. Incomplete ^1H , ^{13}C , and ^{15}N chemical shift assignments have been determined for loop A of the hairpin ribozyme. These assignments have been realized using standard through-space and through-bond correlated spectroscopy experiments with non-labeled as well as uniformly labeled loop A. Ambiguity caused by spectral overlap presents challenges in the assignment process. However, the use of two- and three-dimensional experiments facilitated the resolution of several overlapped peaks. Our observation of imino and amino proton peaks suggest thermally stable folded loop A RNA at least until 25°C, while the observation of a sheared G-A imino base-pair confirmed the formation of the hairpin structure. The observation of guanosine and uridine imino proton peaks was attributed to canonical Watson-Crick base pairs consistent in number and arrangement with our predicted secondary structures (**Figure 3-2**). Imino proton peaks for G8 and U+3 were observed in the mutant U+2C/C+3U sequence as predicted. Using ^1H , ^{15}N -HSQC, imino and amino spin pairs were identified, which led to imino assignments (**Table 3-1**, **Figure 3-11**). New imino and amino proton peaks were detected in extended stem I (U+2C/C+3U loop A, **Figure 3-1 D**) and stem II (Loop A extended, **Figure 3-1 C**) which facilitated the imino assignment process. These new peaks were localized proximal to the internal loop in U+2C/C+3U loop. The partial stability of the loop region in the U+2C/C+3U loop A may

support our proposal that reduced mutant activity could impart decreased flexibility of internal loop residues. To validate this proposition, more mutants with minimum wild type changes, especially mutants that perturb docking, need to be screened. The presence of a G-U mismatch registered strong imino NOEs which served as our starting point for imino assignments. The incorporation of a G-U mismatch in a stem is an important strategy since the G-U NOEs are very strong. We are in the process of assigning all the resonances of the wildtype extended loop A construct.

The spectrum of the non-exchangeable protons in loop A was well dispersed in the aromatic region but not well resolved in the ribose region due to the degeneracy of the ribose protons within their environment. Aromatic C8/C6 and C2 resonances were clearly resolved albeit with a few overlaps especially in the pyrimidine region. This was expected especially due to the degeneracy of C6 protons in uridines and cytosine. The identification of the ribose protons was initiated by a combination of 3D HCCH-TOCSY (34) and HCCH-COSY (35) experiments. The HCCH-TOCSY correlates all protons in each ribose spin system, allowing for complete assignment of ribose ^1H and ^{13}C . We identified twenty one H1'/H2' cross-peaks lying in symmetric quadrants which facilitated the assignment of C2' symmetric resonances based on correlation with H2' chemical shifts. The C2' and C3' resonate at approximately the same frequency, thus required the collection of the well resolved 3D-HCCH-TOCSY. A 3D ^{13}C -edited NOESY was used to extend the 2D ^1H , ^1H NOESY into 3D based on the chemical shift of the ^{13}C attached to one of the protons. This method greatly improves ^1H signal overlap with sufficient ^{13}C chemical shift dispersion. Using the NOESY walk, significant assignments can be made to complete the process. Assignment process can be challenging especially due to resonance overlap and the lack of clear starting points to facilitate the necessary connectivity. Thus far, we

have made significant progress in loop A resonance assignment. The exchangeable imino protons and nitrogen were unambiguously assigned but some work still remains to fully assign aromatic ^1H and ^{13}C resonances. The finished assignments are important for the detailed NMR spin-relaxation experiments of ^{13}C dynamics which has been comprehensively discussed in the next chapter (Chapter 4) of this thesis. Where necessary, we have reported the residue dynamics results by peak number rather than assigned atom, since complete resonance assignments is not necessary for ^{13}C relaxation acquisition but will be useful in comprehensive residue dynamics analysis.

ACKNOWLEDGEMENT

I wish to acknowledge Dr. Minako Sumita for doing the Circular Dichroism (CD) docking work between wildtype loop A and loop B

REFERENCES

REFERENCES

1. Madhani, H. D., and Guthrie, C. (1994) Dynamic RNA-RNA interactions in the spliceosome, *Annu Rev Genet* 28, 1-26.
2. Wahl, M. C., Will, C. L., and Luhrmann, R. (2009) The spliceosome: design principles of a dynamic RNP machine, *Cell* 136, 701-718.
3. Hampel, A., and Tritz, R. (1989) RNA catalytic properties of the minimum (-)sTRSV sequence, *Biochemistry* 28, 4929-4933.
4. 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, *P Natl Acad Sci USA* 83, 8859-8862.
5. Butcher, S. E., Heckman, J. E., and Burke, J. M. (1995) Reconstitution of hairpin ribozyme activity following separation of functional domains, *J Biol Chem* 270, 29648-29651.
6. 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.
7. Buzayan, J. M., Hampel, A., and Bruening, G. (1986) Nucleotide sequence and newly formed phosphodiester bond of spontaneously ligated satellite tobacco ringspot virus RNA, *Nucleic Acids Res* 14, 9729-9743.
8. Symons, R. H. (1992) Small catalytic RNAs, *Annu Rev Biochem* 61, 641-671.
9. Perrotta, A. T., and Been, M. D. (1992) Cleavage of oligoribonucleotides by a ribozyme derived from the hepatitis delta virus RNA sequence, *Biochemistry* 31, 16-21.
10. Lilley, D. M. (2004) The Varkud satellite ribozyme, *RNA* 10, 151-158.
11. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A., and Breaker, R. R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme, *Nature* 428, 281-286.
12. Esteban, J. A., Banerjee, A. R., and Burke, J. M. (1997) Kinetic mechanism of the hairpin ribozyme. Identification and characterization of two nonexchangeable conformations, *J Biol Chem* 272, 13629-13639.
13. Fedor, M. J. (1999) Tertiary structure stabilization promotes hairpin ribozyme ligation, *Biochemistry* 38, 11040-11050.

14. Bevilacqua, P. C., and Yajima, R. (2006) Nucleobase catalysis in ribozyme mechanism, *Curr Opin Chem Biol* 10, 455-464.
15. Doudna, J. A., and Lorsch, J. R. (2005) Ribozyme catalysis: not different, just worse, *Nat Struct Mol Biol* 12, 395-402.
16. Shippy, R., Siwkowski, A., and Hampel, A. (1998) Mutational analysis of loops 1 and 5 of the hairpin ribozyme, *Biochemistry* 37, 564-570.
17. Komatsu, Y., Kumagai, I., and Ohtsuka, E. (1999) Investigation of the recognition of an important uridine in an internal loop of a hairpin ribozyme prepared using post-synthetically modified oligonucleotides, *Nucleic Acids Res* 27, 4314-4323.
18. Perez-Ruiz, M., Barroso-DelJesus, A., and Berzal-Herranz, A. (1999) Specificity of the hairpin ribozyme. Sequence requirements surrounding the cleavage site, *J Biol Chem* 274, 29376-29380.
19. Chowrira, B. M., Berzal-Herranz, A., and Burke, J. M. (1991) Novel guanosine requirement for catalysis by the hairpin ribozyme, *Nature* 354, 320-322.
20. Cai, Z., and Tinoco, I., Jr. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite, *Biochemistry* 35, 6026-6036.
21. Walter, N. G., Chan, P. A., Hampel, K. J., Millar, D. P., and Burke, J. M. (2001) A base change in the catalytic core of the hairpin ribozyme perturbs function but not domain docking, *Biochemistry* 40, 2580-2587.
22. Sargueil, B., Hampel, K. J., Lambert, D., and Burke, J. M. (2003) In vitro selection of second site revertants analysis of the hairpin ribozyme active site, *J Biol Chem* 278, 52783-52791.
23. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis, *Nature* 410, 780-786.
24. 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.
25. Torelli, A. T., Krucinska, J., and Wedekind, J. E. (2007) A comparison of vanadate to a 2'-5' linkage at the active site of a small ribozyme suggests a role for water in transition-state stabilization, *RNA-A Publication of the RNA Society* 13, 1052-1070.
26. Ojwang, J. O., Hampel, A., Looney, D. J., Wong-Staal, F., and Rappaport, J. (1992) Inhibition of human immunodeficiency virus type 1 expression by a hairpin ribozyme, *P Natl Acad Sci USA* 89, 10802-10806.

27. 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.
28. 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 Res* 15, 8783-8798.
29. Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods Enzymol* 180, 51-62.
30. Kim, I., McKenna, S. A., Viani Puglisi, E., and Puglisi, J. D. (2007) Rapid purification of RNAs using fast performance liquid chromatography (FPLC), *RNA* 13, 289-294.
31. Plateau, P., Gueron, M. (1982) Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences, *J. Am. Chem. Soc* 104, 7310-7311.
32. Rupert, P. B., Massey, A. P., Sigurdsson, S. T., and Ferre-D'Amare, A. R. (2002) Transition state stabilization by a catalytic RNA, *Science* 298, 1421-1424.
33. Desjardins, G., Bonneau, E., Girard, N., Boisbouvier, J., and Legault, P. (2011) NMR structure of the A730 loop of the Neurospora VS ribozyme: insights into the formation of the active site, *Nucleic Acids Res* 39, 4427-4437.
34. Nikonowicz, E. P., and Pardi, A. (1992) Three-dimensional heteronuclear NMR studies of RNA, *Nature* 355, 184-186.
35. Bax, A., Clore, G. M., Driscoll, P. C., Gronenborn, A. M., Ikura, M., and Kay, L. E. (1990) Practical Aspects of Proton Carbon Carbon Proton 3-Dimensional Correlation Spectroscopy of C-13-Labeled Proteins, *J Magn Reson* 87, 620-627.

CHAPTER 4

INSIGHTS INTO THE INTERNAL DYNAMICS OF LOOP A USING NMR SPIN RELAXATION

ABSTRACT

Conformational dynamics is an important feature in proteins and RNA. However, there is little information outlining the relationship between dynamics and RNA function. Structural studies of the hairpin ribozyme have reported significant structural rearrangements at the active site upon domain-domain interaction, suggesting that dynamics may facilitate transitions from the ground to active state. In this work, we have used longitudinal (^{13}C T_1) and traverse (^{13}C $T_{1\rho}$) relaxation experiments as well as heteronuclear NOE to interrogate the internal dynamics and disorder within loop A. We have also used relaxation dispersion measurements of power-dependent $T_{1\rho}$ to understand conformational exchange of residues in the μs -ms timescale. Internal motions were observed on a wide variety of timescales within the non-helical regions of loop A suggesting a complex landscape of accessible states, and integrated correlations amongst the observed motions. These results demonstrate the usefulness of NMR spin relaxation measurements in probing dynamics of complex molecules exemplified in ribonucleic acids.

4.1 INTRODUCTION

Over the years, there has been an increasing appreciation of the critical role of conformational dynamics in many facets of molecular function (1, 2). With its ability to act as a site-specific probe of the dynamic properties of biomolecules, heteronuclear NMR spin relaxation has emerged as a leading technique for experimental studies of molecular dynamics. Relationships between spin relaxation data and function may be drawn in several ways. If the target structure of a conformational change is known, the extent and manner in which the molecule fluctuates toward that state can be assessed. More directly, it may be possible to selectively freeze out specific motional variables using mutation or modification and analyze the effects on function. Currently, there is an active debate about the relationship between dynamics and function. One model suggests a direct coupling between dynamic modes throughout the molecule to the chemical catalytic steps using dihydrofolate reductase (DHFR) as an example (3). Motions directly linked to catalysis in DHFR were investigated using NMR and the dynamic modes were found to be on similar timescales as chemical transformations. Similar correlations have also been reported in RNase A (4) and adenylate kinase (5). For RNA molecules, conformational dynamics play an equally important role in events such as RNA catalysis and ligand recognition (6-9). In catalytic RNA the role of dynamics is even more prominent than most protein enzymes. It appears that the conformational fluctuations from a stable ground state to an activated conformer are an inherent part of most ribozymes. One notable example includes the conformational rearrangements observed between crystal structures of precursor and product states in hepatitis delta virus (HDV) with catalytic or inhibitory metal ions binding only to the precursor form (10). Therefore, a detailed understanding of the structure-function relationship of ribozymes requires the characterization of both the static structure and conformational dynamics

of the molecule. Conformational rearrangements have been observed between isolated and docked hairpin ribozyme domain (loop A and B) structures (11-13) and the mechanism of these rearrangements underpin the focus of this study. These features suggest that conformational sampling is not only vital for conformational pre-organization but may also be involved in directing chemistry. Recently, intermolecular loop A and B interaction kinetic rates have been reported by our group (14). The docking association rate reveals a very slow docking process between loop A and loop B (five orders of magnitude below the diffusion limit) despite a tight docking affinity. This slow association rate as well as structural variability between the free and docked forms, especially for the key residues involved in the interaction, informs our hypothesis that isolated loops A and B may independently sample docking-competent states, raising the possibility of a double conformational capture docking mechanism.

The relationship between conformational dynamics and the catalytic cycle of the hairpin ribozyme is not well understood. The major loop rearrangements observed upon docking, as well as the kinetically unfavorable docking process, both argue for a tight and mechanistic role for conformational dynamics in the pre-organization of the catalytically-competent active site of the hairpin ribozyme. NMR spectroscopy presents us with a unique ability to investigate dynamic properties of molecules over a range of different timescales with atomic resolution (15-23) (**Figure 4-1**). The discovery and success of RNA isotope-labeling techniques (24, 25) permitted the use of ^{13}C and ^{15}N nuclei as points of investigating conformational dynamics in RNA (26, 27). Conformational transitions with links to catalytic function in ribozymes will potentially register within the microseconds to milliseconds time scale. Such rearrangements may be thermally-induced leading to exchange between two or more conformational states. Heteronuclear spin relaxation measurements are suitable for probing dynamics on a wide range

of timescale especially within picoseconds to milliseconds. For fast (ps-ns) timescale dynamics, longitudinal (T_1), transverse (T_2) relaxation, and heteronuclear NOE NMR techniques are mostly utilized (2). Also, deuterium NMR has previously been used for the study of fast motions in proteins as well as RNA systems (28). For slower dynamic processes (μ s to ms), transverse and rotating frame relaxation rates (R_2 and $R_{1\rho}$) are ordinarily measured as a function of the effective applied power (**Figure 4-1**). Both of these methods can be used to extract the rate of exchange (k_{ex}) between two distinct states (where k_{ex} is the sum of the rate constants in the equilibrium between the two states) since conformational exchange on this timescale contributes to R_2 and $R_{1\rho}$. For the simplest case of a two-site exchange model of A reversibly interconverting with B, the exchange rate constant, k_{ex} , is the sum of the forward and reverse rate constants, k_1 and k_{-1} .

$$k_{ex} = k_1 + k_{-1} \quad (4-1)$$

$$R_{1\rho} = \frac{1}{T_{1\rho}} = \frac{1}{T_{1\rho}^{\infty}} + P_a P_b (\Delta\omega)^2 \frac{\tau_{ex}}{1 + \omega_1^2 \tau_{ex}^2} \quad (4-2)$$

The nature of the NMR exchange regime depends on the relation between k_{ex} and $\Delta\omega$, the chemical shift difference between the populated states. For a system exchanging between two states, the transverse relaxation rate under spin-locked conditions, $R_{1\rho}$, is described by equation 4-2 (29) where $T_{1\rho}$ is the measured relaxation time, $T_{1\rho}^{\infty}$ is the relaxation time at infinite spin-lock power, P_a and P_b are the fractional populations of conformations a and b (assuming two-site exchange), $\Delta\omega$ is the chemical shift difference between states a and b , $\omega_1 = \gamma B_1$ is the spin-lock power expressed in radians per second, and τ_{ex} is the lifetime for the exchange process calculated as $1/k_{ex}$.

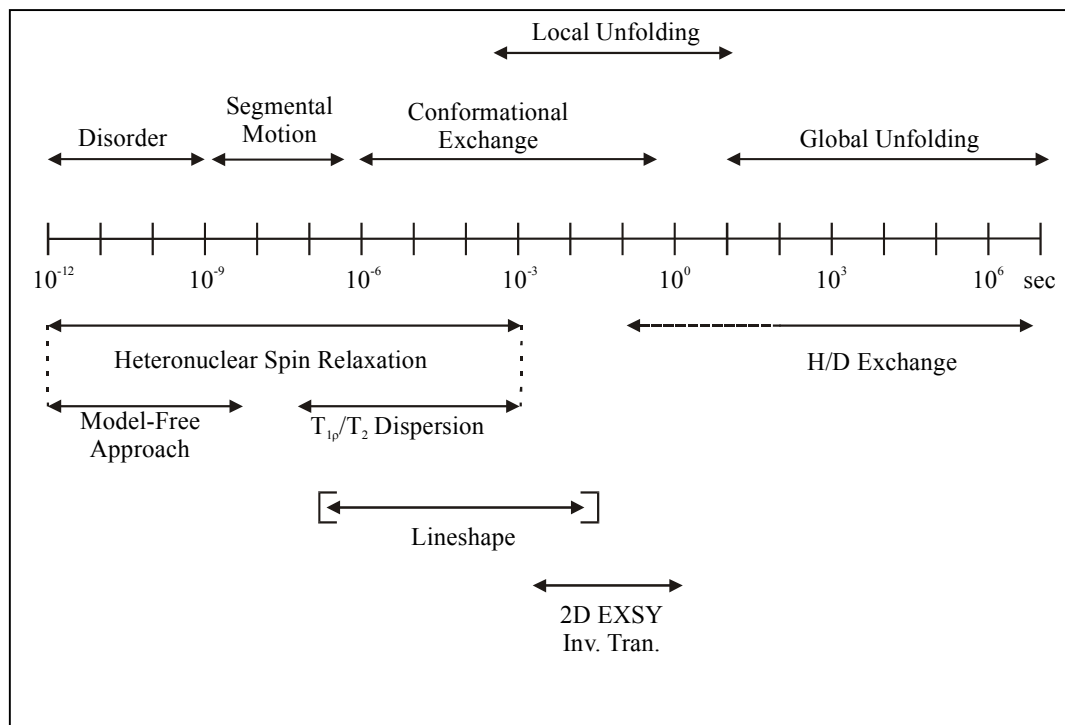


Figure 4-1: Timescale of macromolecular internal motions relevant in biological processes. Various NMR techniques are useful for probing motions from picoseconds to milliseconds.

In this work, we have probed the dynamics of loop A domain of the hairpin ribozyme as an initial step in understanding the active site dynamics in RNA-RNA tertiary interaction. The results complement the calculational work in Chapter 2 and provide insights into the role of conformational sampling of loop A in the formation of tertiary structure in the hairpin ribozyme.

4.2 MATERIALS AND METHODS

4.2.1 RNA preparation

^{13}C and ^{15}N labeled nucleotides were incorporated in the hairpin ribozyme loop A (**Figure 4-2**) by *in vitro* transcription (30, 31) using recombinant phage T7 RNA polymerase (T7 RNAP) enzyme in the presence of a synthetic DNA template. Transcription was initiated with RNAP following addition of DNA template, uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled nucleotide triphosphates (ATP, UTP, CTP, and GTP) and Mg^{2+} . Transcription was continued for 3 hours at 37 °C, stopped by the addition of EDTA, and analyzed by 15% denaturing gel electrophoresis. Transcription product was subjected to ethanol precipitation and RNA precipitate was then reconstituted in H_2O and subjected to purification through a HiLoad 26/60 SuperdexTM 75 column (32) integrated with AKTA Fast Performance Liquid Chromatography (FPLC) monitored at A_{260} with an isocratic flow rate of 3 mL/min sodium phosphate buffer, pH 6.40. Fractions corresponding to loop A were pooled and thoroughly exchanged with H_2O using Amicon Ultra-4 centrifugal filter units, 3000 MWCO. RNA purity was determined by denaturing gel. RNA was then dried under vacuum and reconstituted in 100% D_2O and appropriate NMR buffer conditions (10 mM phosphate buffer, 100 μM EDTA, pH 5.5)

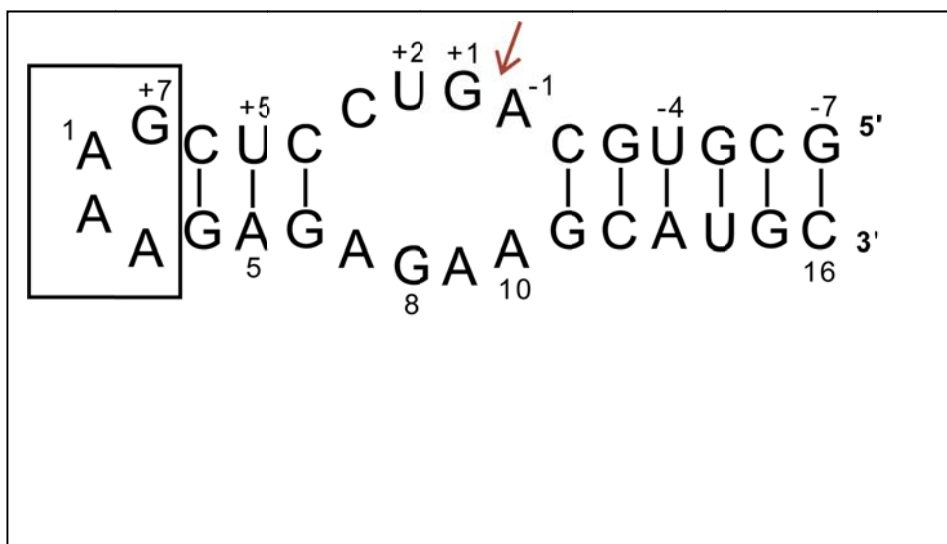


Figure 4-2: The hairpin ribozyme loop A construct used in NMR relaxation studies. Red arrow is the cleavage site. GNRA tetraloop (boxed) was appended at the end of stem I for stabilization .

4.2.2 NMR studies

All NMR data were acquired on a Varian UnityINOVA 600 MHz (^{13}C 150 MHz) spectrometer. The ^1H , ^{13}C HSQC spectrum was acquired with 1024×512 complex points in the t_2 and t_1 dimensions, respectively, with corresponding spectral widths of 6000 and 4524 Hz, and a 1 s recycle delay. The proton RF carrier was centered on the residual HDO signal, and the ^{13}C carrier frequency was set at 145 ppm (including a 35 ppm pulse sequence adjustment) to capture the resonances in the aromatic C8/C6 and adenine C2 region. Fast dynamics T_1 , $T_{1\rho}$, and ^1H - ^{13}C heteronuclear NOE (hNOE) data were acquired using similar parameters as ^1H , ^{13}C -HSQC with minor variations of published pulse sequences (33). T_1 delay times ranged from 50 -1200 ms, while the $T_{1\rho}$ delay times were from 10-100 ms. These relaxation data were acquired over 96 time-points in the second dimension at a constant 2980 Hz spin-lock field. Saturated and non-saturated heteronuclear NOE experiments were acquired in interleaved fashion with a proton irradiation of 3 sec within a total recycle delay of 5 sec in the “saturated” experiment. In relaxation dispersion studies of purine C8, pyrimidine C6 and adenine C2 carbons, $R_{1\rho}$ experiments were acquired as a function of the applied spin lock power, ω_1 . ^{13}C carrier was positioned at 145 ppm with a sweep width of 6000 x 4524 Hz. A total of 106 complex t_1 increments of 64 transients each were collected with a recycle delay of 2 s, in a series of 15 different spin-lock powers ranging from 1.8 to 6 kHz at a constant relaxation delay of 35 ms. All the dispersion data were acquired using on-resonance $R_{1\rho}$ (33). All NMR data were processed using FELIX 2002 (Felix NMR, Inc.). Before Fourier transformation (FT) in t_2 dimension, the FID was offset-corrected at 20%, a 5 Hz exponential line broadening function was applied in t_2 with zero filling. A cosine-squared apodization function was used in t_1 .

4.2.3 Data analysis

Cross-peak integrated volumes were exported to Igor Pro 5.0.4 (WaveMetrics) after peak-picking and volume integration using FELIX 2002. R_1 and $R_{1\rho}^{\text{obs}}$ relaxation rates were extracted by fitting the integrated volume intensities to a single exponential decay using two parameters to limit potential fitting bias (34). The relaxation rates were obtained from the individual fits with the errors reported from the exponential fit. The $^1\text{H}, ^{13}\text{C}$ hNOE was determined from the ratio of the saturated and non-saturated peak volume intensity.

For dispersion analysis, the magnitude of the spin-lock field was considered as the effective field resulting from the applied B_1 field and the offset Ω according to equation 4-3.

$$\omega_{eff} = \sqrt{\omega_1^2 + \Omega^2} \quad (4-3)$$

In addition, the observed $T_{1\rho}$ decay process has contributions from both transverse and longitudinal relaxation, because of tilting the spin-lock axis out of the xy plane. This effect was also considered using measured values of R_1 ($R_1=1/T_1$) (35). The $R_{1\rho}^{\text{obs}}$ rates were calculated using the equation $R_{1\rho}^{\text{obs}} = (1/T) \times -\ln(I(T)/I_0)$, where T is the relaxation delay in seconds, I is the measured intensity, and I_0 is the measured intensity with no relaxation delay (reference relaxation). The data for I_0 was acquired three times and averaged. $R_{1\rho}$ rates were calculated from $R_{1\rho}^{\text{obs}}$ using the equation:

$$R_{1\rho}^{\text{obs}} = R_1 \cos^2 \theta + R_{1\rho} \sin^2 \theta \quad (4-4)$$

where $\theta = \tan^{-1} (\omega_1/\Omega)$ is the angle of the spin-lock axis from vertical, ω_1 is the spin lock power in Hz and Ω is the offset in Hz. Exchange parameters k_{ex} , $\Phi_{ex} (P_a P_b (\Delta\omega)^2)$, and $R_{1\rho}^{\infty}$ were extracted by fitting dispersion curves to $R_{1\rho}$ equation 4-2 (21) using Igor Pro 5.

4.2.4 Model-free analysis

Internal motion, quantified as a generalized order parameter S^2 in the ps-ns timescale using the model-free formalism of Lipari & Szabo (36-38), was performed using R_1 , $R_{1\rho}$, and 1H , ^{13}C hNOE measurements acquired at 600 MHz. All the model-free analysis calculations were performed by Dr. Charles Hoogstraten using the computer program Modelfree 4.1. S^2 , τ_e , and R_{ex} motional parameters were determined for atoms in the hairpin ribozyme. An estimate of the isotropic rotational correlation time was obtained from the $T_1/T_{1\rho}$ ratio for resonances in the helical stem that showed no evidence of exchange contributions when analyzed including dipole-anisotropy relaxation mechanisms (39). Chemical shift anisotropy (CSA) values used were 131, 186.5 and 167 ppm for C8, C6 and C2 atoms respectively with a r_{HC} bond length of 1.1 Å.

4.3 RESULTS AND DISCUSSION

4.3.1 ^{13}C Relaxation measurements

Longitudinal and transverse relaxation rates (R_1 and $R_{1\rho}$) were determined for purine C8 and C2 atoms as well as pyrimidine C6 atoms of hairpin ribozyme loop A to assess the fast internal motion (ps-ns) in loop A. In **Figure 4-3** we have shown representative ^{13}C R_1 and $R_{1\rho}^{obs}$ curves for C8 atoms in the tetraloop (A2) and active site loop region (A9) at 25°C (**Figure 4-3**). Transverse and longitudinal relaxation rates were obtained for all the C8 resonances of purines, the C6 resonances of pyrimidines and the C2 resonances of adenines in in the loop A. There were

a total of 39 probes used to assess base dynamics. Both the T_1 and $T_{1\rho}$ data were fit to single-exponential decays and the relaxation rates extracted. Results are reported in **Figure 4-3** and **Table 4-1**. In general, most of the residues reported typical R_1 and $R_{1\rho}$ values of ~ 2.2 and $\sim 40 \text{ s}^{-1}$ as well as a heteronuclear NOE ratio of 1.2. This is consistent with the relaxation rates observed in helical regions for which most of the residues in loop A are found. This observation suggests that most of the residues in loop A are base-paired within the helical environment as reported in chapter 3 of this thesis. Hoogstraten *et al*, have also reported similar relaxation rates in the helical region of a comparable RNA system, the leadzyme (20). Pyrimidines mostly showed no significant differences observed for relaxation rates in their C6 resonances consistent with previous work done in lead-dependent ribozyme (40). This was expected because most of the pyrimidines are within the helical secondary structure of loop A except U+2 and C+3 which are in the loop region. However, one pyrimidine (assigned to either U+2 or C+3) residue showed an overall increase in R_1 by $\sim 31\%$ which suggests dynamics. This result is consistent with our molecular dynamics simulation studies of base and sugar dynamics reported for U+2 and C+3 in Chapter 2 of this thesis. The helical C8 and C2 resonances displayed no evidence of exchange contributions from the relaxation data based on Model free analysis (**Figure 4-4**, **Table 4-2**). Fast internal motion was observed in various residues as indicated by increase in the R_1 values as well as a decrease in $R_{1\rho}$ values. The R_1 for peak 4 C8 was $\sim 19\%$ higher than the average value of R_1 calculated for helical C8 atoms. Conversely, its $R_{1\rho}$ was dramatically reduced by $\sim 53\%$ relative to the average $R_{1\rho}$ of the helical resonances. Peak 8 reported a 33% increase in R_1 and a similar (32%) decrease in $R_{1\rho}$. Peaks 26 and 35 reported an increase in $R_{1\rho}$ but constant R_1 relative to the average helical values. An increased R_1 together with a decreased R_2 (or $R_{1\rho}$) generally indicates ps-ns disorder, while an increased R_2 (or $R_{1\rho}$) with a constant R_1

suggests exchange contribution. Conformational exchange is usually validated by relaxation dispersion measurements reporting on the residue. It should be noted that peak 26 represents a purine C8 resonance while peak 35 represents an adenine C2 resonance as assigned in the previous chapter (Chapter 3) of this thesis. These dynamics may be reporting on similar residue from the C8 and C2 resonances within a single stranded loop region. In fact, peak 26 was assigned as A1 which resides in the GAAA tetraloop. The tetraloop A1 (peak 26) reported internal dynamics (**Figure 4-4**, **Table 4-2**), based on the S^2 value of 0.87 indicating rapid (ps-ns) disorder. This base was proposed to be disordered from structural data (41). A3 (Peak 15), similarly residing in the tetraloop, reported a 65% enhancement in the C8 R_1 and a 68% reduction in $R_{1\rho}$ relative to the helical regions. This is also consistent ps-ns dynamics. However, previous studies on the leadzyme showed a larger $R_{1\rho}$ for the C8 resonance of A3 (40) with a follow-up study reporting a conformational exchange lifetime of 80 μ s (20). Peak 12 recorded the highest enhancement in $R_{1\rho}$ of > 100% (**Figure 4-4**) relative to the helical rates. Also, this peak showed only modest 32% reduction of R_1 . This suggests that the C8 resonance of this residue undergoes conformational exchange in the μ s-ms timescale. Conformational exchange parameters determined from Model free analysis (**Figure 4-5**) confirmed a C8 conformational exchange with R_{ex} of ~ 140 s^{-1} . However this resonance did not show ps-ns dynamics based on a higher order parameter of 0.903. This suggests that this conformational exchange may play a role in conformational sampling. Our preliminary assignment suggests that this residue may be G8, which plays a catalytic role in the hairpin ribozyme. G8 is base-paired to the dynamic U+2 in the docked structure of hairpin ribozyme (12), suggesting that this interaction may be partly driven by conformational exchange in this residue. Peak 9 C8 relaxation data reported ps-ns dynamics based on its calculated order parameter suggesting that this residue may be located in the loop or

terminal region. Our preliminary assignments discussed in Chapter 3 of this thesis indicate that peak 9 is likely G+1, located in the loop region, or a terminal G. Peak 10 (assigned as tetraloop A2), did not show any proof of dynamics, in the ps-ns timescale. This suggests that A2 is fairly stable due to stacking interactions within the tetraloop. This is consistent with our previous observations made on A2 in the results our MD simulation of loop A (**Figure 4-6**). Similar observations have been made on the adenine, A2, base of a GAAA tetraloop which localizes between residues involved in stacking interactions while being protected from chemical modification upon RNA folding with no evidence for dynamics at the C8 site (40, 42).

4.3.2 Relaxation dispersion

By observing an increase in the measured transverse relaxation rates as well as R_{ex} contributions in the model-free analysis for some residues, relaxation dispersion curves of on-resonance $R_{1\rho}$ experiments were acquired to examine conformational exchange processes within loop A in detail. Motions on the μ s-ms timescales can affect the transverse relaxation. While model-free analysis can propose the presence of exchange (R_{ex}) contributions, a more accurate determination of R_{ex} contributions is in the analysis of relaxation dispersion curves. Relaxation dispersion curves are the dependence of the measured transverse relaxation rate on the effective B_1 field as determined by the applied radiofrequency (RF) field and the resonance offset for $R_{1\rho}$ experiments. Over the years, relaxation dispersion curves have proved useful in investigating conformational exchange processes extensively in proteins (22, 23, 43-46). However, few dynamic exchange studies have been done for RNA (20, 47, 48). Relaxation dispersion curves were fit to equations 4-2 to determine k_{ex} , $R_{1\rho}^{\infty}$, and Φ_{ex} , where $\Phi_{ex} = P_a P_b (\Delta\omega)^2$.

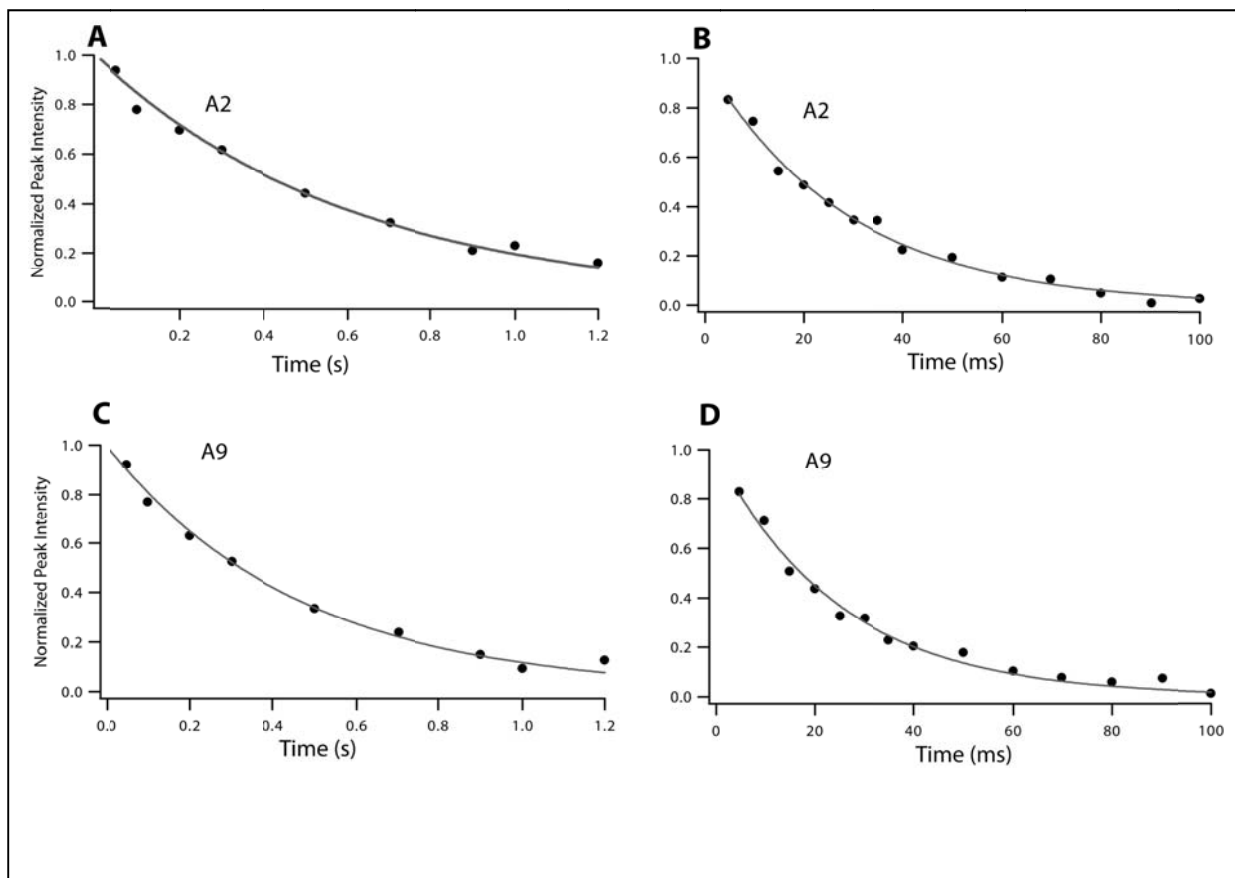


Figure 4-3: Representative ^{13}C R_1 (A and C) and $R_{1\rho obs}$ (B and D) curves for A2 (A and B) and A9 (C and D). Curves represent non-linear least squares two-parameter fits to a single exponential.

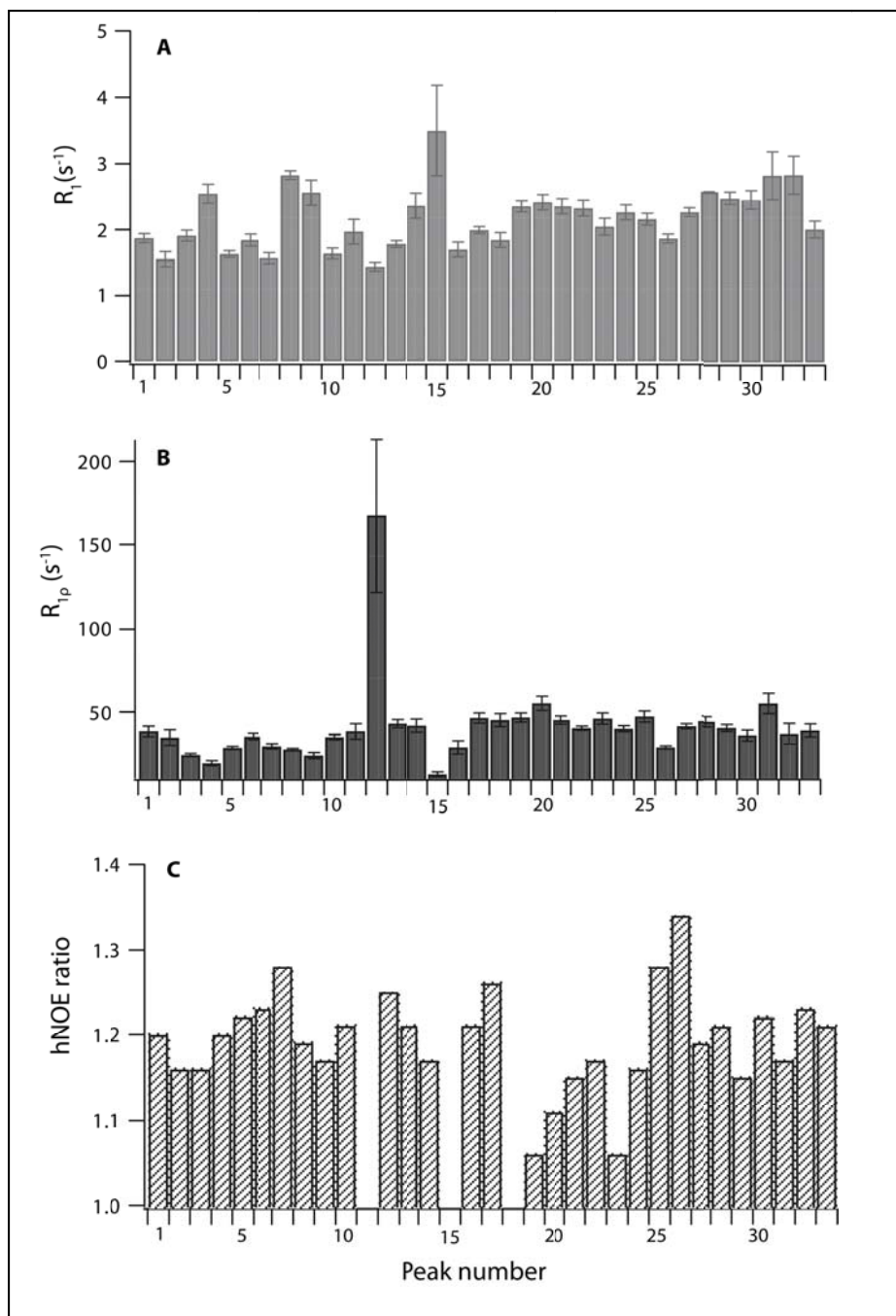


Figure 4-4: Histograms of the measured relaxation rates in loop A are shown for (A) longitudinal relaxation, T_1 (B) transverse relaxation, $T_{1\rho}$, and (C) heteronuclear NOE. The peak numbers correspond to those in Tables 3-2 and 3-3 in Chapter 3.

Table 4-1: ^{13}C R_1 , $R_{1\rho}$ and heteronuclear NOE measurements.

R_1 and $R_{1\rho}$ valued are for C8 (purines), C6 (pyrimidine) and C2 (adenine) atoms of loop A at 600 MHz. Peaks 34-39 are adenine C2 peaks.

Assignment	R_1	$R_{1\rho}$	hNOE
Peak 1	1.87±0.07	38.35±3.10	1.20
Peak 2	1.55±0.11	34.64±4.49	1.16
Peak 3	1.91±0.08	24.53±1.01	1.16
Peak 4	2.54±0.14	19.50±1.47	1.20
Peak 5	1.63±0.05	28.65±1.00	1.22
Peak 6	1.84±0.09	35.48±1.80	1.23
Peak 7	1.57±0.08	29.55±1.43	1.28
Peak 8	2.83±0.07	27.76±0.62	1.19
Peak 9	2.56±0.19	23.99±1.70	1.17
Peak 10	1.64±0.08	35.12±1.53	1.21
Peak 11	1.97±0.19	38.38±4.66	N/A
Peak 12	1.44±0.07	167.44±45.70	1.25
Peak 13	1.78±0.05	43.06±2.46	1.21
Peak 14	2.36±0.19	41.65±4.07	1.17
Peak 15	3.50±0.68	12.74±1.40	N/A
Peak 16	1.70±0.11	28.88±3.87	1.21
Peak 17	1.99±0.05	46.34±3.14	1.26
Peak 18	1.84±0.11	45.12±3.98	N/A
Peak 19	2.35±0.08	46.67±2.68	1.06
Peak 20	2.41±0.11	55.14±4.34	1.11
Peak 21	2.35±0.11	45.08±2.55	1.15
Peak 22	2.32±0.12	40.36±1.23	1.17
Peak 23	2.04±0.13	45.90±3.46	1.06
Peak 24	2.26±0.11	40.05±1.94	1.16
Peak 25	2.16±0.09	47.19±3.44	1.28
Peak 26	1.86±0.07	29.11±0.94	1.34
Peak 27	2.26±0.07	41.55±1.55	1.19
Peak 28	2.57±0.01	44.57±2.90	1.21
Peak 29	2.47±0.10	40.78±2.02	1.15
Peak 30	2.45±0.14	36.32±3.25	1.22
Peak 31	2.82±0.36	55.52±6.15	1.17
Peak 32	2.83±0.29	37.06±6.37	1.23
Peak 33	2.00±0.13	39.39±4.02	1.21
Peak 34	1.73±0.09	41.10±3.15	1.25
Peak 35	1.98±0.24	29.68±4.61	1.20
Peak 36	1.90±0.07	34.91±1.22	1.36
Peak 37	1.84±0.05	44.24±2.01	1.33
Peak 38	1.89±0.13	35.90±2.52	N/A
Peak 39	1.70±0.04	34.58±1.16	1.16

Table 4-2: Model-free analysis of loop A relaxation measurements.

Order parameter (S^2), correlation time (τ_{ex}) and exchange contribution to transverse relaxation (R_{ex}) were extracted from T_1 , $T_{1\rho}$ and hNOE data.

Peak #	S^2	τ_{ex} , ps	R_{ex} , s ⁻¹
1	1.000		
2	0.990		
3	0.000		
4	0.000		
5	0.921	60.2	0
6	0.908	446.3	5.5
7	0.893	74.0	0
8	0.000		
9	0.688	708.6	0
10	0.957	158.1	3.8
12	0.903	0	137.9
13	0.929	433.4	12.5
14	1.000		
16	0.947	317.6	
17	0.859	493.9	17.6
19	1.000		
20	1.000		
21	1.000		
22	0.934	616.3	
23	1.000		
24	0.926	523.8	
25	0.806	540.3	19.9
26	0.870	278.4	0
27	0.000		
28	1.000		
29	0.889	780.4	0
30	0.854	484.0	0
31	1.000	0	0
32	0.779	626.0	0
33	1.000	0	0
34	0.901	80.4	0
35	0.924	0	0
36	0.854	136.2	0
37	1.000	0	6.5
39	0.903	26.3	0
40	0.905	26.8	0
41	0.905	26.8	0

No satisfactory model could be fit for peaks assigned 3, 4, 8, 11, 15 and 18 using the combinations of S^2 , τ_{ex} , and R_{ex} and an isotropic model for overall rotation of the system.

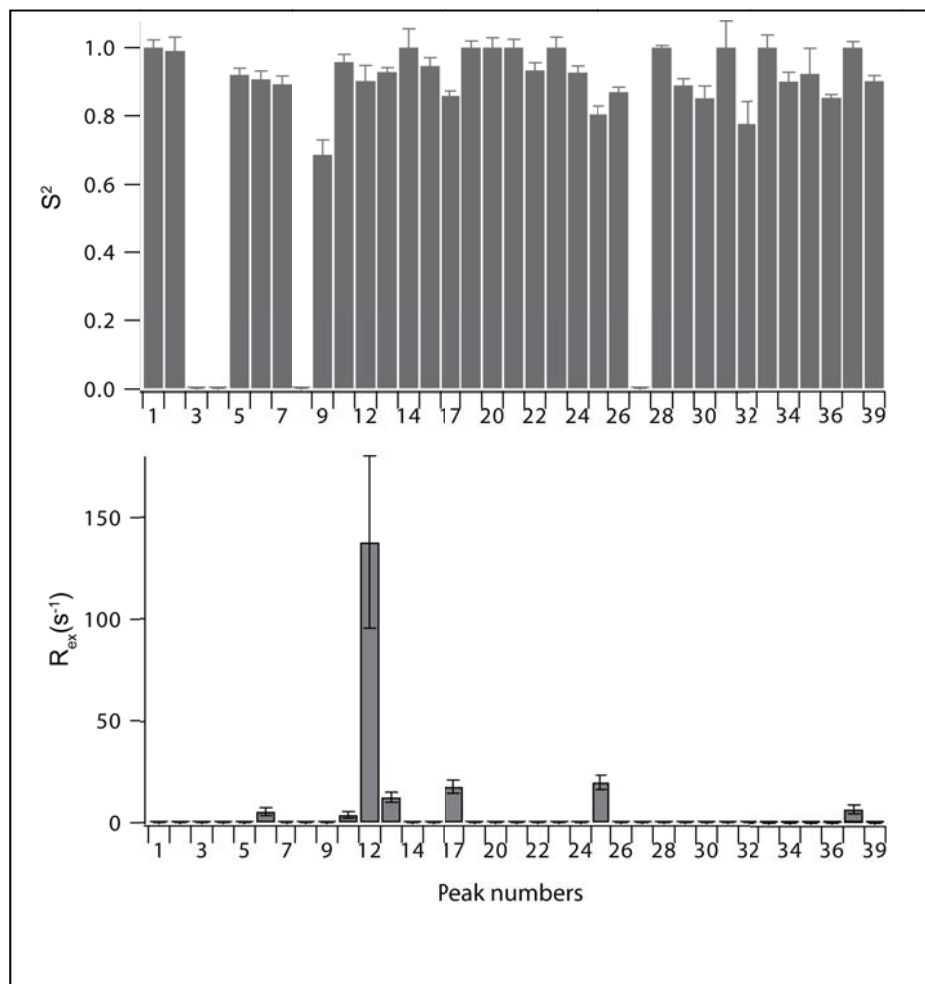


Figure 4-5: Histogram representation of order parameter (S^2) and exchange contribution (R_{ex}) derived from relaxation measurements as analyzed by Model free formalism.

From the $R_{1\rho}$ dispersion experiment we have reported on conformational exchange on several C8/C6 resonances in loop A (**Figure 4-7**, **Table 4-3**). From this data, we report best-fit values of τ_{ex} , as well as minimum values of $\Delta\omega$, the chemical shift difference between the two exchanging populations, calculated by setting $P_a = P_b = 0.5$. C8/C6 resonances represented by peaks 3, 12 and 17 reported conformational exchange as shown in the power-dependent dispersion curves (**Figure 4-7**). The remaining resonances, by contrast, showed transverse relaxation rates that were independent of the applied B_1 power, i.e., an absence of dispersion, as expected from the previous results of Model free analysis. This implies that conformational exchange in the μs -ms timescale only occurs in specific residues, presumably situated in the non-helical regions of loop A. Peak 3 reported an exchange time of 85 μs whereas peaks 13 and 17 both reported an exchange time of 667 μs suggesting that there exchange may be correlated. Details for all the resonances exchange parameters have been tabulated (**Table 4-3**).

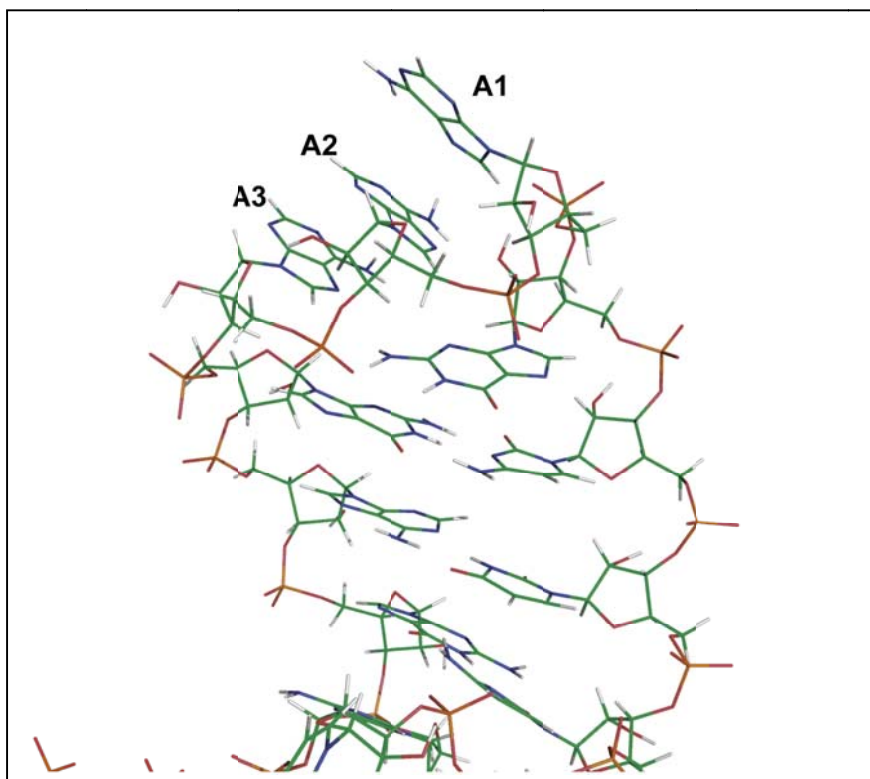


Figure 4-6: Simulation structure of loop A highlighting base stacking in the GAAA tetraloop involving A2 stacked between A1 and A3. These interactions limit A2 dynamics.

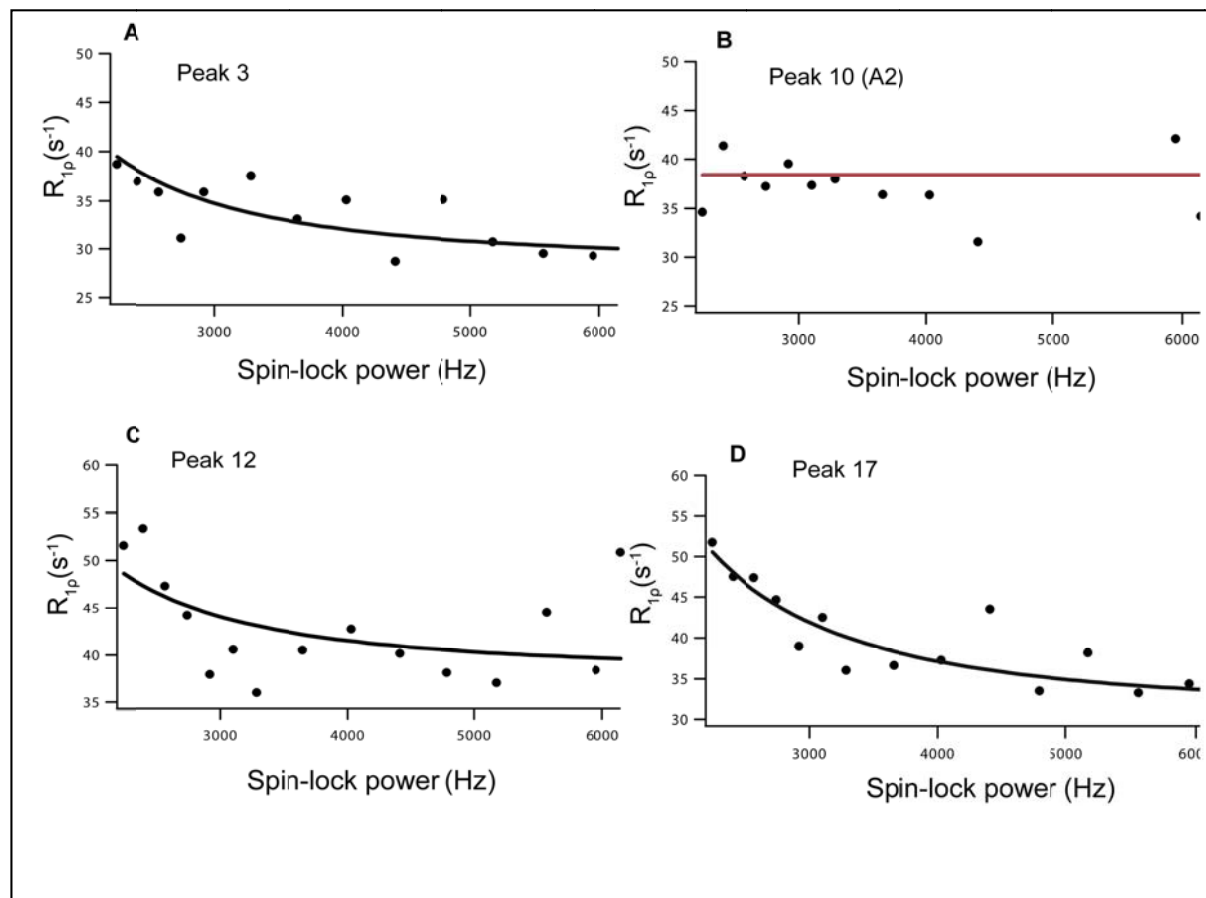


Figure 4-7: Representative relaxation dispersion curves of selected C8 resonances. Peaks 3, 12 and 17 show power-dependent conformational exchange (A, C and D) whereas peak 10 (B) does not undergo exchange.

Table 4-3: Chemical exchange parameters from $T_{1\rho}$ dispersion analysis of loop A.
 $\Delta\omega$ was calculated from $P_a P_b (\Delta\omega)^2$ using $P_a = P_b = 0.5$

Peak #	$P_a P_b (\Delta\omega)^2$	$\tau_{\text{ex}} (\mu\text{s})$	$\Delta\omega_{\text{min}} (\text{ppm})$
1	2.27×10^6	693.0	3.20
2	4.85×10^5	14.2	1.48
3	3.28×10^5	84.5	1.22
4	6.69×10^4	579.7	0.55
5	-	-	-
6	-	-	-
7	1.96×10^5	297.4	0.94
8	4.02×10^5	618.4	1.35
9	-	-	-
10	-	-	-
11	-	-	-
12	1.40×10^6	666.7	2.51
13	1.40×10^6	667.7	2.51
14	1.66×10^6	676.0	2.74
15	1.64×10^5	6.8	0.86
16	1.45×10^6	426.1	2.56
17	2.62×10^6	667	3.44
18	9.53×10^5	40.1	2.07
19	1.76×10^6	833.5	2.82
20	-	-	-
21	-	-	-
22	1.11×10^6	544.8	2.24
23	1.33×10^6	32.6	2.45
24	1.08×10^6	32.6	2.21
25	5.13×10^6	8.3	1.52
26	-	-	-
27	8.39×10^5	386.2	1.94
28	-	-	-
29	4.32×10^5	314.9	1.40
30	-	-	-
31	-	-	-
32	-	-	-
33	1.38×10^6	25.5	2.49
34	1.21×10^6	67.1	2.34
35	-	-	-
36	1.35×10^6	1506	2.47
37	1.76×10^6	595.2	2.82
38	8.46×10^5	284.8	1.95
39	9.49×10^5	318.6	2.07

In an NMR relaxation experiment, conformational exchange in the microsecond to millisecond timescale causes increased line-broadening of NMR signals by an amount (R_{ex}) that contributes to the measured overall transverse relaxation rate (R_2^{eff}). From R_{ex} three physical parameters for a two-site dynamic process can be obtained: rates of interconversion (k_{ex}), the relative populations of the exchanging species (P_a and P_b) and the chemical shifts between the exchanging species ($\Delta\omega$). In this work, we have assessed conformational dynamics of loop A using the structural disorder (ps-ns) parameter extracted using longitudinal relaxation, transverse relaxation and heteronuclear NOE. The dynamics reported for various some loop residues showed a wide variety of timescales ranging from picosecond to millisecond timescale. The diverse timescales reported for loop A residues provide us with insights for understanding the energy landscape of RNA.

ACKNOWLEDGMENT

I would like to thank Max T. Rogers NMR staff Kermit Johnson, Dr. Daniel Holmes and Dr. Li for assistance in using 600 MHz spectrometer. I would also like to thank Dr. James Johnson Jr. for programming the $R_{1\rho}$ equation in Igor Pro (WaveMetric) software. This facilitated the analysis of all the $R_{1\rho}$ reported in this thesis.

REFERENCES

REFERENCES

1. Hammes-Schiffer, S. (2002) Impact of enzyme motion on activity, *Biochemistry* 41, 13335-13343.
2. Al-Hashimi, H. M. (2005) Dynamics-based amplification of RNA function and its characterization by using NMR spectroscopy, *Chembiochem* 6, 1506-1519.
3. Boehr, D. D., McElheny, D., Dyson, H. J., and Wright, P. E. (2006) The dynamic energy landscape of dihydrofolate reductase catalysis, *Science* 313, 1638-1642.
4. Beach, H., Cole, R., Gill, M. L., and Loria, J. P. (2005) Conservation of mus-ms enzyme motions in the apo- and substrate-mimicked state, *J Am Chem Soc* 127, 9167-9176.
5. Wolf-Watz, M., Thai, V., Henzler-Wildman, K., Hadjipavlou, G., Eisenmesser, E. Z., and Kern, D. (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair, *Nat Struct Mol Biol* 11, 945-949.
6. Leulliot, N., and Varani, G. (2001) Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture, *Biochemistry* 40, 7947-7956.
7. Treiber, D. K., and Williamson, J. R. (2001) Beyond kinetic traps in RNA folding, *Curr Opin Struct Biol* 11, 309-314.
8. Fedor, M. J., and Williamson, J. R. (2005) The catalytic diversity of RNAs, *Nat Rev Mol Cell Biol* 6, 399-412.
9. Kuimelis, R. G., and McLaughlin, L. W. (1998) Mechanisms of Ribozyme-Mediated RNA Cleavage, *Chem Rev* 98, 1027-1044.
10. Ke, A., Zhou, K., Ding, F., Cate, J. H., and Doudna, J. A. (2004) A conformational switch controls hepatitis delta virus ribozyme catalysis, *Nature* 429, 201-205.
11. Cai, Z., and Tinoco, I., Jr. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite, *Biochemistry* 35, 6026-6036.
12. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis, *Nature* 410, 780-786.
13. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat Struct Biol* 6, 212-216.

14. 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.
15. Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules, *Chem Rev* 104, 3623-3640.
16. Palmer, A. G., 3rd. (1997) Probing molecular motion by NMR, *Curr Opin Struct Biol* 7, 732-737.
17. Akke, M. (2002) NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis, *Curr Opin Struct Biol* 12, 642-647.
18. Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules, *Chem Rev* 104, 3623-3640.
19. Palmer, A. G., 3rd, and Massi, F. (2006) Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy, *Chem Rev* 106, 1700-1719.
20. Hoogstraten, C. G., Wank, J. R., and Pardi, A. (2000) Active site dynamics in the lead-dependent ribozyme, *Biochemistry* 39, 9951-9958.
21. Palmer, A. G., 3rd, Kroenke, C. D., and Loria, J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules, *Methods Enzymol* 339, 204-238.
22. Korzhnev, D. M., Orekhov, V. Y., Dahlquist, F. W., and Kay, L. E. (2003) Off-resonance R1rho relaxation outside of the fast exchange limit: an experimental study of a cavity mutant of T4 lysozyme, *J Biomol NMR* 26, 39-48.
23. Kim, S., and Baum, J. (2004) An on/off resonance rotating frame relaxation experiment to monitor millisecond to microsecond timescale dynamics, *J Biomol NMR* 30, 195-204.
24. Nikonowicz, E. P., Sirr, A., Legault, P., Jucker, F. M., Baer, L. M., and Pardi, A. (1992) Preparation of ¹³C and ¹⁵N labelled RNAs for heteronuclear multi-dimensional NMR studies, *Nucleic Acids Res* 20, 4507-4513.
25. Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. D., and Williamson, J. R. (1992) Preparation of isotopically labeled ribonucleotides for multidimensional NMR spectroscopy of RNA, *Nucleic Acids Res* 20, 4515-4523.
26. Akke, M., Fiala, R., Jiang, F., Patel, D., and Palmer, A. G., 3rd. (1997) Base dynamics in a UUCG tetraloop RNA hairpin characterized by ¹⁵N spin relaxation: correlations with structure and stability, *RNA* 3, 702-709.

27. Hall, K. B., and Tang, C. (1998) ^{13}C relaxation and dynamics of the purine bases in the iron responsive element RNA hairpin, *Biochemistry* 37, 9323-9332.
28. Vallurupalli, P., and Kay, L. E. (2005) A suite of ^2H NMR spin relaxation experiments for the measurement of RNA dynamics, *J Am Chem Soc* 127, 6893-6901.
29. Blackledge, M. J., Bruschweiler, R., Griesinger, C., Schmidt, J. M., Xu, P., and Ernst, R. R. (1993) Conformational backbone dynamics of the cyclic decapeptide antamanide. Application of a new multiconformational search algorithm based on NMR data, *Biochemistry* 32, 10960-10974.
30. 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 Res* 15, 8783-8798.
31. Milligan, J. F., and Uhlenbeck, O. C. (1989) Synthesis of small RNAs using T7 RNA polymerase, *Methods Enzymol* 180, 51-62.
32. Kim, I., McKenna, S. A., Viani Puglisi, E., and Puglisi, J. D. (2007) Rapid purification of RNAs using fast performance liquid chromatography (FPLC), *RNA* 13, 289-294.
33. Yamazaki, T., Muhandiram, R., and Kay, L. E. (1994) Nmr Experiments for the Measurement of Carbon Relaxation Properties in Highly Enriched, Uniformly $\text{C-}^{13}\text{N-}^{15}$ -Labeled Proteins - Application to C-^{13} (Alpha) Carbons, *J Am Chem Soc* 116, 8266-8278.
34. Viles, J. H., Duggan, B. M., Zaborowski, E., Schwarzsinger, S., Huntley, J. J., Kroon, G. J., Dyson, H. J., and Wright, P. E. (2001) Potential bias in NMR relaxation data introduced by peak intensity analysis and curve fitting methods, *J Biomol NMR* 21, 1-9.
35. Davis, D. G., Perlman, M. E., and London, R. E. (1994) Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the T1 rho and T2 (CPMG) methods, *J Magn Reson B* 104, 266-275.
36. 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, *J Am Chem Soc* 104, 4546-4559.
37. Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .2. Analysis of Experimental Results, *J Am Chem Soc* 104, 4559-4570.

38. Mandel, A. M., Akke, M., and Palmer, A. G. (1995) Backbone Dynamics of Escherichia-Coli Ribonuclease Hi - Correlations with Structure and Function in an Active Enzyme, *J Mol Biol* 246, 144-163.
39. Kay, L. E. (1998) Protein dynamics from NMR, *Nat Struct Biol* 5 Suppl, 513-517.
40. 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.
41. Heus, H. A., and Pardi, A. (1991) Structural Features That Give Rise to the Unusual Stability of Rna Hairpins Containing Gnra Loops, *Science* 253, 191-194.
42. Hoogstraten, C. G., Legault, P., and Pardi, A. (1998) NMR solution structure of the lead-dependent ribozyme: Evidence for dynamics in RNA catalysis, *J Mol Biol* 284, 337-350.
43. Eisenmesser, E. Z., Bosco, D. A., Akke, M., and Kern, D. (2002) Enzyme dynamics during catalysis, *Science* 295, 1520-1523.
44. Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain, *J Am Chem Soc* 123, 11341-11352.
45. Millet, O., Loria, J. P., Kroenke, C. D., Pons, M., and Palmer, A. G. (2000) The static magnetic field dependence of chemical exchange linebroadening defines the NMR chemical shift time scale, *J Am Chem Soc* 122, 2867-2877.
46. Cole, R., and Loria, J. P. (2002) Evidence for flexibility in the function of ribonuclease A, *Biochemistry* 41, 6072-6081.
47. Blad, H., Reiter, N. J., Abildgaard, F., Markley, J. L., and Butcher, S. E. (2005) Dynamics and metal ion binding in the U6 RNA intramolecular stem-loop as analyzed by NMR, *J Mol Biol* 353, 540-555.
48. Johnson, J. E., Jr., and Hoogstraten, C. G. (2008) Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using ¹³C NMR spin relaxation and specific isotope labeling, *J Am Chem Soc* 130, 16757-16769.

CHAPTER 5

SUMMARY, DISCUSSION AND FUTURE WORK

5.1 SUMMARY AND DISCUSSION

The projects pursued in this thesis were inspired by the overarching idea of investigating and correlating dynamics in RNA with possible function. Using molecular dynamics simulation and NMR spin relaxation techniques, we interrogated residue dynamics effects in loop A of the hairpin ribozyme and reported our results herein. The combination of computational (molecular dynamics simulation) and NMR techniques provides a robust dynamics probe for macromolecules.

Molecular dynamics simulations play a significant role in the interpretation of experimental data bearing on the time-dependent properties of biomolecules. MD has increasingly become common and suitable with the improvement of force fields and computer resources. This has improved the size of tractable molecules and the length of time that a molecule can be simulated. Longer simulations can more thoroughly explore conformational sampling and thus the conformational space available for the molecule. Molecular dynamics simulations could be useful in interpreting NMR experiments, including for fast motions seen primarily in NMR relaxation measurements (1, 2). NMR spectroscopy, on the other hand, detects motions across a broad range of molecular motion timescales (3). While studies of molecular motions have been comprehensively done in proteins using combined computer simulations and NMR (4-7), these motions have been less studied in RNA molecules despite their diverse nature. One major challenge that has partly hindered the computational studies of RNA is the force field limitations. Most of the current force fields have not been successful in supporting long simulations of RNA. An all-atom CHARMM36 force field has recently been developed based on cross-validation with NMR data (8). Motions have also been probed using NMR in some RNA molecules with the focus on base dynamics (9-12) and also ribose dynamics

(13). We have integrated computer simulation and NMR techniques to understand broad aspects of RNA dynamics and interaction. The evidence of active site rearrangement in the hairpin ribozyme loops, based on structural data, suggests mobility of certain residues in the activation process. Using combined analysis of simulations and NMR to analyze RNA molecules, the role of dynamics can be demonstrated for RNA molecules during RNA folding or catalysis.

In Chapter Two of this Thesis, MD simulation was used to determine conformational heterogeneity in RNA based on alternate base-pair formation within a subset of residues in the loop region of domain A of the hairpin ribozyme. The MD data was primarily acquired using CHARMM36 force field after force field comparison with amber ff10 force field indicated better structural overlap with CHARMM36 simulation. The observed conformers were extensively mapped and determined to form predominantly in the active site of loop A. Three main conformers and several minor conformations were observed in our simulations as analyzed by the Markov State model analysis. These three major conformers depict the local energy minima wells where these conformers reside within which there exist several minor states that undergo kinetic transitions within these local minima based on subtle structural changes and dynamics within a set of residues. RNA base residues and backbone dynamics played a major role in alternate base-pair formation and conformational heterogeneity in loop A. This suggests that conformational sampling and transition in loop A RNA is a key strategy to avoiding the kinetic traps that localize RNA in non-functional conformations. Kinetic transitions were achieved by loop A sampling a series of kinetic states which function to populate other conformers by accessing kinetic energy barriers that may limit the formation of functional (active) RNA. Of the conformations that were sampled, the most populated conformer, AA/CA, closely sampled conformational properties similar to the activated (docked) loop A conformation, suggesting the

activating role induced by conformational heterogeneity. This observation highlights the role of conformational sampling in pre-organizing and activating loop A (conformational selection) for tertiary RNA-RNA interactions. The slow docking association rates reported for hairpin ribozyme interaction (14) implies that some structural rearrangements occur in the independent loops (loops A and B) that initiate or stabilize the tertiary interactions necessary for docking together. This unique intrinsic base-pair rearrangement within a subset of kinetic states along the potential energy surface supports the description of a rugged but accessible free energy landscape of RNA with mechanistic properties that direct pre-organization to the activated state. The multiple conformers and inter-conformer transitions observed lay a foundation for understanding tertiary RNA interactions in hairpin ribozyme and RNA. Thus far, we have made significant correlations between the dynamics of RNA residues and conformational heterogeneity that we observed in loop A in the context of active site formation.

In Chapter Three, we compared the docking transitions between wild type and mutant loop A using a circular dichroism assay. Docking was abrogated in the U+2C/C+3U mutant, which has additional two Watson-Crick base-pairs, as expected. This suggests the role of dynamics in facilitating docking. Because the U+2C/C+3U mutant is unable to form the docked structure despite docking with a U+2C single mutant (15), it is plausible that the U+2C/C+3U mutant reduces or quenches docking-activating dynamics. Important cytosine and uracil residues which we observed to be dynamic in the chapter 2 of this thesis are immobilized by base-pairing. It is also possible, that the constraint imparted on loop A by additional base-pairs could have inactivated the mutant. We also determined the suitable loop A construct for NMR studies by a series NMR studies on various loop A constructs. The best-behaved loop A construct was used for resonance assignments of ^1H , ^{13}C , and ^{15}N nuclei with their chemical shifts reported in the

results section of this chapter. Using exchangeable and non-exchangeable NMR experiments we have assigned significant resonances of loop A for the NMR-active nuclei.

In Chapter Four, loop A dynamics were studied using the NMR relaxation measurements namely longitudinal relaxation (T_1), transverse relaxation ($T_{1\rho}$) and heteronuclear NOE (hNOE) using a ^{13}C -uniformly labeled sample. Fast internal motions in the order of ps - ns timescale were analyzed using Lipari & Szabo's Model-free approach by (16-18) with data from ^{13}C R_1 , $R_{1\rho}$, and heteronuclear NOE of loop A. Loop A was generally found to be a rigid molecule on this timescale with internal generalized order parameters, S^2 , of at least 0.9 in the helical regions as expected. Some atoms, however, were determined to have internal, effective correlation times indicative of fast motions on the ps to ns timescale, while others reported slow exchange in the μs -ms timescale. A wide variety of internal motions were observed in the non-helical portions of loop A, especially in the GAAA tetraloop and the loop region. The loop region constitutes the cleavage site of the hairpin ribozyme, and therefore it is plausible that these dynamics may play a role in accessing various energy wells in the complex energy landscape of loop A, in parallel with the calculational results described above. These internal motions time scales suggest a complex landscape of accessible states and potential correlations between observed motions.

5.2 FUTURE DIRECTION

In this thesis we have reported significant advancements in the analysis of RNA using integrated computational and NMR approach. However, questions still remain on some aspects of RNA dynamics and how these dynamics relate to RNA function. Below are insights into the future direction that will help us gain a deeper understanding of dynamics-function relationship in RNA.

5.2.1 Base dynamics for the hairpin ribozyme

We have analyzed the base dynamics in the isolated loop A of the hairpin ribozyme using ^{13}C uniformly-labeled sample for the side-chain C8/C6/A-C2 sites. This represents the dynamics happening in the nucleic acid side chains. Molecular dynamics studies suggest certain dynamics reported for loop residues that seem to be correlated with base dynamics. To comprehensively investigate ribose dynamics in these residues, we will adopt the ribose specific 2',4'- ^{13}C labeling technique (19) to assess ribose backbone dynamics. Using model-free analysis of disorder on the ps-ns scale and multiple-field relaxation dispersion analysis of $R_{1\rho}$ and CPMG data will help obtain ^{13}C ribose relaxation rates for loop A residues that have shown chemical exchange. Our relaxation data mostly displays repressed transverse relaxation rates in the internal loop, which supports conformational exchange in this region of the molecule.

The assessment of dynamics in isolated domains can also be extended to loop B, whose structure has previously solved by NMR (20). Maintaining the construct whose NMR structure is available will be important due to the availability of resonance assignments. Similarly, base and ribose dynamics should be obtained by relaxation measurements (T_1 , $T_{1\rho}$ and hNOE) and relaxation dispersion to obtain dynamic rates and parameters in a variety of time regimes (ps-ms). With the ground state dynamics defined, active site perturbations due to docking can be probed using labeled domain A with unlabeled domain B in the presence and absence of $[\text{Co}(\text{NH}_3)_6]^{3+}$. To prevent cleavage, an A38U mutant may be used instead of wild type loop A since it supports docking but not cleavage within the timescale of NMR experiments. Transitions relevant after during and after the docking process can be captured and analyzed in the context of available data. Such experiments also can definitively probe timescale of docking perturbations.

The work detailed in this thesis investigated dynamics in loop A RNA molecules and the results obtained reveal an integrated but complex conformational sampling and transitions throughout the RNA free energy landscape. This complexity presents a clearer picture of the structural versatility of RNA in fast and intermediate timescales relevant in the catalytic cycle of ribozymes.

5.2.2 Functional read-out of dynamics

Previously, it has been difficult to probe directly the functional relevance of a particular RNA conformational change or dynamic region. Our lab has recently developed a specific isotope-labeling scheme that labels alternate ribose carbons specifically to allow for ribose dynamics analysis using NMR relaxation (21). RNA labeled as such can be used to unambiguously probe ribose pucker C3'-endo/C2'-endo inter-conversions. The major question remains how we can determine the functional relevance of such interconversion? Our lab has recently demonstrated that we can preferentially perturb RNAs ribose pseudo-rotation using locked nucleic acids (LNA), which features a methylene linkage between 2' oxygen and 4' carbon, to completely restrict the nucleotide to a C3'-endo, and assay for functional effects of such perturbation (22). This technique has been used with great success to probe three systems with distinct types of conformational change: the fluctuation of an inactive ground state structure of a lead-dependent ribozyme to a low-population active conformation of an unknown structure, the unfolding of a UUCG tetraloop, and the rearrangement of the U1 snRNA hairpin II (U1hpII) from a free conformation to a protein-bound form (22). From these experiments Julien *et al* showed that the imposition of a C3'-endo conformation using LNA at sites U6 of the UUCG tetraloop and C5 of U1hpII (both C2'-endo) resulted in a drastic decrease in the thermodynamic stability and protein binding respectively. By contrast, LNA substitution at G7 of the leadzyme

and A6 of U1hpII showed only minor effects implying that ribose conformational changes at these sites was not essential for the formation of the active state of leadzyme or for protein recognition by U1hpII. There was also a 20-fold increase in catalytic rate in the leadzyme upon LNA substitution at G9 which suggests the conformational activation by dynamics regulation. This phenomenon supports the idea of active and inactive conformational states observed due to dynamics in Chapter Two of this thesis. This technology presents a platform for mapping out conformational changes towards the activation of the hairpin ribozymes. LNAs have also previously been incorporated into helical regions and recognition arms of 10-23 catalytic DNA and the hammerhead ribozymes in an effort to improve targeting of cellular and RNA sequences resulting in dramatic effects upon structural stability and target affinity (23, 24). We can therefore employ LNA at specific dynamic residues to assess function that is coupled with ribose dynamics. This conformational change may be due to the pseudo-rotation angle inversion which couples to the RNA backbone dynamics. In the docked structure of the hairpin ribozyme G+1 is extruded from the helical stack of loop A (25), and is stabilized by burial in its cognate loop B pocket. This constrains the ribose of both A-1 and G+1 into C2'-endo puckers. Despite the C2'-endo pucker retention in G+1 between undocked (26) and docked loop A (25) structures, the ribose pucker of the adjacent nucleotide A-1 dramatically shifted from a C3'-endo to a C2'-endo upon docking. This suggests pucker rearrangement of A-1 to accommodate the docked state conformation. Similar pucker interconversions were observed in G8, C+3 and U+2 residues in loop A. Our simulation studies have also observed complementary pucker transitions in residues A7, G8, A9, G+1, U+2 and C+3, some being coupled to base-flipping. These residues provide us with initial points to begin the systematic functional probing of pucker dynamics using LNA substitution. These perturbations can be investigated by various assays including binding and

kinetic with relevant controls. Mutations or modifications have the limitation of the possibility of functional effects arising from features of the probe other than the intended perturbation. For LNAs, this possibility may be due to the loss of hydrogen-bonding groups on the 2'-hydroxyl or to the steric effect resulting from the bridging methylene group. A partial control for these effects should employ complementary analysis of 2'-deoxy and/or 2'-O-methyl modifications (22). In cases for which either of these modifications is poorly tolerated, the usefulness of LNA probing is reduced. Ideally, this technique would be well complemented by the possibility of locking ribose to the C2'-endo. Data from both C3' and C2' endo-locked residues would probe the ribose dynamics unambiguously. Nevertheless, the ability to lock a single residue in one conformation (C3'-endo) still presents an important step to understanding dynamics in RNA. With the available biochemical, kinetic and structural information for hairpin ribozyme, we can ultimately relate dynamics, structure and function in hairpin ribozyme and more generally small RNA catalysts.

REFERENCES

REFERENCES

1. Lipari, G., Szabo, A., and Levy, R. M. (1982) Protein Dynamics and Nmr Relaxation - Comparison of Simulations with Experiment, *Nature* 300, 197-198.
2. Levy, R. M., Karplus, M., and Wolynes, P. G. (1981) NMR Relaxation Parameters in Molecules with Internal Motion - Exact Langevin Trajectory Results Compared with Simplified Relaxation Models, *J Am Chem Soc* 103, 5998-6011.
3. Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules, *Chem Rev* 104, 3623-3640.
4. Ewald, C., Christen, M. T., Watson, R. P., Mihajlovic, M., Zhou, T., Honegger, A., Pluckthun, A., Caflisch, A., and Zerbe, O. (2015) A Combined NMR and Computational Approach to Investigate Peptide Binding to a Designed Armadillo Repeat Protein, *J Mol Biol* 427, 1916-1933.
5. Liu, Q., Shi, C. W., Yu, L., Zhang, L. H., Xiong, Y., and Tian, C. L. (2015) General order parameter based correlation analysis of protein backbone motions between experimental NMR relaxation measurements and molecular dynamics simulations, *Biochem Bioph Res Co* 457, 467-472.
6. Asami, S., Porter, J. R., Lange, O. F., and Reif, B. (2015) Access to C alpha Backbone Dynamics of Biological Solids by C-13 T-1 Relaxation and Molecular Dynamics Simulation, *J Am Chem Soc* 137, 1094-1100.
7. Allner, O., Foloppe, N., and Nilsson, L. (2015) Motions and Entropies in Proteins as Seen in NMR Relaxation Experiments and Molecular Dynamics Simulations, *J Phys Chem B* 119, 1114-1128.
8. Huang, J., and MacKerell, A. D. (2013) CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data, *J Comput Chem* 34, 2135-2145.
9. Hoogstraten, C. G., Wank, J. R., and Pardi, A. (2000) Active site dynamics in the lead-dependent ribozyme, *Biochemistry* 39, 9951-9958.
10. Blad, H., Reiter, N. J., Abildgaard, F., Markley, J. L., and Butcher, S. E. (2005) Dynamics and metal ion binding in the U6 RNA intramolecular stem-loop as analyzed by NMR, *J Mol Biol* 353, 540-555.
11. Shajani, Z., and Varani, G. (2005) ¹³C NMR relaxation studies of RNA base and ribose nuclei reveal a complex pattern of motions in the RNA binding site for human U1A protein, *J Mol Biol* 349, 699-715.

12. Shajani, Z., and Varani, G. (2008) ^{13}C relaxation studies of the DNA target sequence for hhaI methyltransferase reveal unique motional properties, *Biochemistry* 47, 7617-7625.
13. Johnson, J. E., Jr., and Hoogstraten, C. G. (2008) Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using ^{13}C NMR spin relaxation and specific isotope labeling, *J Am Chem Soc* 130, 16757-16769.
14. 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.
15. Sargueil, B., Hampel, K. J., Lambert, D., and Burke, J. M. (2003) In vitro selection of second site revertants analysis of the hairpin ribozyme active site, *J Biol Chem* 278, 52783-52791.
16. Lipari, G., and Szabo, A. (1981) Nuclear magnetic resonance relaxation in nucleic acid fragments: models for internal motion, *Biochemistry* 20, 6250-6256.
17. 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, *J Am Chem Soc* 104, 4546-4559.
18. Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules .2. Analysis of Experimental Results, *J Am Chem Soc* 104, 4559-4570.
19. 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.
20. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat Struct Biol* 6, 212-216.
21. 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.
22. Julien, K. R., Sumita, M., Chen, P. H., Laird-Offringa, I. A., and Hoogstraten, C. G. (2008) Conformationally restricted nucleotides as a probe of structure-function relationships in RNA, *RNA* 14, 1632-1643.
23. Christiansen, J. K., Lobedanz, S., Arar, K., Wengel, J., and Vester, B. (2007) LNA nucleotides improve cleavage efficiency of singular and binary hammerhead ribozymes, *Bioorg Med Chem* 15, 6135-6143.

24. Donini, S., Clerici, M., Wengel, J., Vester, B., and Peracchi, A. (2007) The advantages of being locked. Assessing the cleavage of short and long RNAs by locked nucleic acid-containing 8-17 deoxyribozymes, *J Biol Chem* 282, 35510-35518.
25. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis, *Nature* 410, 780-786.
26. Cai, Z., and Tinoco, I., Jr. (1996) Solution structure of loop A from the hairpin ribozyme from tobacco ringspot virus satellite, *Biochemistry* 35, 6026-6036.