UNIVERSITY OF CALIFORNIA SAN DIEGO

Investigation of *Tetrahymena* Group I Intron Structural Dynamics

A Thesis submitted in partial satisfaction of the requirements

Master of Science

in

Chemistry

By

Alexa Perez

Committee in charge:

Professor Navtej Toor, Chair
Professor Simpson Joseph
Professor Ulrich Müller

The Thesis of Alexa Perez is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

I dedicate this thesis to…

All of those who made this journey possible

My mentors and role models

My family and friends

those who are with me in spirit

Stella and Sterling, too!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

TET	*Tetrahymena* thermophila group I intron
cryo-EM	Cryogenic electron microscopy
RNA	Ribonucleic acid
IGS	Internal guide sequence
wt	Wild-type
Δ*L9*	L9/P5 mutant
Δ*L9`*	L9/P5 mutant without solubility enhancing sequence
PAGE	polyacrylamide gel electrophoresis
αG	exogenous guanosine
ωG	3` splice site
P.li.	Pylaiella littoralis
O.i.	Oceanobacillus iheyensis

ACKNOWLEDGEMENTS

I want to express my deep appreciation to my PI and the committee chair for accepting me into his lab and helping me grow as a scientist. I am grateful to my defense committee for contributing their time, knowledge, and expertise. This journey would not have been possible without the support from Brian Leigh and the funding from the ACS Bridge Program at UCSD. I want to thank D. Toso at the Cal Cryo facility at UC Berkley for collecting the cryo-EM data used in this study.

I want to acknowledge Daniel Haack for showing me the ropes of the lab, teaching me all he knows, and helping with prepping grids. I am grateful to Jason Hingey for dedicating time to helping me through the rough patches and eating all the baked goods. Thanks to Boris Rudolfs for assistance with data processing, model building, and structural refinement. Many thanks to the rest of the Toor lab for the aid and moral support throughout this endeavor.

Thanks to my friends, especially Lindsay, for the continual emotional support. Lastly, I want to recognize my family, Mom, Dad, and Michael, for their constant support, unconditional love, and always believing in me.

VITA

2021 Bachelor of Arts in Biochemistry, University of San Diego

2023 Master of Science in Chemistry, University of California San Diego, Chemistry & Biochemistry

ABSTRACT OF THE THESIS

Investigation of *Tetrahymena* Group I Intron Structural Dynamics

by

Alexa Perez

Master of Science in Chemistry

University of California San Diego, 2023

Professor Navtej Toor, Chair

Group I introns are a class of catalytic RNAs that can perform a self-splicing reaction. This thesis will examine the large-scale conformational changes of three Tetrahymena group I intron (TET) states. A combination of mutational analysis and cryogenic electron microscopy (cryo-EM) was used to determine the dynamics associated with the L9/P5 mutation (ΔL9) in

which a long-range tertiary contact that affects the structure-function of TET is knocked out. This mutation allowed for the visualization of dynamics that were not seen in the previously published models of the wild-type (*wt)* construct. There are three different states of TET modeled: an intermediate form in which the intron is connected to the 5` exon, but the 3` exon is spliced; a second step (S2) state where the intron is base pairing to the ligated exons, and finally, a post-second step (post-S2) state in which the intron is fully spliced. Understanding the structural changes during the splicing of Δ*L9* TET provides specific insights into the structure-function relationships of group I introns and general mechanisms of RNA catalysis and dynamics. The structural biology methods developed to facilitate these experiments have also expanded the use of cryo-EM as a general tool for studying many different RNAs. Furthermore, by increasing our knowledge of RNA catalysis, we gain a more comprehensive grasp of these essential biological macromolecules and their role in the origins of life on Earth.

INTRODUCTION

Group I introns are a class of RNA that can both store genetic information and display catalytic activity. In the early 1980s, Tom Cech and Art Zaug discovered an RNA intron from the protozoan *Tetrahymena* thermophila that could excise itself from adjacent exons without a protein cofactor. [1] They could piece together the necessary reaction conditions that revealed RNA could catalyze chemical reactions. They coined the term “ribozyme” to describe a new group of catalytically active RNA. [1] Since then, a furious effort has been made to understand ribozymes' structure, functionality, and history.

GROUP I INTRON STRUCTURE AND FUNCTION

Group I introns perform a splicing mechanism in which two successive transesterification reactions occur. During the first step, an exogenous guanosine (αG) acts as the nucleophile that attacks the 5’ splice site and covalently attaches itself to the 5’ end of the intron (see Figure 1). [1, 2, 3] During the second step, the 3’-OH of the newly freed 5` exon attacks the 3’ splice site (ωG), leading to a final product of two ligated exons and the excised intron (see Figure 1). [1, 2, 3, 4]

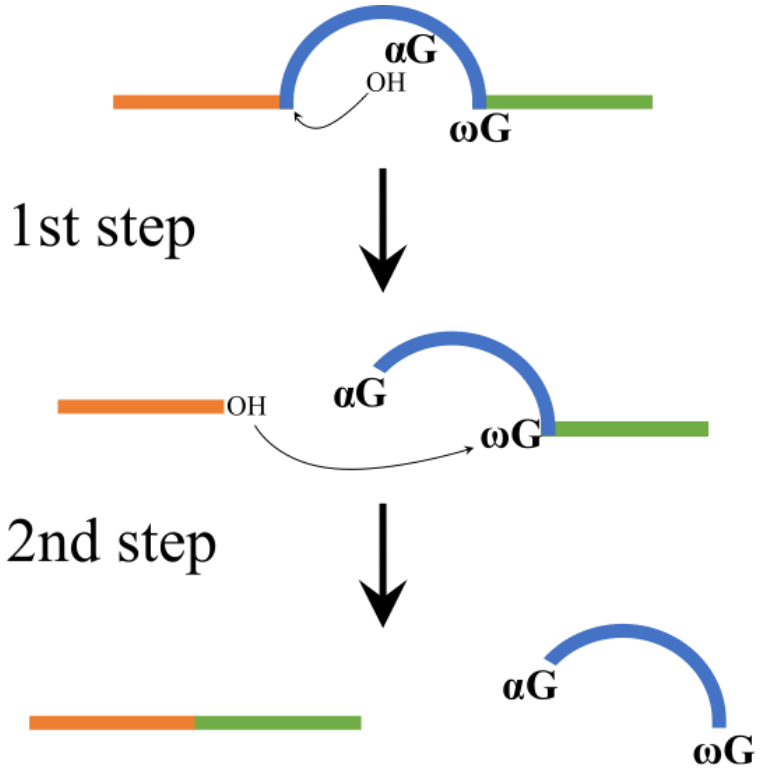

Figure 1: Two-step self-splicing mechanism. The exogenous guanosine attacks the 5'splice site in the first step. In the second step, the 3' OH attacks the 3' splice site, resulting in ligated exons and a linear intron. 5' exon is shown in orange, 3'exon is in green, and the intron is in blue.

Small yet significant pieces of the mechanism were revealed by investigating the conditions needed for intron splicing unveiled. First, the ωG nucleotide at the 3` end of the RNA sequence was not seen in the DNA sequence, suggesting it had a role in intron splicing. [1, 5] Guanosine triphosphate (GTP) and magnesium ions were also shown to be essential for the mechanism to occur. Since then, the precise roles of guanosine and magnesium have been well elucidated. [5, 6]

As previously discussed, two guanosines are significant for intron splicing: the αG and ωG. In the late 80s, the precise structure of the active site bound by guanosine was determined. [5] Since then, mutagenesis and biochemical splicing assays demonstrated the specific nucleotides

of the active site, some of which are directly involved in guanosine binding. [5, 7] These foundational studies led to a more comprehensive understanding of how the catalytic core was formed to facilitate the first step of splicing, transition to, and complete the second step. [8, 9]

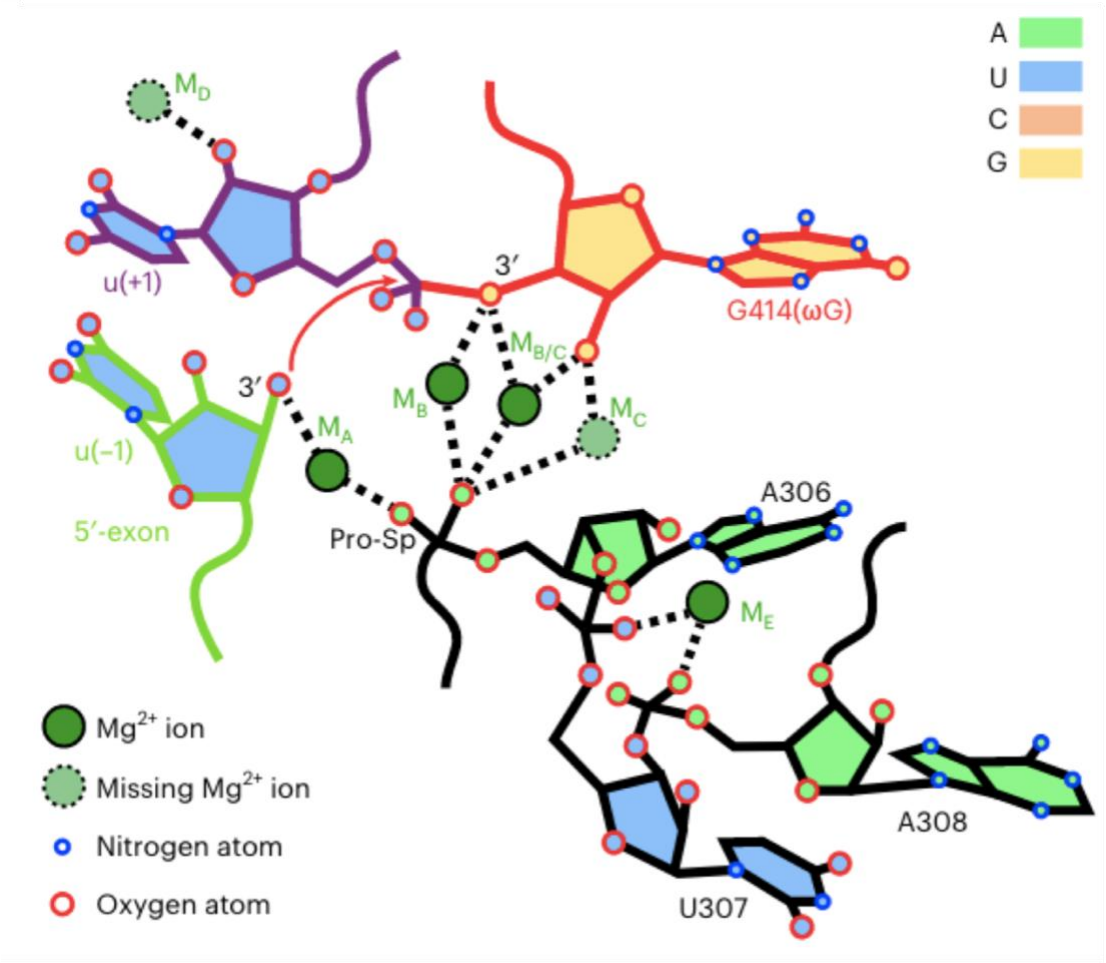

Figure 2: The guanosine binding site of TET during the second step of splicing. The ωG is shown (red) being coordinated by M_B, $M_{B/C}$, and M_C. M_A is coordinating the 5` exon (green). $M_D$ is coordinating the 3` exon (purple). U307 and A308 are being coordinated by M_E. This figure is taken from Luo et al. [10]

Magnesium is essential for the structural and catalytic coordination of the RNA. [11, 6]

Magnesium is the metal used for the two-metal ion mechanism of group I introns. One metal

initiates the nucleophile to attack the electrophile. [12] Another metal enhances the phosphorous atom electrophilicity and stabilizes the structure of the active site. [12] Various metals participate in the ribozyme's active site or peripheral domains during the first and second steps. For example, metals M_A-M_E have been identified as magnesium ions directly associated with splicing (Fig. 2). [13, 10] These positively charged metal ions pull on the negatively charged phosphate RNA backbone and coordinate them to allow for structural arrangements needed for the catalytic activity to occur. Figure 2 shows the metal ion coordination of the five core metals in the active site during the second step of splicing. Similarly, metal ions are also responsible for the active site coordination for the first step of splicing.

Group I introns have been shown to contain conserved sequence motifs and core secondary and tertiary structures. [14, 15] Starting in the early 1980s, various group I introns from fungal species were shown to resemble each other. [2] Similar comparisons were later found between nuclear and mitochondrial introns. [16] Sequence analysis and secondary structure comparisons have revealed that the three domains from Fig. 3 are found in some variety in all group I introns. [2, 17, 18] In addition, similar tertiary interactions between RNA secondary structures are found within this class of RNA. [19] These peripheral contacts help stabilize the overall structure and allow for proper RNA functionality. [20]

The TET intron comprises multiple stem-loops labeled P1-P10 and P13-P14 (Fig 3). The intron can be divided into the scaffold, substrate, and catalytic domains. The P4-P6 sections are shown, which comprise the scaffold domain responsible for initiating the folding of the other domains. [21] The substrate domain contains the internal guide sequence (IGS), P1 (not shown), and P2, which are essential for substrate recognition. [22] The IGS directs the formation of the P1 stem, the beginning sequence of the intron base-pairing to the 5` exon. [23, 24] The IGS is also

responsible for the P10 stem formation, in which the intron base pairs to the 3` exon. [13] P13 and P14 are other segments of the substrate domain that are necessary for structural integrity. P13 is a pseudoknot formed between P2 and P9, and P14 is a pseudoknot peripheral interaction between P2 and P5 (Figure 3). [13] The catalytic domain is made by sections P3 and P7-P9, responsible for forming the guanosine binding site. [25]

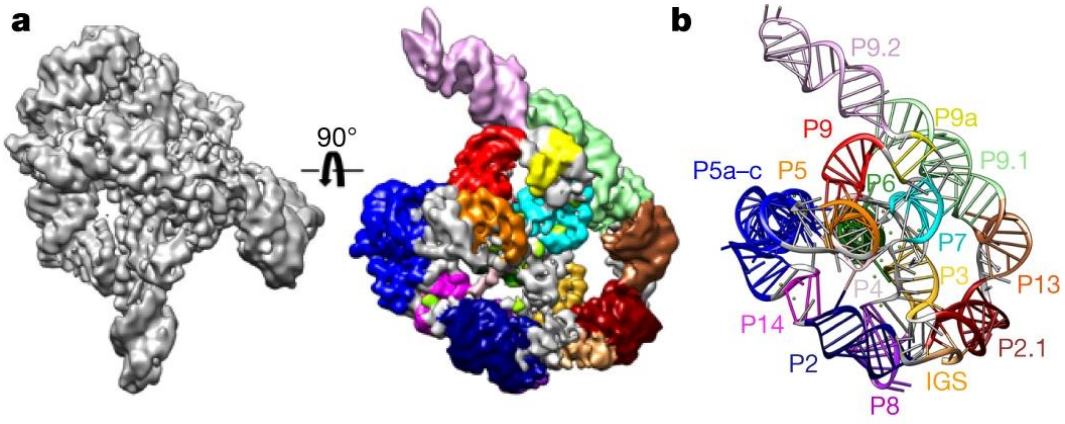

Figure 3: L16 ScaI ribozyme. a) 3D density coloring and labels as shown in b) structure with the helical domains labeled and colored. The P1 and P10 stems are not shown due to the removal of exon sequences from this model. The figure is taken directly from Su, Zhang, Kappel, *et al.*

TET is the model system for studying RNA catalysis and structural dynamics in group I introns. [1, 26] Mutational analysis followed by kinetic studies shows how alteration of TET structure affects the splicing rate at different stages of splicing. [17, 18] Structural studies have provided evidence that explains the structure-function relationship of group I introns and how various functional groups and intermolecular forces between specific bases stabilize different transition states. [18, 19] X-ray crystallography produced the structures of these transition states. [22,30,31] Then, cryo-EM became a significant technique that expanded the resolution and

knowledge of these intermediate structures. [11, 23, 24] Since then, more intermediate structures of TET have been observed at all splicing stages. [12, 25, 26]

GROUP I INTRON EVOLUTION AND HISTORY

Arguably, the most significant implication of the discovery of group I introns as catalytic RNAs is their involvement in the origins of life. The RNA world hypothesis theorizes that life originated with various complex catalytic RNAs that could self-replicate. [36] Ribozymes are thought to play an essential role in the chemical origins of life due to their ability to provide genetic material and carry out catalysis without the need for proteins. [37] Understanding the structure-function relationships of catalytic RNA, such as group I introns, can provide a foundation for explaining how RNA was involved with the beginning of life on Earth.

Additional examples were found after discovering the group I introns, including a new class of splicing RNAs termed group II introns. [38] Like group I introns, group II introns also use a two-step mechanism of successive transesterification reactions. [39] Furthermore, both groups use similar types of tertiary interactions for connecting the various secondary helical paired segments. [40] Figure 4 compares the splicing mechanism for both groups. Both mechanisms result in the ligation of exons and the intron excision; However, the group II intron forms a lariat bond. Another critical feature of group II introns is the internal adenosine branch site that acts as the nucleophile of step 1 versus the exogenous G of group I introns. [41] Despite these differences, the overall mechanism and chemistry of splicing strongly resemble group I introns. [42]

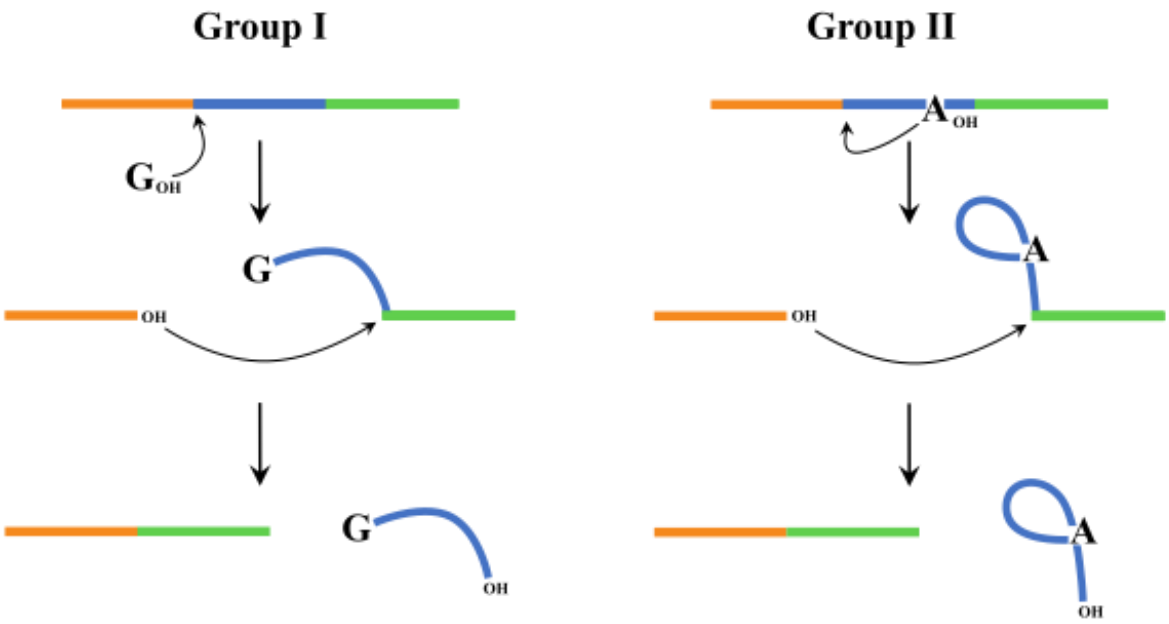

Figure 4: Comparison of the self-splicing mechanism of group I and group II introns. 5` exon is shown in orange, intron in blue, and 3` exon in green.

Tracing the history of group I introns has shown the convergent evolution of similar functions of group II introns. Both classes of introns can catalyze their excision from within introns and function as a retroelement following the completion of the splicing reaction. Despite their similarities, phylogenic investigations have reported that these groups do not share any conserved sequences, structure, or splicing mechanism. [38] Furthermore, there are similarities between these two groups and the spliceosome. Group II introns and the spliceosome have the same splicing mechanism. [43] The main difference is that the spliceosome requires a complex formation of proteins and RNA to splice the nuclear pre-mRNA. [44, 45]

MATERIALS AND METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

The DH5α strain of E. coli cells (Thermo Fisher Scientific) was used to prepare DNA circular plasmid for in vitro transcription. The cells were grown in Luria-Bertani (LB) broth. They were incubated at 37°C overnight and recovered via centrifugation at 5,000 RCF. Post-induced cells were incubated at 22°C for two days. Cells were harvested via centrifugation.

PLASMID CLONING

All genes were synthesized (IDT or Genscript) and cloned into a pUC57 vector using the EcoRV restriction site. The cloned plasmids were then transformed into the DH5α cells. For all cryo-EM experiments, the mutant genes contained a 9-nt 5′ exon and a 9-nt 3′exon followed by a HindIII cut site. All TET constructs used for the in vitro splicing assays contained a 150-nt 5′ exon and 90-nt 3′ exon followed by a HindIII cut site. The Δ*L9`* construct contained the L9/P5 mutation. The ΔL9 construct used for cryo-EM samples had this mutation, and two non-conserved areas, stem-loops P5b and P9.1a, were removed, and an extra sequence was added to increase the solubility of the Δ*L9* construct.

IN VITRO RNA TRANSCRIPTION

The plasmids were linearized using the HindIII restriction site (NEB). Approximately 40 μg of template DNA was added to a total volume of 1 mL of *in vitro* transcription buffer (50 mM Tris-HCl pH 7.5, 25 mM $MgCl_2$, 5 mM DTT, 2 mM spermidine, 0.05% Triton X-100, and 5 mM of each NTP). Under low magnesium conditions, 10 mM $MgCl_2$ was used. T7 RNA polymerase was used to initiate transcription. Thermophilic inorganic pyrophosphatase was added to

minimize the buildup of pyrophosphatase precipitate. The reaction mixture was incubated at 30°C for 3 hrs. $CaCl_2$ was added to a final concentration of 1.2 mM along with Turbo DNase and incubated again at 30°C for 45 minutes to digest the DNA template fully. Proteinase K was added and incubated at 30°C for an additional hour. The resulting solution was centrifuged to remove any precipitate and filtered through a 0.2 μm filter. The filtered solution was buffer exchanged 7 times, each time using 14 mL of filtration buffer (5 mM Na-cacodylate pH 6.5 and 10 mM $MgCl_2$) through a 100 kDa molecular weight cut-off filter. After the final buffer exchange step, the RNA was concentrated to approximately 1 mg/mL for use in downstream experiments.

CRYO-EM SAMPLE PREPARATION AND DATA COLLECTION

3 μl of RNA sample was loaded onto glow discharged (Pelco Easiglow), plasma cleaned Quantifoil R1.2/1.3 copper 300-mesh grids (Quantifoil). The grids were blotted before being plunged into liquid propane/ethane cooled by liquid nitrogen using a manual plunger under > 80% humidity at 4°C. Cryo-EM data was collected on a Titan Krios electron microscope (ThermoFisher Scientific) operating at 300 keV and equipped with a K3 direct electron detector (Gatan) at UC Berkeley. Micrographs were recorded in a Titan Krios microscope (FEI) operating at 300 kV, equipped with a K3 direct electron detector (Gatan). A nominal magnification of 105,000x (0.811 Å per pixel) and a defocus range of −0.5 μm to −2.0 μm was used for imaging the sample using the super-resolution camera setting (0.4055 Å per pixel). Each micrograph was collected as 50 dose-fractionated frames, with a dose rate of 8 e-/pixel/s, a total exposure time of 2.72 s, and a frame exposure time of 0.05 s, resulting in a total dose of ~50e-/Å2. In total, 7516 micrographs were collected using beam shift in SerialEM.

DATA PROCESSING

A total of 7516 micrographs movie stacks were motion corrected and binned to 0.811 Å/pixel using patch motion correction in cryoSPARC 4.2.0. The average defocus level of each aligned micrograph was determined with PatchCTF. 1,243 micrographs were selected for preliminary data processing. 702,001 particles were selected using cryoSPARC's Gaussian blob picker. These particle images were subjected to 2D classification, and 240,962 particles were selected to generate a 3D ab initio model. Projections of this model were then used as templates for auto-picking, which yielded 4,016,904 particles from 4,995 micrographs using the cryoSPARC template picker. The particle images were extracted with 2.5-fold binning and were subjected to one round of 2D classification, followed by two rounds of 3D classification using hetero-refinement. The 543,099 selected particles were re-centered and re-extracted with 1.5-fold binning for a final pixel size of 1.22 A/pixel. These particles were subjected to another round of 3D classification using hetero-refinement, and the resulting 412,382 particles were refined to a 3.83 Å map. Particles were then subjected to defocus correction using local CTF refinement, and non-uniform refinement of these particles yielded a 3.08 Å map. 212,478 of these particles were 3D classified into 10 classes using the clustering mode of 3D variability analysis. Images belonging to three distinct catalytic states, intermediate (23794 particles, 3.9 Å), S2 (27360, 3.8 Å), and post-S2 (23961, 4.0 Å), were selected for refinement and modeling.

MODEL BUILDING AND STRUCTURE REFINEMENT

All modeling was performed in COOT. [46] The coordinates were refined in real space using PHENIX and COOT. [47, 48] Models for the intermediate, S2, and post-S2 maps were initiated from PDBs 7YG9, 7YGB, and 7YGD, respectively. [35] The P9 helix and exons were

remodeled in COOT, and then coordinates were refined in real space using PHENIX. All software was compiled by SBGrid.[49]

CHAPTER 1: CONFORMATIONAL DYNAMICS OF GROUP I INTRON SPLICING

INTRODUCTION

This chapter will focus on the structural dynamics of the Tetrahymena group I intron during catalysis by using both mutational analysis and cryo-EM. This method has been previously used to identify conformational dynamics during the catalysis of a group II intron. [43] In this study, RNA domains were observed engaging in large-scale movements to shuffle substrates and products in the active site between the two steps of splicing. [43] Due to the similarities between these two splicing systems, a similar structure-function relationship was believed to exist for group I introns.

The initial goal of this project was to attempt to capture the pre-catalytic structure of TET. Long-range tertiary contacts impact the structural stability and dynamics of RNA. [27] For TET, five key tertiary interactions are essential for different catalytic roles. [26] These peripheral contacts, when mutated, have been shown to change the rate of individual steps of splicing. [26, 27] Mutation of the L9/P5 (Δ*L9*) interaction in the TET group I intron decreases the binding rate and affinity for guanosine, which allows for a slower reaction rate for the first step of the splicing. [26] Furthermore, the Δ*L9* was chosen to maintain the ribozyme's overall structure and catalytic ability. It was predicted that Δ*L9* could slow down the initial step of splicing to visualize the αG engaging in the nucleophilic attack on the 5` splice site. Due to the dependence on magnesium for efficient group I intron splicing, it was hypothesized that TET, in its pre-catalytic state, could be trapped with the right conditions in conjunction with the mutation. Transcription conditions such as temperature, magnesium concentrations, and different purification methods were used to

get the precursor RNA to be the predominant product. Under the right conditions, we could visualize this outcome on denaturing PAGE for both *wt* and Δ*L9* TET (Fig 6).

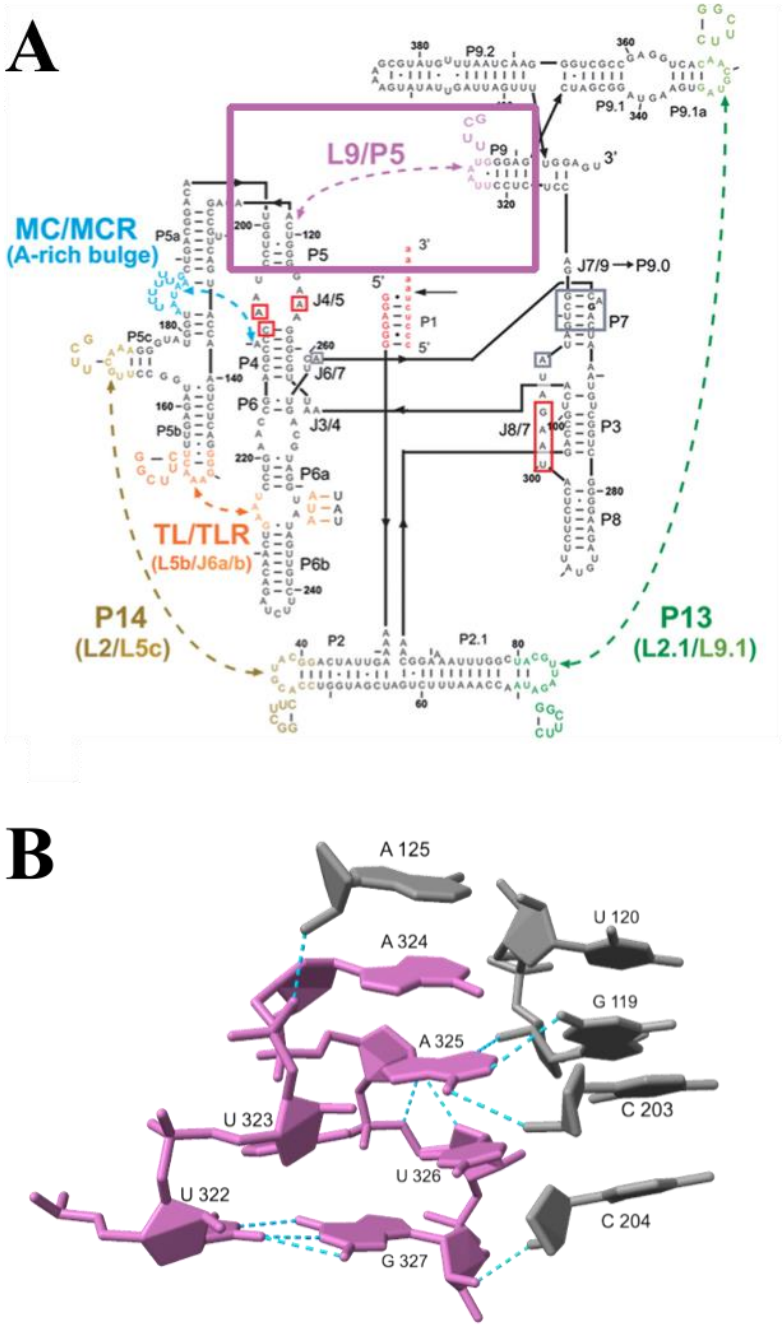

Figure 5: L9/P5 interaction A) TET secondary structure. The L9/P5 interaction is boxed in purple. The mutated stem-loop sequence is shown above. B) Hydrogen bonding (blue) of intact L9/P5 interaction. The coloring matches Figure 5A, taken from Benz-Moy and Herschlag. [26]

Two recent studies using alternative methods to obtain a pre-catalytic structure of wt TET analogs were published during this time. Due to these publications, the project's goal shifted, and the Δ*L9* TET construct, under standard conditions, was prepped and sent for data collection via cryo-EM as described in the methods section. My new goal was to understand the structural changes caused by the L9/P5 long-range interaction mutation. Understanding these dynamics could lead to a possible explanation of what role these peripheral interactions play a significant role in the structure-function paradigm of TET. Furthermore, insight into how this mutation affects the structure could lead to a clearer understanding of previously seen kinetic deviations. [26]

By understanding the specific structural changes of TET, I hoped to provide a structure-function model for the overall mechanism of group I intron catalysis. Due to the similar mechanisms between the various classes of introns, this study will also provide a framework to investigate similar RNA interactions in group II introns and the spliceosome. [1, 50, 51] Therefore, the method development and design used to obtain these structures will expand the possibilities in which cryo-EM can be applied to understanding RNA structure and dynamics.

RESULTS AND DISCUSSION

IN VITRO RNA TRANSCRIPTION OPTIMIZATION

Transcription conditions of TET were optimized to produce the highest yield and purity of precatalytic precursor RNA. After transcription, the RNA was purified and analyzed via denaturing polyacrylamide gel electrophoresis (PAGE). First, the temperature was varied during in vitro transcription. The Δ*L9* and *wt* constructs were transcribed at 30℃, 37℃, and 45℃ (Fig. S1). There was no RNA transcription occurring at 45℃ for both constructs. The higher intensity of the precursor RNA band at 30℃ compared to 37℃ suggests that transcription at a lower temperature was optimal for the desired outcome. Next, the transcription time was varied, but no significant difference in RNA production was observed (data not shown).

Different magnesium concentrations during transcription significantly affected the overall RNA production and the level of splicing products produced. The standard transcription buffer used contains 25 mM $MgCl_2$. At this concentration, determining what catalytic states were present based on size was challenging due to the high amount of alternative splicing products formed (Fig. 6). The magnesium concentration was lowered to determine some of the products being formed. At 10mM $MgCl_2$, fewer alternative splicing products were formed, suggesting that lower magnesium results in inhibited splicing of both *wt* and Δ*L9* (Fig. 6). Interestingly, transcription with 10mM $MgCl_2$ resulted in both the *wt* and Δ*L9* precursor RNA being the primary product. Furthermore, Δ*L9* with lower magnesium concentrations produced fewer splicing products than *wt*. Together, these results support the initial hypothesis that the mutation of the L9/P5 interaction under the right conditions would prevent the splicing of this group I intron.

While the precursor of both constructs was captured, the wt TET structures were already published when these results were produced. [12, 22] Despite the structures of TET analogs displaying pre-S1 states, there has yet to be a structure of an unmodified TET with αG positioned for the first step of splicing. This precursor *wt* RNA could be incubated with magnesium and GTP and immediately prepared for cryo-EM to capture the pre-catalytic state of TET with naturally occurring exons and no intron modifications.

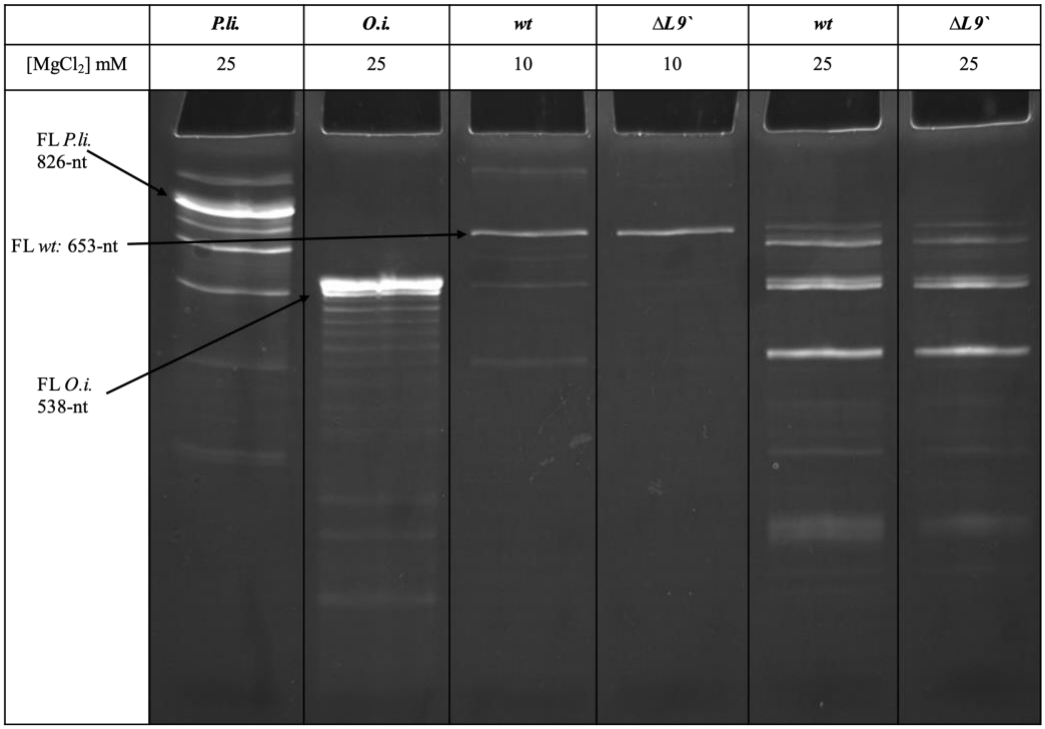

Figure 6: Image of 4% denaturing polyacrylamide gel stained with ethidium bromide. Lanes 1 and 2 show full-length *P.li.* and *O.i.* size markers. Lanes 3-6 show *wt* and Δ*L9* TET RNA transcribed with 10 or 25 mM $MgCl_2$.

To find out if this mutation would preserve overall RNA catalysis, transcription splicing products were visualized via denaturing PAGE. Comparison between *wt* and Δ*L9* TET showed

that the L9/P5 mutation did not inhibit intron splicing under standard transcription conditions, supporting the claim that overall splicing kinetics are unaffected by this mutation (Fig 6). [26]

CRYO-EM DATA COLLECTION AND REFINEMENT

Table 1 shows cryo-EM statistics for this study. The complete data set contained 476,000 particles and produced a consensus map at 3.08Å resolution. The three catalytic states, intermediate, S2, and post-S2, had maps of resolution 3.9Å, 3.8Å, and 4.0Å, respectively. After real space refinement, the maps showed a good correlation with the models, with an FSC map to model resolutions of 4.1, 3.9, and 4.2 angstroms.

Table 1: Cryo-EM data collection and refinement statistics.

Data collection	Full data set	Intermediate	S2	post-S2
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300	300	300
Camera	Gatan K3	Gatan K3	Gatan K3	Gatan K3
Magnification	105,000	105,000	105,000	105,000
Nominal defocus range (μm)	0.5-2.0 um	0.5-2.0 um	0.5-2.0 um	0.5-2.0 um
Exposure time (s)	2.7	2.7	2.7	2.7
Number of frames	50	50	50	50
Dose rate (e-/Pixel/s)	8	8	8	8
Total dose (e-/Å2)	50	50	50	50
Pixel size (Å)	0.811	0.811	0.811	0.811
Micrographs collected	7516	7516	7516	7516
Micrographs processed	4995	4995	4995	4995
Total particles	476,053			
Particles used in final map	476,053	23794	27360	23691
Map resolution global (FSC 0.143)	3.08	3.86	3.75	4.01
Map sharpening B-factor (Å2)	-109	-73.8	-68.9	-88.9
Model composition				
Non-hydrogen atoms		9384	9784	9108
RNA bases		439	458	426
Refinement (PHENIX)				
Refinement package			Real space refinement	
CC (volume)		0.84	0.83	0.84
CC (mask)		0.85	0.84	0.85
CC (peak)		0.8	0.78	0.8
FCS map-to-model (0.5)		4.1	3.9	4.2
Rms deviations				
Bond length (Å)		0.013	0.013	0.011
Bond angles (°)		1.034	1.092	1.124
Validation				
Molprobity score		3.29	3.32	3.42
All-atom clashscore		46.01	49.23	63.04

After collecting data via cryo-EM, seven maps were classified that capture ΔL9 in three different catalytic states. Structures for each state were then modeled into these maps. These states are shown in Figure 7 and can be separated into three categories. Figure 7A represents an intermediate state of splicing where the 5` exon and intron are present. Figure 7B represents the second step of splicing in which the intron is base-paired to the ligated exons. The third state is the post-second step, in which the intron no longer pairs with the exons (Fig. 7C).

The core structures of the Δ*L9* models were comparable to the *wt*, supporting the hypothesis that this mutation would maintain the overall RNA structure. More extensive structural changes seen in the Δ*L9* models seem to be localized to the site of mutation and the remainder of the P9 domain. The fact that the core structure remains intact could explain why the splicing activity remains efficient despite the L9/P5 mutation.

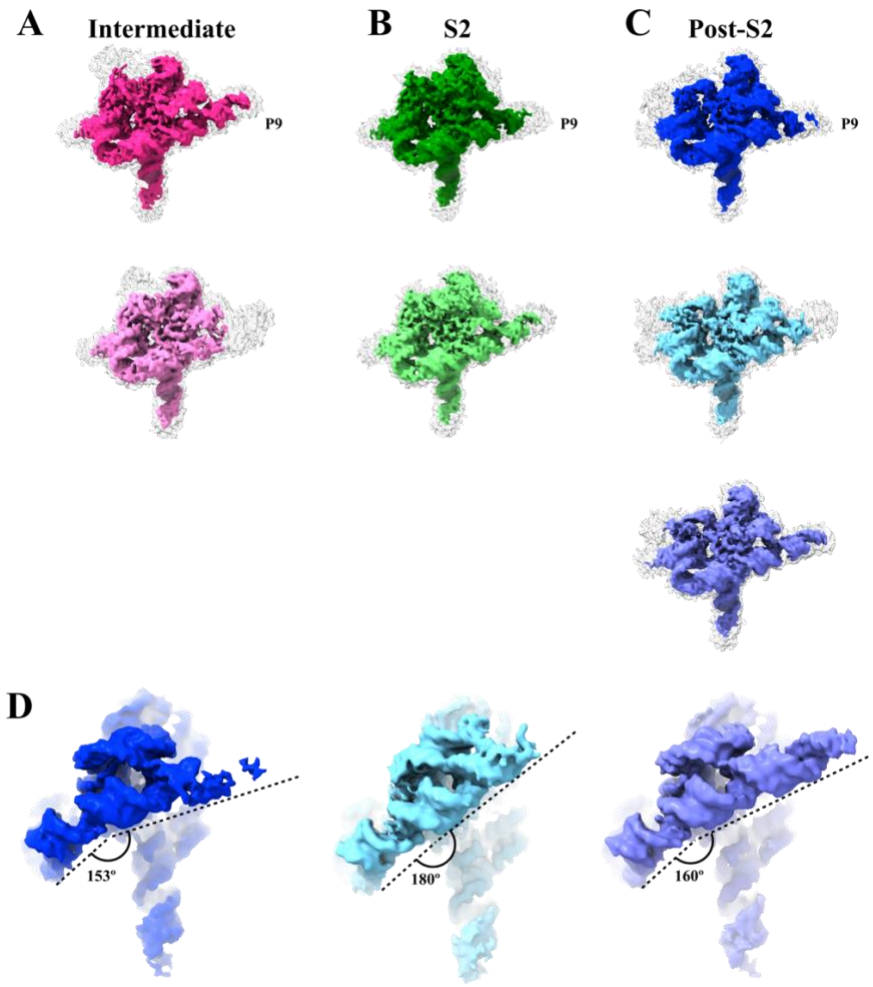

Figure 7: Cryo-EM maps of TET. A) Intermediate state maps of TET with the 5` exon and intron present. B) Second step state maps of TET with the intron and ligated exons present. C) Post second step state maps of TET with just the intron present. D) Measurement of the angles of the P9 stem for post-S2 maps.

The differences between maps of the same state were localized to the P9 stem (Fig. 7). For example, Fig. 7d highlights the different angles of the P9 stem within maps of the same state. The angle between the base of P9 and the stem varies between 153º to 180º while the core of the RNA shows these three models represent pst-S2 of splicing. Since the mutation made was in the P9 domain, this could explain why different stem orientations are seen. The angle of the P9 domain can be a downstream result of the broken connection between the L9 and P5 stem.

These maps provide structural evidence that the mutation has successfully knocked out the peripheral contact. Despite the differences in content, each map lacks connective density where the L9/P5 interaction should exist. In addition, the models in Figure 7 show the distances between the mutated L9 stem-loop and the G119 from P5 for each of the three states. The distances are 11.647, 13.182, and 10.821 Å for the intermediate, S2, and post-S2 states (Fig. 7). The nucleotides are shown to be pointing away from each other, supporting the claim that this interaction no longer exists in the Δ*L9* TET construct. Further modeling needs to be done to determine if there is a correlation between the angle of P9 and the distance of separation between the L9 and P5. Together, this structural data and the results from the gel in Fig. 5 support the hypothesis that this mutation of TET would preserve the overall structure and splicing activity of the RNA.

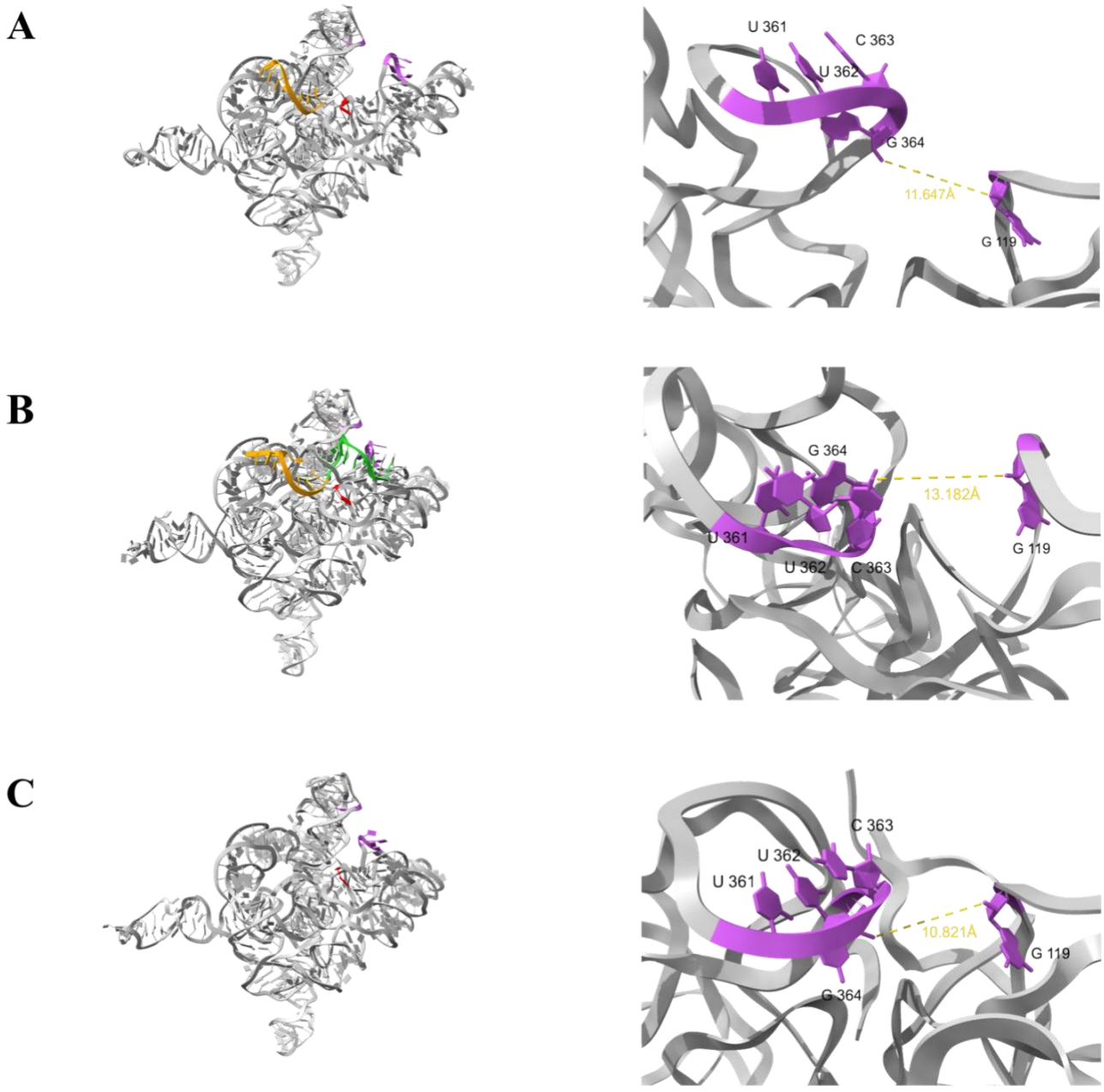

Figure 8: Models of TET and measured distance between L9 and P5. The intron is in grey, L9/P5 interaction is in purple, and the ωG is in red. A) Intermediate intron state with 5` exon (orange). B) S2 state of the intron with 5` exon (orange) and 3` exon (green). C) post S2 state of the intron.

INTERMEDIATE STATE

The intermediate state of Δ*L9* TET captured is noteworthy because the 5` exon is present alongside the intron, but the 3` exon is not (Fig. 7A and 8A). This data shows a never-before-seen structure of TET in which the 3` exon has been spliced from the intron before the 5` exon, which remains bound to the intron in this structure. Traditionally, the 5` exon is spliced in the first step, while the 3` exon remains connected to the 3` end of the intron (Figures 1 and 4). The ωG is required for the excision of the 3` exon in the second step of splicing. It also remains in the intron after complete splicing and release of exons (Fig 1). Interestingly, ωG is the last nucleotide present in the densities of the maps for this intermediate state (Fig. 7A), suggesting that the 3` exon is now free.

There is no connective density between the splice site and where the 3` exon should be. This supports the claim that the exon has been spliced and is no longer attached to the intron. One possibility is that the ωG was hydrolyzed, allowing for the 3` exon excision. In this scenario, water could have catalyzed the independent splicing of the 3` splice site. This could be due to the L9/P5 mutation allowing for stretching the 3` end and potentially increasing solvent accessibility.

SECOND STEP STATE

The second step state of Δ*L9* TET in this data set represents the state where the intron forms the P1 and P10 stems with the ligated 5` and 3` exons, respectively (Fig. 7B and 8B). Figure 9 shows a side-by-side model comparison of Δ*L9* and *wt* TET intermediate states. Our model was directly compared to each second-step model that has been published. The model shown in Fig. 9b was the only other model that displayed the intron fully spliced and base-paired to the ligated exons. [35] Both *wt* and Δ*L9* contain a density that suggests a continuous helix in which the intron is base pairing with both exons (Fig. 9a). Due to the shorter substrates compared to Δ*L9* exons, only part of this helix is seen for the *wt* model (Fig. 9b).

Despite being in the same state, the Δ*L9* active site has some key differences compared to the published *wt* model. [35] In the *wt* model, these substrates differ from the naturally occurring sequences. For example, Fig. 9b shows how the sequence U408, A409, and C410 that comes right before the ωG is missing from the *wt* model. [35] In the Δ*L9* model, the nucleotides in this region occupy a clear density in which U408 points down, parallel to the end of the 3` exon, and A409 and C410 are positioned in a similar plane to ωG (Fig. 9a).

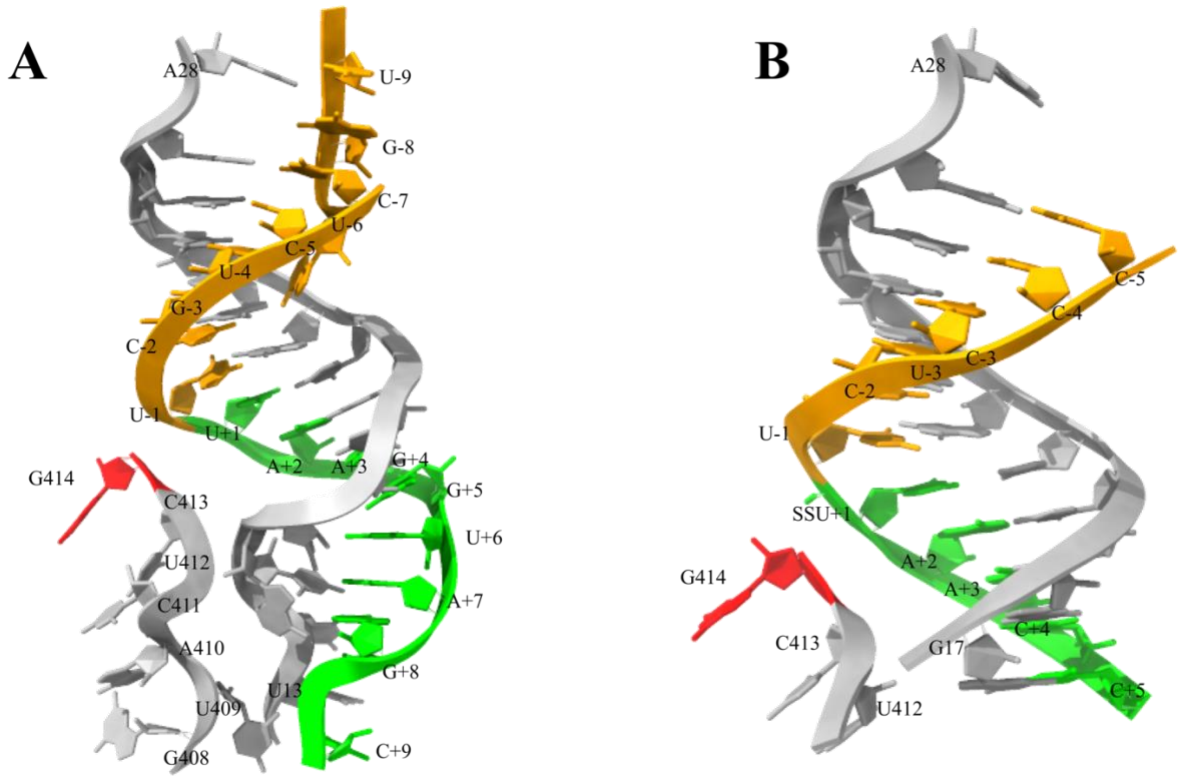

Figure 9: Close-up of the active site during the second step of splicing. Models were renumbered to fit standard numbering rules. The coloring is the same as in Figure 7. A) S2 state model displaying the intron base-paired to both exons. B) Conformation 3 of wt TET from Zhang et al.[34]

POST SECOND STEP STATE

Post S2 state map in Fig. 7C and model in Fig. 8C show the intron after both exons have been ligated and released. Compared to the other states, the 5` end of the intron was more difficult to model due to poor density. The sequence that could be confidently modeled points outward from the rest of the intron, unlike the other models that fold back toward the remainder of the intron (Fig 10a). One possibility is that the base pairing between the intron and the 5` exon restricts the 5` end from moving outward. After the ligated exons are released, the 5` end can move around, becoming more dynamic.

Interestingly, the 3` end of the intron did not follow this same trend. While the 5` end of the post-S2 intron was in a drastically different position compared to the other states, the 3` end of the intron points followed a similar pattern (Fig. 10b). The last three nucleotides U412, C413, and G414 are shown for clarity. The variation in orientation could be due to the different angles of the P9 stem between these states. More modeling would need to be done to compare maps of each state with a similar positioning of the P9 stem to see if there is any correlation with the 3` end of the intron. G414, ωG, is essential for the second step of splicing. Structural refinement at this region will provide information on how Δ*L9* can maintain overall splicing efficiency based on the active site positioning for the second step.

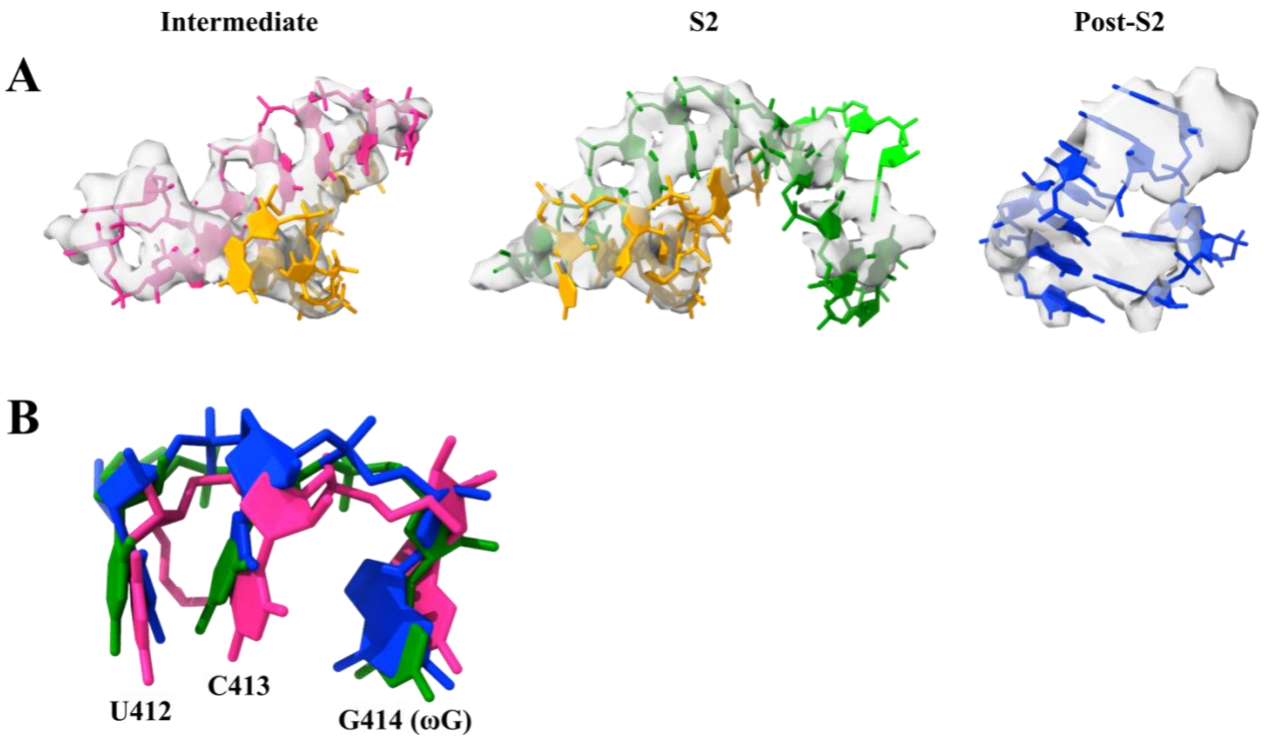

Figure 10. Comparison of 5` and 3` end of the three ΔL9 states. A) Maps and models of the intermediate (pink), S2 (green), and post-S2 (blue) states. The 5` exon is shown in orange, and 3` exon is shown in lime green. B) Overlay of the 3` end of all three states of TET.

FINAL CONCLUSIONS

To summarize, three different stages of splicing of ΔL9 TET were collected from this data set. First was the intermediate state in which the 3` exon was excised before the 5` exon, which is base-paired to the intron (Fig. 7A and 8A). Secondly, the S2 state showed the intron base pairing to the ligated exons (Fig. 7B and 8B). Finally, the fully spliced intron was free of both exons (Fig. 7B and 8B). In addition, these maps in Fig. 7 and the models shown in Fig. 8 provide structural evidence that the mutation successfully knocked out the L9/P5 peripheral contact. It was also hypothesized that this mutation would not cause deleterious changes to the structure or the splicing capabilities of the intron. The data presented above support this idea.

Due to time constraints, the final seven maps were split into three groups based on overall similarities (Fig. 7). One structure was modeled per group using the best map based on resolution and fit (Fig. 8). More modeling needs to be done to obtain a structure for each of the remaining maps. From this point, each model will be evaluated and compared to the published models of the corresponding catalytic state, similar to the abovementioned models. The next step will be a thorough investigation of the effects of the L9/P5 mutation on the structure that could explain the change in the kinetics compared to the *wt* TET. In addition, a closer look will be taken at the relationship between this mutation and the 3` end dynamics of the intron seen in Figures 7 and 8.

One question remains whether the 5` exon in the intermediate state is covalently bound to the intron. The density at the 5` splice site is very weak; therefore, while we can determine the base pairing between the 5` exon and the intron, it is unclear whether the 5` exon has been excised. This leads to the question of whether 5` exon release is dependent on the ligation of the 3` exon. It is clear that this intermediate has released the 3` exon individually. However, the ability of the

5` exon to be released individually remains unknown. The answer to how exactly exons are released from the introns after splicing has yet to be determined. This splicing intermediate could provide new insight into how the exons rely on ligation for proper release.

An interesting conclusion made from this study is the various positions of the P9 segment between and amongst the various states. Our results show that the mutation of the L9/P5 interaction maintains the splicing rate; However, the product distribution varies compared to the wt construct (Fig. 6). For example, the amount of G-independent product formation increases significantly when this mutation is introduced (Fig 6). In the typical splicing mechanism, the first step product is the free 5` exon, and the intron ligated to the 3` exon. This is followed by splicing the 3` exon to produce the second step products of ligated exons and free intron. Therefore, we assume that the products of each step would be the main products visualized on a gel and in the cryoEM data collection. However, based on the cryoEM data collected in this study, we see that the Δ*L9* forms the intermediate state as one of the main products. This intermediate is not typically seen as a main product since it requires the 3` exon to be removed while the 5` exon remains base-paired to the intron. One theory is that the mutation causes the first step to be slowed down, allowing the 3` exon to be hydrolyzed by water. However, the density of the 5` splice site is too weak to determine whether or not it has been spliced.

Additionally, since the release of exons remains unknown, it could be possible that the 5` exon needs to be ligated to the 3` exon to be released if it has already been spliced. Overall, this major product changes the mechanism by decoupling the two steps. Therefore, we can conclude that the L9/P5 mutation is not only responsible for the rate of the first step but also has a role in coupling the first and second steps.

Three main functional domains of group I introns have been previously determined. The scaffold domain is essential for initiating RNA folding, the substrate domain is responsible for substrate recognition, and the catalytic domain is where the substrate binds (Fig. 11). While the mutation ΔL9 is located in the catalytic domain, we see large-scale structural changes in the area of the RNA that is not occupying any of the three domains. Due to the combination of these structural differences and the apparent decoupling of exon splicing and possible exon release, I propose the possibility of a fourth domain called the "coupling domain," which consists of the P9a, P9b, P9.1, and P9.2 segments (Fig. 11). Separating this domain from the rest of the RNA reveals a difference in splicing product distribution, but not splicing efficiency or RNA core structure. Furthermore, we now have a potential explanation for the functional relationship of this part of the ribozyme structure.

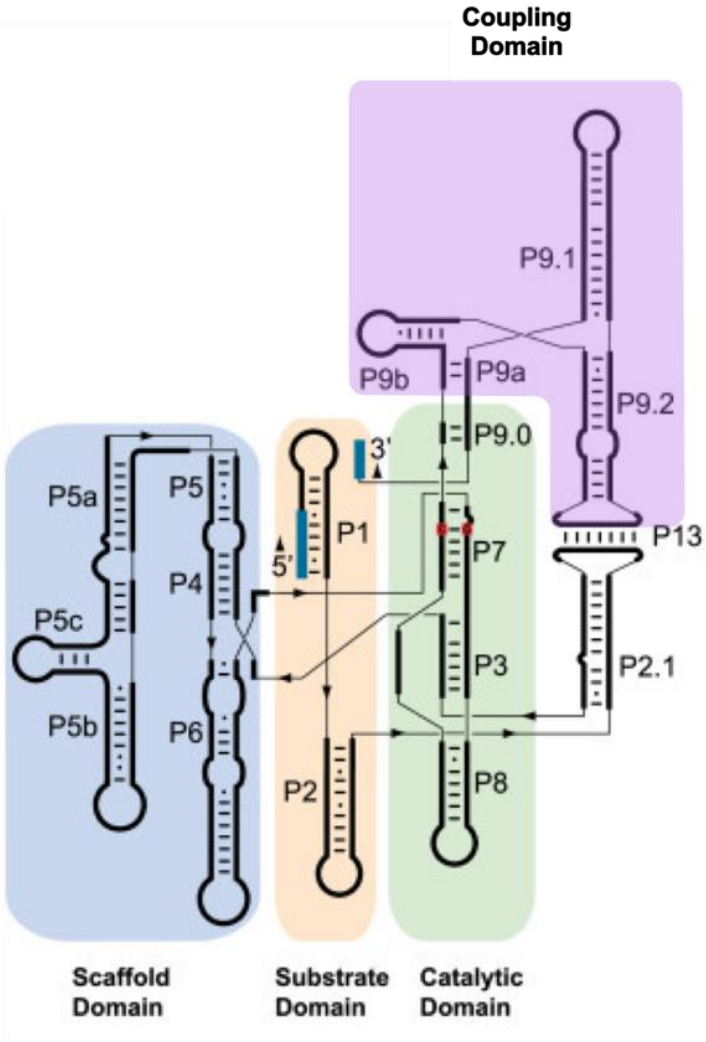

Figure 11: Coupling domain (purple), a potential new domain of group I introns. The three previously determined helical domains of group I introns scaffold domain (blue), substrate domain (orange), and catalytic domain (green). The secondary structure of TET is shown in this figure.

The L9/P5 contact is conserved among group I introns. Therefore, the results we have seen for TET could be extrapolated to other group I introns. Further understanding of TET structure and how this mutation and others alter the splicing rate provides the groundwork for investigating analogous interactions in other self-splicing catalytic RNAs. Cryo-EM structure analysis could also be done for the four remaining long-range interactions described by Benz-

Moy and Hershlag [26] to determine how these mutations affect the kinetics of splicing at various steps.

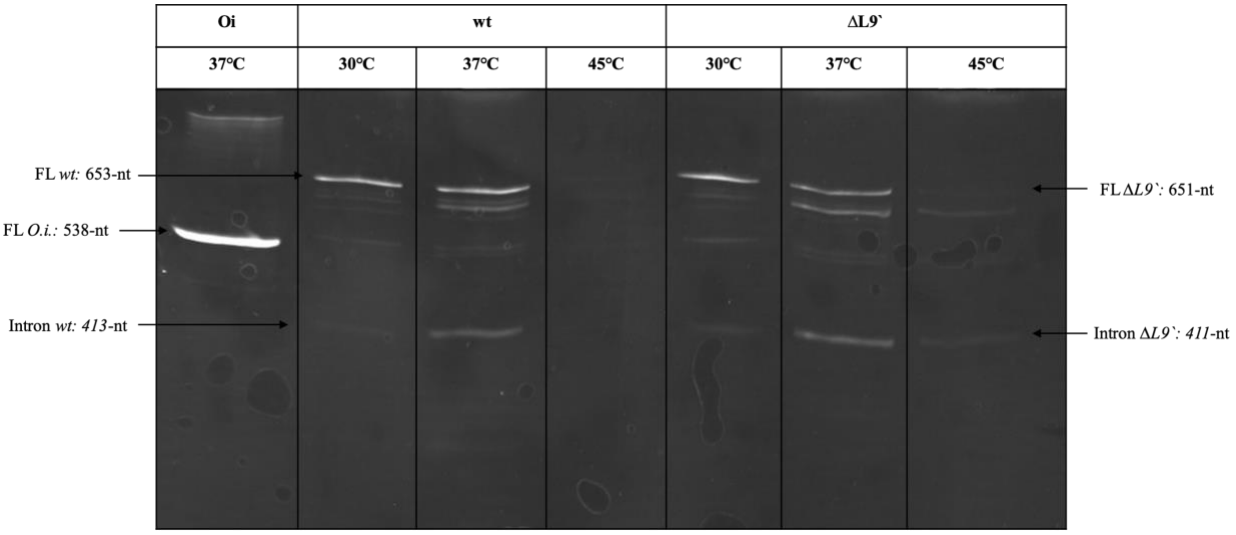

Figure S1. 4% denaturing PAGE stained in ethidium bromide. Transcription of the *wt* and Δ*L9`* were run at 30°C, 37°C, and 45°C. *O.i.* was used as a control and size marker.

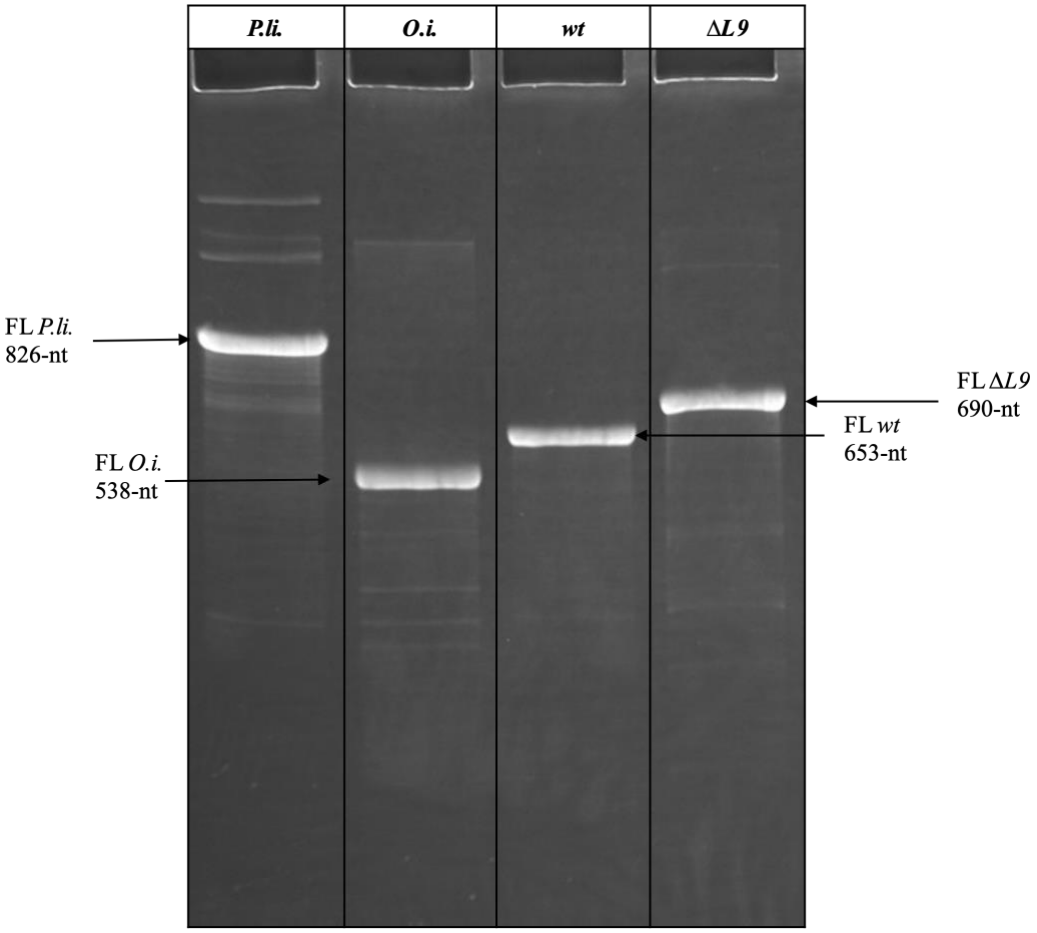

Figure S2. 4% denaturing PAGE stained in ethidium bromide. Transcription products of *wt* and Δ*L9*` with 10 mM $MgCl_2$ during transcription. *P.li.* and *O.i.* are used as size markers.

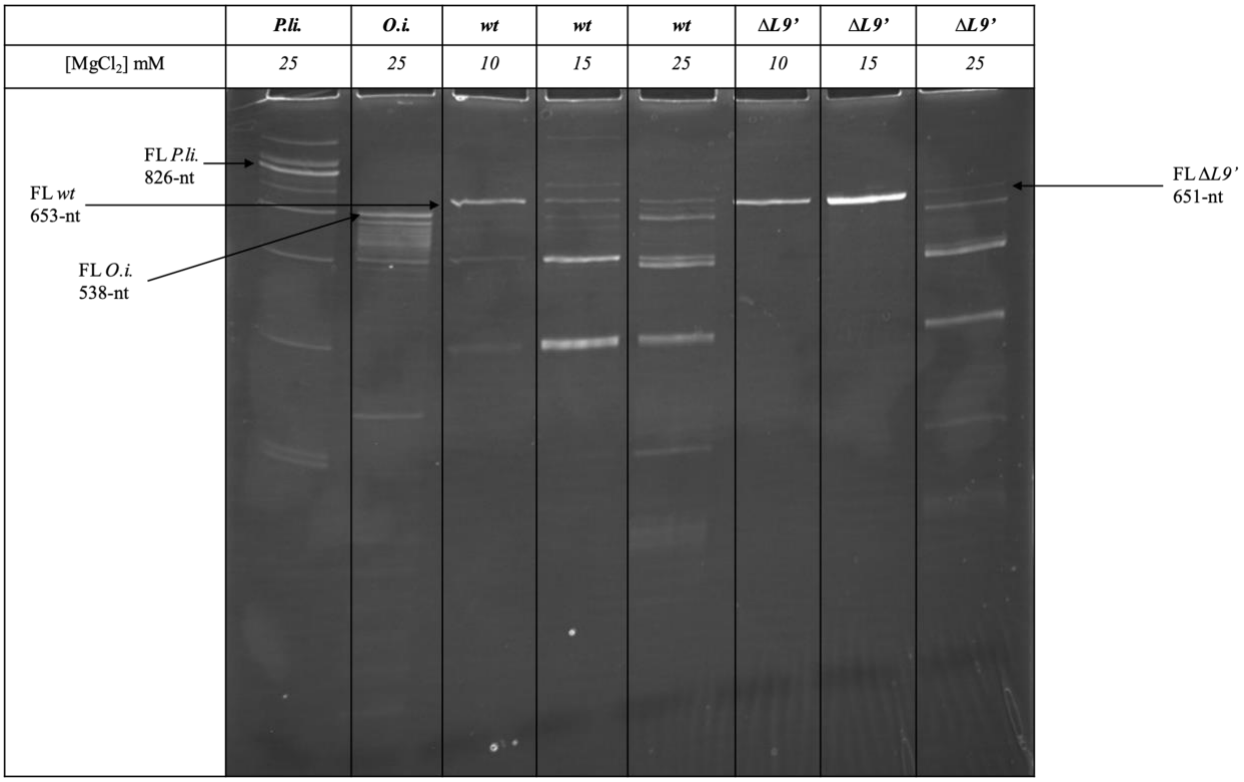

Figure S3. 4% denaturing PAGE stained in ethidium bromide. Transcription products of *wt* and Δ*L9. P.li.* and *O.i.* are used as size markers. RNA was transcribed as described in the methods, with varying concentrations of $MgCl_2$.

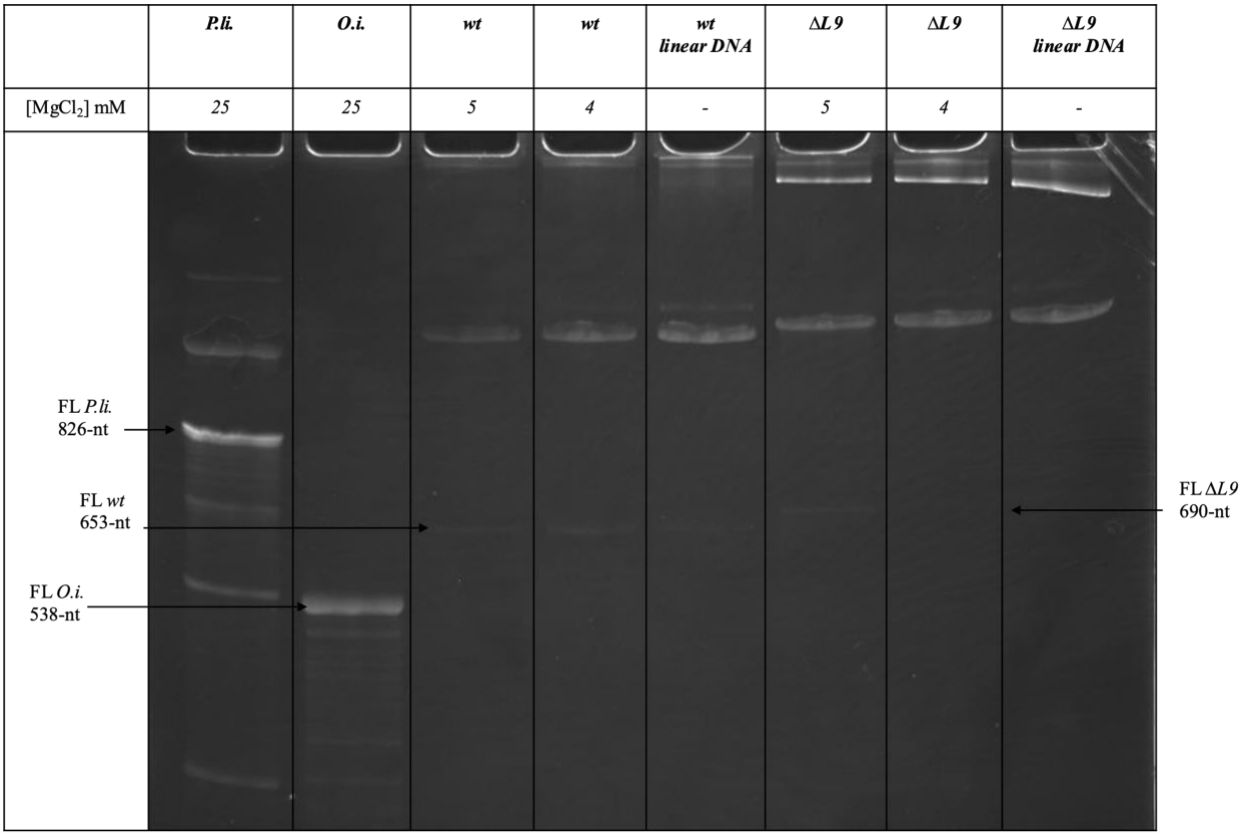

Figure S4. 4% denaturing PAGE stained in ethidium bromide. Transcription products of *wt* and Δ*L9. P.li.* and *O.i.* are used as size markers. Linear *wt* DNA lane is used to assess any DNA contamination. RNA was transcribed as described in the methods, with 4 or 5 mM $MgCl_2$.

REFERENCES

(1) Cech, T. R. Self-Splicing and Enzymatic Activity of an Intervening Sequence RNA from *Tetrahymena*. *Bioscience Reports* **2004**, *24* (4–5), 362–385. https://doi.org/10.1007/s10540-005-2738-3.

(2) Davies, R. W.; Waring, R. B.; Ray, J. A.; Brown, T. A.; Scazzocchio, C. Making Ends Meet: A Model for RNA Splicing in Fungal Mitochondria. *Nature* **1982**, *300* (5894), 719–724. https://doi.org/10.1038/300719a0.

(3) Inoue, T.; Sullivan, F. X.; Cech, T. R. New Reactions of the Ribosomal RNA Precursor of Tetrahymena and the Mechanism of Self-Splicing. *Journal of Molecular Biology* **1986**, *189* (1), 143–165. https://doi.org/10.1016/0022-2836(86)90387-6.

(4) Price, J. V.; Engberg, J.; Cech, T. R. 5′ Exon Requirement for Self-Splicing of the Tetrahymena Thermophila Pre-Ribosomal RNA and Identification of a Cryptic 5′ Splice Site in the 3′ Exon. *Journal of Molecular Biology* **1987**, *196* (1), 49–60. https://doi.org/10.1016/0022-2836(87)90510-9.

(5) Michel, F.; Hanna, M.; Green, R.; Bartel, D. P.; Szostak, J. W. The Guanosine Binding Site of the Tetrahymena Ribozyme. *Nature* **1989**, *342* (6248), 391–395. https://doi.org/10.1038/342391a0.

(6) Stahley, M. R.; Strobel, S. A. Structural Evidence for a Two-Metal-Ion Mechanism of Group I Intron Splicing. *Science* **2005**, *309* (5740), 1587–1590. https://doi.org/10.1126/science.1114994.

(7) Sengupta, R. N.; Van Schie, S. N. S.; Giambaşu, G.; Dai, Q.; Yesselman, J. D.; York, D.; Piccirilli, J. A.; Herschlag, D. An Active Site Rearrangement within the *Tetrahymena* Group I Ribozyme Releases Nonproductive Interactions and Allows Formation of Catalytic Interactions. *RNA* **2016**, *22* (1), 32–48. https://doi.org/10.1261/rna.053710.115.

(8) Mei, R.; Herschlag, D. Mechanistic Investigations of a Ribozyme Derived from the Tetrahymena Group I Intron: Insights into Catalysis and the Second Step of Self-Splicing. *Biochemistry* **1996**, *35* (18), 5796–5809. https://doi.org/10.1021/bi9527653.

(9) Karbstein, K.; Lee, J.; Herschlag, D. Probing the Role of a Secondary Structure Element at the 5′- and 3′-Splice Sites in Group I Intron Self-Splicing: The Tetrahymena L-16 ScaI Ribozyme Reveals a New Role of the G•U Pair in Self-Splicing. *Biochemistry* **2007**, *46* (16), 4861–4875. https://doi.org/10.1021/bi062169g.

(10) Luo, B.; Zhang, C.; Ling, X.; Mukherjee, S.; Jia, G.; Xie, J.; Jia, X.; Liu, L.; Baulin, E. F.; Luo, Y.; Jiang, L.; Dong, H.; Wei, X.; Bujnicki, J. M.; Su, Z. Cryo-EM Reveals Dynamics of Tetrahymena Group I Intron Self-Splicing. *Nat Catal* **2023**. https://doi.org/10.1038/s41929-023-00934-3.

(11) Stahley, M. R.; Adams, P. L.; Wang, J.; Strobel, S. A. Structural Metals in the Group I Intron: A Ribozyme with a Multiple Metal Ion Core. *Journal of Molecular Biology* **2007**, *372* (1), 89–102. https://doi.org/10.1016/j.jmb.2007.06.026.

(12) Steitz, T. A.; Steitz, J. A. A General Two-Metal-Ion Mechanism for Catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90* (14), 6498–6502.

(13) Su, Z.; Zhang, K.; Kappel, K.; Li, S.; Palo, M. Z.; Pintilie, G. D.; Rangan, R.; Luo, B.; Wei, Y.; Das, R.; Chiu, W. Cryo-EM Structures of Full-Length Tetrahymena Ribozyme at 3.1 Å Resolution. *Nature* **2021**, *596* (7873), 603–607. https://doi.org/10.1038/s41586-021-03803-w.

(14) Michel', F.; Westhof', E. Modelling of the Three-Dimensional Architecture of Group I Catalytic Introns Based on Comparative Sequence Analysis.

(15) Michel, F.; Jacquier, A.; Dujon, B. Comparison of Fungal Mitochondrial Introns Reveals Extensive Homologies in RNA Secondary Structure. *Biochimie* **1982**, *64* (10), 867–881. https://doi.org/10.1016/S0300-9084(82)80349-0.

(16) Waring, R. B.; Scazzocchio, C.; Brown, T. A.; Davies, R. W.; Brenner, S. Close Relationship between Certain Nuclear and Mitochondrial Introns. *Journal of Molecular Biology* **1983**, *167* (3), 595–605. https://doi.org/10.1016/S0022-2836(83)80100-4.

(17) Michel, F.; Dujon, B. Conservation of RNA Secondary Structures in Two Intron Families Including Mitochondrial-, Chloroplast- and Nuclear-Encoded Members. *The EMBO Journal* **1983**, *2* (1), 33–38. https://doi.org/10.1002/j.1460-2075.1983.tb01376.x.

(18) Li, Z.; Zhang, Y. Predicting the Secondary Structures and Tertiary Interactions of 211 Group I Introns in IE Subgroup. *Nucleic Acids Res* **2005**, *33* (7), 2118–2128. https://doi.org/10.1093/nar/gki517.

(19) Mitra, S.; Laederach, A.; Golden, B. L.; Altman, R. B.; Brenowitz, M. RNA Molecules with Conserved Catalytic Cores but Variable Peripheries Fold along Unique Energetically Optimized Pathways. *RNA* **2011**, *17* (8), 1589–1603. https://doi.org/10.1261/rna.2694811.

(20) Butcher, S. E.; Pyle, A. M. The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks. *Acc. Chem. Res.* **2011**, *44* (12), 1302–1311. https://doi.org/10.1021/ar200098t.

(21) Doherty, E. A.; Doudna, J. A. The P4-P6 Domain Directs Higher Order Folding of the Tetrahymena Ribozyme Core. *Biochemistry* **1997**, *36* (11), 3159–3169. https://doi.org/10.1021/bi962428+.

(22) ADAMS, P. L.; STAHLEY, M. R.; GILL, M. L.; KOSEK, A. B.; WANG, J.; STROBEL, S. A. Crystal Structure of a Group I Intron Splicing Intermediate. *RNA* **2004**, *10* (12), 1867–1887. https://doi.org/10.1261/rna.7140504.

(23) Wang, J.-F.; Downs, W. D.; Cech, T. R. Movement of the Guide Sequence During RNA Catalysis by a Group I Ribozyme. *Science* **1993**, *260* (5107), 504–508. https://doi.org/10.1126/science.7682726.

(24) Strobel, S. A.; Cech, T. R. Tertiary Interactions with the Internal Guide Sequence Mediate Docking of the P1 Helix into the Catalytic Core of the Tetrahymena Ribozyme. *Biochemistry* **1993**, *32* (49), 13593–13604. https://doi.org/10.1021/bi00212a027.

(25) Tanner, M. A.; Cech, T. R. Joining the Two Domains of a Group I Ribozyme to Form the Catalytic Core. *Science* **1997**, *275* (5301), 847–849. https://doi.org/10.1126/science.275.5301.847.

(26) Benz-Moy, T. L.; Herschlag, D. Structure–Function Analysis from the Outside In: Long-Range Tertiary Contacts in RNA Exhibit Distinct Catalytic Roles. *Biochemistry* **2011**, *50* (40), 8733–8755. https://doi.org/10.1021/bi2008245.

(27) Shi, X.; Bisaria, N.; Benz-Moy, T. L.; Bonilla, S.; Pavlichin, D. S.; Herschlag, D. Roles of Long-Range Tertiary Interactions in Limiting Dynamics of the *Tetrahymena* Group I Ribozyme. *J. Am. Chem. Soc.* **2014**, *136* (18), 6643–6648. https://doi.org/10.1021/ja413033d.

(28) Strobel, S. A.; Cech, T. R. Exocyclic Amine of the Conserved G·U Pair at the Cleavage Site of the *Tetrahymena* Ribozyme Contributes to 5‘-Splice Site Selection and Transition State Stabilization. *Biochemistry* **1996**, *35* (4), 1201–1211. https://doi.org/10.1021/bi952244f.

(29) Strobel, S. A.; Ortoleva-Donnelly, L. A Hydrogen-Bonding Triad Stabilizes the Chemical Transition State of a Group I Ribozyme. *Chemistry & Biology* **1999**, *6* (3), 153–165. https://doi.org/10.1016/S1074-5521(99)89007-3.

(30) Stahley, M. R.; Strobel, S. A. RNA Splicing: Group I Intron Crystal Structures Reveal the Basis of Splice Site Selection and Metal Ion Catalysis. *Current Opinion in Structural Biology* **2006**, *16* (3), 319–326. https://doi.org/10.1016/j.sbi.2006.04.005.

(31) Marcia, M.; Somarowthu, S.; Pyle, A. Now on Display: A Gallery of Group II Intron Structures at Different Stages of Catalysis. *Mobile DNA* **2013**, *4* (1), 14. https://doi.org/10.1186/1759-8753-4-14.

(32) Bonilla, S. L.; Vicens, Q.; Kieft, J. S. *Cryo-EM Reveals an Entangled Kinetic Trap in the Folding Pathway of a Catalytic RNA*; preprint; Biochemistry, 2022. https://doi.org/10.1101/2022.04.05.487152.

(33) Li, S.; Palo, M. Z.; Pintilie, G.; Zhang, X.; Su, Z.; Kappel, K.; Chiu, W.; Zhang, K.; Das, R. Topological Crossing in the Misfolded *Tetrahymena* Ribozyme Resolved by Cryo-EM. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119* (37), e2209146119. https://doi.org/10.1073/pnas.2209146119.

(34) Zhang, X.; Li, S.; Pintilie, G.; Palo, M. Z.; Zhang, K. Snapshots of the First-Step Self-Splicing of *Tetrahymena* Ribozyme Revealed by Cryo-EM. *Nucleic Acids Research* **2023**, gkac1268. https://doi.org/10.1093/nar/gkac1268.

(35) Li, S.; Palo, M. Z.; Zhang, X.; Pintilie, G.; Zhang, K. Snapshots of the Second-Step Self-Splicing of Tetrahymena Ribozyme Revealed by Cryo-EM. *Nat Commun* **2023**, *14* (1), 1294. https://doi.org/10.1038/s41467-023-36724-5.

(36) Joyce, G. F.; Orgel, L. E. Prospects for Understanding the Origin of the RNA World. **1993**.

(37) Heller, L.; Press, C. The RNA World, Second Edition_37. **1999**.

(38) Tourasse, N. J.; Kolstø, A.-B. Survey of Group I and Group II Introns in 29 Sequenced Genomes of the Bacillus Cereus Group: Insights into Their Spread and Evolution. *Nucleic Acids Res* **2008**, *36* (14), 4529–4548. https://doi.org/10.1093/nar/gkn372.

(39) Lambowitz, A. M.; Zimmerly, S. Group II Introns: Mobile Ribozymes That Invade DNA. *Cold Spring Harb Perspect Biol* **2011**, *3* (8), a003616. https://doi.org/10.1101/cshperspect.a003616.

(40) Costa, M.; Michel, F. Frequent Use of the Same Tertiary Motif by Self-Folding RNAs. *The EMBO Journal* **1995**, *14* (6), 1276–1285. https://doi.org/10.1002/j.1460-2075.1995.tb07111.x.

(41) Konarska, M. M.; Grabowski, P. J.; Padgett, R. A.; Sharp, P. A. Characterization of the Branch Site in Lariat RNAs Produced by Splicing of MRNA Precursors. *Nature* **1985**, *313* (6003), 552–557. https://doi.org/10.1038/313552a0.

(42) Peebles, C. L.; Perlman, P. S.; Petrillo, M. L.; Tabor, J. H.; Jarrell, K. A. A Self-Splicing RNA Excises an Lntron Lariat.

(43) Haack, D. B.; Yan, X.; Zhang, C.; Hingey, J.; Lyumkis, D.; Baker, T. S.; Toor, N. Cryo-EM Structures of a Group II Intron Reverse Splicing into DNA. *Cell* **2019**, *178* (3), 612-623.e12. https://doi.org/10.1016/j.cell.2019.06.035.

(44) Frendewey, D.; Keller, W. Stepwise Assembly of a Pre-MRNA Splicing Complex Requires U-SnRNPs and Specific Intron Sequences. *Cell* **1985**, *42* (1), 355–367. https://doi.org/10.1016/S0092-8674(85)80131-8.

(45) Grabowski, P.; Seiler, S.; Sharp, P. A Multicomponent Complex Is Involved in the Splicing of Messenger RNA Precursors. *Cell* **1985**, *42* (1), 345–353. https://doi.org/10.1016/S0092-8674(85)80130-6.

(46) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Cryst D* **2004**, *60* (12), 2126–2132. https://doi.org/10.1107/S0907444904019158.

(47) Afonine, P. V.; Poon, B. K.; Read, R. J.; Sobolev, O. V.; Terwilliger, T. C.; Urzhumtsev, A.; Adams, P. D. Real-Space Refinement in PHENIX for Cryo-EM and Crystallography. *Acta Crystallogr D Struct Biol* **2018**, *74* (Pt 6), 531–544. https://doi.org/10.1107/S2059798318006551.

(48) Adams, R. B.; Hermalin, B. E.; Weisbach, M. S. The Role of Boards of Directors in Corporate Governance: A Conceptual Framework and Survey. *Journal of Economic Literature* **2010**, *48* (1), 58–107.

(49) Morin, A.; Eisenbraun, B.; Key, J.; Sanschagrin, P. C.; Timony, M. A.; Ottaviano, M.; Sliz, P. Collaboration Gets the Most out of Software. *eLife* **2013**, *2*, e01456. https://doi.org/10.7554/eLife.01456.

(50) Toor, N.; Hausner, G.; Zimmerly, S. Coevolution of Group II Intron RNA Structures with Their Intron-Encoded Reverse Transcriptases. *RNA* **2001**, *7* (8), 1142–1152. https://doi.org/10.1017/S1355838201010251.

(51) Haugen, P.; Simon, D. M.; Bhattacharya, D. The Natural History of Group I Introns. *Trends in Genetics* **2005**, *21* (2), 111–119. https://doi.org/10.1016/j.tig.2004.12.007.