Structural Dynamics of the Hepatitis Delta Virus and Hairpin Ribozymes: Implications for Function

by

Kamali N Sripathi

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Medicinal Chemistry)
in the University of Michigan
2014

Doctoral Committee:

Professor Nils G. Walter, Chair
Professor Charles L. Brooks, III
Professor Heather A. Carlson
Professor George A. Garcia

I believe in taking a positive attitude toward the world, toward people, and toward my work. I think I'm here for a purpose. I think it's likely that we all are, but I'm only sure about myself. I try to tune myself into whatever it is that I'm supposed to be, and I try to think of myself as a part of all of us – all of mankind and all of life. I find it's not easy to keep these lofty thoughts in mind as the day goes by, but it certainly helps me a great deal to start out this way.

~ Jim Henson

Use what talents you possess: the woods would be very quiet if no birds sang there except those that sang best.

~ Henry Van Dyke

Life's like a movie, write your own ending
Keep believing, keep pretending
We've done just what we set out to do
Thanks to the lovers, the dreamers, and you.

~ Kermit the Frog and the Muppets

Dedication

To my family, who have felt every setback and every triumph 1,000 times more than I have.

Acknowledgements

I am fortunate to have gained a very strong support system during my graduate school career. I would first like to thank my advisor, Dr. Nils Walter, for taking a chance on a very unconventional student that most would not. With his guidance, I have been able to accomplish things in the lab that I never would have believed possible as a frightened, confused first year. I would also like to thank my committee members, Dr. George Garcia, Dr. Charles Brooks III, and Dr. Heather Carlson, for their valuable feedback during my development towards becoming an independent scientist.

I am also very grateful to all of my collaborators. I thank Dr. Jiří Šponer, Dr. Nadʻa Špačková, and Dr. Kamila Réblová for their patience and guidance in helping me learn the nuts and bolts of molecular dynamics simulations and the AMBER programs. I would also like to thank Dr. Michal Otyepka, Dr. Pavel Banáš, and Mr. Vojtech Mlýnký for all of the extensive conversations about ribozymes and quantum mechanical calculations. I respect and admire all of you as much for your great intelligence as for being such kind and nurturing people. I also thank Dr. Charles Brooks III and Mr. Garrett Goh, for all of the conversations and information exchange surrounding the hairpin ribozyme and constant pH MD simulations. I look forward to generating excellent products with you both and with Nils in the near future.

I would also like to thank the wonderful mentors I have acquired during graduate school. Thank you to Dr. Lisa B. Lewis of Albion College, for our conversations about and her insights into teaching and academic life, and for the opportunity to guest-teach at Albion

College. I would also like to thank Dr. Katiuska Luna Cancalon for her continued guidance on both my personal and professional journeys, and for her support in helping me realize my true self.

I would also like to thank all past and present Walter lab members; I never knew such a wonderful group of simultaneously talented and helpful people existed until I joined the lab. All of you have been fabulous to work, to learn, and to grow with, and I am so grateful to have found this environment without even knowing that I needed it. I would especially like to thank my first mentors in the lab, Dr. Sethuramasundaram Pitchaya for his immense patience and kind guidance at the start of my tenure in the Walter lab, and Dr. Alex Johnson-Buck for his gentle manner, his great help especially during my candidacy preparations, and all of our conversations about life in general. I would also like to thank Ms. Wendy Tay, Mr. Paul Lund, and Mr. Thomas Corey Custer, for being such great resources and wonderful friends.

I consider myself very lucky to have a large group of marvelous friends, both new and old. I thank them all for their everlasting support. I would especially like to thank Dr. Fardokht Abulwerdi, for being such a wonderful member of our cohort. I have loved growing into an independent thinker with you. Thank you, Dr. Beth Girnys-Kuszpit and Dr. Jessica Bell, for being such great role models, and people in the department that I could always take my questions to. Thank you to Ms. Nancy Wu, for her sweetness and for allowing me to help her. I have loved being able to periodically focus on helping you to take my mind off my own jobs, dear. I would also like to thank Ms. Megan Baker, for always being willing to give me a good dose of reality, and for being the first of my students to become a friend. Thank you to Dr. Ningkun Wang, for being such a gentle and detail-oriented resource during my data meeting and thesis preparations. Thank you, Mr. Leshern Karamchand, for our

conversations about life, love and the universe. I have loved our talks, and sharing a year of teaching experiences with you. I would also like to thank my wonderful friends, Ms. Madeleine Ewers, Ms. Sara Worsham, for always having faith in me, being proud of me, helping me laugh at the times I needed it most, and for always thinking more of me than I think of myself. I am so glad that we all got to grow up, and are still growing, together.

Lastly, I would like to thank the fantastic Ms. Larissa Kipa and Dr. Ramya Krishnan for always being my magnificent best friends. Thank you, Lara-bean, for always taking my side, for always understanding my obsessions, and for being such a gentle, intelligent, exemplary, kind, compassionate and loving friend, not only to me, but to everyone you meet. You are a constant inspiration to me in a thousand ways, and I am so grateful to have you in my life. Thank you, darling Ramya, for being my wonderful best friend whether we are in the same country or if you are halfway around the world, for always thinking that I am intelligent and capable, for always listening to all of my issues, for wanting to defend me no matter what, for reminding me that it is okay to get angry, for our talks, our walks and our yam fries. I am so happy that we found each other when we did. Even though I miss you terribly, I know we will be friends forever.

I would also like to thank Dr. Kathleen Nolta for the very special relationship that I depend on so heavily. Thank you so much, Kathleen, for your wisdom, for laughing with me, for being so patient with me, for always seeing me for everything I am, and for everything that you have done and continue to do for me. I respect and admire you so much, and I would not have made it this far if not for you.

Lastly and most importantly, I would like to thank my wonderful family. This degree really belongs to them, because they have felt every setback and triumph at least a thousand

times more than I have. I thank them for keeping me sane and well fed, all while enduring my stressed and unhappy moods. Thank you all for always thinking the best of me and for always being proud of me. Everything I accomplish in life is because of your undying love

Table of Contents

List of Figures

List of Tables

Abstract

Small catalytic RNAs (ribozymes) are a class of RNAs less than 200 nucleotides in length, which are able to catalyze both cleavage and subsequent re-ligation of their own phosphodiester backbones. Cleavage results in formation of 2',3' cyclic phosphate and 5' hydroxyl termini. There are five naturally occurring members of this group: the hepatitis delta virus (HDV) ribozyme, the hairpin ribozyme, the hammerhead ribozyme, the Varkud Satellite (VS) ribozyme, and the *glmS* ribozyme. All of these RNAs employ a distinct active site fold and combination of mechanisms to perform their chemical function. Despite extensive biochemical and structural characterization, the structural requirements of catalysis for the HDV and hairpin ribozymes have not been ascertained. The work in this thesis interrogates the conformational dynamics of these ribozymes to provide deeper insight into their structure-function relationships. In Chapter 2, we utilize a combination of ensemble fluorescence resonance energy transfer (FRET) experiments and molecular dynamics (MD) simulations to outline a portion of the free energy folding pathway of the HDV ribozyme, rationalizing all extant crystal structures. Our MD simulations in Chapter 3 define specific structural roles for the two GU wobble pairs in the HDV ribozyme active site through mutational analysis. We show in Chapter 4, with the aid of newly developed constant pH MD simulations for nucleic acids, that experimental pK_a measurements of an active site residue in the hairpin ribozyme reflect a minor conformational population. Taken together, the results in this thesis reveal the complex network of noncovalent interactions within the HDV and hairpin ribozymes that can have either favor or disfavor catalysis. Our results thus provide

insights into the function of these two catalytic RNAs through interrogations of their structural dynamics. Such knowledge consequently provides a deeper understanding of catalytic RNAs, and all non-coding RNAs in general. A thorough knowledge of these complex biomolecules is of growing importance given the increasing recognition of the crucial roles played by RNA in all three domains of life.

Chapter 1

INTRODUCTION: THE IMPORTANCE AND UBIQUITY OF RNA

1.1 Beyond central dogma

The central dogma of molecular biology describes the progression of information during the propagation of life:[1] DNA is used as a template to transcribe RNA, which is then translated into protein. Proteins consequently fulfill all structural and catalytic roles required for cell birth, growth and death. This progression, formally published by Francis Crick in 1970, describes what he termed "general transfers", at the time believed to occur in all cells (Fig. 1.1, black arrows). He also summarized a few "special transfers" (Fig. 1.1, blue arrows) in this work, some of whose existence had been supported by the discovery that many viruses use RNA as their genetic material.[1] The "special transfers" involving RNA required catalytic ability in this biomolecule, which was not observed until the 1980's. Ribonuclease P (RNaseP)[2] and the *Tetrahymena* group I intron[3] were the first two RNAs discovered that are able to facilitate chemical reactions. The discovery of RNase P as a ribozyme in its ability to process tRNAs for protein translation expanded the catalytic repertoire of RNA.[4] The *Tetrahymena* intron further enhanced this catalytic ability by mimicking the behavior of transposable elements: it self-splices from the rRNA of the *Tetrahymena*.[4] The discovery of these two "ribozymes" thus provided the potential for RNA to function beyond the limitations of its four bases, while still being a vessel for genetic information. With the discovery of catalytic and transposable abilities in RNA, the "RNA World" hypothesis was

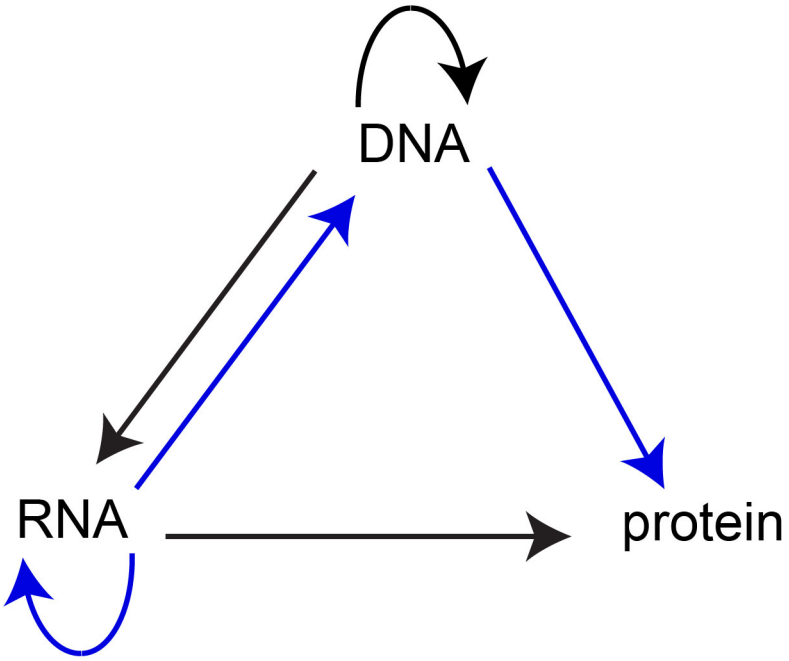

Figure 1.1 The central dogma of molecular biology.

Adapter from [1]

born.[4] Since then, catalytic RNAs have been discovered in all sizes and a plethora of contexts.[5, 6]

1.2 Catalytic strategies of RNA enzymes

Since the discovery of RNase P and the *Tetrahymena* group I intron, catalytic RNAs have been found to catalyze a variety of reactions in a wide array of sizes.[5, 6] The catalytic power of large ribonucleoprotein complexes such as the ribosome and spliceosome have been proven to be due to their RNA components.[5, 6] The reactions catalyzed by both RNase P and the group I intron, the later of which is considered a ribozyme of intermediate size,[7, 8] have been shown to be metal-dependent. Smallest of all, the autolytic ribozymes catalyze site-specific cleavage of their own phosphodiester backbones primarily through the use of nucleobases.[5-8]

The ribosome, one of the largest ribozymes, catalyzes a very specific reaction (Fig. 1.2a).[6] Its role in the formation of peptide bonds is distinct from the phosphoryl transfers of other ribozymes, and results in elongation of nascent polypeptide chains. The amino acids required for polypeptide chain lengthening are brought to the ribosome by charged tRNAs, to which the amino acids are covalently linked.[6] The A(aminoacyl)-site tRNA molecule is aminoacylated on the 3' hydroxyl of the terminal A76, as is the tRNA carrying the growing peptide chain in the P(peptidyl)-site. Peptide bond formation occurs when the α amino group lone pair of the A-site tRNA's amino acid attacks the C-terminal carbonyl group of the growing peptide. This attack effectively transfers the growing peptide chain from the P-site tRNA to the A-site tRNA, one of the steps which allows a new aminoacylated tRNA to enter the A-site.[6]

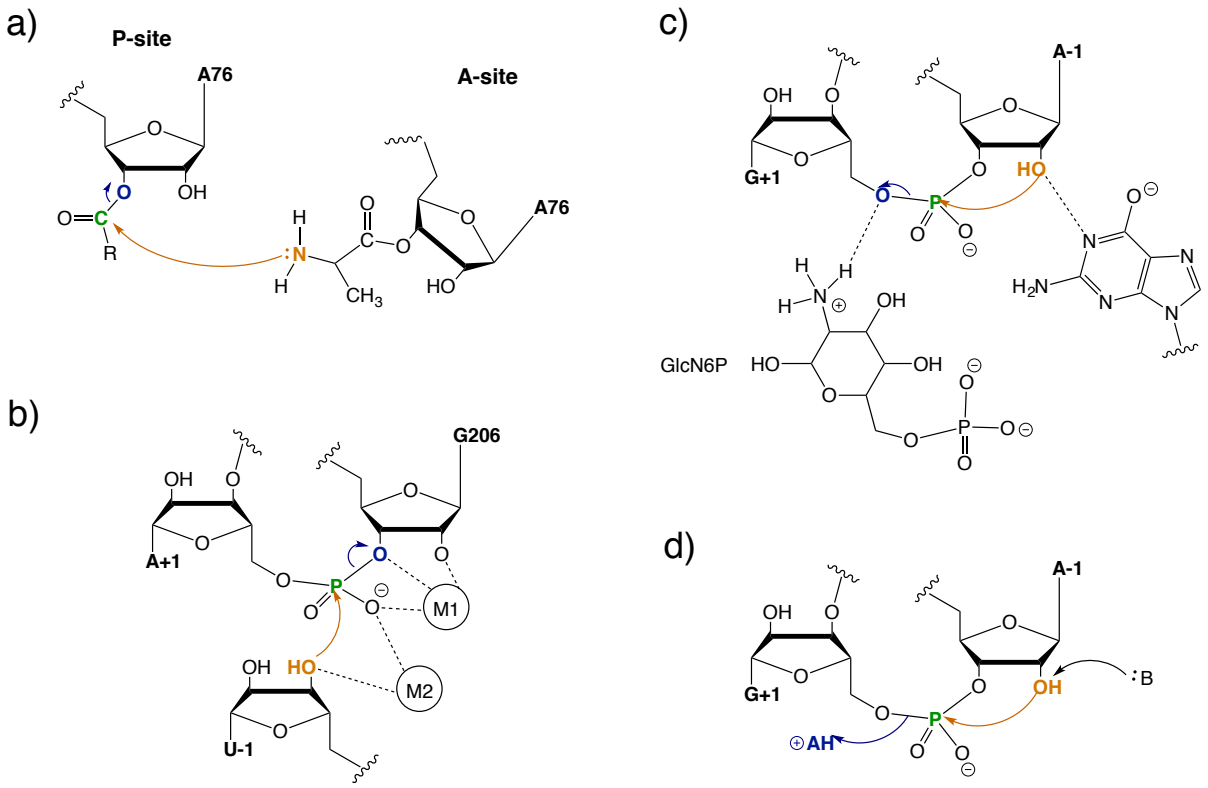

Figure 1.2 Diverse mechanisms of catalytic RNA.

(a) Mechanism of ribosomal peptide synthesis. (b) The second step of splicing in the *Tetrahymena* group I self-splicing intron. (c) Putative cleavage reaction in the *glmS* ribozyme. (d) Mechanism for general acid-base catalysis in the HDV and hairpin ribozymes, the focus of this thesis. Adapted from [6]

Self-splicing group I introns, ribozymes of a more moderate size, have been especially interesting due to their mimicry of transposable element dynamics;[4-6] they excise themselves from mRNA in a two-step mechanism. In the *Tetrahymena* group I intron, the first step consists of a nucleophilic attack in *trans* by guanine at the intron's 5' splice site.[6] In the second step (Fig. 1.2b), the intron is freed upon a second attack, that of the 5' exon at the 3' splice site. This step ultimately results in the ligation of the two exons. Structures of the *Tetrahymena* group I intron have illustrated two metal ions close to the cleavage site, both coordinated to five ligands each during the second step of splicing.[6] As initially hypothesized by Steitz and Steitz, one metal ion activates the 3' attacking hydroxyl of the exon, while the second metal ion activates the leaving group.

These findings lead to the hypothesis that all ribozymes are obligate metalloenzymes.[5-7, 9] Further information obtained on the class of smallest catalytic RNA (ribozymes) proved this hypothesis to be false.[6-8] A unique method of catalysis employed by a ribozyme is that of the *glmS* ribozyme. This ribozyme is also a riboswitch, in that it undergoes conformational change due to the binding of its ligand, glucosamine-6-phosphate (GlcN6P).[6] The *glmS* ribozyme occurs in the 5' untranslated region of the mRNA that codes the enzyme responsible for the synthesis of this small biomolecule, glucosamine-6-phosphate synthetase. The conformational change that occurs upon GlcN6P binding enables the *glmS* ribozyme to cleave its mRNA and downregulate production of GlcN6P.[6] This small catalytic RNA adopts a double-nested pseudoknot conformation to carry out its function.[5, 6] Crystal structures of the *glmS* ribozyme indicate that the position of bound GlcN6P may enable it to aid in catalysis, potentially as a general acid to facilitate the leaving of the 5' hydroxyl leaving group (Fig. 1.2c). Bound GlcN6P may thus serve as a cofactor in *glmS* ribozyme cleavage,

similar to the roles that protein cofactors play.[6] This recently implicated method of ribozyme catalysis further broadens the repertoire of strategies available to RNA to catalyze chemical reactions.

Although unique, the *glmS* ribozyme's catalytic strategy is not the only one employed by small catalytic RNA. The group is composed of four other naturally occurring members: the Hepatitis delta virus (HDV) ribozyme,[10-12] the hairpin ribozyme,[13-17] the hammerhead ribozyme,[18] and the Varkud Satellite (VS) ribozyme. These RNAs, < 200 nucleotides in length, facilitate site-specific autolytic cleavage into 2'3' cyclic phosphate and 5' hydroxyl termini. They adopt a wide variety of conformations and a few discrete methods to carry out site-specific cleavage, as discussed briefly for the *glmS* ribozyme (see above). In contrast to this ribozyme's potential for cofactor-mediated general acid-base catalysis, biochemical and structural data have indicated that the HDV and hairpin ribozymes, the focus of this thesis, employ general acid-base catalysis believed to be facilitated by their own endogenous nucleobases. In the HDV ribozyme, C75 (Fig. 1.3c) is believed to be the catalytic nucleotide that acts either as a general acid or a general base in catalysis[10-12, 19-23] (either AH+ or :B in Fig. 1.2d; discussed further in section 1.3.1). There are several candidates for general acids and bases in the hairpin ribozyme active site (Fig. 1.3f). A38 has been implicated as the general acid (AH+ in Fig. 1.2d),[24-27] while G8 and the nonbridging oxygens of the scissile phosphate have been implicated as the general base (:B in Fig. 1.2d).[28-31] Alternatively, both A38[32] and the nonbridging oxygens[33, 34] have been implied as proton shuttles (i.e., both AH+ and :B in Fig. 1.2d). Both the HDV[35, 36] and hairpin ribozymes[37] are "*cis*-acting" (catalytic nucleotide and scissile phosphate on the same strand) *in vivo*. However, characterization of

both ribozymes has been carried out in *trans* to probe their therapeutic utility and to elucidate their structure function relationships.[38-41]

1.3 Tertiary structures of the hepatitis delta virus and hairpin ribzoymes

1.3.1 The hepatitis delta virus ribozyme

The hepatitis delta virus (HDV) ribozymes are housed within the 1.7 kb genome of the hepatitis delta virus.[35] The virus is a negative-strand RNA virus that is only able to co-infect or super-infect hosts already infected with hepatitis B.[36] The resulting exacerbation of symptoms causes liver cirrhosis and hepatosarcoma among other chronic effects. The virus replicates via double rolling circle replication, resulting in generation of complementary multimeric genomic and antigenomic RNA pairs.[35, 36] These multimers are cleaved into monomers due to action of the genomic and antigenomic ribozymes. Ligation occurs through an as-yet unknown mechanism, and these RNAs are converted to DNAs for viral replication using the host's machinery.[35, 36] Although no HDV ribozymes have ever been observed to re-ligate their cleavage products, this is also a necessary step that must be catalyzed by the ribozyme in order for viral replication to continue.[35, 36] The reverse transcription of the re-ligated viral RNA to DNA is likely coordinated through the action of a converted RNA-dependent RNA polymerase, assumed to be commandeered by the virus from the host.

The genomic and antigenomic HDV ribozymes are thus responsible for a crucial step in viral replication. They adopt a complex double-nested pseudoknot in order to accomplish their cleavage goals (Fig. 1.3a-c). The psuedoknots are composed of 5 helices, P1-P4 and P1.1, which form two coaxial stacks (P2 | P3 and P1 | P1.1 | P4). Several specialized motifs stabilize this structure, such as a pair of A-minor motifs above and a C41 quadruple interaction below the active site.[11, 12, 42] The active site itself contains two G:U wobble pairs,

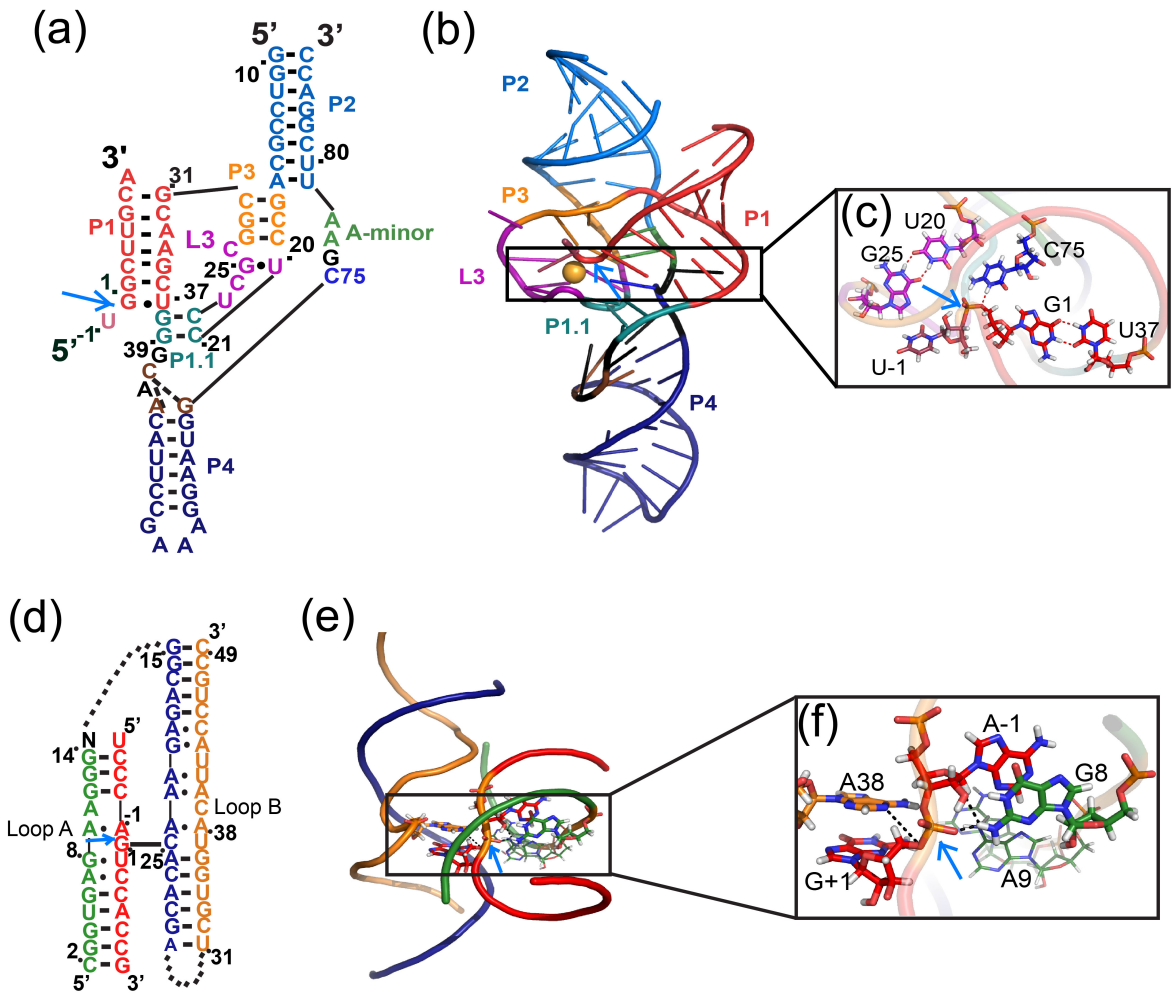

Figure 1.3 Secondary and tertiary structures of the ribozymes discussed in this thesis.

Secondary structure (a), tertiary structure (b) and active site close-up (c) of the latest precursor of the HDV ribozyme.[10] Helices are denoted with distinct colors in panels (a) and (b), that are reflected in the carbons in panel (c). Secondary structure (d), tertiary structure (e), and active site close-up (f) of the junctionless hairpin ribozyme.[16] Individual strands are denoted in distinct colors in panels (d) and (e), and are reflected in the carbons in panels (e) and (f). Atom colors are as follows in panels (c) and (e): nitrogens are in blue, oxygens are in red, hydrogens are in white, and phosphorus is in orange. The blue arrows in all panels denote the scissile phosphorus.

most recently believed to be a *cis* and *trans* wobble, which each lend their unique metal binding capabilities in support of catalysis. Although active in a variety of monovalent and divalent ions, the HDV ribozymes show optimal activity in magnesium ions, which are also necessary for proper folding.[35] This complex tertiary structure forms a deep electronegative active site cleft, and enables the catalytic C75 (in the genomic ribozyme and C76 in the antigenomic form) to come within close proximity to the scissile phosphate.[10-12] Cleavage has been shown to proceed through a general acid-base mechanism, with C75 believed to act as either the general acid or the general base.[20, 21]

1.3.2 The hairpin ribozyme

The hairpin ribozyme was initially discovered in the negative strand of a satellite RNA of the tobacco ringspot virus.[37] Unlike the HDV ribozyme, both cleavage and ligation reactions have been observed in this ribozyme.[37] Its tertiary structure has proven to be simpler than that of the HDV ribozyme, although it catalyzes the same site-specific cleavage transesterification reaction.[7, 8, 43] In its most basic form, the ribozyme consists of two stem-loops, either connected by linkers or separate (Fig. 1.3d-f).[38, 39, 41] Loop A contains the cleavage site and the G8, A9 and A10 nucleotides that have been shown to enhance catalysis.[28, 29, 31, 44] Loop B contains the A38 nucleotide, also shown to play an indispensible role in catalysis.[45] Catalysis occurs when loops A docks onto loop B, forming a series of complex tertiary interactions that are crucial for optimal activity (Fig. 1.3e,f): a ribose zipper, and binding pockets for the G+1 and U42 nucleotides.[46-48]

1.4 Elucidation of ribozyme catalysis using both structural and functional data

1.4.1 Functional hypotheses from HDV ribozyme structures

Despite the differences in geometric complexity of the HDV and hairpin ribozymes, they are both believed to undergo general acid-base catalysis during cleavage.[35] The first crystal structure of the HDV ribozyme, a product, resolved C75 within close proximity to the G1(O5') leaving group, leading to the hypothesis of C75 acting as a general base in the forward cleavage reaction. Initial imidazole rescue experiments[21] and later molecular dynamics (MD)[49] and quantum mechanical/molecular mechanical (QM/MM)[19] corroborated this theory, but a large body of work also countered it or provided ambiguous data. Assessment of C75 ^{13}C-NMR signals in the product ribozyme indicated no significant increase over pK_a of free cytosine base, but did not rule out the possibility that C75's pK_a could be elevated in the transition state.[23] Subsequently, the structure for a precursor HDV ribozyme was solved, which positioned the catalytic nucleotide at equal distances from the U-1 hydroxyl and the G1 5' leaving group.[12] Additional pH-rate profiles and phosphorothioate experiments supported a general acid role for C75,[20,22] further supported by Raman crystallography measurements indicating anticooperative behavior between an active site magnesium and protonated C75.[50] The latest precursor reveals a more product-like global conformation and C75 resolved near the oxygen of the G1 5' hydroxyl, poised to act as the general acid (Fig. 1.3c).[10] Thus, structural data on the HDV ribozyme presents conflicting hypotheses regarding its mechanism of cleavage, underscoring the need for further tests to elucidate all aspects of its function.

1.4.2 The effects of multiple hairpin ribozyme structures on theories for catalysis

The first docked crystal structures of the hairpin ribozyme were not solved until 2001,[14, 15] and were consequently able to deepen understanding of functional studies. Early studies showed that the hairpin ribozyme utilized an ion-independent mechanism of catalysis,[51, 52] which was confirmed by the lack of ions, as well as water, in the aforementioned structures. These studies also revealed that the cleavage mechanism was pH-dependent, which could be at least partially explained by the structure of transition state mimics:[15] A38 was resolved near the G+1 5' hydroxyl group, implying that its pK_a is shifted enough in the transition state to be protonated, potentially also enabling A38 to act as a general acid in catalysis. This observation validated previous results showing a heightened pK_a[25, 53] and an essential role in catalysis[24, 45] for A38, although the specific nature of that role was not ascertained. Additionally, all these structure[14, 15] show that G8 may act as a hydrogen bond donor to the U-1 2' hydroxyl, consistent with experimental observations that the N1 position of guanine is protonated throughout the cleavage reaction.[28, 29, 31] However, the latter study still proposed several potential roles for a protonated G8 in the active site, leaving its role unclear as well. Subsequent crystal structures containing variants at both the G8 and A38 positions attempted to further elucidate their roles in catalysis.[13, 16, 26, 54] Variants at both positions exhibited increased dynamics in the active site, but were not able to distinguish between the different hypothetical pathways employed by the protonated forms of the bases.

Later computational studies sought to further reconcile experimental data, but resulted in the formation of new mechanistic hypotheses. MD simulations suggested the potential of A38 to act as a general base as well as the possibility of its role as a proton shuttle.[55,32] The former studies found, however, that a canonical G8 and protonated A38 were

most consistent with the crystal structure data available at the time. QM/MM studies explored the possibility of nonbridging oxygens to act as proton shuttles as well.[33, 34, 56] Thus, despite a wealth of structural data to rationalize mechanistic studies, several plausible hypotheses still persist regarding the cleavage mechanism of the hairpin ribozyme.

1.5 Using Structure to Elucidate Ribozyme Function

Despite the considerable amount of work conducted on both the HDV and hairpin ribozymes, no universal theories on their method of catalysis exist (as outlined in the previous section). Building upon the long history of structure-function studies upon ribozymes, the work in this thesis seeks to deepen the understanding of ribozyme catalysis using structural methods of interrogation. The results outlined herein provide detailed representations of global and local conformational changes surrounding catalysis in the HDV ribozyme, as well as structural effects on the pK_a values of hairpin ribozyme active site residues.

1.5.1 The power of synergy: experimental and computational methods in concert

Our group has previously demonstrated the combined power experimental and computational methods to study structure-function relationships in ribozymes.[57-59] We have been able to strengthen our hypotheses and results by providing atomic resolution to RNA footprinting and cleavage data, while investigating the real-world effects implied by our observations from MD simulations. We take further advantage of this strategy in Chapter 2, to elucidate potential catalytic pathways of the HDV ribozyme along its free energy landscape. Our experiments utilize ensemble fluorescence resonant energy transfer (FRET), a powerful distance-dependent technique that allows us to monitor global conformational change in a variety of HDV ribozymes during catalysis. Additionally, we synergistically use MD

simulations to study local conformational changes of the HDV ribozyme active site surrounding the cleavage event.

1.5.2 Computational methods for understanding ribozyme active site dynamics

MD simulations are useful for providing a window into the rapid, small-scale dynamics of a molecule not easily accessible by experimental techniques.[60-62] These all-atom simulations depend on a high-resolution NMR or crystal structure and an empirical molecular mechanical force field. [63] Our group has used these simulations to gain insight into conformational dynamics of RNA that underlie experimentally measurable behaviors.[32, 49, 55, 57-59, 62, 64-66] Although computational description of RNA has proved challenging, due, in large part to its high conformational flexibility, special attention in recent years has been given to reducing force field artifacts.[67] Thus, in Chapter 3, we use MD simulations, using these latest force field modifications,[63, 67] to investigate the effects of active site GU wobbles on catalytic fitness of the HDV ribozyme.

Despite the great utility of MD simulations, sampling of different protonation states within a single simulation is not possible. However, a free energy perturbation (FEP) technique has been recently developed and tested on ionizable nucleic acids.[68-71] This technique enables the coupling of the thermodynamic evolution of the λ variable and proton dynamics.[72] Thus, these types of simulations allow sampling of protonated and unprotonated states of specified conjugate acid/base pairs within a single simulation. In Chapter 4, we implement these simulations to test the protonation states of various key bases within a variety of structured RNA and DNA.

The understanding of RNA's roles and functions in cellular life cycles has increased dramatically since the first elucidation of Francis Crick's central dogma of biology.[1] The

discovery of the first ribozymes[2, 3] in the 1980's enabled the ideas of RNA's functions in biology to take a particularly unexpected turn. Since then, RNA's complex structure-function relationships have been extensively studied. Taken together, the results outlined in this thesis contribute significant depth towards a complete understanding of these structure-function relationships in small ribozymes, and as an extension in all highly structured RNA. Such an understanding is crucial, given the increased occurrence of small ribozymes in all kingdoms of life,[73-76] and the growing presence of RNA as therapeutic agents.[77]

Chapter 2

DISPARATE HDV RIBOZYME CRYSTAL STRUCTURES REPRESENT INTERMEDIATES ON A RUGGED FOLDING FREE ENERGY LANDSCAPE[1]

2.1 Introduction

The hepatitis delta virus (HDV) ribozyme (Fig. 2.1) is a small catalytic RNA originally found in the genome and antigenome of the hepatitis delta virus. [35, 36, 78, 79] It has a minimal length of ~85 nucleotides and catalyzes the cleavage of its own phosphodiester backbone in *cis*, resulting in the formation of 5'hydroxyl and 2',3'-cyclic phosphate products. Self-cleavage of the genomic and antigenomic ribozymes, which share ~75% sequence identity, is an essential step in the double rolling-circle replication of the HDV. [35, 36] The overall fold of the HDV ribozyme consists of a double-nested pseudoknot encompassing helices P1, P1.1, P2, P3, and P4 (Figs. 1a,b), [11, 80, 81] which leads to a tight-knit structure with a deeply electronegative active site cleft. This cleft attracts metal ions and promotes cleavage by shifting the pK_a of an active site cytosine at position 75 (in

[1] Adapted from **Kamali N. Sripathi**, Wendy W. Tay, Pavel Banás, Michal Otyepka, Jiří Šponer, and Nils G. Walter, 'Disparate HDV Ribozyme Crystal Structures Represent Intermediates on a Rugged Folding Free Energy Landscape' (submitted to *J. Mol. Biol.*) K.N.S. performed fluorescent EMSAs, steady-state FRET assays and analysis on 1SJ3_U-1+, U-1, dU-1+, and U-1G+ simulations. W.W.T. performed radioactive cleavage assays, time-resolved FRET measurements, and analysis of U-1+, Product, Xtal_U-1+, dUGG+, Xtal_dUGG+, dUGG and Cis+ simulations. K.N.S., W.W.T. and N.G.W. wrote the manuscript.

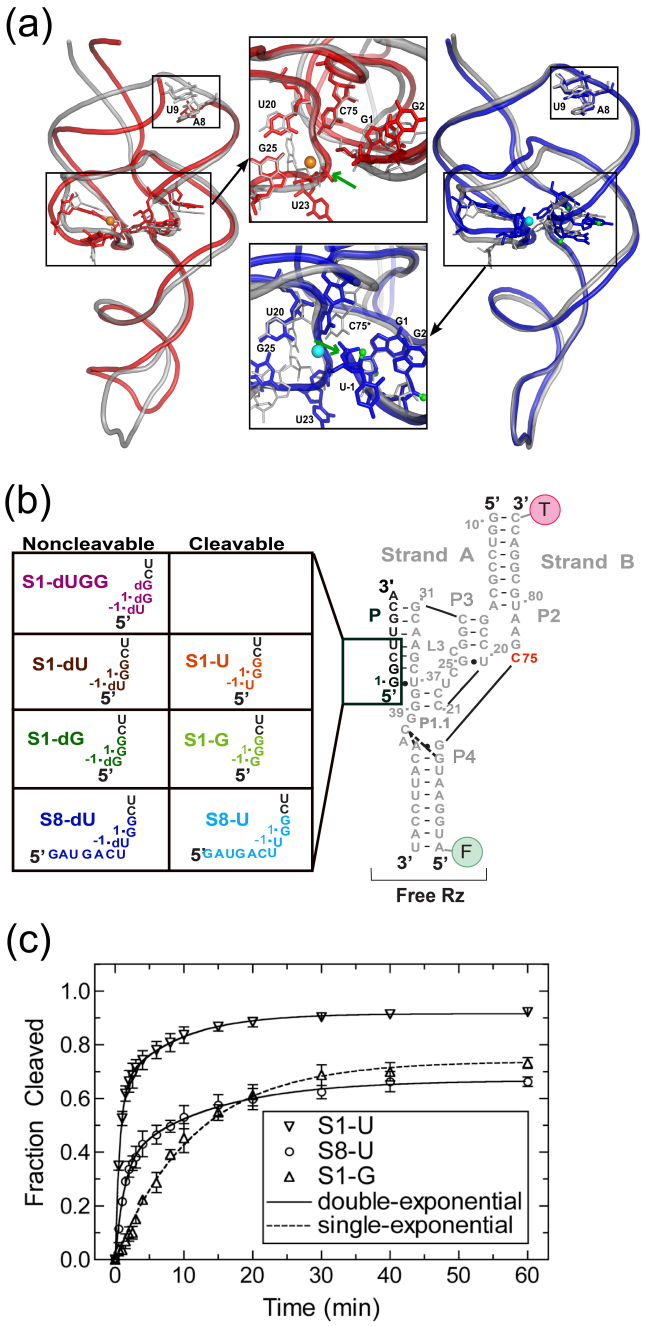

Figure 2.1 The trans-acting HDV ribozyme studied here.

(a) Superpositions of trans-acting (red) [10] and cis-acting (blue) [12] precursor crystal structures on the cis-acting product structure (silver) [11]. Active site magnesium ions are shown in orange and cyan for the trans- and cis-acting ribozymes, respectively. The smaller black boxes on both overlays highlight the region where there is a break in the RNA backbone for the trans-acting ribozyme. The green arrows indicate the cleavage site in both insets. The U-1 nucleotide and scissile phosphate were not resolved in the trans-acting precursor structure. (b) The 8 three-stranded precursor and product ribozymes used in our solution experiments. (c) Cleavage assay time courses for S1-U, S8-U and S1-G. Cleavage assays were conducted at 25 ºC in 40 mM Tris-HCl, pH 7.5, 25 mM DTT and 11 mM MgCl2. Data for S1-G were fit with a single-exponential increase function, while the data for S1-U and S8-U were fit best with a double-

exponential increase function. Standard deviations from at least three independent assays are shown as error bars.

the genomic ribozyme; C76 in the antigenomic form) towards neutrality. [11, 81] Based on these structural constraints, numerous HDV-like ribozymes were recently discovered in the human and other eukaryotic genomes, suggesting that this RNA motif may have widespread functions in RNA processing and gene regulation. [73, 82, 83]

In addition to numerous studies of self-cleavage by the *cis*-acting genomic and antigenomic HDV ribozymes, [20, 21, 35, 50, 57, 58, 81, 84-95] significant work has also been performed to investigate *trans*-acting forms of this ribozyme for enzymology and therapeutic purposes. [22, 23, 40, 80, 96-102] Generally, *trans*-acting HDV ribozymes exhibit a 1-2 order-of-magnitude slower catalytic rate constant compared to their *cis*-acting counterparts, likely due to a less tightly knit structure. [40] Both *cis*- and *trans*-acting ribozymes have been shown to exhibit in solution end-to-end lengthening when the catalytic precursor cleaves to form the product, with the *cis*-acting versions exhibiting more modest changes than the ~15 Å observed for *trans*-acting ribozymes. [80, 95, 99, 103]

Since 1998, a number of crystal structures of the *cis*-acting genomic ribozyme have been solved of both precursor and product forms. [10-12, 104] The emergence of the *cis*-acting precursor structure at up to 2.2 Å resolution was accompanied by the observation of a structural change upon release of the 5' product and catalytic metal ion, as C75 moves deeper into the active site cleft to form hydrogen bonds with C22's nonbridging oxygen and the 5' OH leaving group of G1, while the adjacent G25 rotates in conformation from *anti* to *syn.* [12, 49] These changes result in a collapse of the P1 and P3 helices towards the active site and a more compact catalytic core and concomitant overall lengthening of the product (Fig. 2.1a), [12] consistent with the results from solution probing. [80, 95, 99]

As is typical, the *cis*-acting precursor was modified to prevent self-cleavage during crystallization in these studies, which was accomplished by either a C75U mutation, crystallization in the absence of Mg^{2+}, or incorporation of a single 2'-deoxy modification at the cleavage site, although the latter structure was low in resolution.[12] In 2010, a different team solved a low-pH, 1.9 Å-resolution crystal structure of a *trans*-acting precursor HDV ribozyme in the presence of Mg^{2+}, inactivated with three 2'-deoxy modifications at the -1, 1 and 2 positions around the cleavage site (Fig. 2.1a, bottom inset).[10] Consistent with the fact that molecular replacement with the *cis*-acting product was used to elucidate its structure, the new trans-acting structure features both active-site and global conformations similar to those of the product (Fig. 2.1a). [10, 11] For example, G25 is in the *syn* conformation and the relative positions of the P1 and P3 helices are more similar to those in the *cis*-acting product than the precursor ribozyme (Fig. 2.1a). Such close resemblance to the product suggests that this ribozyme may capture an activated conformer past the conformational change needed to activate all previous ribozyme versions. This hypothesis would predict a fast cleavage rate constant; however, the newly crystallized HDV ribozyme cleaves similarly slowly to other *trans*-acting versions. [10] Additionally, the substrate strand was sufficiently flexible (and/or nicked) so that no electron density was observed for the scissile phosphate and upstream U-1 nucleotide, further increasing resemblance with the cleavage product. The authors modeled these moieties using the cleavage site of the hammerhead ribozyme. [10] Consequently, questions remain concerning the relationship of the various proposed and experimentally observed conformations to each other and to the functional activity of the HDV ribozyme.

To probe the functional dynamics of the newly crystallized *trans*-acting HDV ribozyme in solution, here we have employed a combination of experimental and

computational approaches. Fluorescence resonance energy transfer (FRET) gel shift assays revealed a heterogeneous population of molecules in solution at physiological pH, suggesting that the crystal structure may represent only a single conformation selected by crystallization. Steady-state and time-resolved FRET measurements consistently detected an ~8 Å lengthening of the end-to-end distance along the P1-P4 axis upon cleavage in solution, an elongation between those previously found for *trans*- and *cis*-acting HDV ribozymes. [80, 95, 99, 103] We also performed a total of 1.8 μs of molecular dynamics (MD) simulations that showed that the hammerhead ribozyme model of the substrate conformation in the active site is compatible with a favorable catalytic in-line fitness (a measure of the poise to undergo catalytic transesterification), as long as C75 is protonated. By contrast, intermolecular crystal contacts and 2'-deoxyribose modifications found in the crystal structure near the active site result in unfavorable in-line fitness, offering an explanation for the lack of experimental electron density in this region. A *cis*-acting version of the crystallized ribozyme interconverts more frequently between unfavorable and favorable fitness, consistent with a lowered kinetic barrier towards cleavage. We lastly find that a U-1G mutation results in long-lived states of unfavorable fitness, consistent with its low cleavage activity and the conservation of U-1 in clinical isolates of the virus. Taken together, our data are consistent with a parsimonious model wherein the existing precursor crystal structures, determined for distinct constructs at varying pH, represent alternate RNA conformations on a rugged folding free energy landscape.

2.2 Materials and Methods

2.2.1 Preparation of RNAs

All sequences used here were based on the recent *trans*-acting HDV ribozyme crystal structure. [10] The GAAA loop of the P4 stem was replaced with an extension of the P4 helix to generate a three-stranded version of the ribozyme (Fig. 2.1b), as previously described.[80] Additionally, U55 in the P4 helix was omitted to prevent potential misfolding. All RNA oligonucleotides were purchased from the Keck Biotechnology Resource Laboratory at the Yale University School of Medicine (http://keck.med.yale.edu/oligo). Some strands were ordered already 2'-deprotected and cartridge-(Glen-Pak) purified, while others were 2'-deprotected in-house according to the manufacturer's instructions and then purified as described[105] by denaturing, 8 M urea, 20% polyacrylamide (19:1 acrylamide:bisacrylamide) gel electrophoresis. The bands of interest were cut out, diffusion-eluted into either 1 mM EDTA or "Crush 'n Soak" buffer (0.5 M NH_4OAc, 0.1 mM EDTA), ethanol-precipitated and resuspended in autoclaved double-distilled water, as previously described.[105] Oligonucleotides that were cartridge purified by Keck were analyzed by denaturing, 8 M urea, 20% polyacrylamide gel electrophoresis to ensure high purity. RNA concentrations were calculated from absorption at 260 nm.

Strand B was purchased already labeled with a 5'-fluorescein and a 3'-C7-amino modifier in preparation for further in-house labeling with 5-carboxytetramethylrhodamine (TMR) succinimidyl ester (Invitrogen) according to the manufacturer's protocol, using overnight incubation at room temperature (22 ^{o}C). The doubly-labeled RNA was then purified by denaturing, 8 M urea, 20% polyacrylamide gel electrophoresis. The bands of interest were cut, diffusion-eluted either into 1 mM EDTA or Crush 'n Soak buffer. The

RNA was then ethanol-precipitated and resuspended using autoclaved, double-distilled water, similar to previous protocols.[105] The concentration of the doubly-labeled Strand B was determined from the absorption spectrum and estimated to contain ~35% of singly-labeled starting material.

2.2.2 Radioactive cleavage assays

Radioactive cleavage assays to determine ribozyme activity were performed similar to those previously described.[80] Cleavable substrate strands were 5'-end-labeled using T4 polynucleotide kinase (NEB) and [γ-32P] ATP. All cleavage assays were carried out under single-turnover conditions, where free ribozyme concentration was ~125-fold above the substrate concentration, unless otherwise noted. Standard assay conditions were a buffer of 40 mM Tris-HCl, pH 7.5, 25 mM DTT and 11 mM $MgCl_2$ at 25 ^oC, unless otherwise stated. Ribozymes were preannealed by mixing Strand A and twice the concentration of Strand B (labeled with fluorescein and TMR) under standard assay conditions without magnesium, heating to 90 ºC for 2 min, and then cooling to 25 ^oC for 10 min using a water bath. $MgCl_2$ was added to a final concentration of 11 mM. The free ribozyme was incubated at 50 ^oC for 10 min, followed by cooling to 25 ^oC for 10 min. The final ribozyme concentration, based on the concentration of Strand A, was 750 nM. Cleavage assays were initiated by adding a trace amount (< 6 nM) of 5'-^{32}P-labeled substrate. Alternatively, the conditions of a previous study [94] were followed by mixing a 250-fold excess of ribozyme over the ^{32}P-labeled substrate in 25 mM Tris-HCl, pH 7.0, heating at 90 ^oC for 2 min, and cooling to room temperature for 10 min. The mixture was then incubated at 37 ^oC for 2 min. Magnesium chloride was added to initiate the assay to be carried out at 37 ^oC. In either case, aliquots of 5 µl were taken at appropriate time intervals and quenched with 10 µl of stop solution (75% (v/v) formamide,

150 mM EDTA and 0.1% (w/v) each of bromophenol blue and xylene cyanol dyes). The 5' cleavage product was separated from the uncleaved substrate by gel electrophoresis using denaturing, 8 M urea, 20% polyacrylamide gel electrophoresis at a limiting power of 25 W for 1.5 to 2 h at room temperature. The gel was then exposed to a Phosphorimager screen for 18-22 h and scanned using a Typhoon 9410 Variable Mode Imager. Bands were quantified using ImageJ [106] and the calculated fraction cleaved at each time point was normalized to the sum of the substrate and product bands. Error bars were derived from 3-5 independent assays. Time traces of product formation were fitted with either a single-exponential ($y = A(1 - e^{(-k_{obs}t)})$) or double-exponential increase function ($y = A_1(1 - e^{(-k_{obs1}t)}) + A_2(1 - e^{(-k_{obs2}t)})$), utilizing the Marquardt-Levenberg nonlinear least-squares regression method (GraphPad Prism version 4.0b for Mac, GraphPad Software), where A is the pre-exponential factor and k_{obs} is the observed rate constant for the reaction.

2.2.3 Gel electrophoretic mobility shift assays (EMSAs)

Similar to previous protocols,[80] 50 pmol of substrate or product strand was incubated with 30 pmol of Strand A and 8 pmol of Strand B in standard buffer without magnesium (final concentration of 40 mM Tris-HCl, pH 7.5, 25 mM DTT), supplemented with 10% (v/v) glycerol. The samples were then incubated using the protocol outlined for cleavage assays, with two minor changes: the initial 90 ºC incubation lasted 1 instead of 2 min and the two 10-min incubations were conducted at room temperature on the bench top instead of in a 25 ºC water bath. $MgCl_2$ was added to a final concentration of 11 mM and the solution incubated for at least 10 min at 50 ºC. The electrophoresis unit had been allowed to pre-equilibrate at 4 ºC for at least 2 h. After incubation, the samples were loaded onto a nondenaturing 10%

polyacrylamide (19:1 acrylamide:bisacrylamide) gel containing 11 mM $Mg(OAc)_2$ and 40 mM Tris(OAc), pH 7.5. An electric field of ~10 V/cm was applied to the samples, which were electrophoresed for 12-14 h at 4 ºC. The gel was then imaged using a Typhoon 9410 Variable Mode Imager as described.[80] A dual-color image was generated by the scanner software and opened in ImageQuant. Relative band intensities were determined by analysis using ImageJ;[106] relative band percentages were obtained by summing the intensities from the fluorescein and TMR scans and normalizing them to their sum for each lane.

2.2.4 Steady-state FRET measurements

Solutions of 43-49 nM Strand B (final concentration after full assembly) in standard buffer were annealed as specified in the EMSA protocol. 140 µl of such a solution were added to a Starna quartz fluorescence cuvette, its fluorescein label excited at 470 nm (bandwidth 8 nm) and emissions detected at 520 nm (bandwidth 16 nm) and 585 nm (bandwidth 16 nm) for fluorescein and TMR, respectively. After Strand B signal proved stable, 5 µl Strand A was added in 3.25-fold excess over Strand B for S8-U, S8-dU, S1-U, S1-dU, and S1-dUGG and P runs, and in a 3.8-fold excess over Strand B for S1-G and S1-dG runs. The resulting change in FRET was recorded for 30 min. Once this signal was stable, traces were further followed upon the addition of a substrate or product strand in 5.42-fold excess over Strand B for S8-U, S8-dU, S1-U, S1-dU, S1-dUGG and P runs, and in 5.8-fold excess for the S1-G and S1-dG runs. All substrate and product strands were added in 5-µl volumes, for a total final volume of 150 µl. The resultant signal change was measured in all cases for at least 5,000 seconds. In the case of the free Rz runs, 5 µl of standard buffer was added instead of the third strand, so that the final reaction volume was consistent. The ratio of acceptor to donor signal was calculated and normalized to the starting baseline. Each trace was then smoothed using a

running average with a window of 10, and further normalized to control runs. Average curves of three runs for substrates and control, and four runs for the product, are reported here. Results were plotted as the relative FRET ratio decrease using Prism (GraphPad Prism version 4.0b for Mac, GraphPad Software).

2.2.5 Time-resolved FRET measurements

Preparations of 53 nM of Strand B (6 singly-labeled with fluorescein and 6 doubly-labeled with fluorescein and TMR) were incubated in standard buffer without magnesium for 1 min at 90 ºC, followed by 10 min at 25 ºC. $MgCl_2$ was added to a final concentration of 11 mM and the solution was incubated for 10 min at 50 ºC, followed by 10 min at 25ºC. Strand A was then added so that its final concentration would be 3-fold that of Strand B. Strands A and B were incubated at 25 ºC for 30 min. The third strand, i.e., one of the noncleavable substrate strands or the product strand, or buffer in the case of the free ribozyme sample, was then added to a final concentration that was 50-fold that of Strand B. Each of the 6 pairs of samples was incubated at 25 ºC for 2 h and then kept at 4ºC when not being used to measure decay lifetimes.

An ISS Alba confocal fluctuation microscope (based on an Olympus IX-81 frame) was used to collect fluorescein decay lifetimes. 40 µl of sample were loaded onto a coverslip, which was placed on a 60× 1.2 numerical aperture water immersion objective. A blue-enhanced supercontinuum laser (Fianium SC 400-6-PP, operated at 6 W) excited fluorescein at 488 nm with 100 ps pulse width and 20 MHz pulse picking. Decays were collected at 531 nm using a photomultiplier tube (Hamamatsu H7422P-40) feeding into a time-correlated single-photon counting acquisition card SPCM-830 (Becker & Hickl), with 4096 bins at 12.20 ps/bin. An instrument response function was obtained by measuring the back reflection

off a coverslip without sample. For each sample, decays of donor (fluorescein) only and donor with acceptor (TMR) present were collected. Decays were collected so that the maximum raw peak count was between 40,000 to 60,000 counts. Five consecutive measurement reads were collected for a single sample and there were three inter-day replicates for each sample.

The donor-only traces with fluorescence intensity I_D were fit with a sum of exponentials

$$I_{\mathrm{D}}(t) = \sum_i \alpha_{\mathrm{Di}} \exp(-t/\tau_{\mathrm{Di}})$$

to extract three intrinsic donor lifetimes (τ_{Di} is the intrinsic donor lifetime and α_{Di} is the fractional contribution of each lifetime). Using these donor lifetimes, the doubly-labeled samples decays with fluorescence intensity I_{DA} were fit with the following model for a distance distribution:

$$I_{\mathrm{DA}}(t) = (1-f_{\mathrm{A}})I_{\mathrm{D}}^{0}\sum_i \alpha_{\mathrm{Di}} \exp(-t/\tau_{\mathrm{Di}}) + f_{\mathrm{A}}I_{\mathrm{D}}^{0}\int_0^{\infty} P(r)\sum_i \alpha_{\mathrm{Di}} \exp\left\{-\frac{t}{\tau_i}\left[1+\left(\frac{R_0}{R}\right)^6\right]\right\}\mathrm{d}R$$

where I_D^0 is the donor intensity at time 0, f_A is the fraction of molecules labeled with acceptor (allowing to correct for any singly labeled RNA), R is the distance between donor and acceptor, R_0 is the Förster distance (55 Å for the fluorescein and TMR pair) and $P(r)$ is the distance distribution. Distance distributions assumed a three-dimensional Gaussian function model based on the equation:

$$P(r) = 4\pi R^2 c \exp\left[-a(R-b)^2\right]$$

where a and b are parameters that describe the shape of the distribution and c is the normalization constant.

2.2.6 Molecular dynamics simulations

All MD trajectories were generated using the AMBER10/11 suite of programs [107, 108] and the ff99 force field with recent modifications to χ (essential for RNA) and α/γ torsions (ff99bsc0χ_{OL3} version, current AMBER default).[63, 67] Structures were net-neutralized with Na^+ ions placed at locations of optimal electrostatics using the LEaP module of AMBER10/11 and solvated using a rectangular periodic TIP3P water box model[109] extended to a distance of ≥ 10 Å from any ribozyme atom or Na^+ ion. The ABC protocol[110, 111] was used to minimize and equilibrate structures prior to production runs. The minimization and equilibration protocol as well as the production runs were performed using the PMEMD module of the AMBER suite of programs. For the production runs, the particle mesh Ewald (PME) method was employed with a grid spacing of 1.0 Å, cubic-spline interpolation and a heuristic pair list update with a Lennard-Jones cutoff of 9.0 Å. Production runs were carried out at 300 K with constant-pressure boundary conditions and the Berendsen temperature coupling algorithm[112] with a time constant of 1.0 ps. The SHAKE procedure[113] was applied to constrain bonds involving hydrogen. Parameters for the protonated cytosine were derived using the Cornell et al. procedure.[114, 115] Partial atomic charges were calculated for a protonated N1-methylcytosine using restrained electrostatic potential (RESP) fitting[116] and the HF/6-31G(d) method. Ab initio calculations were done using the program Gaussian03.[117] These charges were then used for the protonated cytosine; phosphate and sugar charges were kept the same as in unprotonated cytidine phosphate. Residual charge was added to C1' to ensure a net charge of zero for the entire nucleotide. The U-1+ starting structure was prepared by globally aligning the hammerhead ribozyme 2OEU and *trans*-acting HDV ribozyme 3NKB[10] structures to copy the active site residues from the former to the latter. The

1SJ3_U-1+ structure was prepared by globally aligning the *cis*-acting 1SJ3[12] and *trans*-acting 3NKB[10] structures. The coordinates for U-1, the scissile phosphate, and G1 were then copied from 1SJ3[12] into 3NKB.[10] The resulting structures were minimized and equilibrated using a protocol that involved first restraining the newly incorporated nucleotides, and subsequently removing those restraints. After these stages, the structure was equilibrated using the ABC protocol. Following minimization and equilibration, the structure was carefully checked for the presence of unfavorable geometry before carrying out production runs. For the U-1G mutation, the U-1 residue was mutated to G-1 by first deleting the atoms unique to uracil that do not overlay with guanine. The AMBER LEaP module was then used to fill in the missing atoms for guanine. The *cis*-acting model of the *trans*-acting ribozyme was created similarly to that of the 1SJ3_U-1+ structure, by modeling in the J1/2 joiner nucleotides from the *cis*-acting product crystal structure (PDB ID 1SJ3).[11] First, the product and *trans*-acting structures were globally aligned, then the J1/2 joiner region was cut and pasted into the *trans*-acting structure using PyMol. The newly created *cis*-acting structure was minimized and equilibrated using a protocol that involved first restraining the newly incorporated J1/2 joiner region, and subsequently removing those restraints. Following minimization and equilibration, the structure was carefully checked for the presence of unfavorable geometry before carrying out production runs. The Sander module was used to set distance restraints for appropriate simulations (Table 2.2).

MD trajectories were analyzed using the AMBER ptraj module and the program Simulaid.[118] The ptraj module was used to calculate heavy-atom distances, angles and radial distribution functions. All snapshots in the figures were generated from 100 ps-averaged structures calculated using ptraj. Simulaid was used to determine hydrogen bonding time

traces using parameters of ≥ 120º donor-hydrogen-acceptor angle and ≤ 3.5 Å donor-acceptor distance. In-line fitness and other time traces were plotted using Matlab (2010b) (The MathWorks Inc.). In-line fitness (F) was calculated using the following equation:[119]

$$F = \frac{(\tau - 45)}{180 - 45} \times \frac{3^3}{d^3_{\mathrm{O2'},\mathrm{P}}}$$

where τ is the angle between the 2' oxygen nucleophile (O2'), the phosphorus electrophile (P) and the 5' oxygen of the leaving group (O5') and *d* is the distance between the 2' oxygen nucleophile and the phosphorus electrophile.

2.3 Results

2.3.1 The trans-acting HDV ribozyme preferentially cleaves a short U-1 substrate with biphasic kinetics

The crystallized *trans*-acting HDV ribozyme consists of two strands: the substrate and ribozyme segments. To incorporate site-specific fluorophore labels for solution probing by FRET, we instead synthesized three RNA strands by opening the non-conserved capping loop on helix P4 and labeling the 5' and 3'-ends of the downstream segment (Strand B) with fluorescein and tetramethylrhodamine, respectively (Fig 1b), following an established strategy.[80] We first tested the catalytic activity of the resulting two-strand ribozyme using radioactive single-turnover cleavage assays at 25 ºC on three cleavable *trans*-substrates (Fig. 2.1b): (i) The short, U-1-containing substrate with the sequence used in the crystal structure, termed here S1-U; (ii) the U-1G mutant substrate S1-G to check whether this mutation is as deleterious as previously observed for other HDV ribozymes;[58] and (iii) the longer, 8-nucleotide substrate S8-U with U-1 and a 5' sequence selected for its predicted low tendency

Simulation Label	C75 Protonation State	Modifications of Starting Structure	Number of Replicates	Length (ns) of Each Replicate
U-1+	protonated	all ribose, with cleavage site modeled in	6	40, 20, 20, 20, 20, 100
1SJ3_U-1+[1]	protonated	all ribose, with cleavage site modeled in using 1SJ3	5	50, 50, 50, 50, 20
U-1	unprotonated	all ribose, with U-1 modeled in	5	100, 50, 50, 50, 20
Product	unprotonated	all ribose, no U-1	5	20, 20, 20, 20, 20
Xtal_U-1+[2]	protonated	all ribose	5	20, 20, 20, 50, 20
dU-1+	protonated	d(U-1)	5	100, 50, 50, 50, 20
dUGG+	protonated	d(U-1)G1G2	5	20, 20, 20, 20, 20
Xtal_dUGG+[2]	protonated	d(U-1)G1G2	5	20, 20, 20, 50, 20
dUGG	unprotonated	d(U-1)G1G2	5	20, 20, 20, 20, 20
Cis+	protonated	all ribose, no U-1	5	20, 20, 20, 20, 20
U-1G+	protonated	all ribose, no U-1	5	50, 50, 50, 20, 20

[1]U-1(N3)-C3(O2P) was restrained during equilibration and for the first 10 ns in replicate 1 and 2, and during equilibration of replicate 5.

[2]Six intermolecular distances were restrained during the entire simulation as indicated in Fig. 5a.

Table 2.1 List of MD simulations performed here

to form alternate secondary structures. For S1-U and S8-U, the time traces of fraction cleaved were fit best with double-exponential functions (Fig. 2.1c). The observed biphasic rate constants k_{obs1} and k_{obs2} for S1-U were 1.5 min^{-1} (amplitude: 0.64) and 0.12 min^{-1} (0.28), respectively. For S8-U, k_{obs1} and k_{obs2} were ~2-fold slower with 0.77 min^{-1} (0.36) and 0.075 min^{-1} (0.30), respectively, consistent with a potential of the ribozyme to form weak alternative secondary structures as predicted by the program mfold.[120] A 5-fold increase in ribozyme concentration resulted in a modest 2-fold increase in the k_{obs} values for both S1-U and S8-U, suggesting that the reported values approach the inherent cleavage rate constant (Fig. 2.2). In the case of S1-G, cleavage time traces were best fit with a single-exponential function, yielding a k_{obs} of 0.087 min^{-1} (Fig. 2.1c). A 5-fold increase in ribozyme concentration did not further increase this low value (Fig. 2.2). The ~17-fold lower activity of the S1-G compared to the S1-U substrate is consistent with previous observations that G-1 is much disfavored relative to U-1,[58] rationalizing the evolutionary conservation of U-1 in clinical isolates of the virus.[121]

For the crystallized parent ribozyme with S1-U substrate, a single k_{obs} of ~6.5 min^{-1} was reported at 37 ºC, in 25 mM Tris-HCl, pH 7.0, 10 mM $MgCl_2$ [94], slightly higher than our k_{obs1}. We therefore asked what impact the higher temperature and slightly different annealing protocol have by repeating our assay under these conditions. We again observed double-exponential cleavage, yielding a k_{obs1} of 2.4 min^{-1} (0.63) and a k_{obs2} of 0.99 min^{-1} (0.30) (Fig. 2.2). The only small (2.7-fold) difference between our k_{obs1} at 37 ºC and the previously reported k_{obs} suggests that our fluorophore-labeled variant is a relevant model system for the crystallized HDV ribozyme.

The observation of double-exponential cleavage time courses and final fractions cleaved of less than unity for the S1-U and S8-U substrates under single-turnover conditions suggests kinetic heterogeneity of the ribozyme population. Conversely, the fact that the S1-G substrate is cleaved in a single-exponential time course does not rule out such heterogeneity, especially in light of the significantly lower-than-unity cleavage extent (Fig. 2.1c). Both multiphasic cleavage and low cleavage extent are reflective of the highly frustrated and uneven folding free energy landscapes of greatly structured RNAs.[122] Although in some cases such as the *Tetrahymena* ribozyme, RNA's inherent conformational heterogeneity manifests as multiple, equally active populations (reviewed in [122]), our cleavage studies on this version of the HDV ribozyme indicate the existence of a comparatively larger proportion of differentially active conformers for this molecule.

2.3.2 The crystallized precursor ribozyme exhibits heterogeneous, catalytically active populations

Fluorescence resonance energy transfer (FRET) is a powerful tool to determine global conformations of RNA in solution.[123-126] To ask whether the observed heterogeneous cleavage kinetics of the crystallized *trans*-acting HDV ribozyme are reflected in folding heterogeneities, we used electrophoretic mobility shift assays (EMSAs). These assays are sensitive to the distinct charges and hydrodynamic radii of RNA conformers and can be monitored by FRET, as previously described.[80, 99] In addition to our three cleavable substrates S1-U, S1-G and S8-U, which are expected to largely convert to product upon ribozyme cleavage during either sample preparation or gel electrophoresis, as well as the product strand P for reference, we generated four non-cleavable substrate versions that are expected to remain in the precursor state (Fig. 2.1b): one version of each substrate with

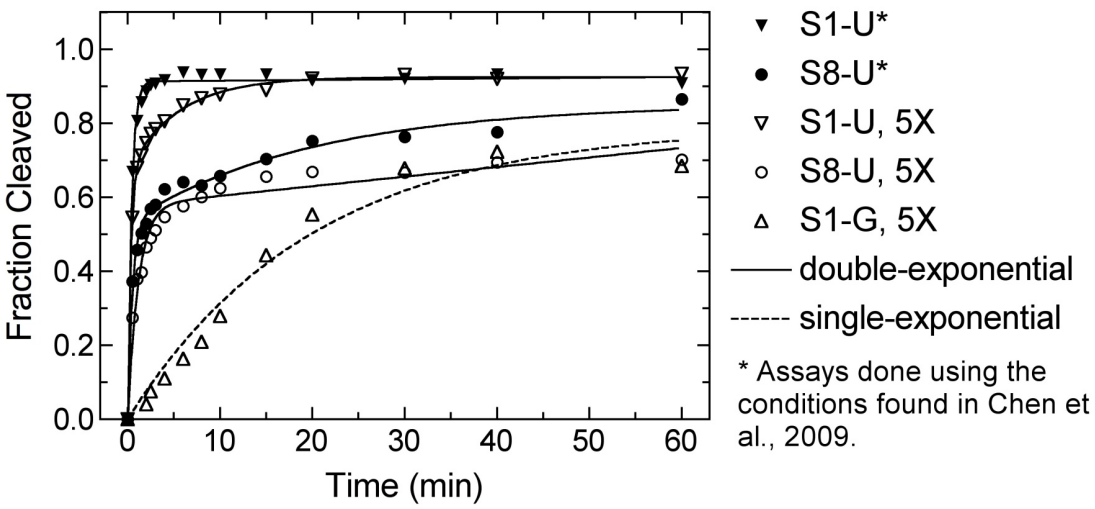

Figure 2.2 Additional cleavage assays at a higher ribozyme concentration.

The asterisk (*) denotes cleavage assays following the protocol by [94], described in the Materials and Methods section of the main text. The 5X denotes the use of ribozyme at a concentration 5-fold greater than that used for assays in the main text.

a single 2'-deoxy modification at the cleavage site (i.e., modifying the U-1 or G-1 nucleotide), termed S1-dU, S1-dG and S8-dU, and one substrate identical to the crystallized one with three 2'-deoxy modifications at the -1, 1 and 2 positions, encompassing the cleavage site (S1-dUGG). We then formed substrate-free ribozyme (Strands A and B annealed, or "Free Rz", Fig. 2.1b) and added each one of the cleavable or non-cleavable substrates, or the product strand, to compare all complexes with Strand B alone on an EMSA gel (Fig. 2.3a).

The cleavable substrate complexes all co-migrate with the single band of the product form of the ribozyme, indicating that they cleave essentially to completeness before or during electrophoresis (Fig. 2.3a). This product band is only slightly-slower migrating than the free ribozyme, suggesting that the addition of the product strand does not significantly increase the hydrodynamic radius of the free ribozyme. By contrast, each of the non-cleavable substrates shows more than one distinct band, where one band migrates similar to the product complex and another is shifted up significantly, with the relative contribution of each band indicated in Fig. 2.3a. The only exception is the crystallized S1-dUGG substrate, whose precursor ribozyme complex runs as a smear spanning the two bands observed for the other non-cleavable substrate complexes (Fig. 2.3a). These observations suggest that each precursor ribozyme complex forms more than one conformation. These conformations either rapidly (compared to the electrophoretic migration time) interconvert or happen to migrate in a continuum of conformations of only slightly different hydrodynamic radius, leading to smearing. Such conformational heterogeneity is consistent with the catalytic heterogeneity observed in our cleavage assays.

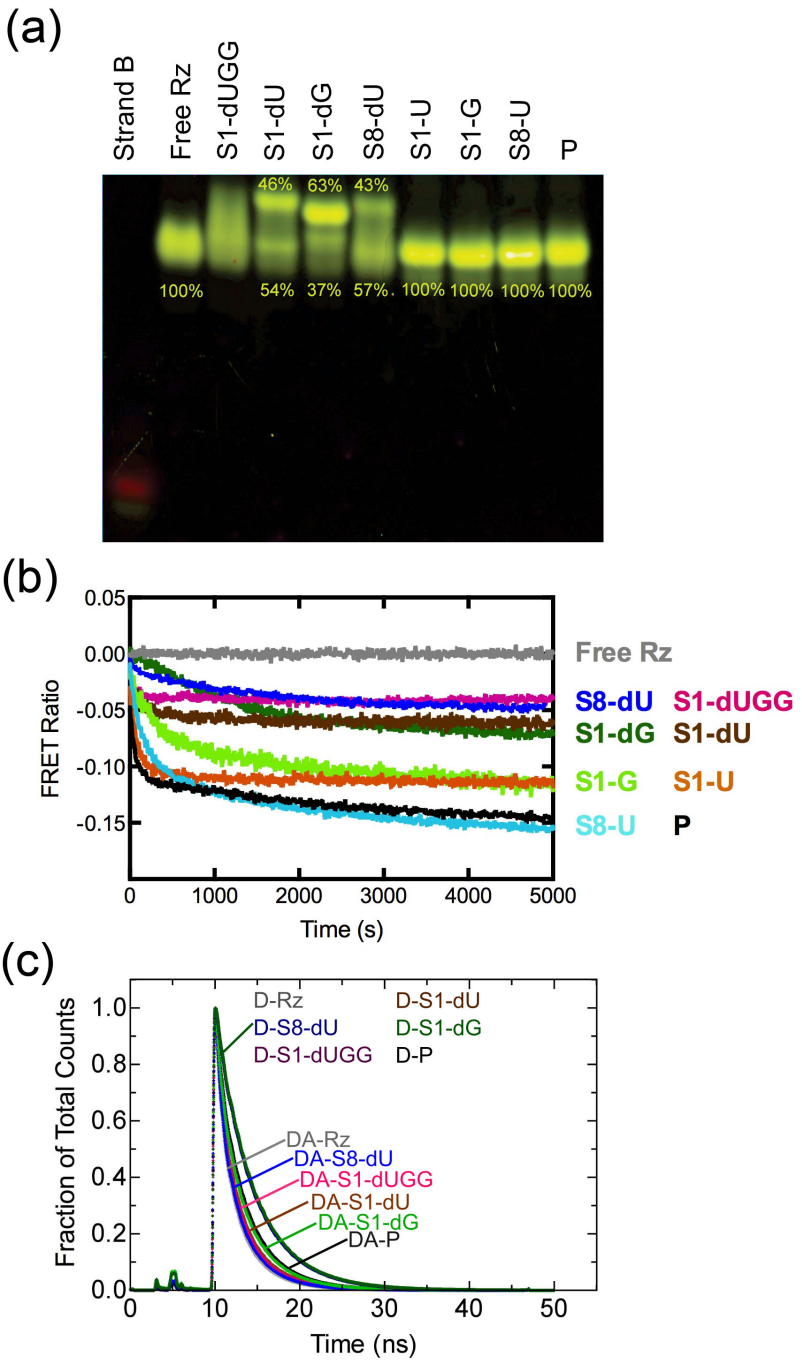

Figure 2.3 FRET probing of our 8 *trans*-acting precursor and product ribozymes.

(a) Fluorescent EMSA of the individual complexes as indicated. Where appropriate, the relative fractions in the main bands of a lane are given. Strand B and free ribozyme (Rz) lanes are shown for reference. (b) Normalized ssFRET time courses upon substrate or product addition, as indicated. A decrease in FRET ratio indicates a lengthening of the fluorophore distance along the longitudinal axis of the ribozyme. Each trace is an average of at least three trials. (c) Representative fluorescein fluorescence decays for free ribozyme (Rz), ribozyme bound to noncleavable substrates, and ribozyme bound to product (P), as indicated. Traces for ribozymes labeled with donor-only (D) and with donor plus acceptor (DA) are plotted. Each trace is one of three replicates, but composed of an average of 5 reads.

The S1-dUGG bound ribozyme appears to exist in solution as a particularly broad set of conformations (Fig. 2.3a), suggesting that crystallization selects for just one of these conformers.[10] This observation, together with the fact that the bottom band in all precursor complexes almost co-migrates with the product complex, may provide an explanation for why the recent precursor crystal structure resembles the previous product crystal structures more than their precursor counterparts: A more product-like conformation was selected for during crystal growth. Consistent with this hypothesis and the notion that the more product-like conformation is catalytically more activated, the non-cleavable S1-dG substrate leads to a smaller fractional population of the product-like band (Fig. 2.3a), as well as to lower catalytic activity (Fig. 2.1c).

2.3.3 All trans-acting ribozymes undergo a global conformational change upon cleavage

To corroborate our EMSA results, we employed a previously developed steady-state FRET (ssFRET) assay, [80, 99] in which we first form the substrate-free ribozyme, then add a cleavable or non-cleavable substrate or the product strand while monitoring in real-time the evolution of the relative FRET ratio. This ratio serves as a measure of changes in distance between the two fluorophores that mark the ends of helixes P4 and P2 (Fig. 2.1b), thus allowing us to quantify the time dependence of global conformational change. As observed for a majority of similar cases before,[99] binding of non-cleavable substrate to the free HDV ribozyme leads to a temporal decrease in FRET ratio, reflecting the substrate binding kinetics (Fig. 2.3b). When adding the product strand P instead, the decrease is more pronounced, suggesting that the P4-P2 end-to-end distance monitored by the FRET pair becomes longer in the product than in any of the precursor structures. The kinetics of the FRET ratio decrease in all cases but the S1-G ribozyme are characterized by a double-exponential decay, providing additional

evidence for heterogeneous folding (Fig. 2.3b). Supporting the notion that formation of product leads to further lengthening of the P4-P2 axis, the combination of binding and subsequent (partial) cleavage upon addition of the cleavable substrates S1-U, S-1G and S8-U leads to FRET ratio decreases close to that associated with binding of product strand P (Fig. 2.3b). The fact that, for the S-1U and S-1G substrates, the overall decrease falls between those of binding the corresponding non-cleavable strands S1-dU/S1-dG and product strand P is further consistent with the observed incomplete conversion of these substrates into product (Fig. 2.1c). Taken together, these observations strongly support the result from our EMSAs that the precursor and product forms of the most recent crystal structure display distinct global conformations. Thus, a global conformational change appears to accompany cleavage, consistent with previous findings on other *trans*-acting HDV ribozymes. [80, 97-99, 101]

To quantify this conformational change and the resulting P4-P2 end-to-end lengthening, we performed time-resolved FRET (trFRET) measurements as described [80, 99, 124-128] on ribozyme free of substrate as well as bound to each one of our non-cleavable substrates or the product strand. As expected, the fluorescein-only singly labeled ribozyme complexes display nearly identical fluorescence decay curves, whereas all fluorescein-tetramethylrhodamine doubly-labeled ribozymes exhibit distinct, faster decays due to varying FRET efficiencies (Fig. 2.3c). Analysis of these decay curves yielded the single Gaussian inter-fluorophore distance distributions summarized in Table 2.1. The free ribozyme exhibits the most compact structure with a mean inter-fluorophore distance of ~51 Å. The precursor ribozymes adopt distances of between 53 Å (S8-dU) and 59 Å (S1-dUGG, Table 2.1), depending on the non-cleavable substrate. The product conformation, with a mean fluorescein-tetramethylrhodamine distance of ~67 Å, is indeed extended, by ~8 Å in the P4-

P2 end-to-end distance, compared to the crystallized S1-dUGG complex (Table 2.1). Such an extension is consistent with our ssFRET results as well as previous studies of other *trans*-acting HDV ribozymes that showed extensions of between 2 and 15 Å upon conversion of the precursor to the product. [80, 99]

Taken together, our EMSA, ssFRET and trFRET data consistently show that the newly crystallized trans-acting HDV ribozyme in solution undergoes a significant conformational change upon cleavage. The ribozyme also displays folding heterogeneity in the precursor form that coalesces upon cleavage. The finding that, in our EMSAs, all of our substrates cleave to completion (Fig. 2.3a) further indicates that all conformational isomers are catalytically active, although to varying degrees as evident from the biphasic kinetics of our cleavage assays (Fig. 2.1c). That the crystal structure shows a relatively well-defined single global conformation further suggests that crystallization may have selected for a more product-like conformer. Consistent with this notion, the P4-P2 end-to-end distance in the crystal is ~70 Å [10], close to that of the product conformer in solution as observed here (~67 Å).

2.3.4 In MD simulations, the modeled hammerhead ribozyme cleavage site results in favorable in-line fitness, as long as C75 is protonated

While a single product-like *trans*-acting HDV ribozyme conformer was crystallized, [10] our FRET studies show that in solution multiple catalytically active global conformations coexist. This discrepancy raises the question of whether the crystal structure itself may hint at such conformational divergence. Indeed, the U-1 and scissile phosphate were not resolved in the otherwise high-resolution electron density and had to be modeled in.[10] The two simplest explanations for this lack of electron density are that a significant fraction of molecules

	Mean Distance (Å)	FWHM (Å)	χ^2
DA-Rz	51 ± 2	30 ± 3	3.1 ± 0.8
DA-S1-dU	58 ± 1	48 ± 8	3.2 ± 1.2
DA-S1-dUGG	59 ± 3	19 ± 25	3.4 ± 1.3
DA-S8-dU	53 ± 3	30 ± 7	3.3 ± 0.9
DA-P	67 ± 4	38 ± 5	2.7 ± 0.3
DA-S1-G	59 ± 2	40 ± 3	3.0 ± 0.4

Table 2.2 Distance distribution parameters obtained from time-resolved FRET data

degraded during crystal growth or that the crystal contains a conformational ensemble with some flexibility and/or disparity around the active site, a possibility favored by the authors.[10] To test this notion and provide a window into the rapid (sub-microsecond), small-scale dynamics of the ribozyme, not easily accessible by experimental techniques, [60] we performed MD simulations on the crystallized *trans*-acting precursor using our established protocols.[49, 57-60, 62, 65, 66] We used the latest parmbsc0$_{\chi OL3}$ [63, 67] variant of the Cornell et al. AMBER force field,[114] which is essential for stable RNA simulations.[129] We analyzed 49 simulations with an aggregate 1.8 μs of simulation time (Table 2.2). With such an extent of simulations available, we focused our analysis on a parameter that monitors the most critical part of the underlying crystal structure, the modeled cleavage site, and is able to relate its conformation to a functional parameter, catalytic activity. To this end, we used the previously described in-line fitness that combines the distance between the attacking 2'-O nucleophile and the phosphorous atom with the in-line attack angle formed by the 2'-O nucleophile, phosphorous and 5'-O leaving group [64, 130] into a single value reflecting the relative poise to undergo catalysis.[119] In essence, a "high" fitness value close to 1 indicates a phosphodiester bond geometry more amenable to cleavage than one described by a "low" fitness value closer to 0.

The cleavage site of the crystallized *trans*-acting HDV ribozyme was proposed to resemble the splayed cleavage site conformation of the crystal structure of the hammerhead ribozyme (PDB ID 2OEU [131]).[10] Incorporating this model into an all-ribose form of the *trans*-acting crystal structure with an N3-protonated C75H$^+$, we performed six simulations termed U-1+ (Table 2.2). We found that half of our simulations predominantly retained a high (> 0.5) in-line fitness, whereas the other half showed low (< 0.5) values with occasional excursions to higher values (Fig. 2.4a). Alternatively, we modeled the cleavage site using an

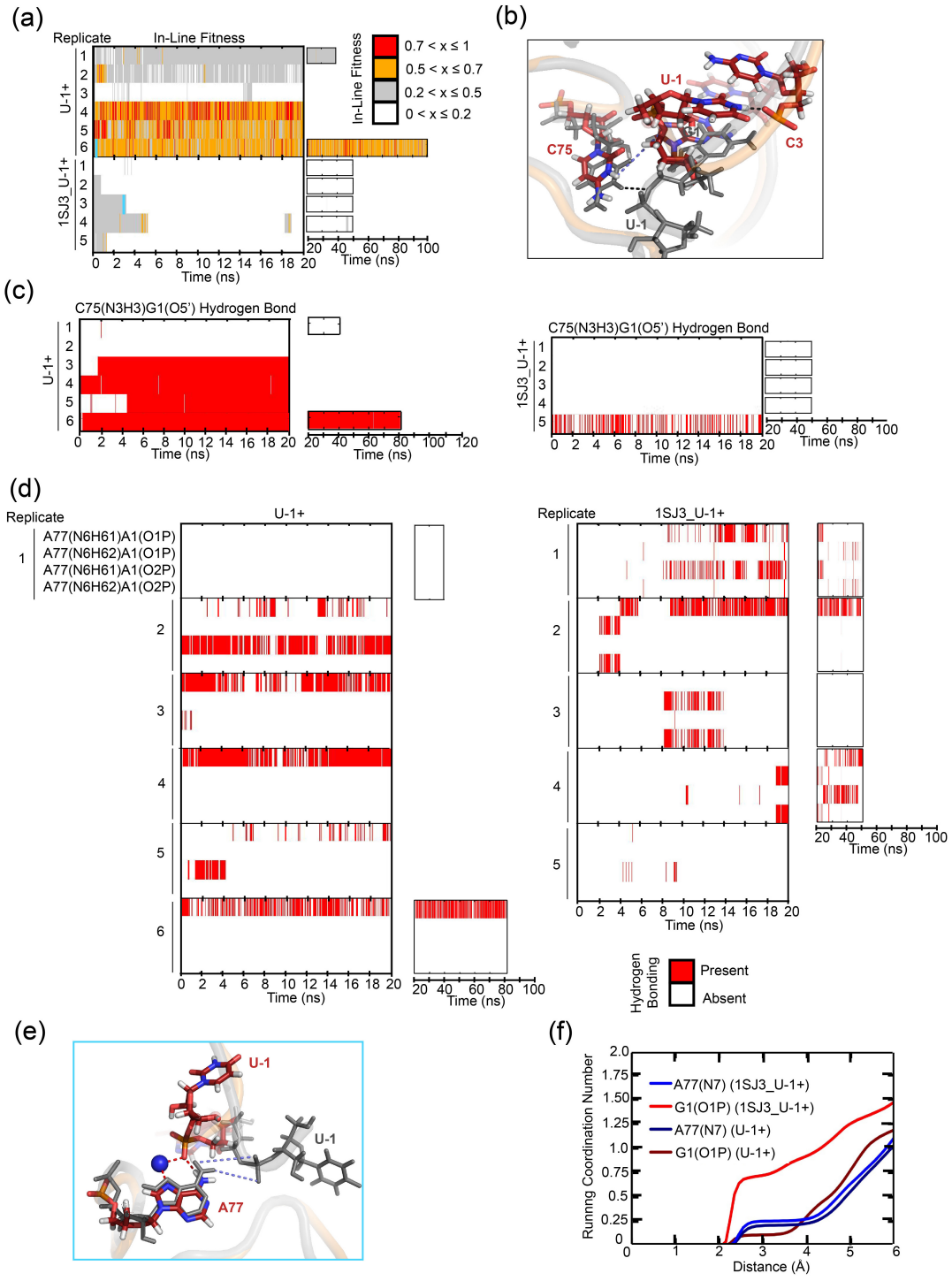

Figure 2.4 Testing the hammerhead ribozyme model for the cleavage site.

(a) In-line fitness of our U-1+ and 1SJ3_U-1+ simulations (Table 2). (b) Overlay of 1SJ3_U-1+ and U-1+ scissile phosphates. The 1SJ3_U-1+ structure is shown in color, while the U-1+ structure is shown in transparent cartoon with relevant residues shown as thickened lines. The U-1(N3)C3(O2P) heavy atom distance restrained in some simulations (Table 2) is denoted with a black dashed line, while the

destabilized C75(N3H3)G1(O5') distance is denoted with a light blue dashed line. For reference, the stable C75(N3H3)G1(O5') distance is denoted in the gray U-1+ structure with a black dashed line. In all stick representations of 1SJ3_U-1+ structures, carbons are colored dark red, nitrogens are blue, oxygens are red, hydrogens are white and sodium ions are blue. (c) Comparing $C75H^{+}$(N3H3)…G1(O5') hydrogen bonding occupancies of U-1+ and 1SJ3_U-1+ simulations. Traces were computed using the Simulaid program and in-house scripts, with cutoffs of ≥ 120° and ≤ 3.5 Å to indicate presence of a hydrogen bond. (d) Monitoring of alternate A77(N6)…G1(O1/O2P) hydrogen bonding interactions for U-1+ and 1SJ3_U-1+ simulations. (e) The alternative ion cage observed in 1SJ3_U-1+ simulations. Overlays of 1SJ3_U-1+ and U-1 are shown in the left-hand portion of the figure. As in panel (b), the 1SJ3_U-1+ structure is shown in color, while the U-1+ structure is in transparent gray cartoon for reference. Relevant residues are shown in sticks and thickened lines, respectively. Red dashed lines indicate hydrogen bonds/ionic interactions that form and are stable in 1SJ3_U-1+ trajectories, while blue dashed lines indicate hydrogen bonds in U-1+ simulations that are destabilized due to the new scissile phosphate conformation of the 1SJ3_U-1+ structure. The cyan lines in panel (a) denotes from which replicate and at what times the snapshots were taken. (f) Coordination number of sodium ions at a given distance from each of the functional groups specified. The plotted curves are weighted averages from all available simulations of a given type.

alternative "U-turn" motif topology that we had identified in the previous *cis*-acting precursor crystal structure (PDB ID 1SJ3; Fig. 2.4b).[58] The resulting five 1SJ3_U-1+ simulations (Table 2.2) exhibited poor in-line fitness, with > 99% of the total simulation time resulting in values of < 0.5 (Fig. 2.4a). In addition, when restraining the U-1(N3)…C3(O2P) distance for the first 10 ns in two simulations (numbers 1 and 2) to that of a hydrogen bond characteristic of the U-turn motif,[58] the in-line fitness never increased beyond 0.5 (Fig. 2.4a). In almost all 1SJ3_U-1+ simulations, $C75H^+$ resided far from G1(O5'), disrupting a hydrogen bond that appears critical for retaining an optimal scissile phosphate geometry (Fig. 2.4c). Often, the scissile phosphate's two nonbridging oxygens accepted a fluctuating, bifurcated hydrogen bond from the two exocyclic amino protons of A77(N6) (Fig. 2.4d). In turn, this interaction created a cation-binding pocket different from that observed in the hammerhead model of the cleavage site, with on average one ion bound to G1(O1P), and a partially (~25%) occupied ion bound to A77(N7) and $C75H^+$(O2) (Figs. 3e,f). Such Na^+ ion binding is often predictive of Mg^{2+} binding while avoiding the complications from insufficient sampling and lack of polarization associated with simulating a tightly binding Mg^{2+}.[60, 65] We conclude that, while the U-turn motif is consistent with the earlier *cis*-acting precursor crystal structure, it does not lend itself to high in-line fitness in the context of the product-like *trans*-acting ribozyme. In addition, combining the simulation data, while still far from full sampling, with our experimental findings suggests that the catalytic core of the HDV ribozyme occupies a rugged folding free energy landscape with multiple low-energy wells.[122, 132, 133]

In our simulations so far we used an N3-protonated $C75H^+$, which we found leads to a $C75H^+$(N3)…G1(5'O) hydrogen bond critical for maintaining a favorable in-line fitness.

Protonation is justified since the crystal structure was obtained at a pH of 5.0,[10] however, since the pK_a of C75 of the HDV ribozyme was measured as ~6.4,[20, 90] at the more physiological pH of 7.5 used in our solution probing a protonated $C75H^+$ will represent only a minor (< 10%) species. To ask what impact an unprotonated C75 has on in-line fitness, we performed five U-1 simulations without protonating C75 in the starting structure (Table 2.2). Indeed, very few excursions to high (> 0.5) in-line fitness occurred (Fig. 2.5a), and without the $C75H^+$(N3)…G1(5'O) hydrogen bond this distance increased to typically 5-10 Å (Fig. 2.5b). Such distances are not typically seen in the U-1+ simulations (Fig. 2.5b). On average, a single Na^+ ion occupied a position close to the unprotonated C75(N3), as well as C75(O2), essentially adopting the position of the H(C75N3) proton in the U-1+ simulations, occupying a pocket of the ribozyme rendered more negative in electrostatic surface potential due to deprotonation of $C75H^+$ (Fig. 2.5c). The fractional residency of this ion was 101 ± 29% for C75(N3) (indicating that occasionally more than one Na^+ binds in the pocket) and 78 ± 21% for C75(O2) (Fig. 2.5d). These observations are consistent with previous experimental studies that suggested anti-cooperativity between protonation of C75 and relatively weak binding of an active-site magnesium ion.[134, 135] Taken together, our data strongly suggest that protonation of C75 to form a $C75H^+$(N3)…G1(5'O) hydrogen bond is critical for the *trans*-acting HDV ribozyme precursor to adopt a favorable in-line fitness value, positioning its scissile phosphate for catalysis.

Further supporting the notion that the $C75H^+$(N3)…G1(5'O) hydrogen bond is a critical anchoring point for a product-like structure, we found that five simulations of a product form of the *trans*-acting crystal structure (with U-1 removed, G1(5'O) protonated and C75 left unprotonated, termed "Product" in Table 2.2) showed only a subtle difference

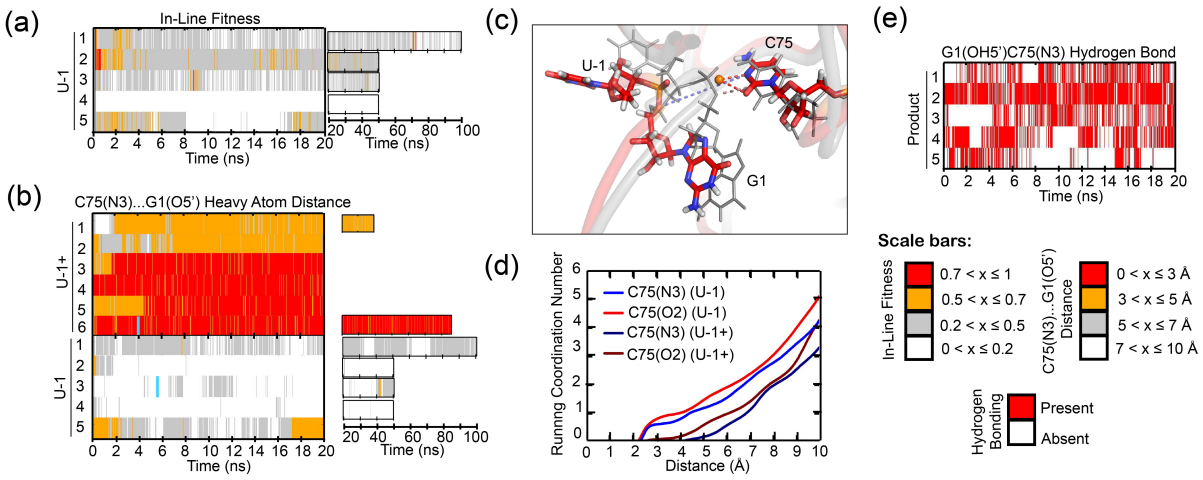

Figure 2.5 A C75(N3)…G1(O5') hydrogen bond supports a product-like active site.

(a) In-line fitness of U-1 trajectories. (b) Comparison of C75(N3)…G1(O5') heavy atom distances in U-1+ and U-1 trajectories. (c) Structural overlay of U-1 and U-1+ active sites. The U-1 structure is in color, while the U-1+ structure is in transparent gray. Relevant residues are shown as sticks and thickened lines, respectively. Red dashed lines indicate stable coordination of the orange sodium ion. Light blue dashed lines indicate the C75H$^+$(N3H3+)…G1(O5') hydrogen bond lost in the U-1 structures. Colored atoms in stick representations are colored as indicated in Fig. 2.4. The cyan lines in panel (b) denote from which replicate and at what times the snapshots were taken. (d) Running coordination numbers of U-1+ and U-1 ion-binding residues. The plotted curves indicate the number of sodium ions coordinated by a given residue at a range of distances, and are weighted averages from all available simulations of a given type. (e) Time traces for the reverse G1(OH5')…C75(N3) hydrogen bond in product trajectories.

within the active site relative to our U-1+ simulations. In this case of a less crowded active site without U-1, an even more stable, flipped G1(OH5')…C75(N3) hydrogen bond forms. This hydrogen bond is present in all five Product simulations, with occupancies ranging from 30%-80% (Fig. 2.5e).

2.3.5 Intermolecular crystal contacts and 2'-deoxy modifications decrease in-line fitness and stacking of U-1 on U23

Upon inspection of the crystal structure, we noticed an intermolecular contact between the closing GAAA tetraloop of P4 of one molecule in the unit cell and the active site in another (Fig. 2.6a).[10] While the several hydrogen bonds from this contact may have fortuitously aided in stabilizing crystals, we sought to understand whether they may have an impact on the cleavage site. To this end, we prepared a starting structure of a ribozyme together with the adjacent P4 stem-loop from a second ribozyme in the unit cell, using six distance restraints to keep the latter P4 stem-loop from changing its relative positioning (Fig. 2.6a). The resulting five simulations termed Xtal_U-1+ (Table 2.2) exhibited considerably poorer fitness than our U-1+ simulations (Fig. 2.6b). A closer look revealed that the loop's formation of hydrogen bonds with U23 and C22 resulted in increased distances of U-1 from U23 and C22 (Fig. 2.6a). These increased distances hinder formation of a U-1(O2')…U23(O1P) hydrogen bond, which in turn destabilizes a specific U-1 conformation wherein it stacks on U23, a stacking interaction that occurred frequently in our U-1+ simulations (Fig. 2.6a). The crystal packing contact thus may contribute to the lack of U-1 electron density [10] since it confers a less well-defined positioning of U-1.

Another unusual feature of the crystal structure is the presence of three 2'-deoxy modifications in U-1, G1 and G2 flanking the cleavage site (Fig. 2.1a). While a single such

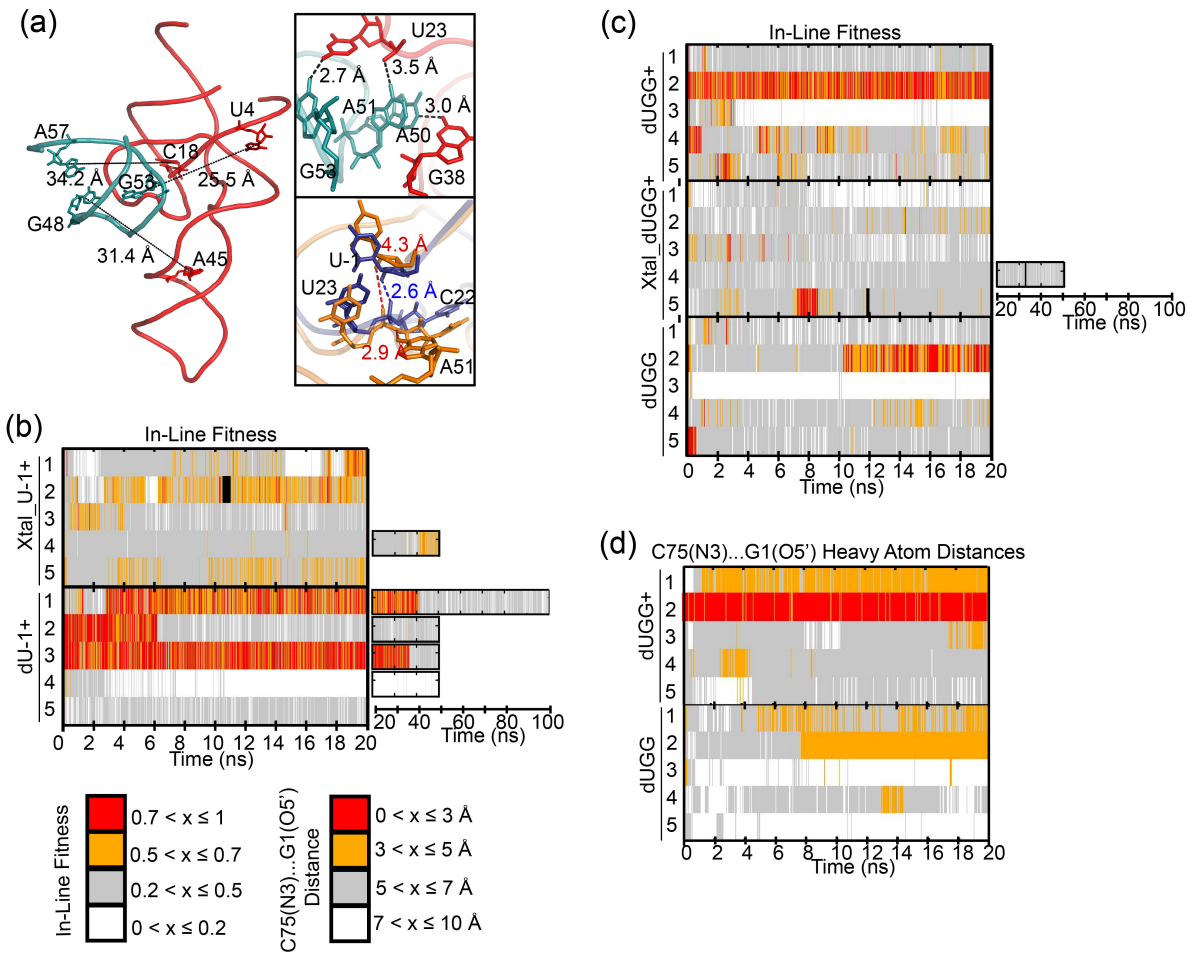

Figure 2.6 Effects of crystallization parameters on active site geometry.

(a) Whole-molecule view of our crystal structure model, with length of long-range restraints and participating residues indicated. Top inset: Short-range restraints and participating residues in our Xtal simulations. Bottom inset: A 100-ps-averaged snapshot from a Xtal_U-1+ simulation (gold) is overlaid with a 100-ps-averaged snapshot from a U-1+ simulation (navy). The adjacent P4 loop interacts with U23 and C22 resulting in increased distances of these residues from U-1. As a result, interactions important for stabilizing favorable fitness cannot easily form. (b) Comparison of in-line fitness values in Xtal_dU-1+ and dU-1+ simulations. (c) Comparison of in-line fitness values in dUGG+, Xtal_dUGG+ and dUGG simulations. (d) Comparison of C75(N3)…G1(O5') heavy atom distances in dUGG+ and dUGG simulations.

modification of the N-1 nucleotide is often used to render a ribozyme substrate non-cleavable by removing the 2'-OH nucleophile that serves as a nucleophile during cleavage, multiple 2'-deoxy modifications are less common. To ask what effect these extended modifications may have on the local structural dynamics of the cleavage site, we first performed five dU-1+ simulations with a single 2'-deoxy modification at U-1 (Table 2.2) and compared them to our all-ribose U-1+ simulations. Because there is no U-1(2'O) in the dU-1+ simulations, an in-line fitness "equivalent" was calculated by using U-1(2'H) in place of U-1(2'O). In two out of our dU-1+ simulations, in-line fitness started off at favorable values, but later sampled unfavorable values as well (Fig. 2.6b). In two other dU-1+ simulations, fitness started with and remained at unfavorable values (Fig. 2.6b). This behavior resembles that of the all-ribose U-1+ simulations (Fig. 2.4a).

We then performed five dUGG+ simulations that carried all three 2'-deoxy modifications of the crystal structure (Table 2.2). In this case, less frequent adoption of favorable in-line fitness was observed and, if adopted, high in-line fitness values were often rapidly lost (Fig. 2.6c). These observations suggest that, as with the intermolecular crystal contacts, 2'-deoxy modifications near the cleavage site may reduce catalytic poise. Specifically the absence of a G1(2'OH)…U-1(5'O) interaction seems to affect this difference, whereas hydrogen bonding by G2(2'OH) does not appear to have distinct effects. Interestingly, the dynamic exchange between high and low fitness values appears to be accelerated for the dUGG+ simulations (Fig. 2.6c), consistent with the large number of closely related (i.e., almost co-migrating) conformations we observed in our EMSAs for S1-dUGG ribozymes (Fig. 2.3a). Such an enhanced flexibility may have further contributed to the lack of experimental electron density for U-1 and the cleavage site in the 2'-deoxy-

modified crystal structure.[10] When we performed simulations with the crystal contacts and all three 2'-deoxy modifications at the active site ("Xtal_dUGG+", Table 2.2), we observed a greater dominance of in-line fitness values < 0.4 in each replicate compared to Xtal_U-1+ or dUGG+ (Fig. 2.6c). Finally, when performing five dUGG simulations with an unprotonated C75 we observed behavior similar to that of the C75 unprotonated dU simulations with particularly low in-line fitness. We also observed generally unfavorable C75(N3)…G1(5'O) heavy atom distances in both the dUGG+ and dUGG simulations (Fig. 2.6d), indicating that the incorporation of all three 2'-deoxy modifications contributes to further active site destabilization beyond simply unfavorable in-line fitness.

2.3.6 A cis-acting precursor model exhibits greater fluctuations in fitness whereas multimodal stacking of a U-1G mutation with U23 results in decreased fitness

A wealth of experimental evidence has indicated that *cis*-acting HDV ribozymes generally cleave ~10-100 times faster than their *trans*-acting counterparts.[57, 80, 94, 136, 137] It has been hypothesized that the loss in catalytic activity of *trans*-acting ribozymes is due to a less poised or stable cleavage site architecture as a result of the distal discontinuity of the RNA backbone.[136, 137] To test this hypothesis in the context of the *trans*-acting crystal structure, we modeled into it two J1/2 joiner nucleotides (A8 and U9) from the previous *cis*-acting product crystal structure [11] to arrive at a corresponding, C75(N3)-protonated (or C75H$^+$) *cis*-acting precursor model (Cis+, see Table 2 and Fig. 2.8a). While generally similar in probability distribution of in-line fitness values (Fig. 2.7), we found this model to exhibit greater dynamics in fitness than our U-1+ simulations, as all five of our Cis+ simulations converted multiple times between high and low in-line fitness values (Fig. 2.8b). This finding contrasts with our U-1+ simulations that show either mostly high or mostly poor in-line fitness over

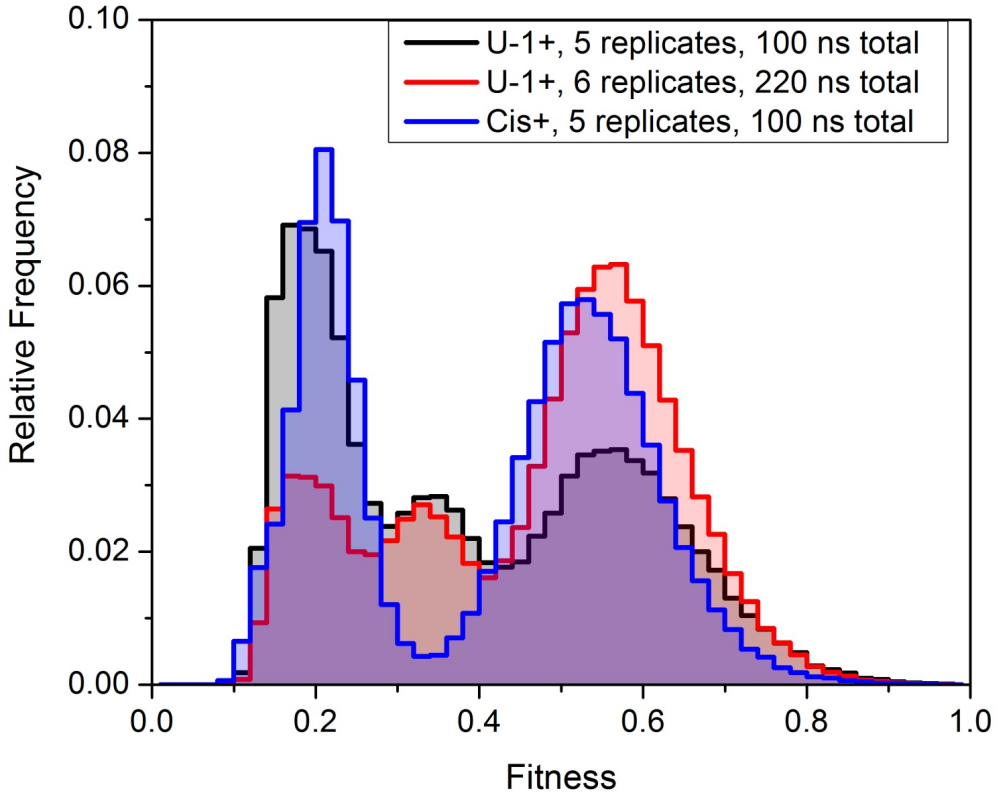

Figure 2.7 Probability density histograms of in-line fitness values for U-1+ and Cis+ simulations.

For U-1+, histograms were made for the first 5 replicates (firt 20 ns from each) as well as for all 6 trajectories together (including trajectory 6 of 120 ns length). For Cis+, there are 5 replicates of 20 ns each and all of these data were used to generate the histogram

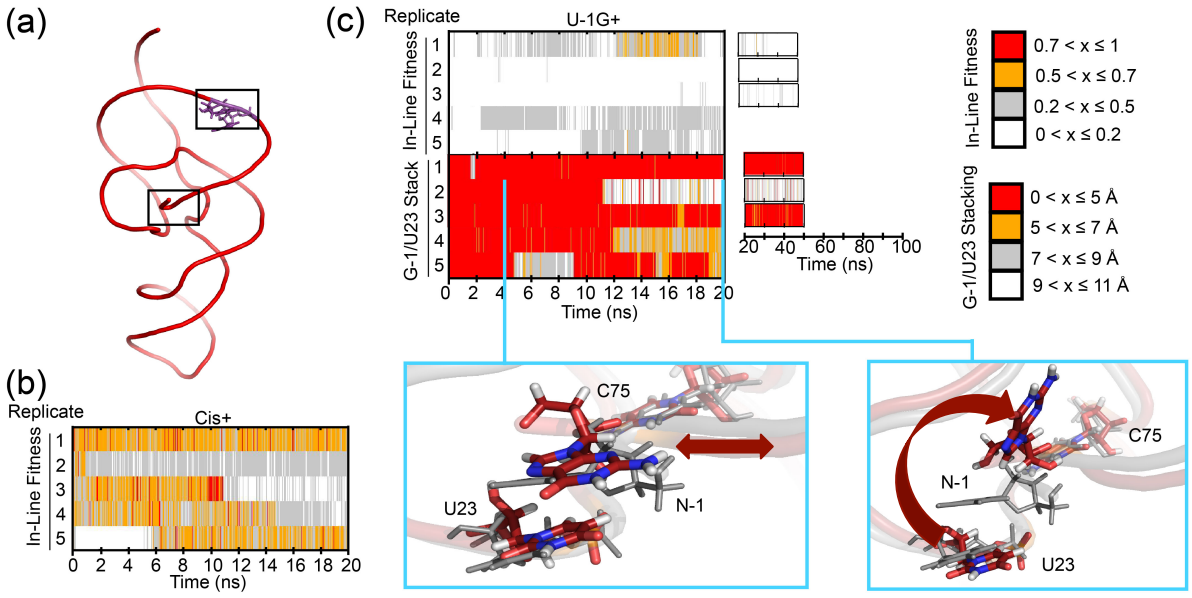

Figure 2.8 Simulating a *cis*-acting precursor model and a U-1G mutant.

(a) Global structure of our Cis+ simulations. Modeled-in residues are denoted in purple. (b) In-line fitness of our Cis+ simulations. (c) In-line fitness (top), stacking behavior (middle) and representative stacking snapshots (bottom) of our U-1G+ simulations. Stacking distances were calculated as the distance between the center of mass of U23 heavy ring atoms, and the center of mass of G-1 atoms C4 and C5. Both snapshots overlay a snapshot from the U-1+ simulation to contrast the stacking behavior. The red double-headed arrow in the left snapshot indicates the ability of G-1 to slide. The curved arrow in the right snapshot depicts the "flipping up" movement of G-1 as it unstacks. The U-1G+ snapshots are in color, while the U-1+ snapshot is in transparent gray. Relevant residues are shown in sticks and thickened lines. Atom colors are as described in Fig. 2.4. The cyan lines in the top graph denotes from which replicate and at what times the snapshots were taken.

entire 20 ns simulations (Fig. 2.4a). These more rapid fluctuations of Cis+ in-line fitness compared to that of the U-1+ simulations are consistent with a lower barrier to transitioning between inactive and active local conformations once C75 is protonated and may rationalize the generally higher catalytic activity of *cis*-acting HDV ribozymes.

Conversely, previous studies have shown that a G-1 upstream of the scissile phosphate results in a decrease in catalytic rate relative to the U-1 found conserved in clinical isolates of HDV.[58, 121] To ask whether the *trans*-acting ribozyme structure may offer an explanation for this effect as well, we analyzed five simulations with a G-1 in place of the U-1, in the presence of an N3-protonated C75H$^+$ (U-1G+, Table 2.2). We found 98.5% of the total frames of these trajectories to adopt in-line fitness < 0.5 (Fig. 2.8c). This trend was mainly determined by unfavorable in-line attack angles, which in turn appeared to be due variable stacking of G-1 with U23; in addition to a sliding motion of U23 on the larger G-1, significant unstacking of the two bases was observed in 3 out of 5 trajectories (Fig. 2.8c). These observations rationalize the decrease in catalytic activity upon U-1G mutation of *trans*-acting HDV ribozymes [58] and further support the functional relevance of our MD simulations.

2.4 Discussion

Disparate crystal structures of the cleavage precursor of the HDV ribozyme have raised questions about the validity of the particular modifications needed for each specific construct to crystallize. In the case of the *cis*-acting precursor, either the catalytic C75 was mutated to U or divalent metal ions were omitted,[12] both of which may have led to trapping of an off-pathway conformation. In the case of a more recent *trans*-acting precursor structure,[10] three 2'-deoxy modifications surrounding the cleavage site and the establishment of intermolecular

contacts during crystal growth may have led to the poorly defined cleavage site, which lacks electron density for the U-1 nucleotide and scissile phosphate. In addition, this structure was grown at pH 5.0, protonating C75 into an ionized species that would be rare at physiological pH given C75's pK_a of ~6.4;[20, 90] this protonation, in turn, likely trapped a rare, more product-like conformation that would be slow to react, as cleavage is disfavored at low pH due to the scarcity of the required general base.[20, 21]

Our data support the unifying model summarized in Fig. 2.9 that resolves these questions. We find that the crystallized *trans*-acting ribozyme shows multiple distinct conformations in solution at physiological pH as evidenced by biphasic cleavage kinetics (Fig. 2.1c) and a smear upon native gel electrophoresis (EMSA, Fig. 2.3a). The observed biphasic cleavage kinetics and the less than 100% cleavage indicate diverse populations of varying activity for each type of ribozyme, and are reflective of the highly frustrated RNA free energy folding landscape. The fact that the EMSA smear is converted into a single product band upon cleavage in the gel (Fig. 2.3a) suggests that all of these conformations eventually cleave and hence are catalytically active even if with varying rate constants. Upon cleavage in solution at physiological pH, the crystallized *trans*-acting HDV ribozyme undergoes a significant conformational change, as observed by ssFRET and trFRET in the form of a significant lengthening along the longitudinal axis (Figs. 2b,c). This finding is similar to observations for all previous *trans*- and *cis*-acting ribozymes,[80, 95, 99, 103] suggesting that it represents an immutable, common feature of the HDV ribozyme in general. Taken together, these observations suggest that occasional, stochastic protonation of C75 at

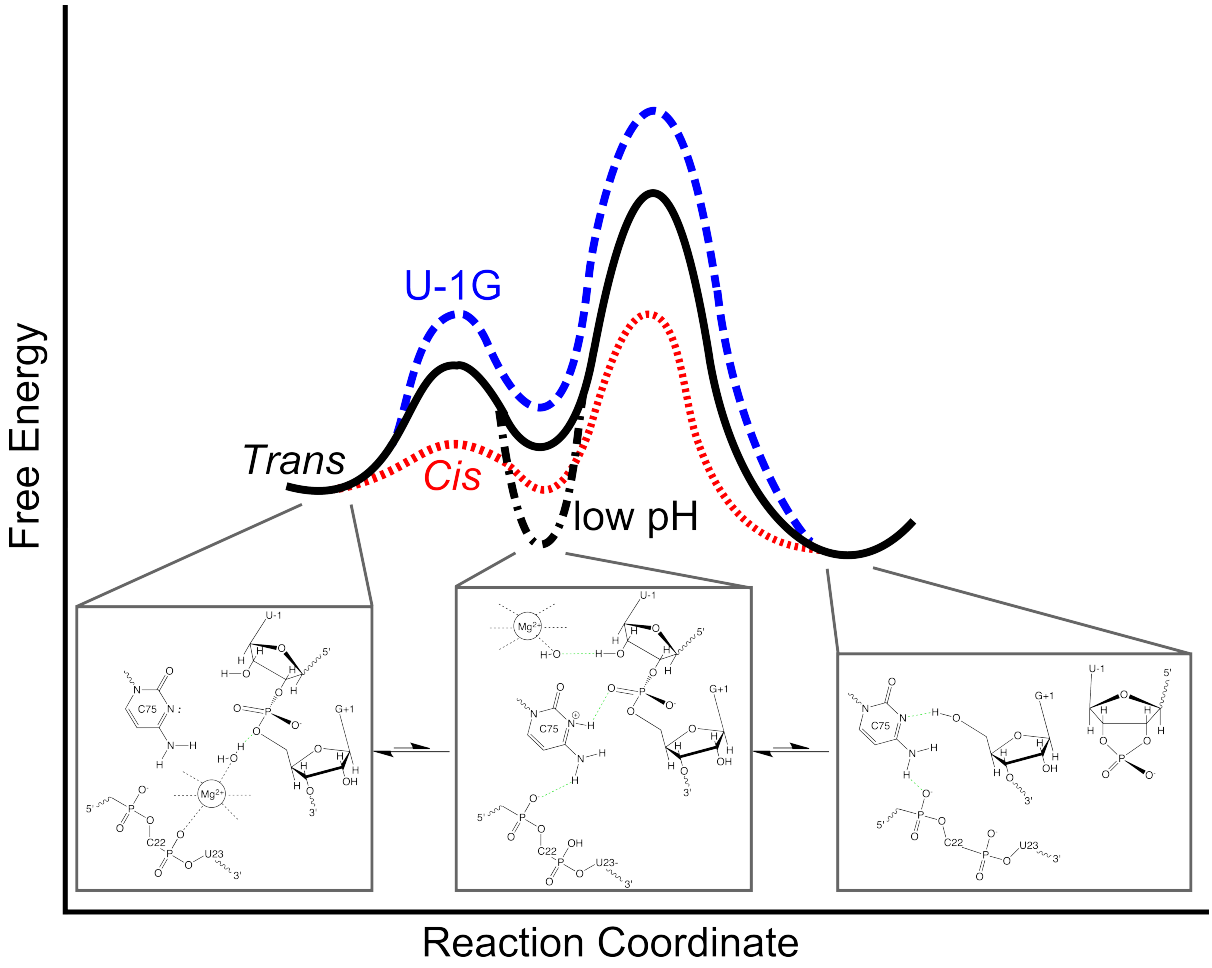

Figure 2.9 A parsimonious, unifying model of the HDV ribozyme folding free energy landscape along the catalytic reaction coordinate, indicating the divergence of the *trans*-acting, *cis*-acting and U-1G mutant ribozymes.

Below each of the main on-pathway intermediates, structural models as derived from the available crystal structures are depicted. The *cis*-acting precursor crystal structure[12] resembles most closely an early intermediate dominant at physiological pH (left). This intermediate still needs to undergo a conformational change that positions the C75H+ for general acid catalysis and was trapped at low pH in the *trans*-acting precursor structure[10] (middle). If a magnesium hydroxide is bound to the active site simultaneously, cleavage into the product leads to loss of the 5'-precursor sequence and magnesium ion (right).

physiologic pH leads to the transient adoption of an ionized state thatmoves $C75H^+$ closer to the cleavage site so that its N3 donates a hydrogen bond to the G1(O5') leaving group (while its exocyclic amine donates a hydrogen bond to one of C22's nonbridging oxygens; it remains unclear whether the adjacent G25 rotates from *anti* to *syn* before, during or after this event). This conformational adjustment sets the stage for the adoption of favorable in-line fitness. If concomitantly a Mg^{2+} ion bound to the adjacent electronegative cleft carries a deprotonated water molecule in its inner-coordination sphere, cleavage can proceed with $C75H^+$ and $Mg(OH)^+$ serving as the general acid and base, respectively (Fig. 2.9). Since the opposing protonation states of $C75H^+$ and $Mg(OH)^+$ are rare and transient, cleavage is slow compared to protein-based ribonucleases.[138-140] According to this model, the conformational change accompanying catalysis[12] is induced by the – at physiological pH – rare protonation of C75. The resulting, transient adoption of an active site conformation similar to that modeled into the *trans*-acting ribozyme structure was enriched by crystallization at low pH. Conversely, our solution probing at physiological pH detects a dominant conformation more similar to the *cis*-acting precursor crystal structure, prior to the conformational change, which represents a legitimate alternate conformation and on-pathway catalytic intermediate. (We note that there must also be other, more globally distinct conformations present in solution to explain the broad smear and multiple bands of the various *trans*-acting ribozyme observed during EMSA, as well as the biphasic cleavage kinetics, which points to deeply trapped conformations[122, 132, 141]).

Energetically, we have evidence that the *cis*-acting HDV ribozyme has a lower barrier towards catalysis than the *trans*-acting version (Fig. 2.9). First, all *cis*-acting ribozymes are 1-2 orders of magnitude faster in cleavage. [40] Second, in our MD simulations a *cis*-acting

precursor model appears to exhibit more rapid fluctuations between conformations of high and low fitness (Fig. 2.8b), suggesting that the barrier between the pre- and post-conformational change states and, as a consequence, also the overall energetic barrier of catalysis is lowered (Fig. 2.9). However, at physiological pH the pre-conformational change state is still thermodynamically favored, leading to the observed pre-conformational *cis*-acting crystal structure. Conversely, a U-1G mutation leads to generally low in-line fitness and thus a higher catalytic barrier by disfavoring the conformational change needed to adopt a suitable active site conformation (Fig. 2.9), at least partially due to trapping of G-1 in a shifted stacking interaction with U23 (Fig. 2.8c). Our data and resulting parsimonious model thus offer rationalizations for a number of independent experimental observations. In addition, they further advance a widely held view of RNA folding free energy landscapes as particularly rugged (compared to, for example, those of proteins).[122, 142]

Our MD simulations provide a glimpse of how subtle some of the conformational rearrangements may be that control a catalytically critical parameter such as the in-line fitness, which needs to reach a value of > 0.5 for RNA cleavage to become significant.[119] Efficient ribozymes are likely to stabilize a favorable in-line fitness, as well as position general acid and base catalysts. The *cis*-acting HDV ribozyme is one of the most efficient self-cleaving ribozymes known, yet is still many orders of magnitude slower than RNase A.[140] One reason may be the electrostatic competition between C75 protonation and Mg^{2+} binding to the active site.[134] Furthermore, while $C75H^+$ is a rare protonated species needed for catalysis, the hydrated Mg^{2+} ion with its high pK_a of ~12 needs to adopt a rare deprotonated state to act as the general base, contrasting events that will have a low probability of coincidence. Protonation of C75 in turn leads to a C75H+(N3)…G1(O5')

hydrogen bond that aids in adopting a favorable in-line fitness. It appears that it is the required temporal coincidence of these partially competing, dynamic low-probability protonation and conformational change events that slows RNA self-cleavage to at best some 10's to 100's min^{-1}.[58, 59] Future studies will likely reveal more such examples for the complex relationship between folding dynamics and function of RNA. Linking results from experimental structural and functional probing with those from computational MD simulations, as accomplished here through the use of in-line fitness as a parameter transcending the gap in accessible timescales, is likely to pave the way for more advances in this direction.

2.5 Acknowledgements

The authors thank Tristan Tabouillot, Kaushik Gunrunathan and May Daher for protocols, training and assistance with the time-resolved FRET experiments. We also thank Nad'a Špačková and Kamila Réblová for protocols and assistance with MD simulations. The authors acknowledge funding from NIH grant R01 GM062357 TO N.G.W., from grant P301/11/P558 from the Grant Agency of the Czech Republic to J.S., and a Canadian NSERC Postgraduate Scholarship to W.W.T.

CHAPTER 3

ROBUSTNESS OF THE ACTIVE SITE ARCHITECTURE OF THE HDV RIBOZYME[2]

3.1 Introduction

The hepatitis delta virus (HDV) ribozyme is a member of the family of small catalytic RNAs. In addition to the HDV ribozyme, this family comprises several well-characterized members such as the hairpin and hammerhead ribozymes.[7, 9, 102, 143] Despite adopting distinct architectures, all small ribozymes catalyze the same reaction: site-specific cleavage of their phosphodiester backbone. This reaction proceeds through S_N2-type in-line attack, in which a nucleophilic 2' hydroxyl directly upstream of the scissile phosphate attacks the electrophilic phosphorus. Ribozyme backbone cleavage results in 3' and 5’ products with 5' hydroxyl and 2'-3' cyclic phosphate groups, respectively.[7, 9, 143] The HDV ribozyme is of special importance because it is the only known member of this class with versions found in both a human pathogen and the human genome.[7, 9, 73, 75, 102, 143] Its reaction is believed to proceed through general acid-base catalysis, with cytosine 75 (C75) long identified as a catalytic nucleotide.[20, 21]

---

[2]Adapted from **Kamali N. Sripathi**, Pavel Banáš, Kamila Réblová, Jiří Šponer, and Nils G. Walter, ‘Robustness of the Active Site Architecture of the HDV Ribozyme’ (in preparation for *PCCP*). Ms. Wendy Tay carried out four of the six Wildtype_Na simulations; P. B. and Mr. Pavel Polakovič contributed generated starting structures for the Wildtype_Na and Wildtype_Mg simulations, and carried out one of the Wildtype_Na and two of the Wildtype_Mg simulations. All remaining simulations and analyses were performed by K. N. S.

The HDV ribozyme in its minimal, 85-nt form adopts a complex tertiary structure, consisting of five helical regions numbered P1 – P4 and P1.1.[35] These helices are arranged in a double-nested pseudoknot, resulting in the formation of two coaxial stacks[11, 35] (Fig. 3.1a,b) and a deep, narrow cleft in which the active site residues interact. The active site is bounded above by a pair of type I and type II A-minor interactions (involving the two green A residues in Fig. 3.1a).[49, 58] Its lower bound is two-fold, consisting first of the 2-base-pair helix P1.1, and below it a unique set of quadruple interactions involving a protonated C41 residue.[11, 12, 42] These latter two features have been shown to be critical hallmarks of the native state: Both the presence of the P1.1 helix and correct formation of the C41 quadruple are crucial for ribozyme activity[11, 42, 85] and architecture.[49]

Several crystal structures have been solved to first ascertain the structure of the active site, and later to elucidate the role of the catalytic C75.[11, 12, 144] The initial structure, a reaction product, revealed the mini-helix P1.1 and the wealth of functional groups within close proximity in the HDV ribozyme active site. Later, several reaction precursor structures, also from Doudna and coworkers,[12] showed a *syn* orientation to the G25 base, in contrast to its *anti* position in the product structure. Additionally, the position of C75 was resolved between U-1(O2') and G1(O5'), indicating that it could act as either a general base or acid, respectively. Global comparison of the product and precursor structures showed a collapse of the P1 and P3 helices towards the active site in the product structure.[11, 12] All of these observations indicated significant conformation changes upon cleavage at both the whole-molecule and active site levels. A more recent HDV ribozyme crystal structure is of another precursor, which exhibits more product-like global and local conformation.[10] G25 is again in the *anti* conformation, and C75 has the potential to hydrogen bond with G1(O5'). This

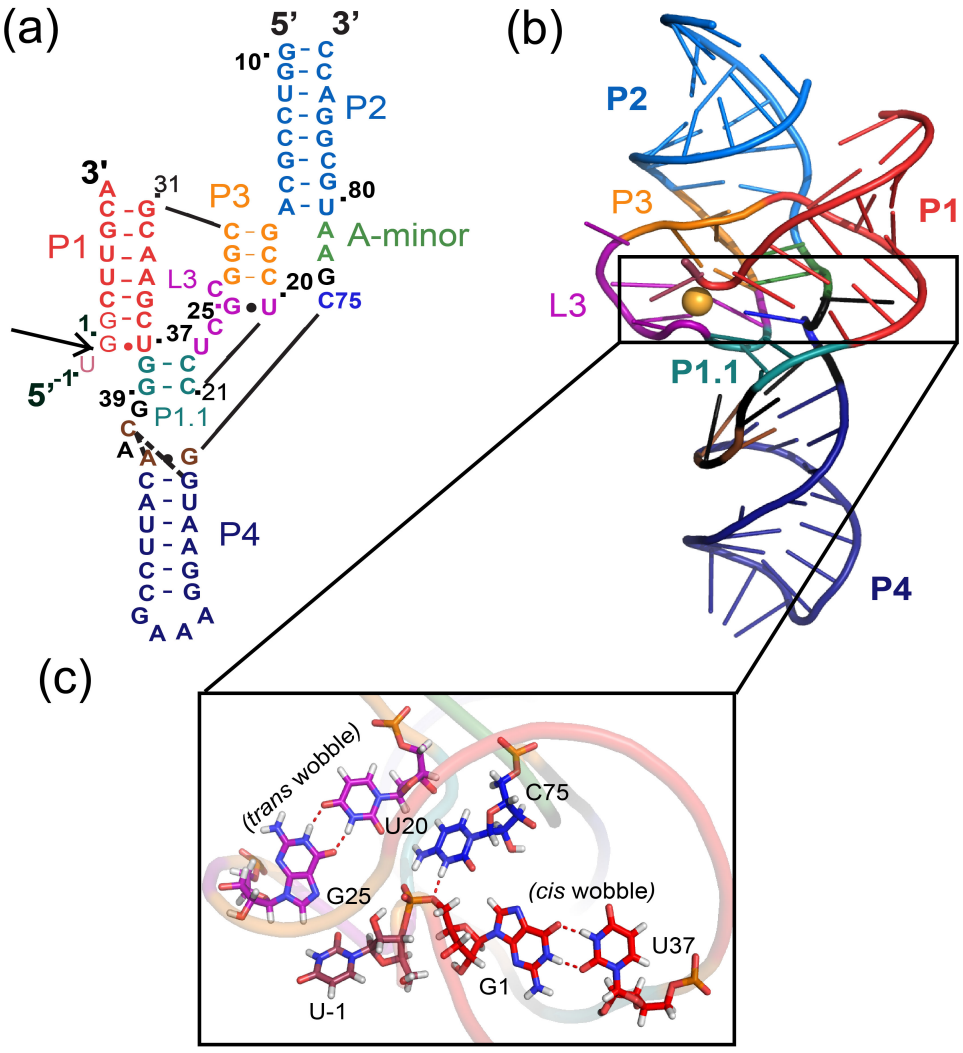

Figure 3.1 Secondary and tertiary structure depiction of the HDV ribozyme

Noteworthy secondary (a) and tertiary (b) interactions between motifs are colored as indicated. The active site is outlined in a black bock in the tertiary structure depiction (left), and the magnesium ion from structure 3NKB is shown in its active site binding pocket in gold. (c) Active site residues of the HDV ribozyme, with *cis* and *trans* GU wobbles indicated.

structure was solved at pH 5 to ensure protonation of C75; however, such protonation is disfavored at the physiological pH at which the ribozyme is active, and at which typical *in vitro* experiments are conducted. Thus, this catalytically crucial conformation represents a minor, yet functionally important, population in standard experimental conditions.

These crystal structures also provided atomic-level resolution on the two GU wobble pairs flanking the HDV ribozyme active site. The most recent crystal structure of a reaction precursor[10] shows them as a *cis* G1U37 wobble adjacent to the scissile phosphate, and a *trans* G25U20 wobble somewhat more distal in sequence (Fig. 3.1c). In general, GU wobble pairs play a special role in RNA biology due to their function in rendering the genetic code promiscuous, as underscored by their high abundance and frequent conservation in functionally important sites of RNAs.[145]

Initial biochemical characterization of the HDV ribozyme indicated that mutation of either of its two GU wobbles resulted in significant decrease in catalytic activity;[146-149] however, more recent evaluation of the G1U37 wobble directly adjacent to the scissile phosphate appeared to suggest more tolerance for alternate base pairing at this position.[91] Recent molecular dynamics (MD) simulations indicate that each wobble plays a distinct role in metal binding; the *cis* G1U37 wobble attracts more diffuse ions, while the *trans* G25U20 wobble exhibits a better-defined binding site to attract longer-residence ions.[150, 151] Subsequent studies on a G25AU20C mutant showed disruption of this metal binding, further highlighting the crucial nature of these wobbles.[144] However, no systematic tests have been performed to explore the atomic-level structural contributions of these conformational landmarks to HDV ribozyme active site architecture and stability.

In order to investigate this question, we performed mutational analyses on the more recent precursor structure via MD simulations. We chose this structure as the basis for our studies because, as discussed in Chapter 2, it represents a conformation especially poised for catalysis. We were thus able to assess the effects of our mutations in the most catalytically relevant context. Our results show that the GU wobbles are stable in the presence of either sodium or magnesium at the active site, but that the identity of the active site ion leads to varying in-line fitness of the scissile phosphate. We also find that in-line fitness is further tuned by interactions between the G28 and U-1 hydroxyl groups in simulations containing sodium. Our mutations of the G25U20 (Table 3.1) wobble exhibit disrupted in-line fitness, C75(N3H3+)G1(O5') general-acid hydrogen bonding, and G1U37 hydrogen bonding. Conversely, alternative hydrogen bonds with nearby moieties of R1Y37 mutants destabilize active site structure due to low in-line fitness and disruption of the general acid hydrogen bond. Further investigation of this site by means of Y1R37 mutants shows comparable, and often increased, base pair and active site stability compared to R1Y37 mutants, indicating the active site's tolerance of pyrimidine-purine pairs at this position. We also find that active site

Type of Simulation	Number of Simulations	Duration of Simulations
Control Simulations		
Wildtype	6	40ns, 3x20ns, 83ns
Mg^b	4	120ns, 78ns, 2x20ns
A25Y20 Mutants		
A25C20	3	3x20ns
A25U20	3	3x20ns
R1Y37 Mutants		
A1+C37	3	3x20ns
A1ºC37	3	3x20ns
A1U37	3	3x20ns
G1C37	3	3x20ns
Y1R37 Mutants		
C1G37	3	3x20ns
U1G37	3	3x20ns

Table 3.1 List of Simulations

disorder present in either wild-type or mutant ribozymes is prevented from destabilizing global conformation due to the type I A-minor and P1.1 motifs above and below the active site, respectively. Taken together, our results define specific structural rationales by which each type of active site GU wobble supports a catalytically proficient active site. We further suggest that the complex tertiary structure of the HDV ribozyme provides significant and redundant layers of protection against global structural degradation. These insights pave the path to fully understanding structure-function relationships within the HDV ribozyme, as well as within the newly discovered family of HDV-like ribozymes.[73, 75]

3.2 Materials and Methods

All structures were prepared and simulated based on the most recent *trans*-acting precursor structure, 3NKB,[10] using the AMBER10/AMBER11 suites of programs. [107, 152] Wildtype_Na and all mutant structures (Table 3.1) were net-neutralized with sodium atoms placed by the LEaP module at the points of greatest electrostatic favorability. Wildtype_Mg structures (Table 3.1) contain a single magnesium ion near the active site retained from structure 3NKB, and were consequently neutralized with sodium. A25N20 and R1Y37 mutations were made using an equilibrated structure with the C75(N3)G1(O5') "general-acid" heavy-atom distance of 3.1 Å and an in-line fitness[119] (as described in Chapter 2) of 0.66. Y1R37 mutations were made in a ~10ns snapshot from a G1C37 simulation containing C75(N3)G1(O5') distance of 3.0 Å and an in-line fitness[119] value of 0.46. For all mutant structures, harmonic restraints on the C75(N3)G1(O5') and Watson-Crick heavy atom distances of the mutated base pair were maintained during equilibration, and released during simulation. We maintained restraints additionally for the first 5 ns of all A1ºC37 simulations. The PMEMD-Sander module of the AMBER10 and AMBER 11[107, 152] suite of programs was

used for equilibration and production runs, using the bsc0 force field[63] with updated χ torsion angles.[67] Structures were either equilibrated using in-house protocols or the previously defined ABC protocol.[111] Hydrogen bonds were analyzed using the Simulaid suite of programs,[118] and in-line fitness components were calculated using the ptraj module in AMBER10/11.[107, 152]

3.3 Results

3.3.1 The HDV Ribozyme Active Site is Stable Regardless of Active Site Metal Ion

The HDV ribozyme has shown distinct behaviors in monovalent and divalent metal ion environments.[20, 88] Although cleavage is optimal in the presence of Mg^{2+}, it also occurs in high (molar) concentrations of Na^{+}.[20, 88] pH-rate studies indicate a potential change in mechanism from monovalent to divalent metal ion conditions.[20] To test whether these biochemical differences correlate with structural differences at either or both of the GU wobbles, as well as at the scissile phosphate itself, we conducted comparative Wildtype_Na (neutralized with sodium) and Wildtype_Mg simulations (containing one magnesium ion and otherwise neutralized with sodium, Table 3.1). The magnesium ion in our Wildtype_Mg simulations was resolved in the crystal structure,[10] and we maintained its position during our calculations. As hypothesized by the authors, its location may hint at its ability to act as a general and/or Lewis acid during catalysis.[10]

Given their ability to capture fast, microscopic biomolecular dynamics at atomic resolution,[60, 66] we found molecular dynamics (MD) simulations ideal for our purposes. We used the most recent *trans*-acting (two-stranded) precursor structure[10] as a basis for our studies; our results from Chapter 2 indicate that this structure represents a conformationally "activated" structure, despite the fact that it is solved at pH 5 where the ribozyme is

minimally active. Although such a structure, involving a protonated C75, occurs in a very small percentage of molecules in solution experiments at physiological pH, the captured conformation nonetheless appears to approximate a necessary intermediate in the ribozyme's reaction pathway. Thus, our Wildtype_Na, Wildtype_Mg, and all mutant simulations are based on this structure so that we are able to distinguish between dynamical differences in the most catalytically relevant context.

In order to correlate the geometry of the scissile phosphate to potential effects on catalysis, we employed the previously described in-line fitness parameter.[119] Briefly, this parameter combines the nucleophile's in-line attack angle and the distance between nucleophile and electrophile into a single parameter that describes the propensity of that phosphodiester bond to be cleaved (see Chapter 2 for equation). The use of this parameter thus provided us with a straightforward metric that we can compare to retention of both GU wobble pairs and other active site hydrogen bonds throughout this work.

Both Wildtype_Na and Wildtype_Mg simulations showed almost complete retention of the G1U37 and G25U20 wobbles throughout simulation (Fig. 3.2a). Despite exhibiting stable lateral boundaries, however, we found the scissile phosphate able to sample a wide

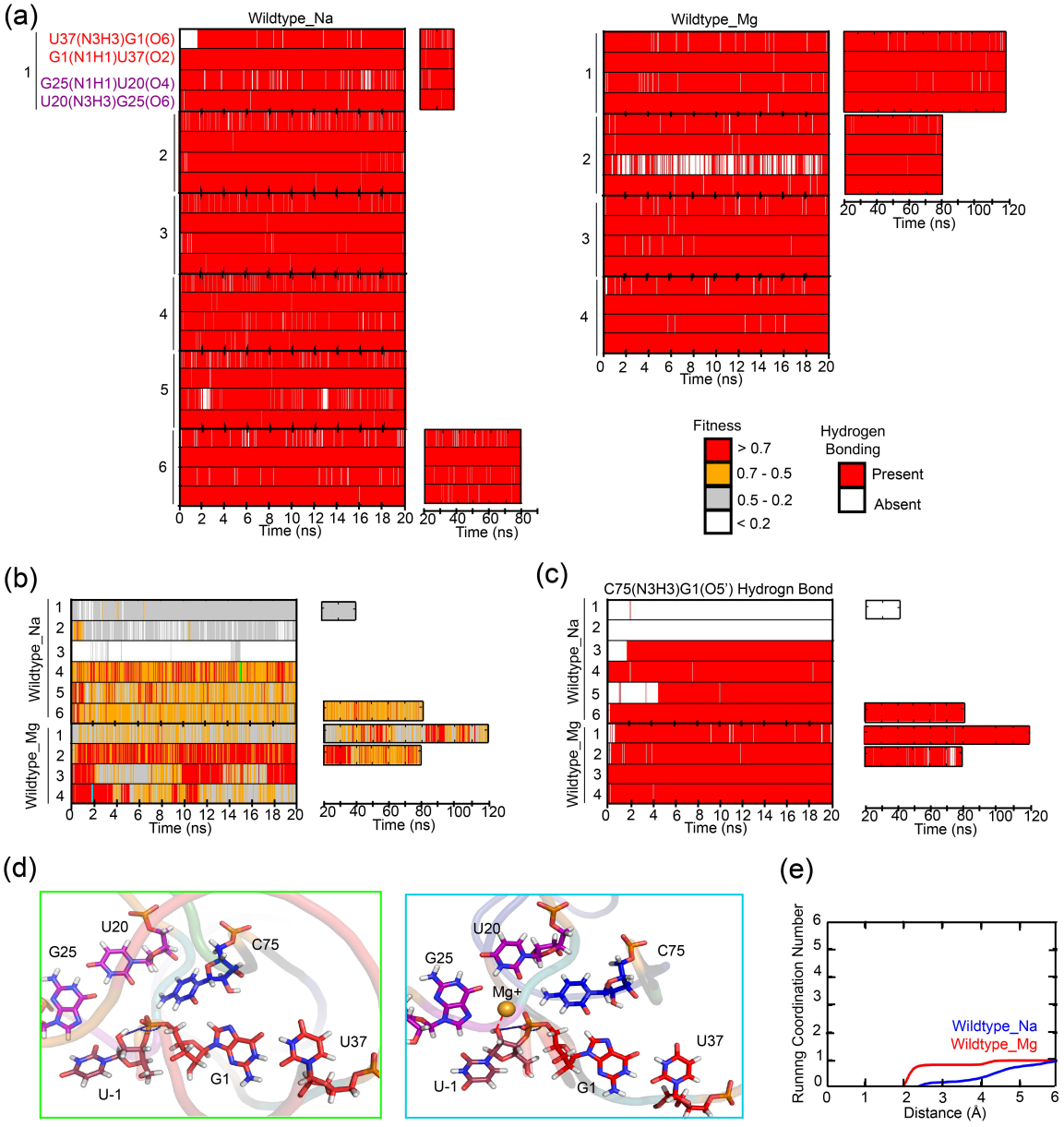

Figure 3.2 Active site dynamics of the HDV ribozyme in the presence of Mg^{2+} versus Na^{+}.

(a) Hydrogen bonding patterns of the cis G1U37 and trans G25U20 wobbles, (b) in-line fitness parameters, and (c) C75(N3H3+)G1(O5') general-acid hydrogen bond retention in the HDV ribozyme active site in Wildtype and Mg. (d) Representative conformations of active site residues in the absence (left panel) and presence (right panel) of the long-residence L3 Mg2+ ion. Residues are colored as in Figure 3.1, with oxygens in red, nitrogens in blue, hydrogens in white and phosphorus in orange for clarity. Images are made from 101ps- or 102ps-averaged snapshots. The bright green and cyan lines in panel (b) indicate from which replicate and at which time the snapshots were taken from the Wildtype and Mg simulations, respectively. Dark blue-and-red dashed lines and dark blue arcs at the scissile phosphates indicate the parameters needed to calculate in-line fitness. Dashed bright red lines in the left panel indicate coordination of the U-1(O2') to the Mg ion. (e) Comparison of ion retention at the U-

1(O2') between Wildtype and Mg simulations. Running coordination numbers indicate the number of ions at a given distance from the U-1(O2'). Curves are a weighted average of all six Wildtype replicates and all four Mg replicates, respectively.

range of conformations based on metal ion identity (Figure 3.2b). As we observed in Chapter 2, Wildtype_Na simulations did not exhibit ubiquitously excellent in-line fitness, and revealed generally static behavior; typically a simulation started with favorable or unfavorable values and maintained that trend throughout a single simulation. Wildtype_Mg simulations experienced more dynamics in in-line fitness, with a large percentage of values ≥ 0.7. We thus observed that the presence of an active-site magnesium ion enables the scissile phosphate to more often sample conformations optimal for catalysis. Our simulations are able to provide a potential structural role for a magnesium ion that so far has only been strongly implicated to support catalysis. [10, 144, 150, 151]

After elucidating the effects that varying L3 ion types have on in-line fitness, we studied these effects on the C75(N3H3+)G1(O5') general-acid hydrogen bond (Figure 3.2c), an additional important marker of active site stability. In our high in-line fitness Wildtype_Na simulations (replicates 4-6 in Fig. 3.2b), we found a C75(N3H3+)G1(O5') occupancy of ≥ 80% (Fig. 3.2c). In rare cases, however, trajectories with low in-line fitness (replicates 1-3 in Fig. 3.2b) were also able to retain the general acid hydrogen bond (Fig. 3.2c). Such an exception is replicate 3, which exhibited unfavorable in-line fitness and a largely retained C75(N3H3+)G1(O5') hydrogen bond. Conversely, in our Wildtype_Mg simulations, the C75(N3H3+)G1(O5') interaction was always retained. These results provide further structural roles for the catalytically implicated magnesium ion. [10, 144, 150, 151]

Previous mechanistic characterizations of HDV ribozyme ion binding sites indicate that structural metal ions are bound strongly with innersphere coordination, while catalytic metal ions are bound weakly through outersphere coordination.[88] In the majority of our simulations, the active site magnesium ion showed short distances to U-1(O2') (Figure 2e),

indicating innersphere coordination. Retention of this ion's position in all our simulations further supports the possibility of its role as a general base during catalysis; however, it is retained too closely to the U-1(O2') to be the outersphere catalytic ion described by Bevilacqua and coworkers.[88] This strong magnesium retention may, alternatively, be due to force field artifacts inherent in simulations containing divalent ions.[60] Despite these possible artifacts, our simulations provide atomistic explanations for better support of catalysis with magnesium compared to sodium, namely through support of higher in-line fitness and higher occupancy of the C75(N3H3+)G1(O5') general acid bond. Our Wildtype_Na simulations are nevertheless better suited to act as control simulations in the current work due to the more realistic force field description of sodium ions, and consequently more ideal behavior, of monovalent ions in simulation.

In summary, our Wildtype_Na and Wildtype_Mg simulations show equal retention of active site GU wobble hydrogen bonds. Our Wildtype_Mg simulations show more stable general acid hydrogen bonding and in-line fitness values compared to magnesium-free Wildtype_Na simulations, providing a structural perspective on metal ion preferences in the HDV ribozyme active site.

3.3.2 G28-U-1 Interactions Tune In-Line Fitness in Wildtype_Na Simulations

During investigation of additional, external influences on in-line fitness, we observed alternating interactions between the 2' hydroxyl groups of U-1 and G28 in Wildtype_Na versus Wildtype_Mg simulations (Fig. 3.3). In most Wildtype_Na replicates of high in-line fitness, the U-1 2' hydroxyl typically pointed towards the scissile phosphate (Fig. 3.3b), enabling the formation of a hydrogen bond between the G28(HO2') hydrogen and U-1(O2') (Figure 3a). The exception to this was replicate 5, in which high in-line fitness was

maintained, but the G28(HO2')U-1(O2') interaction was disrupted due to transient ionic interactions with U-1(O2'). Despite this one exception, we generally observed this interaction to have a stabilizing effect on in-line fitness, partially compensating for the lack of structural stability induced by the presence of a magnesium ion in this vicinity.

A different G28-U-1 interaction occurs in Wildtype_Mg simulations. With the active site metal ion closely coordinated to U-1(O2') and G1(O2P) (Figure 3.2d; Fig. 3.3b), the U-1(HO2') typically pointed to the outer reach of the L3 loop. In this conformation, the U-1 2' hydroxyl became the hydrogen bond donor to the G28(O2'). This bond is often disrupted, however, when sodium ions interact transiently with the U-1(O5') group and "steer" the U-1(O2') group away from G28(O2'). Thus, this hydrogen bond is likely more of a consequence of the Wildtype_Mg U-1(O2') conformation rather than a tuning mechanism of in-line fitness as it is in the Wildtype_Na simulations.

3.3.3 A25N20 Mutants Disrupt Active Site Stability

Having ascertained that the HDV ribozyme's GU wobbles are stable regardless of ion identity (Fig. 3.2a), we performed a thorough probing of each GU wobble. We began with the *trans* G25U20 wobble, due to recent characterization of its metal binding

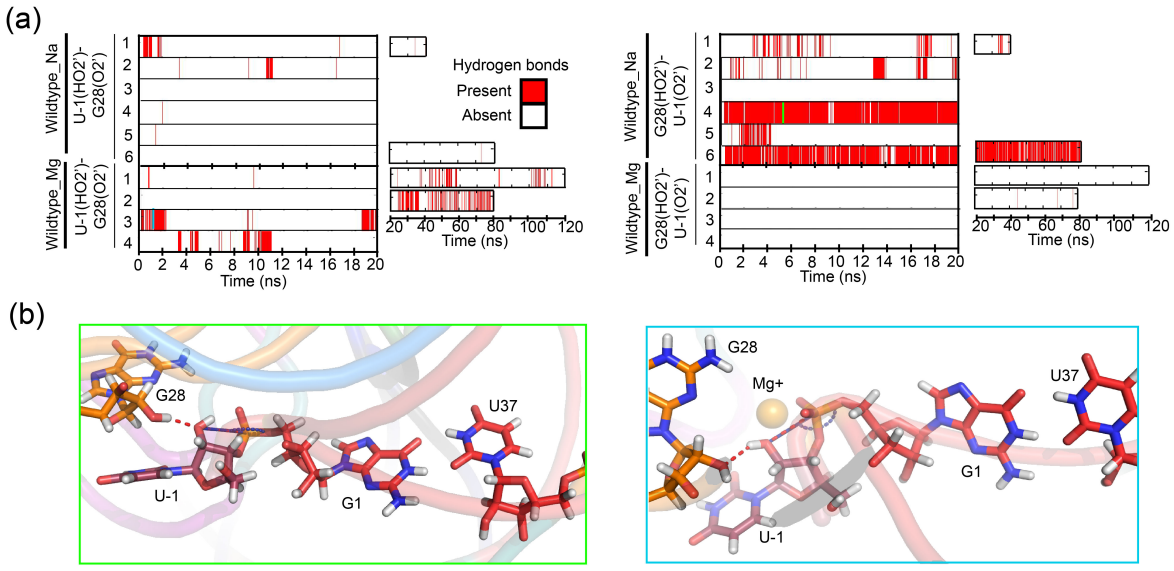

Figure 3.3 Effects of G28 on scissile phosphate conformation.

(a) Occurrence of the two possible hydrogen-bonding interactions between the 2'hydroxyls of G28 and U-1. Note that the G28(2OH')U-1(O2') hydrogen bonding patterns of the Wildtype simulations often match periods of high in-line fitness as observed in Figure 3.2(b). Cyan and bright green lines indicate from which replicate and at which time the snapshots in panel B were taken. (b) Snapshots of G28-U-1 2' hydroxyl interactions in Mg (left) and Wildtype (right) simulations. Snapshots are from structures averaged over 101ps or 102ps of the indicated trajectory. Red dashed lines indicate hydrogen bonds formed in the snapshot. Dark blue-and-red dashed lines and dark blue arcs at the scissile phosphate indicate measured parameters used to calculate in-line fitness.

properties.[144, 150, 151] Earlier mutation studies also showed that this wobble is crucial for effective catalysis.[149] Additionally, this wobble has proven to bind a magnesium ion that is implicated as a Lewis-type acid, a Bronsted-type base, or both, during catalysis.[10, 144] Further work has shown that an A25C20 mutation inhibits the ability of the L3 loop to bind this catalytic magnesium ion.[144] However, no recent investigation has been conducted on the geometric effects of the *trans* G25U20 wobble on active site catalytic ability. We thus used our A25N20 sets of simulations (Table 3.1) to answer this question.

Our A25C20 mutants, isosteric to the native G25U20 wobble, exhibited a stable mutant base pair, but often showed low in-line fitness and loss of other active site hydrogen bonds (Fig. 3.4). Among our replicates, one trajectory exhibited high in-line fitness (replicate 2), and two trajectories displayed in-line fitness predominantly < 0.5. Only replicate 2 had high C75(N3H3+)G1(O5') hydrogen bond occupancy. In low in-line fitness replicates 1 and 3, active site geometry was disrupted likely due to increased L3 dynamics as a result of lack of a nearby ion at the start of the trajectory. These dynamics enable alternative hydrogen bonds to form between C75 and the scissile phosphate's nonbridging oxygens, stabilizing the new conformation (Figure 3.4a). Conversely, although an ion was also not present at the start of replicate 2, the general acid C75(N3H3+)G1(O5') hydrogen bond was retained. Despite the differences in in-line fitness and hydrogen bond occupancies among replicates, the G1U37 wobble stayed mostly intact except for a 4 ns-interruption in the U37(N3H3)G1(O6) distance in replicate 1. This general retention may be attributable to the isostericity of the A25C20 Watson-Crick pair with the *trans* G25U20 wobble.

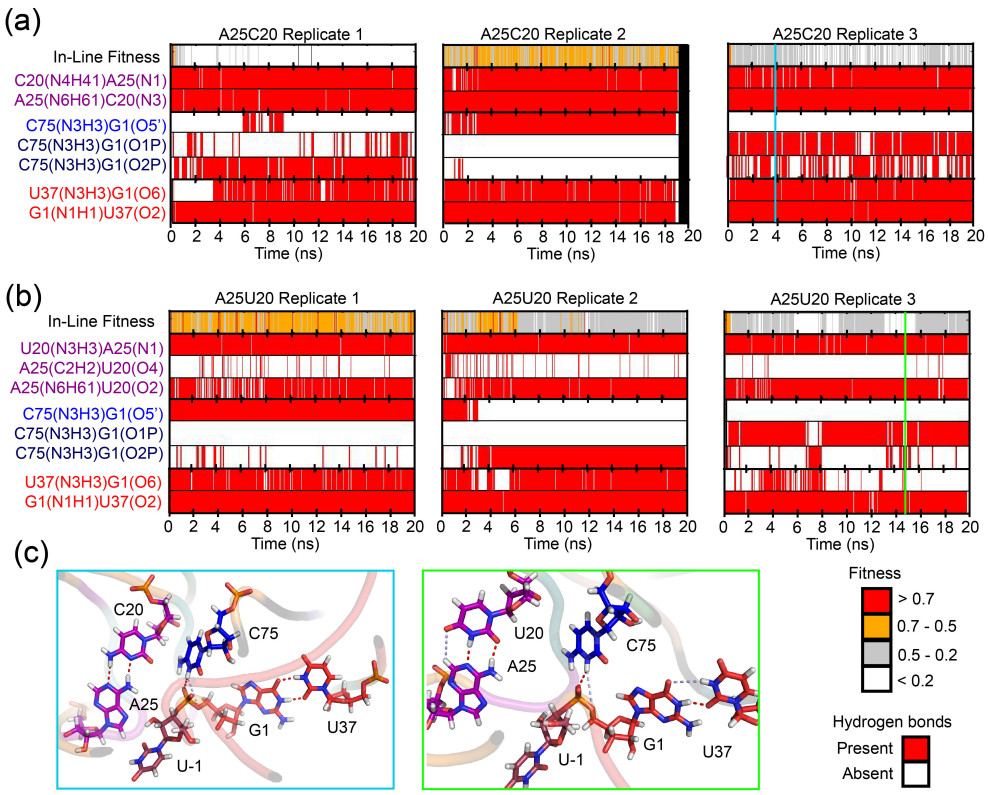

Figure 3.4 Active site dynamics of the A25N20 mutants.

In-line fitness and active site hydrogen bonds of the (a) A25C20 and (b) A25U20 mutants. Text colors are consistent with Figure 3.1: purple indicates the A25N20 wobble; blue indicates canonical, and dark blue indicates noncanonical, hydrogen bonds made by C75; red indicates G1U37 hydrogen bonding interactions. Cyan and bright green lines indicate from which replicate and at which time the snapshots in panel C were taken. (c) Snapshots of representative active site dynamics in the A25C20 (left) and A25U20 (right) mutants. Red dashed lines indicate hydrogen bonds stable during the time of snapshot, while light blue dashed lines indicate hydrogen bonds that are destabilized during the length of the trajectory. Colors of nucleotides are consistent with Figure 3.1, and oxygens are red, nitrogens are blue, hydrogens are white, and phosphorus is orange for clarity.

Although the A25U20 mutant was able to sample high in-line fitness more often than the A25C20 mutant, significant active site disruption still occurred (Fig. 3.4b). All three replicates for this mutant showed at least some high in-line fitness. However, retention of the C75(N3H3+)G1(O5') general acid bond was comparable to that observed of the A25C20 mutant (only one out of three replicates), and similar alternative hydrogen bonds between C75 and G1's nonbridging oxygens were formed in long residence (Figure 3.4b). Interestingly, we saw greater dynamics in the G1U37 wobble pair compared to the A25C20 mutants (replicates 2 and 3, Fig. 3.4b). Such increased active site disruption may be due to the non-isostericity of A25U20 to the native G25U20, which may result in greater L3 dynamics that are consequently transferred to the G1U37 wobble.

In conclusion, we find that both canonical and isosteric wobble mutations to the G25U20 *trans* wobble disrupt active site dynamics. In-line fitness is poorer compared to Wildtype_Na simulations, and canonical C75 and G1U37 hydrogen bonds are disrupted. The observed active site instability in both A25N20 mutants provides a structural rationale which, together with previously described ion binding deficiencies of the A25C20 mutant,[144] explain the experimentally measured decreases in catalytic rate of these mutants.[149]

3.3.4 R1Y37 Mutants Disrupt Local Architecture

To determine if mutations at the G1U37 wobble affected G25U20 stability, analogous to the A25U20 disruptions of G1U37 hydrogen bonding described in the last section, we next simulated our R1Y37 mutants (Table 3.1). All of these structures showed a decrease in active site stability compared to Wildtype_Na simulations, in addition to a less stable mutant pair than observed for the A25N20 mutants (Fig. 3.5, 3.6). The lack of stability in mutant base

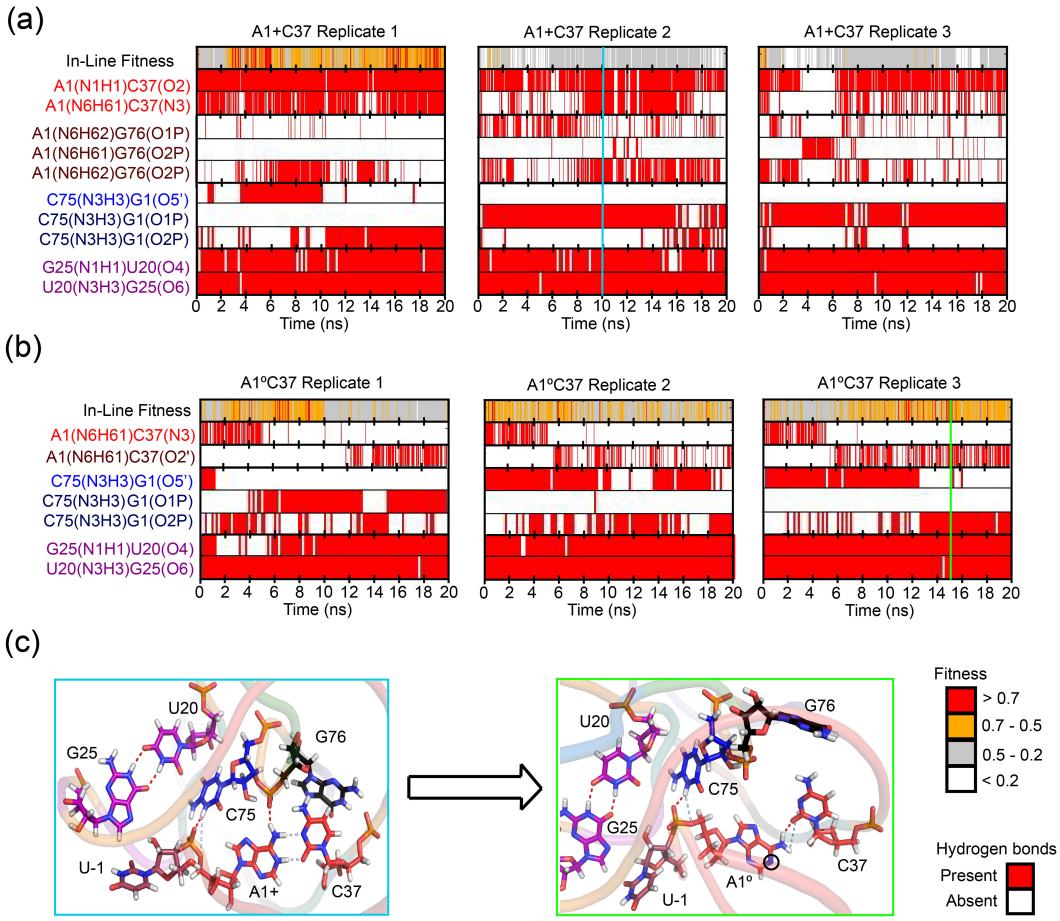

Figure 3.5 Active site dynamics of the A1C37 wobble.

In-line fitness and active site hydrogen bonds of the (a) A1+C37 and (b) A1ºC37 mutants. Colors correspond to Figure 3.1; canonical and noncanonical A1C37 hydrogen bonds are labeled in red and dark red respectively; canonical and noncanonical C75 interactions are in blue and dark blue respectively; and G25U20 interactions are in purple. (c) Conformational differences between protonated (left) and unprotonated (right) A1C37 wobbles. The arrow indicates a change in pH. The black circle in the right panel denotes the position where the hydrogen would be at lower pH. Colors of the nucleotides are retained from Figure 3.1, while important heavy atoms are colored as in previous figures.

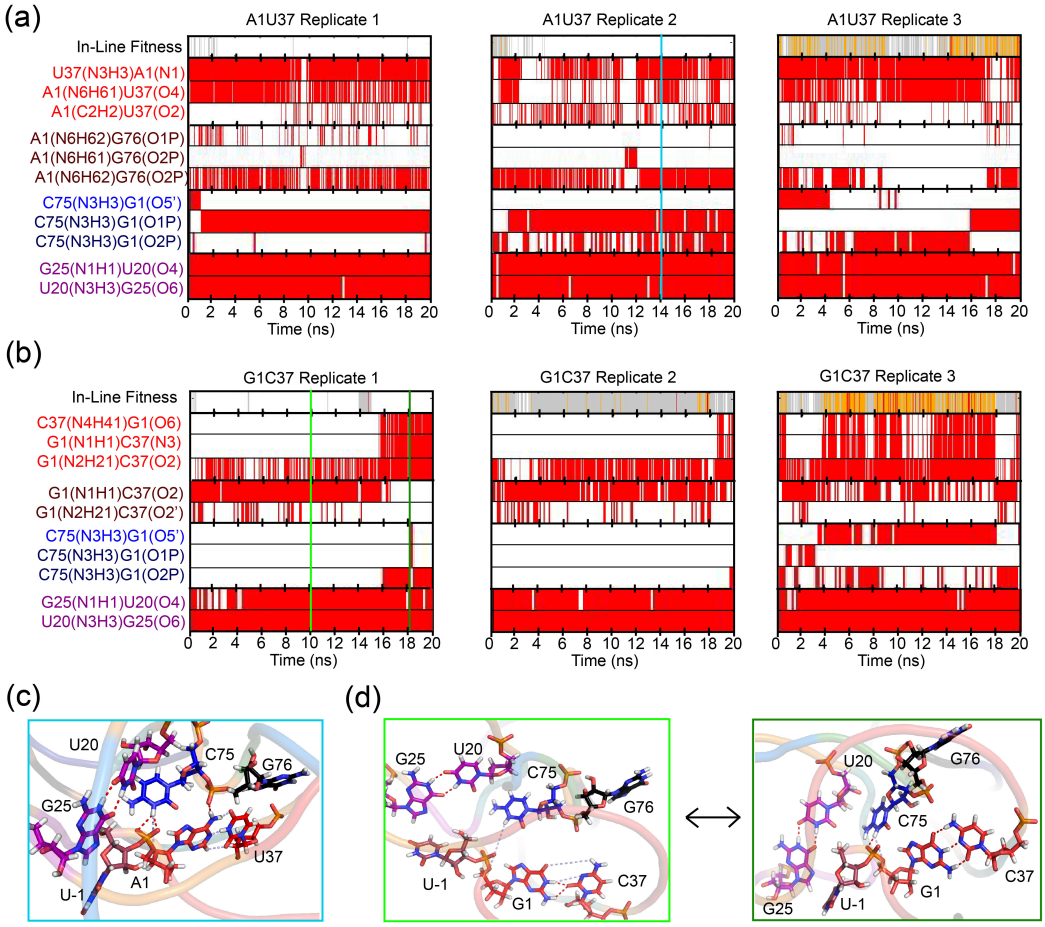

Figure 3.6 Active site dynamics of Watson-Crick Y1R37 mutants.

In-line fitness and active site hydrogen bonds of (a) A1U37 and (b) G1C37 mutants. Coloring of labels and plots is in keeping with previous figures. An additional colored line is added in panel (b) to indicate the replicate and time from which the second snapshot of the G1C37 mutant in panel (d) is taken. Snapshots of (c) A1U37 and (d) G1C37 representative structures. As in previous figures, strongly retained hydrogen bonds are shown in red dashed lines, while canonical, dynamic hydrogen bonds are shown in blue. In panel (d), the shifted (left) and canonical (right) Watson-Crick interactions of the G1C37 mutant are both shown, to indicate their interconverting nature.

pair was due to the formation of alternative hydrogen bonding interactions. Most A1Y37 simulations exhibited interactions between the exocyclic amine of A1 with the nonbridging oxygens of nearby G76 (Figures 3.5a, 3.6a,c), while the dynamics in the G1C37 base pair were due to shifting Watson-Crick interactions (Fig. 3.6b,d). No Y1R37 mutants disrupted the G25U20 wobble to the same extent as the A25U20 mutants disrupted the G1U37 wobble (compare Fig. 3.4-3.6).

We simulated protonated and unprotonated versions of the A1C37 wobble to model the effects of this mutant at low and high pH, respectively (Fig. 3.5). Although the A1+C37 base pair represents an isosteric replacement to the G1U37 wobble, this mutant showed decreased stability of active site interactions compared to Wildtype_Na simulations (Fig. 3.5a,c). Similar to the A25C20 simulations, only one of three replicates exhibited significant high fitness (Fig. 3.5a). This replicate only showed partial retention of the C75(N3H3+)A1(O5') hydrogen bond, however, and indeed even exhibited the noncanonical C75(N3H3+)A1(O2P) hydrogen bond in the same trajectory (Fig. 3.5c).

We also simulated the A1ºC37 wobble, to emulate the structural effects of an unprotonated AC wobble as would be experienced at pH values > 6. The A1ºC37 wobble, interestingly, showed increased high in-line fitness and active site stability compared to the protonated version. Although the single Watson-Crick A1(N6H61)C37(N3) hydrogen bond was not retained after restraints were released (see Materials and Methods), alternative Hoogsteen-Sugar Edge[153] interactions between the A1 exocyclic amine and C37's 2' hydroxyl oxygen appeared to aid in stabilization of the rearranged conformation (Fig. 3.5b,c).

Our two Watson-Crick Y1R37 mutants, A1U37 and G1C37, showed greater decrease in various active site landmarks than observed in the A1C37 mutants (Fig. 3.6). For the

A1U37 mutants, the alternative A1(N6)G76(O1/O2P) interactions appeared to neither stabilize nor destabilize the base pair, but were simply a result of base pair dynamics (Fig. 3.6a,c). These trajectories exhibited the lowest percentage of favorable in-line fitness of all mutants, with not even an entire trajectory exhibiting high values. C75(N3H3+)A1(O5') distances were also present for only a small percentage of one trajectory, with the alternative C75-A1 interactions stabilizing non-Wildtype_Na conformations (Fig. 3.6a,c).

G1C37 mutants show similar behavior to the A1U37 mutants in terms of active site stability (Fig. 3.6b,c). For these mutants, dynamics were due to a shifting of the Watson-Crick interactions: We observed dynamic exchanges between the canonical (C37(N4H41)G1(O6), G1(N1H1)C37(N3)) and noncanonical (G1(N1H1)C37(O2)) hydrogen bonds, often within a single trajectory (Fig. 3.6b,d). This shifting resulted in loss of high in-line fitness and lack of C75(N3H3+)G1(O5') interactions in replicates 1 and 2 (Fig. 3.6b). Our results thus indicate that no type of purine-pyrimidine canonical Watson-Crick base pair is tolerated at the active site, due to decrease in in-line fitness and the destabilization of canonical active site interactions and of the base pairs themselves. These mutants are, however, able to form alternative hydrogen bonding interactions that may aid in preventing active site disorder from spreading throughout the whole molecule.

A wealth of experimental data exists for a variety of *cis*- and *trans*-acting genomic and antigenomic mutants of the G1U37 wobble pair.[144, 146, 147, 149-151] Early studies on A1ºC37 *trans*-acting mutants showed decreased cleavage rate, but an almost wild-type extent of cleavage.[146] These trends may be explained by our observations of high in-line fitness, yet disrupted C75(N3H3+)A1(O5') general acid bonds in this mutant. Cleavage assays for both the *cis*-[147] and *trans*-acting[146] A1U37 mutants imply that the structure should experience at

least partially high fitness and active site stability, which our simulations do not exhibit. It is thus likely that in this case, our simulations captured a subset of population significantly low in fitness, and that high sampling will reveal more catalytically fit active site conformations. The low percentage of catalytically fit G1C37 trajectories is more consistent with trends of *cis*- rather than *trans*-acting mutants.[146, 147] We believe this to be due to the selection process used to obtain the *trans*-acting ribozyme; the G1C37 mutation in an already fast-acting ribozyme enhanced the rate further still, and the resulting construct was used by the group in later studies.[154] Thus we believe that this construct is biased towards a G1C37-dependent cleavage mechanism and its kinetic results are not fully reflective of the mutant's behavior. A1+C37 rate constants were only measured in later pH-rate studies.[91] These results reported near-wild-type behavior along the pH range in which the A1C37 pair was protonated and, indeed, reported such behavior for the other R1Y37 mutants in the study. However, similarly to the work conducted by Nishikawa *et al,* [146] all mutations were made in a specialized, fast-folding construct that may thus affect cleavage results. [91]

In conclusion, all of our R1Y37 mutants display low in-line fitness, decreased occupancy of the general acid C75(N3H3+)N1(O5') hydrogen bond and dynamic Watson-Crick mutated base pairing. The greatest occurrence of in-line fitness for these mutants is in A1C37 wobbles, but is still not comparable to that seen in Wildtype_Na structures. These results indicate that only a G1U37 wobble results is able to foster a high fraction of favorable in-line fitness, as well as retained active site interactions in the HDV ribozyme. We also find that alternative hydrogen bonding interactions occur to a large extent in this group of mutants, which may help contain active site disorder. These alternative interactions occur to a greater extent in our Y1R37 mutants than in our A25N20 mutants, indicating that the

G1U37 wobble has a greater effect on local conformational retention, while the G25U20 wobble effectively stabilizes elements more distal to its position.

3.3.5 Y1R37 Mutants Show Greater Active Site Stability Compared to R1Y37 Mutants

An interesting feature of GU wobbles is that they are not isosteric to the inverse UG wobbles, due to the different glycosidic angles, and relative positions of the C1' atoms. [145, 153, 155, 156] We thus next investigated U1G37 mutant "reverse" wobbles, to study their effects on in-line fitness and active-site geometry. Interestingly, we found all three of our U1G37 replicates to exhibit ideal active site stability, greater than even our Wildtype_Na simulations (Fig. 3.7a,c). The mutant base pair was stable and C75(N3H3+)U1(O5') hydrogen bonds were retained (Fig. 3.7a). This is in direct contrast to experimental results, which predict a severe loss in activity for U1G37 mutants in both *cis*- and *trans*-acting, genomic and antigenomic HDV ribozymes.[146, 147, 157] This discrepancy may be due to differences in construct between experimental studies and our simulations: Studies by Nishikawa *et al*[146] demonstrated significant miscleavage of this mutant at the -1 position, indicating high in-line fitness at this position rather than at the +1 position. This trend was not present in our simulated ribozyme (Fig. 3.7a). However, the constructs in the cited mutagenesis studies contain a significantly longer 5' leader sequence than does our simulated ribozyme.[146, 147, 157] The presence of this leader sequence likely distorts the 5' end of the P1 helix in such a way that the scissile phosphate experiences low in-line fitness, while the G2 phosphate experiences high in-line fitness.

Due to the highly poised in-line fitness of our U1G37 mutant models, we next tested the active site conformational stability of C1G37 mutants to see if similar trends occurred.

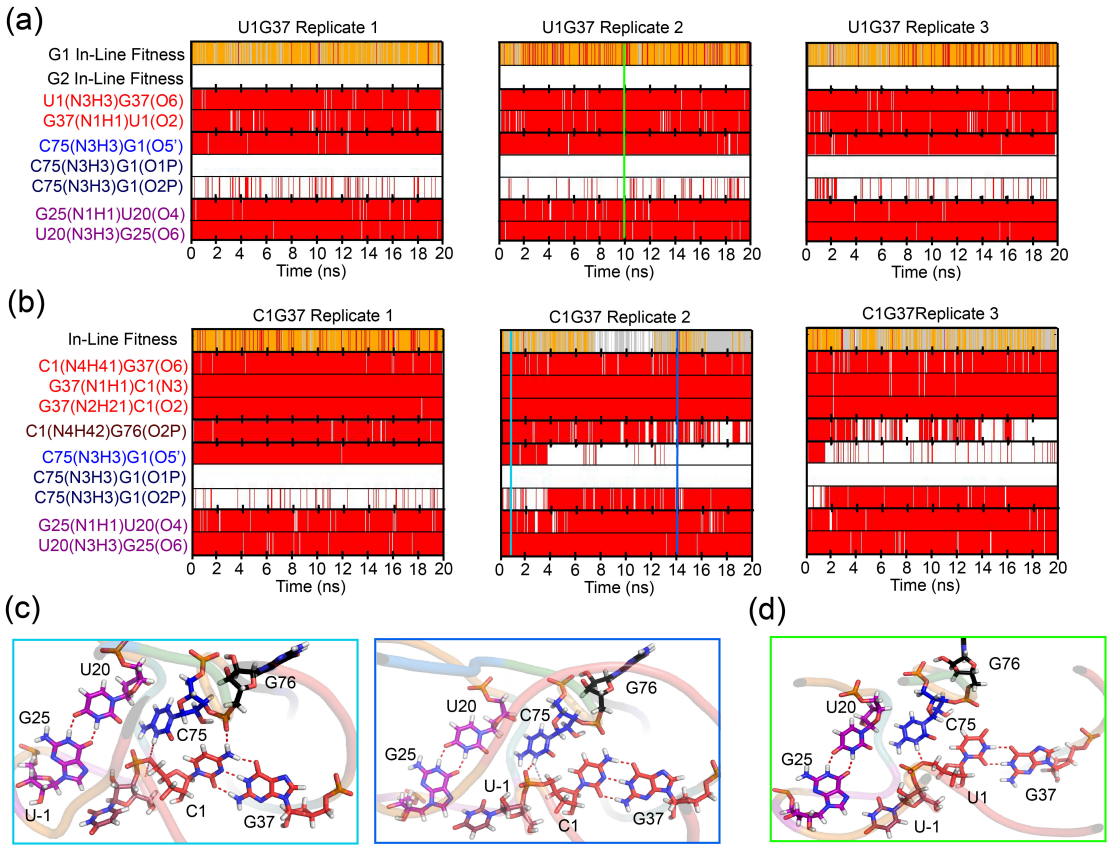

Figure 3.7 Active site conformations of the Y1R37 mutants.

In-line fitness, Y1R37, C75 interactions and G25U20 hydrogen bonds of (a) U1G37 and (b) C1G37 mutants. Snapshots of relevant conformations of (c) C1G37 and (d) U1G37 mutants. Coloring is similar to previous figures. An extra dark blue line in the C1G37 trajectories defines a snapshot with alternative hydrogen bonding interactions.

Comparable to our U1G37 mutants, all of our C1G37 replicates showed some occurrence of high in-line fitness (≥ 0.5). The observed favorable values were not always dependent on retained C75(N3H3+)C1(O5') hydrogen bonding, but in contrast to the R1Y37 mutants, the mutated base pair is always stable (Fig. 3.7b). Thus, the decreased activity observed in mutational studies is likely more dependent on disruption of the general acid bond than on unfavorable in-line fitness. In further contrast to the R1Y37 mutants, the alternate C1(N4H42)G76(O2P) interaction is so high in occupancy (Fig. 3.7b) that it may act as a stabilizing, rather than destabilizing, feature for the mutated base pair. The observed fluctuations in in-line fitness, coupled with stable active site geometry, explain the high cleavage percentage, yet low catalytic rate, observed in some studies for both genomic and antigenomic ribozymes.[146, 157]

We thus find that, contrary to experimental trends, the U1G37 mutants exhibit exceptional in-line fitness and active site stability. Additionally, our C1G37 mutants display higher in-line fitness than any of our R1Y37 mutants, consistent with some previous experimental results.[146, 157] Taken together, our results indicate that Y1R37 mutations are tolerated at the G1U37 wobble better than are R1Y37 mutants.

3.3.6 Active Site Disorder is Contained by Structurally Inert Motifs

The HDV ribozyme's complex tertiary structure is held together by several specialized motifs: type I and type II A-minor interactions with P3 stabilize coaxial stacking of P2 onto P3 and form the top active site boundary (Fig. 3.1, Fig. 3.8a).[11, 12, 158] Additionally, the 2-base-pair P1.1 minihelix bolsters the bottom of the active site (Fig. 3.8a).[11, 12] Because we observed active site disorder in all but our Wildtype_Mg simulations, we investigated how each of these motifs is affected by that disorder (Fig. 3.8). Disruption of either or both of

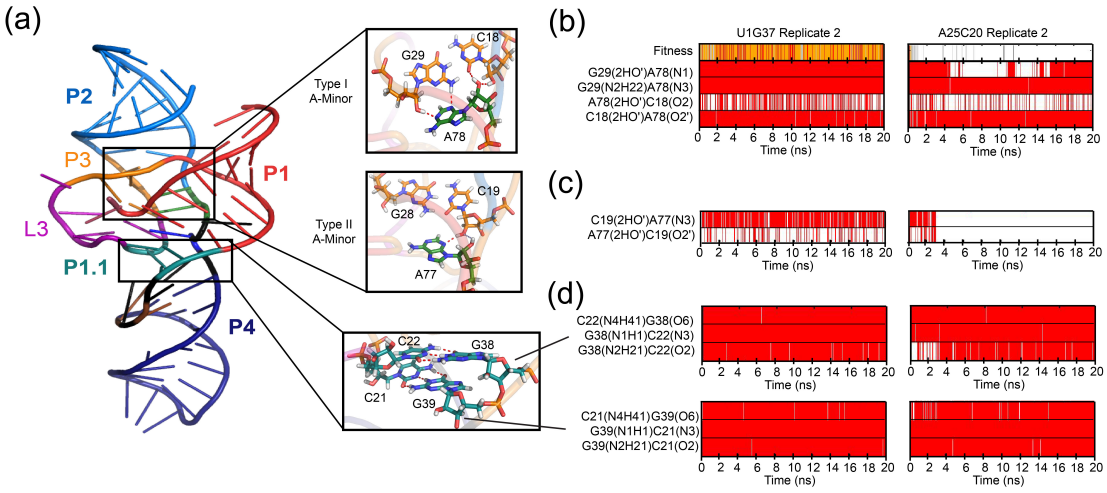

Figure 3.8 Stability of type I A-minor and P1.1 interactions

(a) Global and close-up views of type I and type II A-minor interactions and the P1.1 helix. (b) Representative time traces of A-minor interactions and P1.1 helix hydrogen bonds for high in-line fitness (U1G37, left) and low in-line fitness (A25C20, right) trajectories.

these motifs and consequent global structural disruption could provide an additional explanation for experimentally observed loss in activity for some of these mutants.[144, 146, 147, 149, 157]

A-minor interactions are a well-known type of interactions found in many highly structured RNAs.[158] Upon their initial discovery, their characterization diversified into four categories based on the number of hydrogen bonds they employ. The HDV ribozyme contains two such A-minor interactions: a set of type I A-minor interactions between A78 of the J4/2 joiner and base pair C18G29 of the P3 helix, and a set of weaker, type II A-minor interactions between A77 of the J4/2 joiner and base pair C19G28 of the P3 helix (Fig. 3.8a, top two panels). During our investigation of the active site, we observed that a significant number of replicates in almost every type of simulation showed disruption or complete abolishment of type II A-minor interactions (Fig. 3.8b, middle panel). However, in almost all cases, the type I A-minor interactions remained intact, and reformed if ever disrupted (Fig. 3.8b, top panel). One typical reason for disruption of type I A-minor interactions were dynamic ions allowed within close proximity in a highly perturbed active site. We believe that the overall greater stability of the type I versus type II interactions was due to the comparatively weaker nature of the latter. [158] Additionally, and consistent with previous observations,[49] A77 is only able to form interactions with C19 due to the open nature of the active site (Fig. 3.8a, middle). We noticed that this attribute of the HDV ribozyme's type II A-minor interactions resulted in disruption of these hydrogen bonds whenever the active site was even minimally disturbed. However, because A78 is more fully connected to both partners of the C18:G29 hydrogen bond, the type I interactions are more stable (Figure 3.8a,

top). They thus provide an upper boundary that encases most active site disorder and prevents it from spreading to other portions of the ribozyme.

The P1.1 minihelix was not accounted for in initial pseudoknot predictions, and was first discovered upon resolution of the product HDV ribozyme crystal structure.[11] Mutational experiments proved that it is essential for optimal catalysis of the HDV ribozyme.[85] Because the top base pair of the P1.1 helix, C22G38, forms the junction of the coaxial stack with P1, we tested if disruptions at the scissile phosphate and N1N37 base pair result in further destabilization of the P1.1 helix. In all cases, the P1.1 helix was fully stable regardless of G1U37 or G25U20 mutations (Fig. 3.8b, bottom). We thus find that, similar to the type 1 A-minor interactions above the active site, the P1.1 helix forms a lower bound that prevents active site disorder from spreading to the P4 helix below.

3.4 Discussion

GU wobble pairs result in many structural and chemical effects not observed in canonical Watson-Crick base pairs.[145, 156] Their unique combination of ion- and water-binding functional groups, propagation of up- and downstream backbone distortion, and nonisostericity have made them key recurring players in functional RNAs across all kingdoms of life.[145, 156] This ubiquity of GU wobbles is likely due to their thermodynamic and geometric near-equivalence to standard Watson-Crick base pairs, in addition to their dissimilarities with these base pairs. Two GU wobble of the HDV ribozyme active site, one *cis* (standard) and one *trans* (reverse) wobble, have been shown to be crucial for optimal catalysis. Mutational studies of the *trans* G25U20 wobble indicate that it is crucial for catalysis and correct folding of the ribozyme.[144, 149, 159] This wobble has also been implicated in chelated ion binding through molecular dynamics studies.[150, 151] Additionally, the G1U37

wobble has been hypothesized to bind ions more relevant to structural support, and some mutagenesis studies have indicated that it is crucial for optimal catalysis.[147, 160] Although MD simulations have proven effective in elucidating the crucial nature of the two GU wobbles for ion binding, no in-depth studies have been conducted to study the conformational dependence of active site dynamics on each of the bounding GU wobbles. Here we employ the use of a total of 921 ns of MD simulations to study the rapid (sub-microsecond), small-scale conformational effects of each GU wobble on the HDV ribozyme active site.

Our work in Chapter 2 has shown that the most recent precursor structure[10] represents an activated intermediate in the HDV ribozyme folding landscape that is crucial for catalysis, but rare at physiological pH. We used this structure again in the current work as a basis for our mutations, to probe the roles of each wobble in the context of an active site poised for cleavage. The resulting simulations indicate that, in such an activated ribozyme, each wobble plays a distinct set of structural roles to promote optimal catalysis. Our analyses showed that either Watson-Crick or wobble pairs at the 20-25 position disrupt retention of the general-acid C75(N3H3+)G1(O5') hydrogen bond, as well as retention of the G1U37 wobble (Fig. 3.4). These increases in active site dynamics relative to our Wildtype_Na simulations (Fig. 3.3) likely contribute strongly to the decreased in-line fitness we observe. Thus, in addition to roles in ion binding, our studies show that the *trans* G25U20 wobble acts to stabilize conformational dynamics in regions somewhat distal to the L3 loop. Our further mutations of the G1U37 wobble reveal that this motif acts more locally to stabilize, specifically, the general acid hydrogen bond and in-line fitness. These effects are in contrast to the farther-reaching influences of the G25U20 wobble; indeed, the *trans* wobble's propensity for chelated ion binding may be the mechanism by which it is able to effect comparatively

longer-range conformational changes in this particular biopolymer. Taken together, our results provide specific conformational functions for each type GU wobble in the HDV ribozyme active site. Although our data provide greater depth to the extensive body of knowledge regarding the necessity of GU wobbles and their structural variations in highly structured RNAs, further work must be conducted to study the effects of more subtle changes (such as *syn* versus *anti* bases) on the function of these RNAs.

Chapter 4

EVIDENCE THAT pKa VALUES OF NUCLEOBASES IN STRUCTURED RNAS CAN REFLECT THE ENVIRONMENT OF LESS POPULATED CONFORMATIONS [3]

4.1 Introduction

Compared to proteins, RNA and DNA contain a very limited array of functional groups.[97] However, several studies have displayed the possibility of both RNA and DNA to harbor protonated versions of adenine and cytosine bases as necessary elements in structural stability, enabling a larger variety of functionalities in these biomolecules.[42, 93, 161-166] Such structurally important features are, as a result, highly sensitive to pH, and can potentially serve a variety of roles besides simply structural support. These roles include electrophiles and general acid-base catalysts,[98] and result in shifted pK_a values of the respective bases. The tertiary structures of many structured RNAs do consequently provide pockets of negative electrostatic potential[98] or highly stable structure[167] with the potential to shift nucleobase pK_a's. For example, A25 within the active site of the leadzyme exhibits such a shifted pK_a that it is found in a stable base pair with C6 in the structured-loop active site of this ribozyme (Fig. 4.1a). [168, 169] Additionally, the pH-dependent retention of an A79+C67 wobble within a small internal bulge in spliceosomal U6 RNA is coupled to the conformation of an adjacent U80, shown to be crucial for metal binding at this site (Fig. 4.1b).[170] Protonation of cytosine

[3] Kamali Sripathi performed partial charge calculations. Mr. Garrett Goh translated AMBER force fields for the CHARMM framework and performed interaction energy and pK_a calculations. Mr. Alex Dickson formulated the WEXPLORE method of enhanced sampling.

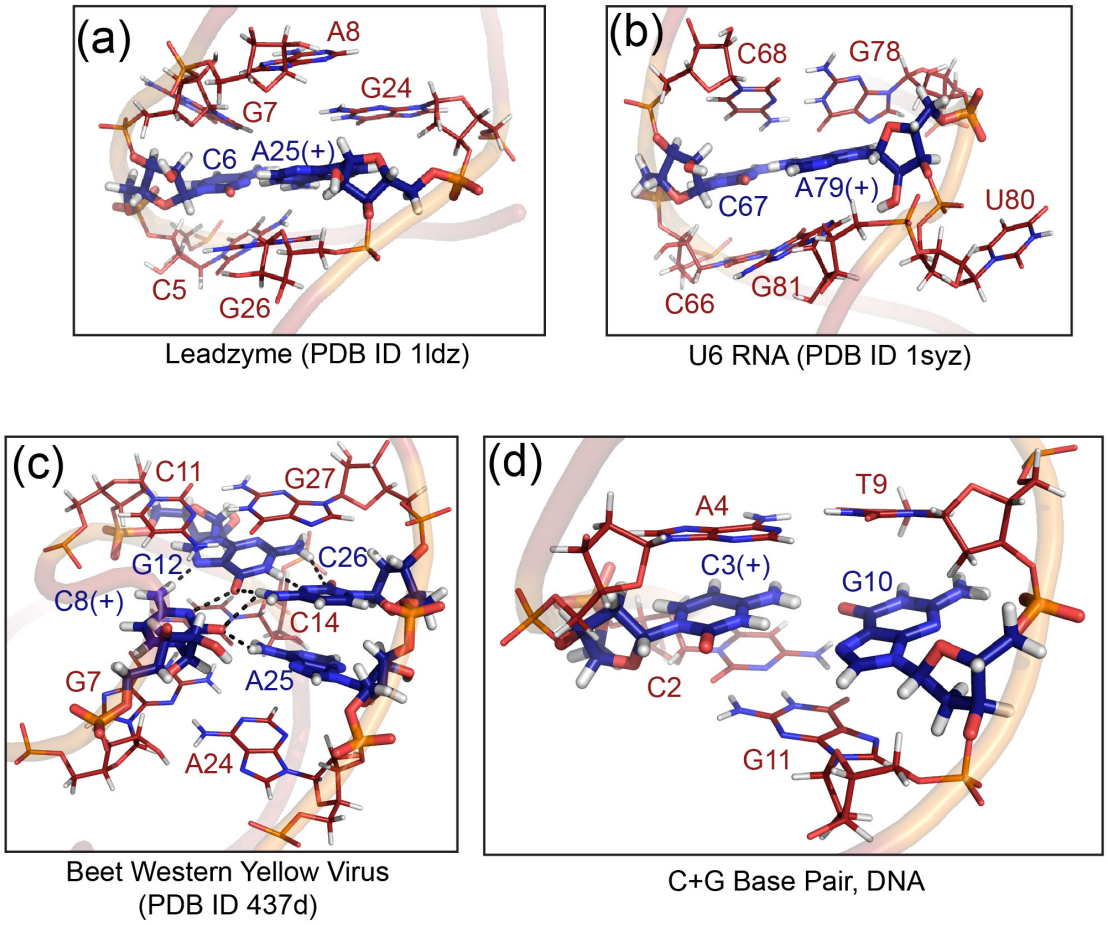

Figure 4.1 Highly structured nucleic acids that provide evidence for residues with shifted pK_a values.

Panels A and B depict structures containing protonated adenines, while panels C and D depict structures with prontated cytosines. Residues with putative protonated are denoted with "(+)"; the carbon atoms of these residues and any others interacting in relevant motifs. Neighboring residues are denoted in dark red. Oxygens are in red, nitrogens are in blue, hydrogens are in white, and phosphorus are in orange.

is also likely favored in a rare quadruple interaction in Beet Western Yellow Virus RNA (Fig. 4.3c)[171] and, to a lesser extent, in GC Hoogsteen pairs in DNA.[172, 173] (Fig. 4.1d)

A particularly interesting example of nucleobases with shifted pK_a's can be found in the hairpin ribozyme (Fig. 4.2). The hairpin ribozyme is found in the negative strand of tobacco ringspot virus satellite RNA and similar viroids,[37] and is a member of the family of the small catalytic RNAs.[7, 9] Like the other naturally occurring catalytic RNAs, the hairpin ribozyme catalyzes cleavage of its own phosphodiester backbone, resulting in 2'3' cyclic phosphate and 5' hydroxyl termini. Found embedded in a 4-way junction (4WJ) in nature, the ribozyme retains comparable activity in 2-way junction and junction-less forms.[38, 39] Catalysis occurs when the substrate binds and the internal loops dock (loops A and B, Fig. 4.2a), forming a series of complex docking interactions that envelop the cleavage site.[37] Active site residues in both loops A and B appear to contribute to cleavage of the phosphodiester bond between A-1 and G+1 on the 5' arm of loop A (Fig. 4.2a, blue arrow). Crystal structures have shown that the active site adopts similar conformations regardless of the type of junction.[13-16]

Studies on the metal- and pH-dependence of the hairpin ribozyme indicate that cleavage occurs via general acid-base catalysis, implying shifts in the pK_a's of active site nucleobases.[51, 52] However, the specific players and their roles are still unclear. A38 of loop B and G8 of loop A appear to be essential for catalysis,[24, 29, 31, 45, 174] with A9 and A10 of loop A potentially playing lesser, supportive roles. Many studies have attempted to isolate microscopic pK_a's of A38 ad G8 to determine if they are sufficiently shifted to act as general acids or bases during catalysis.[24, 25, 29-31, 175] Initial mutagenesis studies by Fedor and coworkers indicated that A-like and G-like moieties are preferred at both of those positions,

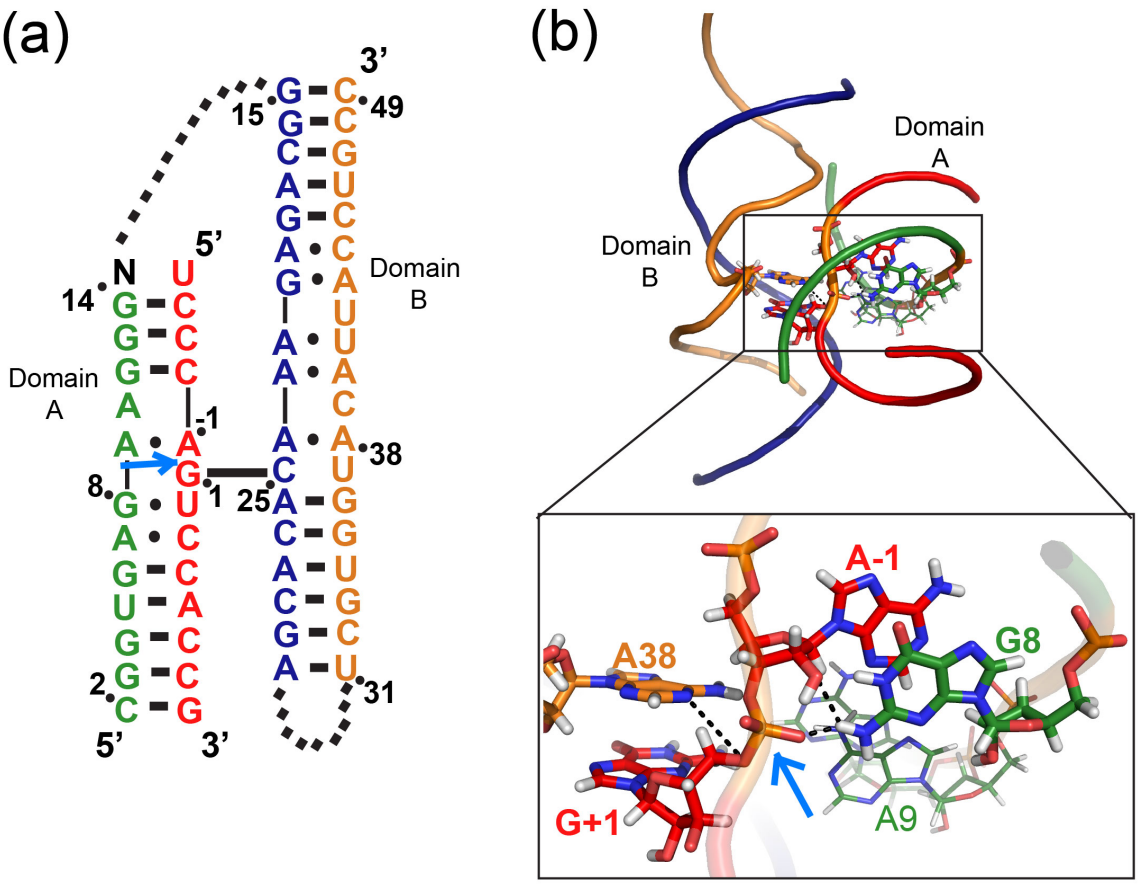

Figure 4.2 Secondary and tertiary structure of the hairpin ribozyme.

(a) Secondary structure of the junction-less hairpin ribozyme crystal structure used in our calculations, adapted from[16]. Separate strands are denoted with separate colors, and a bright blue arrow designates the cleavage site. (b) Tertiary structure and active site (inset) of the hairpin ribozyme. Colors are consistent with (a). Active site residues and relevant hydrogen bonding interactions (black dashed lines) are denoted in the inset. Carbons are colored consistent with (a), nitrogens are blue, oxygens are red, phosphorus is orange and hydrogens are white.

and may either play specific roles in general acid-base catalysis, or act indirectly as electrostatic stabilizers of developing negative charge.[24, 29] Raman crystallography was later used to measure the pK_a of A38 as ~ 5.46; such a value suggests protonation during catalysis and may imply participation in general acid-base chemistry, among other roles.[25] Several fluorescent studies also implied protonation of A38 and G8 during catalysis.[30, 175] Later molecular dynamics (MD) and quantum mechanical/molecular mechanical (QM/MM) hybrid studies provided atomic-level insight into multiple possible pathways,[32, 34, 56] such as the potential of A38[32] or nonbridging oxygens[34, 56] to act as proton shuttles.

Several methods have been employed to measure pK_a shifts, particularly in structured RNA. pH-rate profiles in a variety of systems are able to report the macroscopic pK_a's of chemical reactions; however there often exists more than one interpretation of this type of profile.[176] Thus, there is a great need for pK_a measurement at the microscopic level. NMR spectroscopy has proven a useful tool in probing the pK_a's of folded residues.[169] pH-dependent fluorescent analogues have also been developed to site-specifically probe changes in local electrostatics in a folded RNA.[30, 175, 177] Raman spectroscopy has also been used to detect shifts in pK_a of internal RNA residues.[13, 25, 90] Molecular dynamics simulations have been used to investigate conformational changes in the presence of potential ionizable groups and their conjugate acids or bases.[32, 49, 178] Such studies have provided insight into the conformationally allowed ionization state in each case, although conventional MD simulations are unable to track proton trajectories between various donors and acceptors. Despite the wealth of techniques available to probe active sites containing ionizable groups, none are able to thoroughly survey the occupancies of titratable groups within the active site, their effects on the pK_a's of other ionizable groups, and the resultant effects on local

structural integrity. This chapter presents work towards the calculation of microscopic pK_a's within the active site of hairpin ribozyme, using the recently developed constant pH molecular dynamics (CPHMD)[68, 70, 71] methods. Our tests, using the leadzyme as a model system, indicate that different phases, the quantum mechanical level of theory, and basis sets result in largely similar pK_a values. Our two compared methods of charge derivation applied to a variety of RNAs and DNAs result in pK_a values within 0.5 units of each other, but sometimes show significant variation from experimentally measured pK_a's. Further tests on the hairpin ribozyme using an enhanced sampling method demonstrate that this deviation likely reflects the presence of a minor population not accounted for in our initial calculations. Taken together, our results indicate that various methods of deriving partial charges result in consistent calculated pK_a values, and that experimentally measured pK_a's may often be the result of little-sampled molecular conformations.

4.2 Materials and Methods

4.2.1 Structure preparation and Simulation methods

The AMBER parmbsc0 force field[63] with χ corrections[67] was used for all simulations, and prepared for the CHARMM framework according to previous protocols.[179] Structures were prepared similar to earlier work.[68, 70] Appropriate patches were applied for terminal ends of the various RNA fragments consistent with AMBER nomenclature. The structure used for the hairpin ribozyme consisted of four strands, and the appropriate patches were applied to the 5' and 3' termini for each of those strands. The CHARMM suite of programs (version 36a6)[180] was used for all thermodynamic integration calculations and constant pH MD simulations. The simulation protocol was similar to that used by Goh *et al*.[68, 70] Hydrogen bond lengths were constrained by the SHAKE algorithm.[113] The Leapfrog Verlet algorithm

was implemented, utilizing a 2 fs time step of integration. 12 Å was the nonbonded cutoff, and an electrostatic force switching function and a van der Waals switching function between 10-12 Å were used. λ dynamics were enabled in our simulations *via* the CHARMM BLOCK facility, using the MSλD framework. The functional form of λ was defined as $\lambda^{N_{\exp}}$, [71] enabling titration of multiple sites on the RNA. All energy terms except bonds, angles, and dihedrals were linearly scaled by λ. These three exceptions were allowed to retain full strength to preserve structural geometry, regardless of λ value. A threshold of $\lambda \geq 0.8$ was used to define a fully protonated state ($\lambda = 1$). The hybrid potential energy function was modified by addition of the variable λ-dependent bias, F^{var}, and the associate force constant (k_{bias}) was adjusted to increase sampling of the two protonation states at and above the λ threshold. A Langevin heat bath was coupled to the system to maintain a temperature of 298 K, with a frictional coefficient of 10 ps^{-1}. The structures were minimized for 100 steps of steepest descent, then for 1000 steps of adopted basis Newton-Raphson. The systems were then heated for 1ns, equilibrated of 1ns, and production runs were conducted for 50ns for each structure.

4.2.2 Analysis

Our protocol enabled us to estimate the fraction of unprotonated species (S^{unprot}) [68, 70] based on the populations of protonated (N^{prot}) and unprotonated (N^{unprot}) species. These were taken from three independent simulation runs started from unique seed values. Specific S^{unprot} fractions were calculated at each pH tested, and then combined using the following equation:

$$S^{unprot}(pH) = \frac{1}{1+10^{-n(pH-pK_a)}} \quad [68, 70]$$

where *n* represents the Hill coefficient. Unless reported otherwise, pK_a values and their errors were computed from three separate runs of 5 ns each.

4.2.3 Charge derivation methods

Two methods were compared to calculate partial atomic charges for N9-methyladenine and N1-methylcytosine in the tests in section 4.3.2 (Tables 4.3,4.4). Both methods began with geometry optimization using Gaussian 09,[181] and the RESP method of charge derivation.[116, 182, 183] Method 1 utilized the R.E.D. Server,[184, 185] which automates geometry optimization (by direct connection to Gaussian 09[181]) and RESP fitting.[182] The R.E.D. Server also allows the user to conduct charge derivations on multiple conformations and/or orientations of the same molecule. Method 2 involves calculating geometry optimization and ESP charge derivation manually using Gaussian 09, and calculating RESP charges using the Antechamber suite of programs.[186] Method 1 was employed for the tests conducted in sections 4.3.1, and methods 1 and 2 were compared in section 4.3.2 (see below).

4.3 Results

MD simulations have been used over the last several decades for their ability to probe atomic-level, sub-microsecond molecular movements. Until recently, however, their ability has been limited to investigating fixed protonation states of systems with very dynamic internal conformational changes. We applied recently developed constant pH MD with multisite λ dynamics ($CPHMD^{MS\lambda D}$)[68, 70, 71] using the CHARMM suite of programs[180] to our studies on microscopic pK_a's within structured RNAs. Within the CHARMM framework, we implemented the AMBER parmbsc0 force field[63] with recent χ corrections.[67]

4.3.1 Different basis sets and levels of theory give largely similar results

Because A38 is a recognized position of titration within the hairpin ribozyme, we began calculations focused on titratable adenine. A wide variety of basis sets, levels of theory, and types of solvent were available to us. We investigated three main quantum mechanical levels of theory, and a variety of basis sets in several solvents (Table 4.1). In order to compare effects in a relatively simple system, we calculated the interaction energy of our protonated model adenine with cytosine in the A25+ and C6 of the leadzyme[168] (Table 4.1). We found that the interaction energies were fairly consistent across different solvents. As such, we chose the gas phase as the base for our calculations, in keeping with previous AMBER parameterizations.[114, 116, 182]

Having defined our solvent requirement for parameterization, we calculated pK_a values in the leadzyme using different basis sets and levels of theory in the gas phase (Table 4.2). Our results gave us a standard deviation of ±1 pK_a unit, within the error reported for previous studies.[68, 70] Because our pK_a calculations were relatively consistent with each other, we decided to implement the Hartree-Fock (HF) level of theory and the 6-31G* basis set as the standard for our calculations. This choice is again consistent with the standard parameterization protocol of AMBER.[114, 116, 182]

		Gas	Ether	Water
Hartree-Fock	**6-31G***	-39.86299	-41.99117	-43.26515
	6-31+G*	-40.67655	-42.72306	-43.86291
	cc-pVDZ	-38.99071	-41.6234	-43.36223
	aug-cc-pVTZ	-40.46561	-41.99591	-43.29981
	cc-pVTZ	-38.91571	-41.77309	-41.9168
	aug-cc-pVTZ	-38.80471	-41.79766	-42.54641
B3LYP	**6-31G***	-38.47356	-40.73211	-41.87123
	6-31+G*	-38.79399	-41.31451	-42.57002
	cc-pVDZ	-38.16056	-39.74756	-40.81446
	aug-cc-pVTZ	-38.21353	-40.45188	-42.13833
	cc-pVTZ	-37.79251	-40.2338	-41.23396
	aug-cc-pVTZ	-37.70566	-40.47112	-42.07257
MP2	**6-31G***	-39.86299	-41.99117	-43.26515
	6-31+G*	-40.67655	-42.72306	-43.86291
	aug-cc-pVTZ	-40.46561	-41.99591	-43.29981
	cc-pVTZ	-38.91571	-41.77309	-41.9168
	aug-cc-pVTZ	-38.80471	-41.79766	-42.54641

Table 4.1 Interaction energies for various basis sets, levels of theory and solvents for the A25+C6 base pair in the leadzyme

		pK_a
Hartree-Fock	**6-31G***	6.7
	6-31+G*	6.9
	cc-pVDZ	7.5
	aug-cc-pVTZ	6.5
	cc-pVTZ	6.2
	aug-cc-pVTZ	6.4
B3LYP	**6-31G***	6.8
	6-31+G*	7.0
	cc-pVDZ	6.1
	aug-cc-pVTZ	6.6
	cc-pVTZ	6.7
	aug-cc-pVTZ	6.8
MP2	**6-31G***	6.8
	6-31+G*	6.9
	aug-cc-pVTZ	7.2
	cc-pVTZ	6.6
	aug-cc-pVTZ	7.4

Table 4.2 pK_a values for gas phase calculations testing different basis sets and levels of theory in the leadzyme

4.3.2 Two related methods of charge derivation yield comparable pK_a values

After determining our criteria for quantum mechanical calculations, we compared pK_a values of protonated adenines in our model leadzyme and the hairpin ribozyme. Our leadzyme pK_a value was within 0.5 pK_a units of experimental measurements, consistent with previous results (Table 4.3).[70] However, our results in the hairpin ribozyme gave pK_a's shifted ~ 3 units from experimental values.[24, 52, 175] We thus employed another established method of charge derivation within the AMBER community, method 2,[181, 182, 184, 185] to test if our protocol could be the source of the discrepancy. Method 2 is a more manual approach, compared to method 1, in which the user generates inputs for Gaussian 09 and is able to specify the type of charge derivation conducted by the Antechamber[186] module of AMBER (see Materials and Methods). This method gave pK_a results within 0.5 units of method 1 (Table 4.3).

To determine whether our charge methods' behavior in either the leadzyme or the hairpin ribozyme were more typical, we calculated pK_a values for adenine in a A+C wobble pair in the U6 RNA[170] system using both methods of charge derivation (Table 4.3). Results for adenine in this system consistently overestimated pK_a values, similar to our values for the hairpin ribozyme. However, the U6 RNA experiences a pH-dependent conformational change: the U80 base adjacent to the A79+C67 base pair "flips out" of stacking with A79 at a pH lower than 7.0 (Fig. 4.1b), while stacking with A79 at pH 7.0.[170] We did not take into account the U80-stacked conformation in our calculations, which may partially explain the shifted pK_a we observed.

To further assess the reproducibility of our charge methods on other systems, we also evaluated both charging methods on protonated cytosine in different compounds (Table 4.4).

pK_a	Leadzyme (exp)	Leadzyme (calc)	Hairpin Rz (exp)	Hairpin Rz (calc)	U6 (exp)	U6 (calc)
Method 1	6.5	6.9	6.5	9.5	6.5	8.1
Method 2	6.5	6.6	6.5	9.7	6.5	9.9

Table 4.3 Experimental and calculated pK_a values for both methods of partial charge derivation for protonated adenine in a variety of systems.

pK_a	BWYV (exp)	BWYV(calc)	HG(exp)	HG(calc)
Method 1	8.2	13.5	7.2	6.4
Method 2	8.2	13.1	7.2	6.1

Table 4.4 Experimental and calculated pK_a values for both methods of partial charge derivation for protonated cytosine in a variety of systems.

Results were less consistent for the cytosine, with pK_a's overestimated in a C+AGC quadruple interaction in the beet western yellow virus (BWYV) RNA,[171] and within one pK_a unit of a protonated C+G base pair in an idealized DNA double helix[172, 173] (Fig. 4.1c,d and Table 4.4). The BWYV RNA exhibited a much more significant shift of calculated pK_a than did the idealized C+G base pair.

4.3.3 Minor molecular conformations in simulations may explain experimentally measured pK_a values

There are several possible explanations for the trends we observed in the previous section. One reason may be that the various structural environments each ionizable base experiences; recent experimental results have shown that immediate base pairing environments can have a substantial effect on a nucleobase's pK_a value. [167] A second reason may be, based on the evidence for multiple conformations in the U6 RNA,[170] that our current sampling methods do not take into account alternative conformations that may affect a nucleobase's electrostatic environment. Since previous CPHMD simulations on the leadzyme had exhibited similar trends within a single simulation,[70] we tested for the second possibility at pH 7 using the WEXPLORE method (protocol not yet published) developed in the Brooks group. We used an order parameter of A38(N1)…G+1(P) distance (Figure 4.3(a), insets). The major population from WEXPLORE sampling exhibited a A38(N1)…G+1(P) distance of ~ 5 Å (Fig. 4.3a, left inset), similar that seen in the starting structure. A smaller population was discovered using this method, in which the A38(N1)…G+1(P) distance was ~ 9 Å (Fig. 4.3a, right inset), which would greatly increase the acidity of the A38 N1 nitrogen. Consequently, in pK_a calculations of 50ns involving this minor-population structure, we found the pK_a of A38 to decrease to ~ 7 compared to values of 9.5 in simulations without this

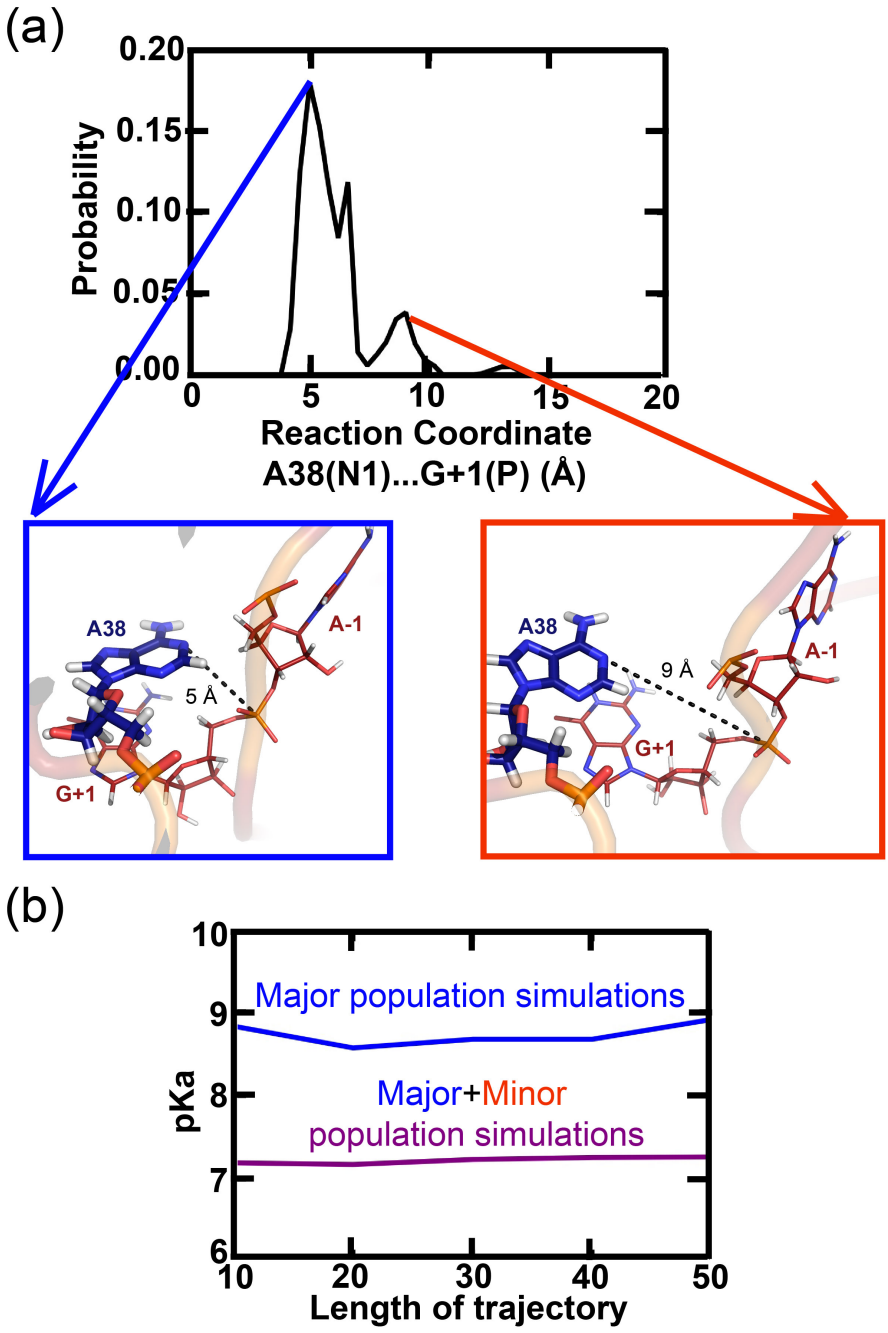

Figure 4.3 The effect of minor conformations on pK_a calculations

(A) Results of enhanced sampling on conformations of the hairpin ribozyme reveal a crystal-like population (left inset) and a minor population with increased reaction coordinate distance. (B) Comparisons of pK_a values from simulations utilizing only major conformaers (blue) and from simulations utilizing a combination of major and minor conformers (purple).

structure. A pK_a of 7 is much more consistent with experimental measurements, and may indicate a trend of experimentally determined pK_a . Our results thus suggest that - experimentally measured pK_a's reflect minor populations that are not always represented by crystal structures.

4.4 Discussion

The ability of the highly negatively charged environment within an RNA to shift pK_a's of buried residues, and the ability of these residues to participate in general acid-base catalysis similar to their amino acid counterparts, has been studied for over a decade.[20, 21, 98] Extensive biophysical and computational characterization of titratable residues in the leadzyme has been conducted,[70, 187] making it an ideal model system. The catalytic mechanism of the hairpin ribozyme has been of particular interest as well, due to early evidence that it does not require metal ions for catalysis,[51, 52, 188] and for the later-revealed conformational similarities of its active site to that of RNase A.[14, 15, 37, 140] Several biochemical studies have indicated the necessity for protonated A38 and G8 residues in the active site,[24, 25, 28-31, 175] and multiple attempts have been made to characterize the microscopic pK_a of these residues to identify their roles in catalysis. [25, 30, 175]

Here we present preliminary work to answer this question using constant pH MD simulations with multisite λ dynamics ($CPHMD^{MS\lambda D}$).[68-71] Our tests reveal that solvent types, basis set and level of theory have little effect on interaction energies of protonated adenine with its environment, while different levels of theory and basis sets only result in a standard deviation of at most ±1 pK_a unit in our gas-phase calculations (Tables 4.1, 4.2). Our calculations of A25 and A38 pK_a in the leadzyme and hairpin ribozyme precursor, respectively, demonstrate values consistent with experiment for the leadzyme,[70] but shifted

significantly into the basic range for the hairpin ribozyme (Table 4.3). This discrepancy is observed in our data regardless of partial charge derivation method, and in a variety of systems containing protonated adenine or cytosine (Tables 4.3, 4.4).

One explanation for our observed trends may be due to recent studies reflecting the effects of environment on base pK_a values.[167] These studies found that more stable base pairing above and below A+C wobble pairs resulted in greater shifts in pK_a values, presumably because they stabilize the stacked (loop-internal) conformation of the wobble. The leadzyme exhibits one of the lowest measured pK_a's in our study, and contains one stable C5G26 base pair below, and potential G28-G7 or G28-A8 mismatches above, its A+C wobble (Fig. 4.3a). Conversely, the U6 RNA has a significantly shifted measured pK_a, and in the conformation used in our calculations is bounded by two very stable GC base pairs (Fig. 4.3b). If our calculations were able to incorporate the U80-stacked conformation, we predict that our measured pK_a would likely shift back towards neutrality due to disruption of the A79+C67 base pair by the stacking of A79 with U80. Our pK_a values for the hairpin ribozyme may similarly be shifted past neutrality due to A38's strong interactions with the scissile phosphate, and with its stable stacking with G+1 below (Fig. 4.4a, left inset). Consequently, the BWYV RNA with potentially protonated C8 exhibits the most shifted pK_a of all, which can also be explained by its environment: C8 itself is involved in a set of quadruple interactions with nearby A25, C26, and G12 (Fig. 4.3c). In addition to this strong driving force for protonation, the C8 quadruple interaction is bounded above and below by two stable GC base pairs, explaining the drastic shift towards basicity we observe in C8's pK_a. Likewise, the idealized DNA helix's C3 may experience reduced shifts in pK_a due to only partial stabilization of a CG base pair below and an AT base pair above it (Fig 4.3d).

Additional reasons for these inconsistent trends may be that our sampling methods do not capture all conformations with the potential to shift the pK_a of the nucleobase in question, as may be the case with the U6 RNA (Fig. 4.3b). As such, we used the recently developed WEXPLORE method to look for other conformations at pH 7 that may contribute to the experimentally measured pKa. Our enhanced sampling methods found a minor population in which A38 is removed by ~9 Å from the scissile phosphate (Fig. 4.4a, right inset), resulting in a significantly more acidic pK_a than our initial calculations showed (Fig. 4.4b). This minor population explains the discrepancy we observe between our calculated pK_a and the experimental pK_a for the hairpin ribozyme. These results also provide additional insight into differences between structural and experimental data. The minor population outlined by the WEXPLORE sampling method provides an alternate structure from crystal geometry that may be more favored in solution studies, and thus dominate determination of experimental pK_a. Thus, experimental pK_a measurements may often, but not always, be dominated by a structure that represents a conformationally minor population in the context of structural studies. Taken together, our results show that calculated pK_a values in complex nucleic acids follow experimental trends with sufficient sampling. Additionally, we reveal that experimental pK_a measurements may be dominated by minor populations sampled in simulations. Our work thus provides a basis for exploring the behavior of A38 and other moieties in the precleaved, transition state and product forms of the hairpin ribozyme.

Chapter 5

SUMMARY AND FUTURE DIRECTIONS

5.1 Summary

The small catalytic RNAs accomplish a single common catalytic goal, site-specific autolytic cleavage, using diverse 3-dimensional folds. Measurements of their activity, as indications of correctly folded molecules, have long been used to further elucidate structure-function relationships in structured RNAs, [122] while their site-specific catalysis and propensity to be functionally refined have made them attractive starting points for the development of therapeutic agents.[77] As understanding for catalytic requisites in small ribozymes increased, a strong connection between global conformational dynamics and catalysis was also observed.[57-59, 80, 99, 126] Of the five members of this group, this thesis focuses on the structure-function relationships of two that have been shown to employ general acid-base catalysis: the hepatitis delta virus (HDV) and hairpin ribozymes. The work in this thesis provides deeper insight into these ribozymes' codependent and synergistic structural interactions, and the implications of these interactions on the mechanistic details of these complex molecules.

5.2 Heterogeneity as a real-world reflection of the free energy folding landscapes of riboymes

The folding free energy landscape of RNA is much more uneven and frustrated compared to that of proteins.[122] This is due to several unique features of RNA. Its backbone is highly flexible compared to a protein's peptide bonds, consisting as it does of six dihedral angles per

phosphodiester linkage.[189] Additionally, because base pairing is not stringently sequence-dependent (for example, an A can bind all U's within an RNA regardless of their surrounding sequences) and duplexes are highly stable, misfolding of the RNA is more often experienced compared to that of proteins.[122] These attributes result in a relatively more rugged folding landscape for structured RNAs, which manifests in *in vitro* populations exhibiting significant structural and functional heterogeneity.[122] Experimental and computational techniques can be used independently or in concert to further understand RNA folding landscapes.[80, 190, 191] The results in Chapter 2 outline a unified folding landscape (Fig. 2.9, black line) in which all existing HDV ribozyme crystal structures reflect the inherent heterogeneity of this structured RNA. This putative set of pathways is derived based on our experimental and MD simulation results. Our FRET results reveal heterogeneous populations present in ground-state precleaved ribozymes, all of which are able to undergo cleavage. MD simulations allow us to probe this heterogeneity and rationalize the different behaviors we observe. *Cis*-acting model ribozymes reveal greater active-site dynamics compared to our *trans*-acting simulations, indicating a lower energy barrier between active and inactive conformations on the folding pathway for the former. Our *trans*-acting precursor simulations exhibit a catalytically fit active site when C75 is kept protonated. This catalytically competent conformation represents a further wrinkle in the HDV ribozyme folding landscape centered around cleavage; the conformation containing a protonated C75, a minor population at physiological pH, is nonetheless required for optimal cleavage. A magnesium ion is required to bind close to the protonated C75 almost simultaneously for cleavage to occur. The synchronization of these opposing catalytic details requires time and energy at physiological pH, thus explaining the comparatively slow cleavage of the HDV ribozyme.

The model defined in Chapter 2 is, however, limited to existing structural data. In the future, enhanced sampling methods, such as umbrella sampling, may be used between the known HDV ribozyme intermediates to sample potential transition states, as well as other intermediates, and provide deeper insight into the folding landscape as related to this complex RNA's catalysis. Additional investigations could build upon the extensive hybrid QM/MM studies on ground-state precursor conformations of the HDV ribozyme.[19, 64, 192, 193] I also propose that constant pH MD (CPHMD) simulations similar to those conducted in Chapter 4 for the hairpin ribozyme may provide further mechanistic insight into how each intermediate in our free energy landscape fosters functional forms of C75 and the U-1 2'hydroxyl, respectively. These CPHMD simulations may further reveal heterogeneous populations for each of these intermediates, thus augmenting this local portion of the HDV ribozyme folding free energy landscape.

Several earlier studies have investigated heterogeneity of the HDV ribozyme with a variety of 3' and 5' overhangs of assorted lengths.[194-197] These studies outlined the context-dependent inhibition or enhancement of cleavage observed in each of these cases, further highlighting the ruggedness of the HDV ribozyme free folding landscape. Early studies by Perrotta and Been[196] defined an alternative pairing interaction involving the P2 helix. The effect of this alternate pairing, P2a, on cleavage depended on the concentration of monovalent salt in which the ribozyme was incubated prior to the addition of magnesium in assays. These results led to the conclusion that conditions of high and low sodium could mimic a co-transcriptional event that could act as an on-off switch during ribozyme folding,[196] reflecting further furrows the free energy landscape. Later studies by the Bevilacqua group investigated extensive addition of sequences up- and downstream of the

ribozyme's minimal cleavage motif.[194, 195, 197] Their characterizations of upstream cleavage sequences with the ability to inhibit and accelerate cleavage,[195] as well as the effects of downstream sequences to foster alternative secondary structure motifs both within and outside of the ribozyme[194, 197] further outlined various potential folding intermediates and kinetic traps within the HDV RNA in *in-vivo*-like conditions. These previous studies provided a broad-scale look at the folding free energy landscape of HDV RNA beyond the minimal ribozyme sequence. Our work, in contrast, focuses on a much more specific portion of the free energy folding pathway: in using a more product-like ribozyme, we focus on a structure that has passed many of the kinetic traps discussed above. Our combined approach further enables us to directly probe the folding pathways adopted by a ribozyme primed for cleavage.

Our studies present the first set of results documenting heterogeneity in an activated version of this RNA, and the resulting model contributes another example of the demonstrated heterogeneity and consequently rugged folding free energy landscape of structured RNAs. The *Tetrahymena* group I intron exhibits conformational heterogeneity that interestingly does not affect its activity.[122] These conformations, classified by their varying docking free energies, are shown to redistribute after de- and renaturation. Because of the few kcal/mol difference in free energy for the various conformers and the fact that they share the same cleavage kinetic, all of these conformations may be globally closely related similar to the smear we see for our S1-dUGG ribozyme (Fig. 2.3b, 2.3). In contrast, studies on the conformational heterogeneity of the hairpin ribozyme indicate higher free energy barriers between subpopulations.[122] Evidence for such increased barriers lie in the demonstrated "molecular memory" of ribozymes over long incubations, and the persistent formation of

faster- and slower-docking species even after denaturation. Our results for the HDV ribozyme thus enhance the current understanding of molecular heterogeneities in structured RNA by providing an example of closely related conformers of a ribozyme poised for catalysis. Despite our focus on a comparatively narrow portion of the HDV ribozyme free energy folding landscape, we still find our related populations to exhibit significant heterogeneity (Fig. 2.1c,2.2). Thus, as previously summarized,[122] such structural variation may not be simply be a troublesome aspect of working with RNAs, but rather a necessary hallmark of a biologically fit population.

5.3 Extended roles and selection pressures on GU wobbles in structured RNAs

Since the earliest crystal structure of the hammerhead ribozyme, a complex network of noncovalent interactions within small catalytic RNAs has been observed.[18, 59, 141] Such networks allow a certain amount of conformational flexibility, which may be one reason for the closely related global conformers we observe in Chapter 2. These networks may also be one way in which ribozyme active sites achieve catalytic efficiency only 1-2 orders of magnitude below that of RNase A.[198] In Chapter 3, we probe the intricate hydrogen bonding network of the HDV ribozyme active site by studying structural ramifications of two crucially placed GU wobble pairs. Our results demonstrate that these wobble pairs each play specific roles to support against local disorder. The *cis* or standard G1U37 wobble adjacent to the cleavage site (Fig. 3.1) acts locally to stabilize in-line fitness and optimal geometry between catalytic C75 and the scissile phosphate, while the *trans* or reverse G25U20 wobble is additionally able to stabilize the G1U37 wobble. Furthermore, the more proximal type I A-minor and P1.1 interactions of the HDV ribozyme's global network act as a resilient barrier against active site disturbance, also preventing global structural degradation.

Although Chapter 3 reveals the ability HDV ribozyme motifs to retain global structure under pressure from active site dynamics, further studies can be conducted to test the stability of these motifs against further internal and external disturbances. For example, it is still unclear how or if the conformation of G25 affects active site fitness. Further MD simulations can be conducted comparing in-line fitness of precursor structures, as well as stability of active-site motifs and changes in electrostatic potential for all ribozyme structures. To test the robustness of the active site core to external disorder, active site geometry can be monitored in the presence of mutations that weaken the type I A-minor interactions. Similar to previous studies,[57, 58] footprinting experiments of these mutants can be conducted in parallel to confirm the extent to which disruption of A-minor interactions interrupts the native protection of active site residues. To supplement previous mutational studies on the P1.1 helix, [85] footprinting experiments and MD simulations could be conducted on similar mutants to monitor active site stability in the event of these types of perturbations. Studies on the P1.1 helix would have additional repercussions for understanding cleavage mechanisms of other members of the HDV-like ribozyme family, since a 2-base-pair P1.1 helix is not always a requirement.[73, 75]

GU wobble pairs are found in a wide variety of contexts, often in structurally crucial positions.[145] Besides the HDV ribozyme, GU wobbles can also be found in rRNA, group I introns and often constitute the well-known third "wobble" position during codon-anticodon pairing in protein translation.[145, 156] The *cis* GU wobble's ubiquity in a variety of contexts is likely due to its structural and energetic similarity to canonical Watson-Crick base pairs; however, its distinct structural and chemical differences often play crucial roles in such biological processes as induced fit, and ion and water binding.[144, 145, 156] Moreover, the

wobble's distinct array of functional groups allows it to participate in a variety of stable three stranded motifs. These motifs and their phylogenetic variation have been assayed using a combination of bioinformatics and MD simulation studies.[199] The authors found that GU wobbles in rRNA from a variety of organisms were more highly conserved when the wobbles were involved in P-interactions (sugar-edge interactions between the given GU wobble and a nearby base pair) and some interactions between the GU wobble and nearby phosphates.[199] Both of the HDV ribozyme's GU wobbles are relatively invariant among isolates, with a G1C37 Watson-Crick pair as the most common mutation,[200] but they do not appear to mimic this trend, despite a DGNP (Deep Groove Not in Pocket) interaction between the G1U37 wobble and nearby G76 (snapshot panels, Fig. 3.5-3.7). This conservation despite lack of selective tertiary structure may be reflective of the highly structured, somewhat degenerative nature of the HDV ribozyme active site. The G1U37's presence at the juncture of the P1/P1.1 coaxial stack may be sufficient selective pressure for the high retention of this wobble in isolates.[200] The stable GC base pairs above and below may compensate for a lack of consistent tertiary structure formation. (Fig 3.1a) Likewise, rather than any stabilizing interactions involving base pairs, perhaps the ions bound by the L3 loop [10, 49, 94, 144] act as an evolutionary conservation mechanism for the G25U20 *trans* wobble. The magnesium in our relevant simulations accordingly shows high occupancy near the scissile phosphate (Fig. 3.2d, right), in a similar position to phosphate or 2' hydroxyl groups in SGP (Shallow Groove in Pocket) interactions.[199] Thus, our work in Chapter 3 on the activated structure of the HDV ribozyme expands the rationale for phylogenetic retention of GU wobbles within structured RNAs. These redundant selection pressures are likely particularly true for HDV isolates as

well as HDV-like ribozymes,[73, 75] but may also have implications for all highly structured RNAs.

5.4 Effects of stable motifs on local pK_a values

The comparative stability of RNA versus protein secondary structure depends strongly on the combined effects of numerous hydrogen bonds employed by such common nucleic acid motifs as double helices.[122] Despite their weakness compared to other noncovalent interactions, intricate networks of hydrogen bonds can synergistically form stable tertiary motifs beyond simple Watson-Crick base pairs, such as A-minor, kissing-loop, pseudoknot and triple helix interactions.[201] These motifs form the basis for the complex global architectures of highly structured RNA molecules, enabling them to bind ions for catalytic and structural purposes, and to site-specifically interact with other biomolecules. A result of the successful tertiary folding of RNA is the configuration of a varying surface of electrostatic potential along the molecule. Such a surface contains pockets of high negative electrostatic potential with the ability to shift the pK_a's of nucleotides, resulting in noncanonical protonation states that add diversity to RNA's catalytic and structural capabilities.

In Chapter 4, we revealed the ability of adjacent secondary and tertiary structure to affect the pK_a of nearby titratable nucleobases. We find that, consistent with previous experimental studies, our systems with protonated adenine or cytosine exhibit shifted pK_a's when particularly stable secondary and tertiary structures flank the ionizable group. Indeed, the pK_a values from our simulations seem to positively correlate with increasing extent of structure (Tables 4.3,4.4). Both structures consisting of a single duplex (protonated adenine in the leadzyme, and protonated C in the idealized DNA helix) show moderate shifts in the

pK_a values of their protonated bases such that our calculations reflect experimental values. Although the proximity of the base pairing partner in either case provides incentive for the proton to remain at the site of ionization, the lack of tertiary constraints likely prevent these two pK_a values from shifting beyond neutrality. The next lowest pK_a we calculate is that of the U6 RNA intramolecular stem-loop (ISL), with a value of 8.1 – 9.9, much shifted from experimental values.[202, 203] This can likely be explained by the pH-sensitive dynamics of adjacent U80, however. At pH values < 6, A79 is typically found in a dynamic base pair with C67, a conformation in which U80 is flipped out of the helical plane, and the base-paired regions form a canonical A-type RNA duplex (Fig.4.1b).[170, 203] At pH > 6, A79 is largely unprotonated, and U80 is able to base pair with C67.[203] Because A79C67 base pairing is disfavored at higher pH, the pK_a of A79 is likely lower than it would be were A79 to participate in the base pair. Our simulations only included the former structure solved at acidic pH[170] in our simulations however; this structure contains stable CG base pairs directly adjacent to the internal loop. Such seemingly simple constraints have been experimentally proven to strongly affect pK_a shifts of ionizable bases, and likely explain the elevated pK_a we observe for A79. Additionally, U80 is shown to be reasonably dynamic when extruded from the resulting helix,[170] and its dynamics adjacent to A79 may act as an additional structural pressure for A79 to maintain the wobble pair. We predict that if we were to incorporate the higher pH structure[203] into our calculations, the pK_a values for A79 would more closely reflect experimental measurements. Our highest measured pK_a is for C8 in a quadruple interaction in the frameshifting pseudoknot of the beet western yellow virus.[171] C8 in this structure is resolved in an elaborate quadruple interaction (Fig 4.1c). The stable hydrogen bonds and the proximity of G12's O6 moiety likely result in the highly stable proton at

C8(N3). The experimental pK_a of C8 (Table 4.4) reflects such a stability of the quadruple, although not to the same extent as our calculated pK_a. This discrepancy may be due to an overstabilization of stacking as a result of the AMBER force fields, but our trends remain consistent with those observed in experiment.

Consequently, the flanking motifs of ionizable A38 in the hairpin ribozyme active site explain the high pK_a we observe for this residue. Our enhanced sampling methods indicate, however, that an alternative conformation is possible in which A38 is further removed from the scissile phosphate. The presence of this minor population appears to dominate experimental trends, while the higher pK_a is more readily predicted based on simulations on available precursor structures. Currently, pK_a calculations are being conducted using various precursor, transition-state-mimic and product structures to see if similar minor populations occur along key steps in the hairpin ribozyme cleavage pathway. Our results in Chapter 4 thus provide further rationalization for pK_a shifts in a variety of structured RNAs. Thus, these simulations provide the ability to test the effects of structure on functionally important factors such as shifted pK_a values of key residues.

In the last several decades, ribozymes have been shown to display a structural diversity and catalytic range previously unheard of in nucleic acids.[5, 6, 9, 143] The power of small catalytic RNAs has been crucial for the development of novel therapeutic agents, enabled largely due to the increase of our in-depth understanding of these molecules' structure-function relationships.[77] The results and future experiments outlined in this thesis provide new insights on RNA structure-function relationships by underscoring the multiple layers of structural complexity that form the basis for these interactions. Such complexity is

likely present in all highly structured RNAs, and must be taken into account in order to obtain a thorough understanding of these molecules.

Appendix A

SCRIPTS USED FOR ANALYZING DATA IN CHAPTERS 2 AND 3

For scripts, comments (not read by the programs) are denoted by "#". For usage and syntax, see the resource denoted at the beginning of every script.

A.1 20ns_trajectory.ptraj

```
#For usage and syntax, see the AMBER manual
#This is a script to generate stripped trajectories from raw, hydrated
#production files

#Steps: 1) Call in coordinate (*.crd, *.rst) and raw trajectory
#(*.traj.gz) files as required
trajin test.crd
trajin ../eq1.rst
trajin ../eq05.rst
trajin ../production0.rst
trajin ../production1.traj.gz
#...Until all *.traj.gz files are loaded

#Step 2: Specify output stripped trajectory (change "xxxx.traj" to
#desired name
trajout xxxx.traj nobox

#Step 3: Center the ribozyme and image appropriately
center :1-73 mass origin
image origin center

#Step 4: Strip water
strip :WAT
```

A.2 In_Line_Fitness.ptraj

```
#For usage, see AMBER manual.
#This is a script that extracts the in-line attack distance and in-line
#attack angle for calculation of the in-line fitness. See Chapter 2,
#Materials and Methods, and
#1999, Soukup et al. (RNA) for equation to calculate in-line fitness.

#Step 1: Call in stripped trajectory (xxxx.traj)
```

```
trajin xxxx.traj

#Step 2: center the ribozyme and image the trajectory
center :1-73 mass origin
image origin center

#Step 3: calculate desired distances and angles
#In-line attack distance: U-1(2O')-A1+phosphate
distance dist10 :1@O2' :2@P out IAD.out time 1.0

#In-line attack angle: U-1(2O'), N1(P), N1(O5')
angle dist11 :1@O2' :2@P :2@O5' out IAA.out time 1.0

#Step 4: Use Excel or similar program and the in-line fitness equation
to generate in-line fitness values
```

All analysis of hydrogen bonding in Chapters 2 and 3 was conducted using the Simulaid program.[118] This process is a multi-stage process.

A.3 Stage 1: simulaid.ptraj

```
#Stage 1 of Simulaid hydrogen bonding analysis: generate an exemplar
#pdb file from the trajectory to be analyzed, using AMBER modules.
#Typically any frame should suffice, although it is recommended to use
#a frame closer to the beginning of the trajectory before too many
#structural dynamics occur.
#For usage and syntax, see the AMBER manual

trajin xxxx.traj 2 2
trajout simulaid.pdb pdb

strip :WAT
```

A.4 Stage 2: Using the Simulaid program

Once `simulaid.pdb` is generated, analyze the trajectory and the exemplar `simulaid.pdb` using the Simulaid program.[118] Downloading information and documentation are found at http://inka.mssm.edu/~mezei/simulaid/, as well as in.[118] Running the Simulaid program will generate `simulaid.hbn`, used in the next sections.

A.5 Stage 3: run_new

```
#Written by Ms. Wendy Tay and modified for batch analysis by Ms. Kamali
#Sripathi
#Stage 3 of Simulaid hydrogen bonding analysis: This is an exemplar
#script that illustrates how to initially extract hydrogen bonding per
#frame from a simulaid.hbn file.
#Usage: ./getframe simulaid.hbn H Rd A Ra
#where H is the hydrogen in the bond, Rd is the residue number of which
#H is a part, A is the acceptor atom, and Ra is the residue number #of
#which A is a part.
#Modify the donors and acceptors as needed
#For further notes on syntax, see 2010, Mezei, M. (J. Comput. Chem)

#!/bin/bash

#G1U37
echo "Working on N1Plot_A"
./getframe simulaid.hbn H3 37 O6 2 > N1Plot_A
echo "Working on N1Plot_B"
./getframe simulaid.hbn H1 2 O2 37 > N1Plot_B
```

This script simply outputs a list of only those frames containing the hydrogen bonds in question in the "x" column, with a list of corresponding "0"'s in the "y" column. The next stage is to fill in the gaps, so that the end products is a list of "1"s and "0"s denoting the time-dependent hydrogen bond presence or absence. In the relevant figures in Chapters 2 and 3, a present hydrogen bond ("0") is red, while an absent hydrogen bond is white ("1").

A.5.1 getframe

This script is called in the run_new script in section A.5 above.

```
#!/bin/csh

#Written by Ms. Wendy Tay.

#This is the script called repeatedly by the run_new script in A.5.
#USAGE: ./getframe SimulaidOutput.hbn Atom1 Res1 Atom2 Res2
#
set file = $1
set atom1 = $2
set res1 = $3
set atom2 = $4
set res2 = $5

#set basename = "snapshot"
#awk '{if ($1=="Configuration") frame=$3}{if ($2=="H1" && $3=="2" &&
```

```
#$7=="O2" && $8=="37") print frame}' $file

awk '\
       {     if ($1=="Configuration"){\
                    frame=$3\
            }if ($2==atom1 && $3==res1 && $7==atom2 && $8==res2){\
                  printf "%5d %s\n", frame, "\t" 0\
                  }\
       }\
' atom1=$atom1 res1=$res1 atom2=$atom2 res2=$res2 $file

#awk '{if ($1=="Configuration") frame=$3}{if ($2==atom1 && $3==res1 &&
$7==atom2 && $8==res2) print frame}' atom1=$atom1 res1=$res1
atom2=$atom2 res2=$res2 $file > tmp.list

#mkdir hbond_files_$atom1\_$res1\_$atom2\_$res2
#foreach C ( ` cat tmp.list ` )
#   cp files/$basename\_$C\.pdb
#hbond_files_$atom1\_$res1\_$atom2\_$res2\/.
#end
```

A.6 Stage 4: get_extra_frame

```
#!/bin/bash

#Written by Ms. Wendy Tay
#This script is to insert the frames not picked up by the getframe_w
#awk script. This script inserts the missing frames and a '1'.
#To be used with all40000.dat or other file containing a complete list
#of frame numbers.

#Change the elements in ‘array’ below based on how many NPlot_* files
#are generated in the run_new script above

array=(A B)

for i in "${array[@]}"

do
      echo "Working on N1Plot_$i"
      awk '{printf "%5d\n", $1}' < N1Plot_$i > temp.1
      awk '{printf "%5d\n", $1}' < all40000.dat > temp.2
      comm -1 -3 temp.1 temp.2 > N1Plot_$i.comm
      cat N1Plot_$i N1Plot_$i.comm > combined

      sort -n combined -o N1Plot_$i.sorted

      rm temp.1 temp.2

      awk '{\
            if ($2 == 0)\
            {\
                  print $0\
            }\
```

```
            if ($2 != 0)\
                {\
            printf "%5d %s\n", $1, "\t" 1\
          }\
        }\
    ' < N1Plot_$i.sorted > N1Plot_$i.out

    #rm N1Plot_$i
    rm N1Plot_$i.comm
    rm combined
    rm N1Plot_$i.sorted
done
```

References

1. Crick, F. (1970) Central Dogma of Molecular Biology, *Nature 227*, 561-&.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The Rna Moiety of Ribonuclease-P Is the Catalytic Subunit of the Enzyme, *Cell 35*, 849-857.
3. Zaug, A. J., and Cech, T. R. (1986) The Intervening Sequence Rna of Tetrahymena Is an Enzyme, *Science 231*, 470-475.
4. Gilbert, W. (1986) ORIGIN OF LIFE - THE RNA WORLD, *Nature 319*, 618-618.
5. Scott, W. G. (2007) Ribozymes, *Curr Opin Struc Biol 17*, 280-286.
6. Strobel, S. A., and Cochrane, J. C. (2007) RNA catalysis: ribozymes, ribosomes, and riboswitches, *Current Opinion in Chemical Biology 11*, 636-643.
7. Doudna, J. A., and Lorsch, J. R. (2005) Ribozyme catalysis: not different, just worse, *Nat Struct Mol Biol 12*, 395-402.
8. Fedor, M. J., Kuzmin, Y. I., Da Costa, C. P., and Cottrell, J. W. (2007) Catalytic mechanism of the hairpin ribozyme, *Abstr Pap Am Chem S 233*.
9. Fedor, M. J., and Williamson, J. R. (2005) The catalytic diversity of RNAs, *Nature reviews. Molecular cell biology 6*, 399-412.
10. Chen, J. H., Yajima, R., Chadalavada, D. M., Chase, E., Bevilacqua, P. C., and Golden, B. L. (2010) A 1.9 A crystal structure of the HDV ribozyme precleavage suggests both Lewis acid and general acid mechanisms contribute to phosphodiester cleavage, *Biochemistry 49*, 6508-6518.
11. Ferre-D'Amare, A. R., Zhou, K., and Doudna, J. A. (1998) Crystal structure of a hepatitis delta virus ribozyme, *Nature 395*, 567-574.
12. 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.
13. 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, *Journal of the American Chemical Society 134*, 16933-16936.
14. 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.
15. 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.
16. 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.
17. Grum-Tokars, V., Milovanovic, M., and Wedekind, J. E. (2003) Crystallization and X-ray diffraction analysis of an all-RNA U39C mutant of the minimal hairpin ribozyme, *Acta Crystallogr D 59*, 142-145.
18. Pley, H. W., Flaherty, K. M., and Mckay, D. B. (1995) 3-Dimensional Structure of a Hammerhead Ribozyme, *J Cell Biochem*, 204-204.

19. Banas, P., Rulisek, L., Hanosova, V., Svozil, D., Walter, N. G., Sponer, J., and Otyepka, M. (2008) General base catalysis for cleavage by the active-site cytosine of the hepatitis delta virus ribozyme: QM/MM calculations establish chemical feasibility, *J Phys Chem B 112*, 11177-11187.
20. Nakano, S., Chadalavada, D. M., and Bevilacqua, P. C. (2000) General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme, *Science 287*, 1493-1497.
21. Perrotta, A. T., Shih, I., and Been, M. D. (1999) Imidazole rescue of a cytosine mutation in a self-cleaving ribozyme, *Science 286*, 123-126.
22. Das, S. R., and Piccirilli, J. A. (2005) General acid catalysis by the hepatitis delta virus ribozyme, *Nat Chem Biol 1*, 45-52.
23. Luptak, A., Ferre-D'Amare, A. R., Zhou, K., Zilm, K. W., and Doudna, J. A. (2001) Direct pK(a) measurement of the active-site cytosine in a genomic hepatitis delta virus ribozyme, *J Am Chem Soc 123*, 8447-8452.
24. Kuzmin, Y. I., Da Costa, C. P., Cottrell, J. W., and Fedor, M. J. (2005) Role of an active site adenine in hairpin ribozyme catalysis, *Journal of Molecular Biology 349*, 989-1010.
25. Guo, M., Spitale, R. C., Volpini, R., Krucinska, J., Cristalli, G., Carey, P. R., and Wedekind, J. E. (2009) Direct Raman Measurement of an Elevated Base pK(a) in the Active Site of a Small Ribozyme in a Precatalytic Conformation, *Journal of the American Chemical Society 131*, 12908-+.
26. Spitale, R. C., Volpini, R., Heller, M. G., Krucinska, J., Cristalli, G., and Wedekind, J. E. (2009) Identification of an Imino Group Indispensable for Cleavage by a Small Ribozyme, *Journal of the American Chemical Society 131*, 6093-+.
27. Cottrell, J. W., Kuzmin, Y. I., and Fedor, M. J. (2007) Functional analysis of hairpin ribozyme active site architecture, *Journal of Biological Chemistry 282*, 13498-13507.
28. Pinard, R., Hampel, K. J., Heckman, J. E., Lambert, D., Chan, P. A., Major, F., and Burke, J. M. (2001) Functional involvement of G8 in the hairpin ribozyme cleavage mechanism, *Embo Journal 20*, 6434-6442.
29. Kuzmin, Y. I., Da Costa, C. P., and Fedor, M. J. (2004) Role of an active site guanine in hairpin ribozyme catalysis probed by exogenous nucleobase rescue, *Journal of Molecular Biology 340*, 233-251.
30. Liu, L., Cottrell, J. W., Scott, L. G., and Fedor, M. J. (2009) Direct measurement of the ionization state of an essential guanine in the hairpin ribozyme, *Nature Chemical Biology 5*, 351-357.
31. Lebruska, L. L., Kuzmine, I. I., and Fedor, M. J. (2002) Rescue of an abasic hairpin ribozyme by cationic nucleobases: Evidence for a novel mechanism of RNA catalysis, *Chem Biol 9*, 465-473.
32. 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 15*, 560-575.
33. Nam, K., Gao, J. L., and York, D. M. (2008) Electrostatic interactions in the hairpin ribozyme account for the majority of the rate acceleration without chemical participation by nucleobases, *Rna-a Publication of the Rna Society 14*, 1501-1507.

34. Nam, K. H., Gaot, J. L., and York, D. M. (2008) Quantum mechanical/molecular mechanical simulation study of the mechanism of hairpin ribozyme catalysis, *Journal of the American Chemical Society 130*, 4680-4691.
35. Been, M. D. (2006) HDV ribozymes, *Curr Top Microbiol Immunol 307*, 47-65.
36. Lai, M. M. (1995) The molecular biology of hepatitis delta virus, *Annu Rev Biochem 64*, 259-286.
37. Ferre-D'Amare, A. R. (2004) The hairpin ribozyme, *Biopolymers 73*, 71-78.
38. Butcher, S. E., Heckman, J. E., and Burke, J. M. (1995) Reconstitution of Hairpin Ribozyme Activity Following Separation of Functional Domains, *Journal of Biological Chemistry 270*, 29648-29651.
39. Shin, C., Choi, J. N., Sang, S. I., Song, J. T., Ahn, J. H., Lee, J. S., and Choi, Y. D. (1996) The loop B domain is physically separable from the loop A domain in the hairpin ribozyme, *Nucleic Acids Research 24*, 2685-2689.
40. Tinsley, R. A., and Walter, N. G. (2007) Long-range impact of peripheral joining elements on structure and function of the hepatitis delta virus ribozyme, *Biol Chem 388*, 705-715.
41. 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.
42. Wadkins, T. S., Shih, I., Perrotta, A. T., and Been, M. D. (2001) A pH-sensitive RNA tertiary interaction affects self-cleavage activity of the HDV ribozymes in the absence of added divalent metal ion, *J Mol Biol 305*, 1045-1055.
43. Ferre-D'Amare, A. R., and Scott, W. G. (2010) Small self-cleaving ribozymes, *Cold Spring Harb Perspect Biol 2*, a003574.
44. Ryder, S. P., Oyelere, A. K., Padilla, J. L., Klostermeier, D., Millar, D. P., and Strobel, S. A. (2001) Investigation of adenosine base ionization in the hairpin ribozyme by nucleotide analog interference mapping, *Rna-a Publication of the Rna Society 7*, 1454-1463.
45. Siwkowski, A., Shippy, R., and Hampel, A. (1997) Analysis of hairpin ribozyme base mutations in loops 2 and 4 and their effects on cis-cleavage in vitro, *Biochemistry 36*, 3930-3940.
46. Klostermeier, D., and Millar, D. P. (2002) Energetics of hydrogen bond networks in RNA: Hydrogen bonds surrounding G+1 and U42 are the major determinants for the tertiary structure stability of the hairpin ribozyme, *Biochemistry 41*, 14095-14102.
47. Wilson, T. J., Zhao, Z. Y., Maxwell, K., Kontogiannis, L., and Lilley, D. M. J. (2001) Importance of specific nucleotides in the folding of the natural form of the hairpin ribozyme, *Biochemistry 40*, 2291-2302.
48. Earnshaw, D. J., Hamm, M. L. H., Piccirilli, J. A., Karpeisky, A., Beigelman, L., Ross, B. S., Manoharan, M., and Gait, M. J. (2000) Investigation of the proposed interdomain ribose zipper in hairpin ribozyme cleavage using 2 '-modified nucleosides, *Biochemistry 39*, 6410-6421.
49. Krasovska, M. V., Sefcikova, J., Spackova, N., Sponer, J., and Walter, N. G. (2005) Structural dynamics of precursor and product of the RNA enzyme from the hepatitis delta virus as revealed by molecular dynamics simulations, *J Mol Biol 351*, 731-748.
50. Gong, B., Chen, J. H., Bevilacqua, P. C., Golden, B. L., and Carey, P. R. (2009) Competition between Co(NH(3)(6)3+ and inner sphere Mg2+ ions in the HDV ribozyme, *Biochemistry 48*, 11961-11970.

51. Hampel, A., and Cowan, J. A. (1997) A unique mechanism for RNA catalysis: the role of metal cofactors in hairpin ribozyme cleavage, *Chem Biol 4*, 513-517.
52. Nesbitt, S., Hegg, L. A., and Fedor, M. J. (1997) An unusual pH-independent and metal-ion-independent mechanism for hairpin ribozyme catalysis, *Chem Biol 4*, 619-630.
53. Ravindranathan, S., Butcher, S. E., and Feigon, J. (2000) Adenine protonation in domain B of the hairpin ribozyme, *Biochemistry 39*, 16026-16032.
54. Macelrevey, C., Salter, J. D., Krucinska, J., and Wedekind, J. E. (2008) Structural effects of nucleobase variations at key active site residue Ade38 in the hairpin ribozyme, *Rna-a Publication of the Rna Society 14*, 1600-1616.
55. 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, *Journal of Physical Chemistry B 114*, 6642-6652.
56. 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, *Journal of Physical Chemistry B 115*, 13911-13924.
57. Sefcikova, J., Krasovska, M. V., Spackova, N., Sponer, J., and Walter, N. G. (2007) Impact of an extruded nucleotide on cleavage activity and dynamic catalytic core conformation of the hepatitis delta virus ribozyme, *Biopolymers 85*, 392-406.
58. Sefcikova, J., Krasovska, M. V., Sponer, J., and Walter, N. G. (2007) The genomic HDV ribozyme utilizes a previously unnoticed U-turn motif to accomplish fast site-specific catalysis, *Nucleic Acids Res 35*, 1933-1946.
59. Mcdowell, S. E., Jun, J. M., and Walter, N. G. (2010) Long-range tertiary interactions in single hammerhead ribozymes bias motional sampling toward catalytically active conformations, *Rna-a Publication of the Rna Society 16*, 2414-2426.
60. 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, *Acc Chem Res 43*, 40-47.
61. Reblova, K., Sponer, J., and Lankas, F. (2012) Structure and mechanical properties of the ribosomal L1 stalk three-way junction, *Nucleic Acids Res 40*, 6290-6303.
62. 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, *Proc Natl Acad Sci U S A 103*, 13380-13385.
63. Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham, T. E., 3rd, Laughton, C. A., and Orozco, M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers, *Biophys J 92*, 3817-3829.
64. Lee, T.-S., Giambasu, G. M., Harris, M. E., and York, D. M. (2011) Characterization of the Structure and Dynamics of the HDV Ribozyme in Different Stages Along the Reaction Path, *Journal of Physical Chemistry Letters 2*, 2538-2543.
65. Krasovska, M. V., Sefcikova, J., Reblova, K., Schneider, B., Walter, N. G., and Sponer, J. (2006) Cations and hydration in catalytic RNA: molecular dynamics of the hepatitis delta virus ribozyme, *Biophys J 91*, 626-638.
66. 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.

67. Zgarbova, M., Otyepka, M., Sponer, J., Mladek, A., Banas, P., Cheatham, T. E., 3rd, and Jurecka, 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.
68. Goh, G. B., Knight, J. L., and Brooks, C. L. (2012) Constant pH Molecular Dynamics Simulations of Nucleic Acids in Explicit Solvent, *Journal of Chemical Theory and Computation 8*, 36-46.
69. Goh, G. B., Knight, J. L., and Brooks, C. L. (2013) Toward Accurate Prediction of the Protonation Equilibrium of Nucleic Acids, *Journal of Physical Chemistry Letters 4*, 760-766.
70. Goh, G. B., Knight, J. L., and Brooks, C. L. (2013) pH-Dependent Dynamics of Complex RNA Macromolecules, *Journal of Chemical Theory and Computation 9*, 935-943.
71. Knight, J. L., and Brooks, C. L. (2011) Multisite lambda Dynamics for Simulated Structure-Activity Relationship Studies, *Journal of Chemical Theory and Computation 7*, 2728-2739.
72. Knight, J. L., and Brooks, C. L. (2009) lambda-Dynamics Free Energy Simulation Methods, *Journal of Computational Chemistry 30*, 1692-1700.
73. Salehi-Ashtiani, K., Luptak, A., Litovchick, A., and Szostak, J. W. (2006) A genomewide search for ribozymes reveals an HDV-like sequence in the human CPEB3 gene, *Science 313*, 1788-1792.
74. Talini, G., Branciamore, S., and Gallori, E. (2011) Ribozymes: Flexible molecular devices at work, *Biochimie 93*, 1998-2005.
75. Webb, C. H. T., Riccitelli, N. J., Ruminski, D. J., and Luptak, A. (2009) Widespread Occurrence of Self-Cleaving Ribozymes, *Science 326*, 953-953.
76. Jimenez, R. M., Delwart, E., and Luptak, A. (2011) Structure-based Search Reveals Hammerhead Ribozymes in the Human Microbiome, *Journal of Biological Chemistry 286*, 7737-7743.
77. Burnett, J. C., and Rossi, J. J. (2012) RNA-Based Therapeutics: Current Progress and Future Prospects, *Chem Biol 19*, 60-71.
78. Flores, R., Ruiz-Ruiz, S., and Serra, P. (2012) Viroids and hepatitis delta virus, *Semin Liver Dis 32*, 201-210.
79. Taylor, J. M. (2012) Virology of hepatitis D virus, *Semin Liver Dis 32*, 195-200.
80. Pereira, M. J., Harris, D. A., Rueda, D., and Walter, N. G. (2002) Reaction pathway of the trans-acting hepatitis delta virus ribozyme: a conformational change accompanies catalysis, *Biochemistry 41*, 730-740.
81. Shih, I. H., and Been, M. D. (2002) Catalytic strategies of the hepatitis delta virus ribozymes, *Annu Rev Biochem 71*, 887-917.
82. Webb, C. H., Riccitelli, N. J., Ruminski, D. J., and Luptak, A. (2009) Widespread occurrence of self-cleaving ribozymes, *Science 326*, 953.
83. Webb, C. H., and Luptak, A. (2011) HDV-like self-cleaving ribozymes, *RNA Biol. 8*, 719-727.
84. Been, M. D., and Wickham, G. S. (1997) Self-cleaving ribozymes of hepatitis delta virus RNA, *Eur J Biochem 247*, 741-753.
85. Wadkins, T. S., Perrotta, A. T., Ferre-D'Amare, A. R., Doudna, J. A., and Been, M. D. (1999) A nested double pseudoknot is required for self-cleavage activity of both the genomic and antigenomic hepatitis delta virus ribozymes, *RNA 5*, 720-727.

86. Perrotta, A. T., and Been, M. D. (2006) HDV ribozyme activity in monovalent cations, *Biochemistry 45*, 11357-11365.
87. Perrotta, A. T., Wadkins, T. S., and Been, M. D. (2006) Chemical rescue, multiple ionizable groups, and general acid-base catalysis in the HDV genomic ribozyme, *RNA 12*, 1282-1291.
88. Nakano, S., Cerrone, A. L., and Bevilacqua, P. C. (2003) Mechanistic characterization of the HDV genomic ribozyme: classifying the catalytic and structural metal ion sites within a multichannel reaction mechanism, *Biochemistry 42*, 2982-2994.
89. Chadalavada, D. M., Cerrone-Szakal, A. L., and Bevilacqua, P. C. (2007) Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions, *Rna-a Publication of the Rna Society 13*, 2189-2201.
90. Gong, B., Chen, J. H., Chase, E., Chadalavada, D. M., Yajima, R., Golden, B. L., Bevilacqua, P. C., and Carey, P. R. (2007) Direct measurement of a pK(a) near neutrality for the catalytic cytosine in the genomic HDV ribozyme using Raman crystallography, *J. Am. Chem. Soc. 129*, 13335-13342.
91. Cerrone-Szakal, A. L., Chadalavada, D. M., Golden, B. L., and Bevilacqua, P. C. (2008) Mechanistic characterization of the HDV genomic ribozyme: the cleavage site base pair plays a structural role in facilitating catalysis, *RNA 14*, 1746-1760.
92. Gong, B., Chen, Y., Christian, E. L., Chen, J. H., Chase, E., Chadalavada, D. M., Yajima, R., Golden, B. L., Bevilacqua, P. C., and Carey, P. R. (2008) Detection of innersphere interactions between magnesium hydrate and the phosphate backbone of the HDV ribozyme using Raman crystallography, *J. Am. Chem. Soc. 130*, 9670-9672.
93. Chen, G., Kennedy, S. D., and Turner, D. H. (2009) A CA(+) Pair Adjacent to a Sheared GA or AA Pair Stabilizes Size-Symmetric RNA Internal Loops, *Biochemistry 48*, 5738-5752.
94. Chen, J. H., Gong, B., Bevilacqua, P. C., Carey, P. R., and Golden, B. L. (2009) A catalytic metal ion interacts with the cleavage Site G.U wobble in the HDV ribozyme, *Biochemistry 48*, 1498-1507.
95. Harris, D. A., Tinsley, R. A., and Walter, N. G. (2004) Terbium-mediated footprinting probes a catalytic conformational switch in the antigenomic hepatitis delta virus ribozyme, *J. Mol. Biol. 341*, 389-403.
96. Been, M. D. (1994) Cis- and trans-acting ribozymes from a human pathogen, hepatitis delta virus, *Trends Biochem. Sci. 19*, 251-256.
97. Harris, D. A., Rueda, D., and Walter, N. G. (2002) Local conformational changes in the catalytic core of the trans-acting hepatitis delta virus ribozyme accompany catalysis, *Biochemistry 41*, 12051-12061.
98. Harris, T. K., and Turner, G. J. (2002) Structural basis of perturbed pK(a) values of catalytic groups in enzyme active sites, *Iubmb Life 53*, 85-98.
99. Jeong, S., Sefcikova, J., Tinsley, R. A., Rueda, D., and Walter, N. G. (2003) Trans-acting hepatitis delta virus ribozyme: catalytic core and global structure are dependent on the 5' substrate sequence, *Biochemistry 42*, 7727-7740.
100. Tinsley, R. A., Harris, D. A., and Walter, N. G. (2003) Significant kinetic solvent isotope effects in folding of the catalytic RNA from the hepatitis delta virus, *J. Am. Chem. Soc. 125*, 13972-13973.

101. Tinsley, R. A., Harris, D. A., and Walter, N. G. (2004) Magnesium dependence of the amplified conformational switch in the trans-acting hepatitis delta virus ribozyme, *Biochemistry 43*, 8935-8945.
102. Walter, N. G., and Perumal, S. (2009) The Small Ribozymes: Common and Diverse Features Observed through the FRET Lens, *Springer Ser. Biophys. 13*, 103-127.
103. Tanaka, Y., Tagaya, M., Hori, T., Sakamoto, T., Kurihara, Y., Katahira, M., and Uesugi, S. (2002) Cleavage reaction of HDV ribozymes in the presence of Mg2+ is accompanied by a conformational change, *Genes Cells 7*, 567-579.
104. Ferre-D'Amare, A. R., and Doudna, J. A. (2000) Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module, *J. Mol. Biol. 295*, 541-556.
105. Walter, N. G. (2002) Probing RNA structural dynamics and function by fluorescence resonance energy transfer (FRET), *Curr. Protoc. Nucleic Acid Chem. 11*, 11.10.11-11.10.23.
106. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis, *Nat Methods 9*, 671-675.
107. Case, D. A., Darden, T. A., Cheatham, I., T. E.,, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., D.H. Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., and Kollman, P. A. (2008) AMBER 10 University of California, San Francisco, CA.
108. Case, D. A., Darden, T. A., Cheatham, I., T. E., , Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2010) AMBER 11, University of California, San Franscisco, CA.
109. 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.
110. Beveridge, D. L., Cheatham, T. E., 3rd, and Mezei, M. (2012) The ABCs of molecular dynamics simulations on B-DNA, circa 2012, *J Biosci 37*, 379-397.
111. Lavery, R., Zakrzewska, K., Beveridge, D., Bishop, T. C., Case, D. A., Cheatham, T., 3rd, Dixit, S., Jayaram, B., Lankas, F., Laughton, C., Maddocks, J. H., Michon, A., Osman, R., Orozco, M., Perez, A., Singh, T., Spackova, N., and Sponer, J. (2010) A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA, *Nucleic Acids Res 38*, 299-313.
112. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath, *J Chem Phys 81*, 3684-3690.
113. 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, *Journal of Computational Physics 23*, 327-341.

114. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M. J., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J Am Chem Soc 117*, 5179-5197.
115. Banas, P., Sklenovsky, P., Wedekind, J. E., Sponer, J., and Otyepka, M. (2012) Molecular Mechanism of preQ1 Riboswitch Action: A Molecular Dynamics Study, *The Journal of Physical Chemistry B 116*, 12721-12734.
116. Cornell, W. D., Cieplak, P., Bayly, C. I., and Kollman, P. A. (1993) Application of RESP charges to calculate conformational energies, hydrogen bond energies and free energies of solvation., *J Am Chem Soc 115*.
117. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. (2004) Gaussian 03, Gaussian, Inc., Wallingford, CT.
118. Mezei, M. (2010) Simulaid: a simulation facilitator and analysis program, *J Comput Chem 31*, 2658-2668.
119. Soukup, G. A., and Breaker, R. R. (1999) Relationship between internucleotide linkage geometry and the stability of RNA, *RNA 5*, 1308-1325.
120. Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction, *Nucleic Acids Res 31*, 3406-3415.
121. Wadkins, T. S., and Been, M. D. (2002) Ribozyme activity in the genomic and antigenomic RNA strands of hepatitis delta virus, *Cell. Mol. Life Sci. 59*, 112-125.
122. Marek, M. S., Johnson-Buck, A., and Walter, N. G. (2011) The shape-shifting quasispecies of RNA: one sequence, many functional folds, *Phys. Chem. Chem. Phys. 13*, 11524-11537.
123. Walter, N. G., and Burke, J. M. (2000) Fluorescence assays to study structure, dynamics, and function of RNA and RNA-ligand complexes, *Methods Enzymol 317*, 409-440.
124. Walter, N. G. (2001) Structural dynamics of catalytic RNA highlighted by fluorescence resonance energy transfer, *Methods 25*, 19-30.
125. 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.

126. Walter, N. G., Harris, D. A., Pereira, M. J. B., and Rueda, D. (2001) In the fluorescent spotlight: Global and local conformational changes of small catalytic RNAs, *Biopolymers 61*, 224-242.
127. Walter, N. G., Burke, J. M., and Millar, D. P. (1999) Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction, *Nat. Struct. Biol. 6*, 544-549.
128. Rueda, D., Wick, K., McDowell, S. E., and Walter, N. G. (2003) Diffusely bound Mg2+ ions slightly reorient stems I and II of the hammerhead ribozyme to increase the probability of formation of the catalytic core, *Biochemistry 42*, 9924-9936.
129. Banas, P., Hollas, D., Zgarbova, M., Jurecka, P., Orozco, M., Cheatham, T. E., 3rd, Sponer, J., and Otyepka, M. (2010) Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins, *Journal of Chemical Theory and Computation 6*, 3836-3849.
130. Lee, T. S., Giambasu, G., Harris, M. E., and York, D. M. (2011) Characterization of the Structure and Dynamics of the HDV Ribozyme at Different Stages Along the Reaction Path, *J. Phys. Chem. Lett. 2*, 2538-2543.
131. Martick, M., and Scott, W. G. (2006) Tertiary contacts distant from the active site prime a ribozyme for catalysis, *Cell 126*, 309-320.
132. Zhuang, X., Kim, H., Pereira, M. J., Babcock, H. P., Walter, N. G., and Chu, S. (2002) Correlating structural dynamics and function in single ribozyme molecules, *Science 296*, 1473-1476.
133. Rueda, D., Bokinsky, G., Rhodes, M. M., Rust, M. J., Zhuang, X., and Walter, N. G. (2004) Single-molecule enzymology of RNA: essential functional groups impact catalysis from a distance, *Proc Natl Acad Sci U S A 101*, 10066-10071.
134. Nakano, S., Proctor, D. J., and Bevilacqua, P. C. (2001) Mechanistic characterization of the HDV genomic ribozyme: assessing the catalytic and structural contributions of divalent metal ions within a multichannel reaction mechanism, *Biochemistry 40*, 12022-12038.
135. Krasovska, M. V., Sefcikova, J., Reblova, K., Schneider, B., Walter, N. G., and Sponer, J. (2006) Cations and hydration in catalytic RNA: molecular dynamics of the hepatitis delta virus ribozyme, *Biophys. J. 91*, 626-638.
136. Shih, I. H., and Been, M. D. (2002) Catalytic strategies of the hepatitis delta virus ribozymes, *Annu. Rev. Biochem. 71*, 887-917.
137. Tinsley, R. A., and Walter, N. G. (2007) Long-range impact of peripheral joining elements on structure and function of the hepatitis delta virus ribozyme, *Biol. Chem. 388*, 705-715.
138. Bevilacqua, P. C. (2003) Mechanistic considerations for general acid-base catalysis by RNA: revisiting the mechanism of the hairpin ribozyme, *Biochemistry 42*, 2259-2265.
139. Bevilacqua, P. C., Brown, T. S., Nakano, S., and Yajima, R. (2004) Catalytic roles for proton transfer and protonation in ribozymes, *Biopolymers 73*, 90-109.
140. Walter, N. G. (2007) Ribozyme catalysis revisited: is water involved?, *Mol. Cell 28*, 923-929.
141. Rueda, D., Bokinsky, G., Rhodes, M. M., Rust, M. J., Zhuang, X. W., and Walter, N. G. (2004) Single-molecule enzymology of RNA: Essential functional groups impact catalysis from a distance, *Proceedings of the National Academy of Sciences of the United States of America 101*, 10066-10071.

142. Woodson, S. A. (2010) Taming free energy landscapes with RNA chaperones, *RNA Biol. 7*, 677-686.
143. Ditzler, M. A., Aleman, E. A., Rueda, D., and Walter, N. G. (2007) Focus on function: Single molecule RNA enzymology, *Biopolymers 87*, 302-316.
144. Chen, J., Ganguly, A., Miswan, Z., Hammes-Schiffer, S., Bevilacqua, P. C., and Golden, B. L. (2013) Identification of the catalytic Mg(2)(+) ion in the hepatitis delta virus ribozyme, *Biochemistry 52*, 557-567.
145. Varani, G., and McClain, W. H. (2000) The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems, *EMBO reports 1*, 18-23.
146. Nishikawa, F., Fauzi, H., and Nishikawa, S. (1997) Detailed analysis of base preferences at the cleavage site of a trans-acting HDV ribozyme: a mutation that changes cleavage site specificity, *Nucleic Acids Res 25*, 1605-1610.
147. Wu, H. N., Lee, J. Y., Huang, H. W., Huang, Y. S., and Hsueh, T. G. (1993) Mutagenesis analysis of a hepatitis delta virus genomic ribozyme, *Nucleic Acids Res 21*, 4193-4199.
148. Wu, H. N., Wang, Y. J., Hung, C. F., Lee, H. J., and Lai, M. M. (1992) Sequence and structure of the catalytic RNA of hepatitis delta virus genomic RNA, *J Mol Biol 223*, 233-245.
149. Tanner, N. K., Schaff, S., Thill, G., Petit-Koskas, E., Crain-Denoyelle, A. M., and Westhof, E. (1994) A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses, *Current biology : CB 4*, 488-498.
150. Veeraraghavan, N., Ganguly, A., Chen, J. H., Bevilacqua, P. C., Hammes-Schiffer, S., and Golden, B. L. (2011) Metal binding motif in the active site of the HDV ribozyme binds divalent and monovalent ions, *Biochemistry 50*, 2672-2682.
151. Veeraraghavan, N., Ganguly, A., Golden, B. L., Bevilacqua, P. C., and Hammes-Schiffer, S. (2011) Mechanistic strategies in the HDV ribozyme: chelated and diffuse metal ion interactions and active site protonation, *J Phys Chem B 115*, 8346-8357.
152. D.A. Case, T.A. Darden, T.E. Cheatham, I., C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and Kollman, P. A. (2010) AMBER 11, University of California, San Franscisco, CA.
153. Leontis, N. B., Stombaugh, J., and Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices, *Nucleic Acids Res 30*, 3497-3531.
154. Nishikawa, F., Kawakami, J., Chiba, A., Shirai, M., Kumar, P. K. R., and Nishikawa, S. (1996) Selection in vitro of trans-acting genomic human hepatitis delta virus (HDV) ribozymes, *European Journal of Biochemistry 237*, 712-718.
155. Stombaugh, J., Zirbel, C. L., Westhof, E., and Leontis, N. B. (2009) Frequency and isostericity of RNA base pairs, *Nucleic Acids Research 37*, 2294-2312.
156. Masquida, B., and Westhof, E. (2000) On the wobble GoU and related pairs, *Rna-a Publication of the Rna Society 6*, 9-15.

157. Perrotta, A. T., and Been, M. D. (1996) Core sequences and a cleavage site wobble pair required for HDV antigenomic ribozyme self-cleavage, *Nucleic Acids Res 24*, 1314-1321.
158. Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B., and Steitz, T. A. (2001) RNA tertiary interactions in the large ribosomal subunit: the A-minor motif, *Proc Natl Acad Sci U S A 98*, 4899-4903.
159. Kawakami, J., Kumar, P. K. R., Suh, Y. A., Nishikawa, F., Kawakami, K., Taira, K., Ohtsuka, E., and Nishikawa, S. (1993) Identification of Important Bases in a Single-Stranded Region (Ssrc) of the Hepatitis-Delta (Delta) Virus Ribozyme, *European Journal of Biochemistry 217*, 29-36.
160. Nakano, S., and Bevilacqua, P. C. (2001) Proton inventory of the genomic HDV ribozyme in Mg(2+)-containing solutions, *J Am Chem Soc 123*, 11333-11334.
161. Gao, X. L., and Patel, D. J. (1987) NMR-STUDIES OF AC MISMATCHES IN DNA DODECANUCLEOTIDES AT ACIDIC PH - WOBBLE A(ANTI)-C(ANTI) PAIR FORMATION, *Journal of Biological Chemistry 262*, 16973-16984.
162. Asensio, J. L., Lane, A. N., Dhesi, J., Bergqvist, S., and Brown, T. (1998) The contribution of cytosine protonation to the stability of parallel DNA triple helices, *Journal of Molecular Biology 275*, 811-822.
163. Jang, S. B., Hung, L. W., Chi, Y. I., Holbrook, E. L., Carter, R. J., and Holbrook, S. R. (1998) Structure of an RNA internal loop consisting of tandem C-A(+) base pairs, *Biochemistry 37*, 11726-11731.
164. Bink, H. H. J., Hellendoorn, K., van der Meulen, J., and Pleij, C. W. A. (2002) Protonation of non-Watson-Crick base pairs and encapsidation of turnip yellow mosaic virus RNA, *Proceedings of the National Academy of Sciences of the United States of America 99*, 13465-13470.
165. Morse, S. E., and Draper, D. E. (1995) PURINE PURINE MISMATCHES IN RNA HELICES - EVIDENCE FOR PROTONATED G-CENTER-DOT-A PAIRS AND NEXT-NEAREST-NEIGHBOR EFFECTS, *Nucleic Acids Research 23*, 302-306.
166. Durant, P. C., and Davis, D. R. (1999) Stabilization of the anticodon stem-loop of tRNA(LyS,3) by an A(+)-C base-pair and by pseudouridine, *Journal of Molecular Biology 285*, 115-131.
167. Wilcox, J. L., and Bevilacqua, P. C. (2013) pK(a) Shifting in Double-Stranded RNA Is Highly Dependent upon Nearest Neighbors and Bulge Positioning, *Biochemistry 52*, 7470-7476.
168. Hoogstraten, C. G., Legault, P., and Pardi, A. (1998) NMR solution structure of the lead-dependent ribozyme: Evidence for dynamics in RNA catalysis, *Journal of Molecular Biology 284*, 337-350.
169. Legault, P., and Pardi, A. (1994) IN-SITU PROBING OF ADENINE PROTONATION IN RNA BY C-13 NMR, *Journal of the American Chemical Society 116*, 8390-8391.
170. Reiter, N. J., Blad, H., Abildgaard, F., and Butcher, S. E. (2004) Dynamics in the U6 RNA intramolecular stem-loop: A base flipping conformational changes, *Biochemistry 43*, 13739-13747.
171. Su, L., Chen, L. Q., Egli, M., Berger, J. M., and Rich, A. (1999) Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot, *Nature Structural Biology 6*, 285-292.

172. Nikolova, E. N., Goh, G. B., Brooks, C. L., and Al-Hashimi, H. M. (2013) Characterizing the Protonation State of Cytosine in Transient G center dot C Hoogsteen Base Pairs in Duplex DNA, *Journal of the American Chemical Society 135*, 6766-6769.
173. Nikolova, E. N., Kim, E., Wise, A. A., O'Brien, P. J., Andricioaei, I., and Al-Hashimi, H. M. (2011) Transient Hoogsteen base pairs in canonical duplex DNA, *Nature 470*, 498-U484.
174. Grasby, J. A., Mersmann, K., Singh, M., and Gait, M. J. (1995) Purine Functional-Groups in Essential Residues of the Hairpin Ribozyme Required for Catalytic Cleavage of Rna, *Biochemistry 34*, 4068-4076.
175. Cottrell, J. W., Scott, L. G., and Fedor, M. J. (2011) The pH Dependence of Hairpin Ribozyme Catalysis Reflects Ionization of an Active Site Adenine, *Journal of Biological Chemistry 286*, 17658-17664.
176. Loudon, G. M. (1991) Mechanistic Interpretation of Ph-Rate Profiles, *J Chem Educ 68*, 973-984.
177. Viladoms, J., Scott, L. G., and Fedor, M. J. (2011) An Active-Site Guanine Participates in glmS Ribozyme Catalysis in Its Protonated State, *Journal of the American Chemical Society 133*, 18388-18396.
178. Veeraraghavan, N., Bevilacqua, P. C., and Hammes-Schiffer, S. (2010) Long-Distance Communication in the HDV Ribozyme: Insights from Molecular Dynamics and Experiments, *Journal of Molecular Biology 402*, 278-291.
179. University of Maryland. (2007-2014) University of Maryland,, University of Maryland,.
180. Brooks, B. R., III, C. L. B., Jr, A. D. M., 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, *Journal of Computational Chemistry 30*, 1545-1614.
181. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, N. J., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009) Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT.
182. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model, *J Phys Chem-Us 97*, 10269-10280.

183. Cieplak, P., Cornell, W. D., Bayly, C., and Kollman, P. A. (1995) Application of the Multimolecule and Multiconformational Resp Methodology to Biopolymers - Charge Derivation for DNA, Rna, and Proteins, *Journal of Computational Chemistry 16*, 1357-1377.
184. Dupradeau, F. Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., Lelong, D., Rosanski, W., and Cieplak, P. (2010) The R.ED. tools: advances in RESP and ESP charge derivation and force field library building, *Phys Chem Chem Phys 12*, 7821-7839.
185. Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., Cieplak, P., and Dupradeau, F. Y. (2011) RED Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments, *Nucleic Acids Research 39*, W511-W517.
186. Wang, J. M., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?, *Journal of Computational Chemistry 21*, 1049-1074.
187. Legault, P., and Pardi, A. (1997) Unusual dynamics and pK(a) shift at the active site of a lead-dependent ribozyme, *Journal of the American Chemical Society 119*, 6621-6628.
188. 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.
189. Murray, L. J. W., Arendall, W. B., Richardson, D. C., and Richardson, J. S. (2003) RNA backbone is rotameric, *Proceedings of the National Academy of Sciences of the United States of America 100*, 13904-13909.
190. Rinnenthal, J., Buck, J., Ferner, J., Wacker, A., Fuertig, B., and Schwalbe, H. (2011) Mapping the Landscape of RNA Dynamics with NMR Spectroscopy, *Accounts Chem Res 44*, 1292-1301.
191. Hyeon, C., and Thirumalai, D. (2011) Capturing the essence of folding and functions of biomolecules using coarse-grained models, *Nature Communications 2*.
192. Thaplyal, P., Ganguly, A., Golden, B. L., Hammes-Schiffer, S., and Bevilacqua, P. C. (2013) Thio Effects and an Unconventional Metal Ion Rescue in the Genomic Hepatitis Delta Virus Ribozyme, *Biochemistry 52*, 6499-6514.
193. Wei, K., Liu, L., Cheng, Y.-H., Fu, Y., and Guo, Q.-X. (2007) Theoretical examination of two opposite mechanisms proposed for hepatitis delta virus ribozyme, *Journal of Physical Chemistry B 111*, 1514-1516.
194. Chadalavada, D. M., Senchak, S. E., and Bevilacqua, P. C. (2002) The folding pathway of the genomic hepatitis delta virus ribozyme is dominated by slow folding of the pseudoknots, *Journal of Molecular Biology 317*, 559-575.
195. Chadalavada, D. M., Knudsen, S. M., Nakano, S., and Bevilacqua, P. C. (2000) A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme, *J. Mol. Biol. 301*, 349-367.
196. Perrotta, A. T., and Been, M. D. (1998) A toggle duplex in hepatitis delta virus self-cleaving RNA that stabilizes an inactive and a salt-dependent pro-active ribozyme conformation, *J. Mol. Biol. 279*, 361-373.

197. 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.
198. Cochrane, J. C., and Strobel, S. A. (2008) Catalytic strategies of self-cleaving ribozymes, *Accounts Chem Res 41*, 1027-1035.
199. Mokdad, A., Krasovska, M. V., Sponer, J., and Leontis, N. B. (2006) Structural and evolutionary classification of G/U wobble basepairs in the ribosome, *Nucleic Acids Research 34*, 1326-1341.
200. Chadalavada, D. M., Cerrone-Szakal, A. L., and Bevilacqua, P. C. (2007) Wild-type is the optimal sequence of the HDV ribozyme under cotranscriptional conditions, *RNA 13*, 2189-2201.
201. Butcher, S. E., and Pyle, A. M. (2011) The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks, *Accounts Chem Res 44*, 1302-1311.
202. 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, *Journal of Molecular Biology 353*, 540-555.
203. Venditti, V., Clos, L., Niccolai, N., and Butcher, S. E. (2009) Minimum-Energy Path for a U6 RNA Conformational Change Involving Protonation, Base-Pair Rearrangement and Base Flipping, *Journal of Molecular Biology 391*, 894-905.