Fall 12-5-2020

Site Tolerance of the RadA Intein

Chase Taul

Christopher Lennon Dr
Murray State University

Murray State University Honors College

HONORS THESIS

Certificate of Approval

Site Tolerance of the RadA Intein

Chase Martin Taul

November 2020

Approved to fulfill the
requirements of HON 437
Professor

_____
Dr. Christopher Lennon, Assistant Professor
Biology

Approved to fulfill the
Honors Thesis requirement
Of the Murray State Honors
Director

_____
Dr. Warren Edminster, Executive
Honors College Director

Examination Approval Page

Author: Chase Martin Taul

Project Title: Site Tolerance of the RadA Intein

Department: Biology

Date of Defense: November 18th

Approval by Examining Committee:

_____ _____
(Dr. Christopher Lennon, Advisor) (Date)

_____ _____
(Dr. Sterling Wright, Committee Member) (Date)

_____ _____
(Dr. Ricky Cox, Committee Member) (Date)

Site Tolerance of the RadA Intein

Submitted in partial fulfillment

Of the requirements

Of the Murray State University Honors Diploma

Chase Martin Taul

November 2020

ABSTRACT

Inteins (intervening proteins) invade genes at the DNA level and splice out at the protein level. Once thought of as only a parasitic type of a mobile genetic element, recent work suggests a mutualistic relationship has formed in some cases within bacterial and archaeal hosts. After translation, a precursor protein is formed with the intein between two exteins. The intein is catalytic and can excise itself out through protein splicing. Intein insertion is biased towards the active site of the protein and is thought to cause inactivation of the host protein prior to splicing. Intein splicing is responsive to a number of environmental cues, suggesting that conditional protein splicing may serve as a novel form of post-translational regulation. The *Pyrococcus horikoshii* (*Pho*) RadA intein will splice in response to substrate single-stranded DNA (ssDNA). Splicing of this protein is inhibited when ssDNA is not present through intein-extein interactions. In this study, the *Pho* RadA intein was moved from its native position to alternative locations within the RadA sequence. In all alternative positions tested, the intein could no longer splice, even though predictive methods suggest it should. The results provide a greater understanding of intein site selection and point towards a rather complex evolutionary relationship between the intein and its hosts.

Table of Contents

Abstract……..…..i

Table of Contents……ii

Index of Figures and Tables……...…iii

Introduction ……...……..1

Experimental Design……...…….....6

Methods……8

Transformation…….....8

Cell Growth and Induction…….……...8

Protein Expression and Purification…….……..…….....9

Splicing Assays …….…….....10

SDS PAGE Electrophoresis……..……10

Results…….……...12

Discussion…….……17

References…….……19

Index

Figure 1. Splicing mechanism of class 1 inteins……..3

Figure 2. +1 Nucleophile Locations of a variety of mutants……...7

Table 1. Table of threonine position of +1 nucleophile, -1 residue, predicted splicing efficiency, and secondary structure of intein……...12

Figure 3. SDS-PAGE gel of T353 splicing assays…….....13

Figure 4. T353 equivalent position on *Pyrococcus furiosus* RadA……...13

Figure 5. SDS-PAGE gel of T408 splicing assays……14

Figure 6. T408 equivalent position on *Pyrococcus furiosus* RadA……...14

Figure 7. SDS-PAGE gel of T138 and T477 splicing assays…….....15

Figure 8. T138 equivalent position on *Pyrococcus furiosus* RadA……...15

Figure 9. T477 equivalent position on *Pyrococcus furiosus* RadA……..….16

Figure 10. T325 equivalent position on *Pyrococcus furiosus* RadA……..…16

INTRODUCTION

Inteins are mobile genetic elements capable of invading host genes at the DNA level and then self-splicing out at the protein level. Inteins are found in all domains of life, with particular abundance in prokaryotes (~25% of bacteria and ~50% of archaea) (Novikova, et al. 2015). Every intein is generally similar in structure and has related mechanisms of splicing (Eryilmaz, et al. 2014). An intein is capable of inserting itself into the genome at the DNA level and is then expressed into a protein through transcription and translation. Inteins were historically thought to be nothing more than molecular parasites that did not benefit the host in any way (Lennon et al. 2017). However, recent evidence suggests some inteins are not purely selfish and that they have evolved regulatory purposes with their protein hosts. In many active sites, ATP cannot bind, and the protein is unable to function when an intein is present. An intein regulates essential functions within its host so it is not surprising that over 70% of all inteins are found within ATPase proteins and over 60% of them are localized in proteins that are involved in DNA replication, recombination, and repair (Novikova, et al. 2015). An intriguing post-translational method of protein regulation is obtained from the strategic placement of the inteins in the active site because splicing can be dependent on environmental cues or even substrates of the host protein.

Inteins have generated interest in a variety of fields including microbiology, protein engineering, and medicine. As of now, affinity tags mediated by self-cleaving inteins are used for a more efficient purification process of proteins and enhanced protein labeling. In this process the intein is used to fuse an affinity tag onto the protein of interest. Next, cleavage, rather splicing, of the intein is induced by changing the pH or temperature conditions. The use of inteins opens up new opportunities to develop revolutionary means of protein engineering by the combining of two separate polypeptide segments. This can be achieved with the use of *trans*-

splicing inteins that are expressed as separate polypeptides and splice upon reassembly (Wood et al. 2014). Currently, the CRISPR/Cas9 system is a good example of the utility of the split-intein design. It is being utilized in genetic research with the potential of treating a variety of genetic diseases in the future. The use of the CRISPR/Cas9 is somewhat limited because the delivery of the adeno-associated virus is hindered by the relatively large size of the Cas9. However, biologists have been delivering the Cas9 encoding sequence into its destination by separating it into two parts. Intein-mediated *trans*-splicing is then used to ligate them together. After its induction into the cell, the Cas9 protein is able to locate the target gene in the organism. The target gene is then cut in order to edit its genome. This technique is utilized to give researchers knowledge on the target gene function (Truon et al, 2015).

Inteins are present in a large variety of bacterial pathogens, but they are not present in humans or other multicellular eukaryotic hosts. This particular trait gives inteins great potential as drug target. *Mycobacterium tuberculosis* is a bacterium that has one of its four inteins within the RecA protein (Zhang et al. 2010). This particular bacterium is a highly contagious pathogen that infects individuals in all parts of the world and is the global leading cause of deaths by an infectious disease. Recently, it has been found that cisplatin (a common chemotherapy drug) represents an exciting lead compound for intein-specific inhibitors. Based on numerous lines of evidence, cisplatin binds to the RecA intein which in turn prevents intein splicing (Chan et al. 2016). This discovery opens the possibility of using alternative methods to fight bacterial infections. Notably, this knowledge could potentially be efficiently utilized to combat the growing antibiotic resistance crisis. In fact, bacteria such as *Coxiella burnetii* and *Mycobacterium leprae,* as well as the fungus *Cryptococcus neoformans,* all contain essential functioning proteins that house inteins. This means diseases such as Q-fever, leprosy, and

cryptococcal meningitis could potentially be treated using intein splicing to induce inhibition. These are examples of the multitude of human diseases that could have promising treatment results using inteins (Zhang et al. 2009).

An extein is a polypeptide sequence that flanks the intein. A biologically active protein contains two exteins that are linked together after the self-splicing of the intein. The reaction of splicing takes place in a mechanism with four steps. Numbering is intein-centric, with -1 residue being the last residue of the N-extein, the first residue of the intein being the 1 residue, and the first residue of the C-extein being the +1 residue. First, the N- terminus amino acid of the intein (Cys or Ser) attacks the preceding peptide bond. Secondly, a thioester linkage that forms from the first step is attacked by the first residue of the C-extein (Cys, Ser, or Thr) in a nucleophilic reaction. This creates a branched intermediate that releases the intein which leaves a thioester bond between the exteins as a result. Lastly, the thioester bond is self-rearranged into a normal peptide bond. Therefore, the N-extein and C-extein are linked together at the end as a result. This leaves no evidence that the intein was ever present in the host protein (Lennon et al. 2017). Off-pathway cleavage reactions can also occur depending on environmental conditions, where the N- or C-extein is cleaved prior to ligation.

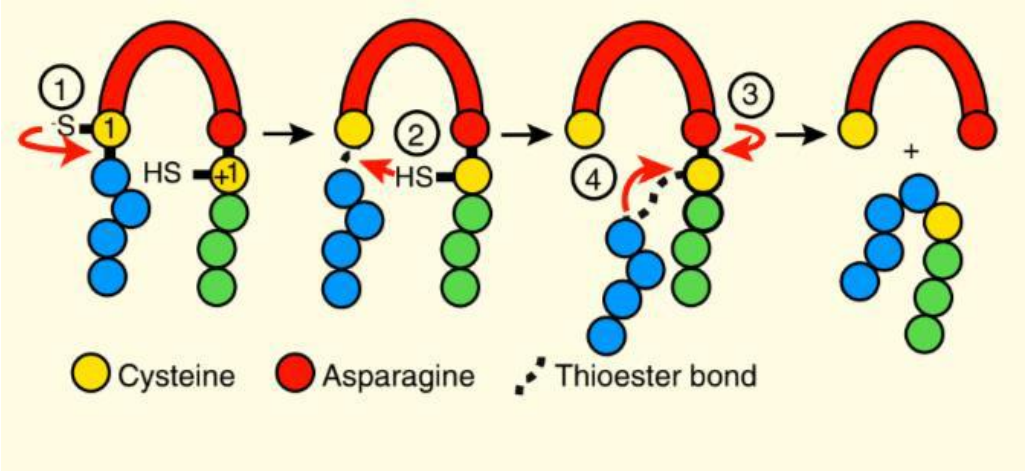

Figure 1: Class 1 mechanism of intein splicing. The red region section the intein, the blue section represents the N-extein, and the green section represents the C-extein. Directly right of the intein is the +1 nucleophilic cysteine while the yellow cysteine on the left is the first residue of the intein.

In this study, we focus on an intein within the homologous recombinase RadA from the archaeon *Pyrococcus horikoshii*. RadA is a DNA-dependent ATPase that has the ability to catalyze homologous DNA pairing and strand exchange. This serves to repair DNA damage throughout archaeal genomes and increases genetic variation (Lennon et al. 2016). This recombinase protein is homologous to bacterial RecA and eukaryotic Rad51 and is thought to be regulated by mainly post-translational means. Inteins are not found in humans however despite these conserved homologous proteins. RadA can bind onto single-stranded DNA (ssDNA) before splicing, but the intein in the active site blocks ATP from binding (Lennon et al. 2016).

The RadA recombinase intein comes from archaeon *Pyrococcus horikoshii* (*Pho*), which is found in deep sea thermal vents. This particular recombinase intein is regulated by long-distance intein-extein interactions that block splicing (Lennon and Belfort 2017). This suggests that the RadA protein is translated into an inactive form and then activated post-translationally. Interestingly, ssDNA accelerates protein slicing by ~50-fold. ssDNA, a natural substrate of RadA, is also a signal of DNA damage. In this elegant example, the very condition where the host protein (e.g. RadA) is needed by the cell induces protein splicing. This finding provides compelling evidence for the role the inteins can play in post-translational regulation (Lennon et al 2016).

There are significant observational differences between the splicing behavior of *Pho* RadA when it is in a native host protein versus a non-native host protein. Interestingly, while splicing is blocked in the native exteins in the absence of ssDNA, the intein splices with

remarkable speed and efficiency in non-native exteins. Despite this behavioral difference, it has not been studied how an intein can react when inserted into various non-native positions within the host protein, RadA. The only definite restriction to intein positioning within a protein is the presence of a nucleophile in the +1 position. This study looks at the behavior of the RadA intein when using alternative threonine residues, the natural +1 residue, in the native sequence. This research helps to more fully understand the generality of the *Pho* RadA system, the use of inteins as post-translational modifiers, and possible effects of inteins in genetically modified systems.

EXPERIMENTAL DESIGN

For the *Pho* RadA protein to splice, a nucleophile is needed in the +1 position. The amino acid in this position is the first amino acid in the C-extein sequence. Cystine, serine, and threonine are the three amino acid residues that can be utilized as a +1 nucleophile (Oeemig et al. 2012). It is thought that the presence of these +1 nucleophiles is the only strict requirement needed to initiate splicing of the intein even though many factors can have an effect on the accuracy and speed of intein splicing (Topilina et al. 2015). Threonine is the amino acid that is utilized as the +1 nucleophile in wild type *Pho* RadA. The *Pho radA* intein was moved from its native position to ten other locations within the *radA* gene preceding other threonines within the primary sequence (seen in Fig. 2). The goal was to observe how an intein's position within its *radA* sequence will alter the splicing ability of the intein. This was designed to test the theory that *Pho* RadA has coevolved with its host to utilize the intein as a method of post-translational modification. In this experiment *Pho* RadA WT (control) was used along with the ten variants of intein position using alternative +1 residues previously cloned by Dr. Lennon. These variants in turn have different extein junction positions. Also, a variant of *Pho* RadA without an intein and another variant with a deactivated intein were utilized. This allowed the experiment to have 13 variations of *Pho radA* cloned separately within the pET vector.

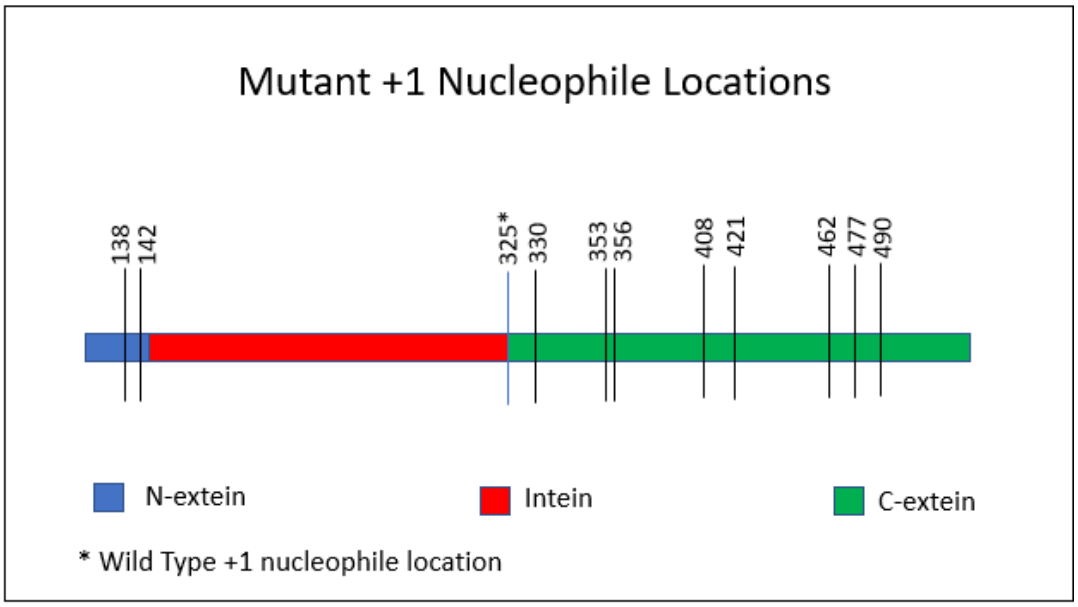

Fig 2. Location of the inserted mutant RadA +1 nucleophiles. Each position represents an alternate threonine's location that has the potential to be used as a +1 nucleophile for splicing after relocation of the intein. The threonines are T138, T142, T330, T353, T356, T408, T421, T462, T477 and T490. The red region represents the intein's wild type location and the blue region represents the wild type N-extein. The green region represents the wild type C-extein.

The vector also encodes for Beta-Lactamase, a gene that provides antibiotic resistance. This allowed for a special selectivity within the experiment because only cells containing the desired plasmid could be cultured in the presence of ampicillin. Agar plates containing ampicillin were also used because they inhibit the growth of cells that lack the Beta-Lactamase gene. This allowed for a selectivity and helped to ensure that only the desired protein was expressed.

These proteins were expressed in *E. coli* before being purified using nickel chromatography. The pET vector also codes for six sequential histidines that precede the N-extein. This His-tag is what allows the protein to be purified later in the experiment by the use of nickel chromatography. LacI regulates the expression of these genes to allow for the target gene to only be expressed when desired. Purified proteins were then tested for their individual competencies of splicing.

Materials and Methods

The experimental step that occurred first was a transformation. The tubes containing competent cells (for transformation) remained on ice except where indicated. Using a micropipette, 1µl of the desired plasmid was then inserted into the competent cells and allowed to remain on ice for 10 minutes. The microcentrifuge tube was then transferred to a water-filled heat bath that was set to 42℃. This is a process known as heat shocking. It allows the uptake of the added plasmid by the isolated colony bacteria in the solution. The solution was held in the heat bath for ~10 seconds before being returned to ice for two minutes. Next, one mL of LB broth was added to the test tube that held the plasmid containing *E. coli* cells. This new solution was incubated at 37℃ for 1 hour. Then, 100 µl of this solution pipetted onto previously prepared agar plates and spread. They were placed in an incubator that was set to 37℃ and allowed to remain there overnight. These plates were prepared with LB and ampicillin. This serves as a selective marker to prevent contamination and get rid of the competent cells that did not take up the plasmid. The plasmid contains a gene called Beta-Lactamase that gives it antibiotic resistance. All the steps were performed repeatedly for the control groups and the mutants.

The second experimental step that occurred was cell growth and induction. A 5 mL starter culture was inoculated with the isolated colonies previously attained. This solution was incubated overnight at 37℃ while shaking at 250 RPM in the shaking incubator. This incubator keeps conditions within it constant. The starter culture was then pipetted into the growth media at a volume ratio of 1:100. This gave a total of 400mL LB growth media with ampicillin before the beakers were placed in a shaking incubator. It was set to remain at 37℃ and have 250 rpm. The solution was allowed to remain in the shaking incubator for 30-60 minutes while optical density

measurements were taken. The optical density measurements were taken using a spectrophotometer. Small samples of the solution were taken and placed into the cuvette. A new optical density reading was taken roughly every 30 minutes to monitor the progression of the bacteria into its log phase of growth. Log phase was determined when the solution reached an optical density reading of about 0.5. The log phase is when cellular ribosomes expressing the desired protein are at the highest concentrations. Optical density readings were stopped once log phase was attained and IPTG (isopropyl **β**-d-1-thiogalactopyranoside) was added. IPTG is used to induce overexpression of the protein of interest. IPTG is a molecular mimic of allolactose that binds to the Lac repressor, which in turn induces transcription and ultimately protein overexpression.

The third experimental step was expression and purification of the desired protein. The cells containing the expressed RadA protein needed to be separated from the spent media to obtain the cells for protein purification. The growth media solution was poured into centrifuge tubes and centrifuged at 3000xG for ten minutes. This allowed a pellet containing the cells and desired protein to form at the bottom. The spent media was carefully decanted from the tube, so the pellet was all that remained within it. Nickel Buffer A (20 nMTris pH 8, 500mM NaCl, 10 mM Imidazole) was added on top of the pellet. The new solution containing the pellet was resuspended before being frozen at -20℃ overnight. This prepared the cells for sonication. Sonication is the use of ultrasound waves that agitate and disrupt the cell membranes of cells. These high-frequency waves cause the implosion of the cells which leads to successful lysing. The frozen solution was allowed to thaw before being placed in large beakers of ice. This ice prevented overheating from the sonication process therefore safeguarding against protein denaturation and premature splicing. The sonication probe was added to the solution. The

sonicator was set to 30% and ran at 30 second intervals (on and off) for 30 minutes. After a successful lysing, the solution was centrifuged at 20,000x for 20 minutes. This separates the desired proteins and cell debris into different layers. The supernatant, which contains the desired protein, was decanted in preparation for purification.

The desired protein was obtained using an affinity chromatography. A column of Ni2+-resin was used. This column binds to the His-tag that is encoded on the RadA amino-terminus. After the binding of the desired protein, the column was rinsed with the washing solution. This pushed everything in the column through except our tightly bound protein of interest. Following this the column was then washed with an elution buffer. The Ni2+-resin has higher affinity for the elution buffer than the His-tag. The elution buffer contains imidazole which replaces the side chain of histidine within the column. The above steps were repeated multiple times for the different variants of proteins.

The fourth experimental step was the use of splicing assays. These were used to determine the conditions that the inteins were able to splice, if at all. This was accomplished by the changing of two variables: addition of single-stranded DNA (ssDNA) and variability in temperature. ssDNA is a factor that can the RadA intein to splice. Each sample was heated to 63℃, 75℃, and 87℃ for 15 minutes. This was done first without the presence of ssDNA. The process was repeated once more with the same temperature conditions but had 188ng/µl of ssDNA in addition.

The fifth experimental step was the utilization of SDS-Page (sodium dodecyl sulphate-polyacrylamide gel electrophoresis). This is a technique used to separate proteins based on their molecular weights. The combination of the polyacrylamide gel and SDS will allow separation

based on molecular weight because it allows the structure and charge of the proteins to be ignored. Electrophoresis will denature the proteins by exposing regions that are normally buried within the protein. In the electrophoresis unit, 1x Tris-Glycine-SDS buffer was loaded. The SDS in the buffer helps keep the denaturing proteins linear by coating the denatured protein with surfactant and has a pH range of 7-9 which makes it useful for biological systems. First, 5 µl of ladder was loaded into the gel. This gave us our template to determine the molecular weights of our samples. Next, 20 µl of the resuspended protein samples (15 µl of sample and 5 µl of dye) were loaded into the adjacent gel slots using a micropipette. The power supply was set to 200 volts and ran for 30 minutes. Afterwards, the gel case was cracked and removed which gave an exposed gel.

Staining of protein with Coomassie dye occurred next using an automated stainer. The gel was placed in between two pieces of thick paper that had been moistened. It was then closed into the gel cassette holder and inserted into the machine. An 11-minute process of staining and destaining was run until the gel had been sufficiently stained. A gel picture was taken next using a gel imager. The technique used to photograph the gel was epi-illumination. A colored and black and white picture was taken and documented.

RESULTS

WT RadA (T325) serves as the control for the experiment while the intein was moved to several alternative threonine locations within the host sequence. The alternative locations of the intein tested were T138, T477, T353, and T408. The control's comparison allows the experiment to be able to quantify the splicing within the RadA mutant variations. All samples were first tested without ssDNA or heating to see their ability to splice.

Prior to experimentation, a variety of characteristics of the modified mutants were predicted. They were the location of the intein within 3D space, splicing efficiency of -1 residue, threonine position of the +1 nucleophile, and secondary structure of the intein (Table 1).

Threonine Position Used as +1 Nucleophile	-1 Residue	Predicted -1 Residue Splicing Efficiency	Secondary Structure of Intein
T138	E	40%	Loop
T477	P	23%	Unstructured
T353	N	95%	Loop
T408	N	95%	Loop
T325	K	97%	Loop-Native site

Table 1: Table showing the threonine positions of the +1 residue, -1 residue, predicted splicing efficiency and secondary structure of the intein.

Upon data analysis, the mutant T353 precursor disappears, but does not splice as evidenced and there are no ligated exteins that are generated (Figure 3) This protein was expressed using the mentioned methods, a nickel pulldown was performed, and then the strain underwent two additional forms of treatment. These forms of treatment were heated without the presence of ssDNA while the other was heated in the presence of ssDNA. Both treatment groups

were heated to 63℃, 75℃, and 87℃ for 15 minutes a piece.

Further analysis of the expressed proteins by electrophoresis showed that our *Pho* RadA control spliced. The electrophoresis gave the results of the splicing assay. However, none of the manipulated mutants displayed significant splicing in the initial attempt (Fig 3 & 5).

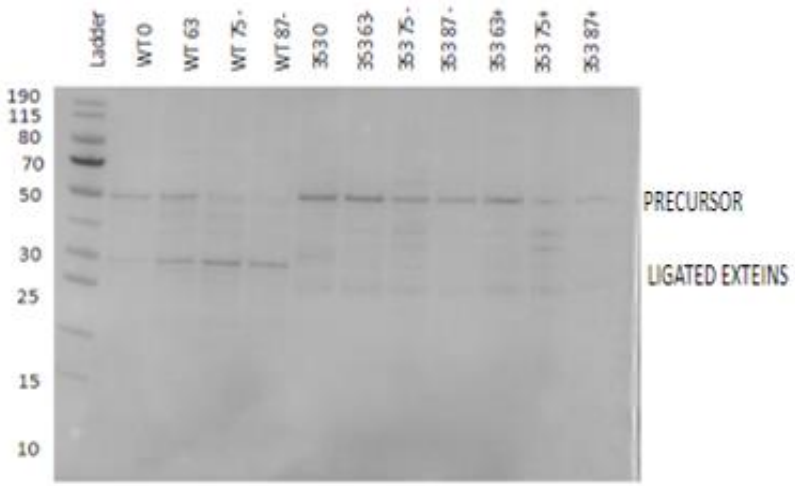

Figure 3: Splicing assay for the T353 mutant. The gel shows variations of this mutant compared to RadA WT. The ladder in the leftmost column serves as a reference.

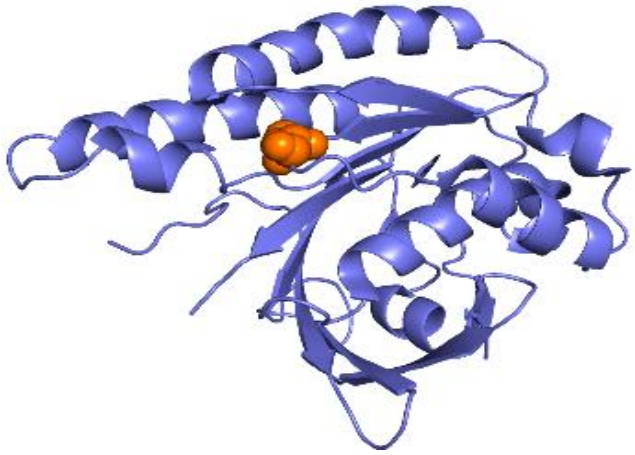

Figure 4. T353 equivalent position on *Pyrococcus furiosus* RadA.

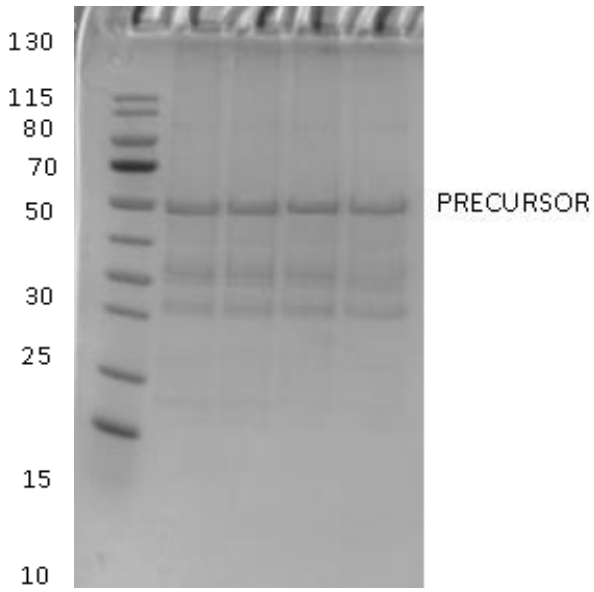

Figure 5: Splicing assay for the T408 mutant. The gel shows the variation of the mutant compared to the ladder.

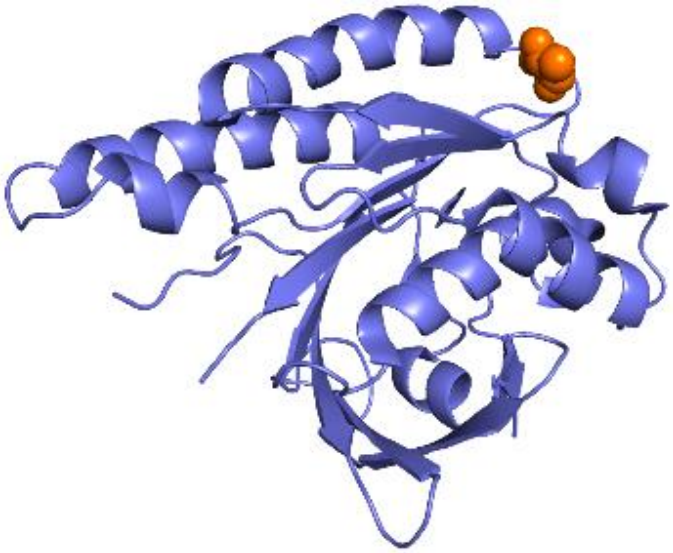

Figure 6. T408 equivalent position on *Pyrococcus furiosus* RadA (native +1).

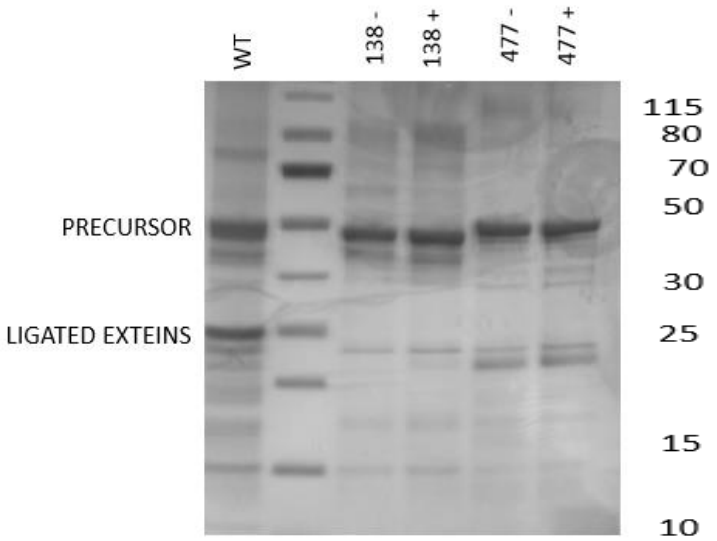

Figure 7: Splicing assay for the T138 and T477 mutants. The WT control serves as a reference point to compare the ligated exteins.

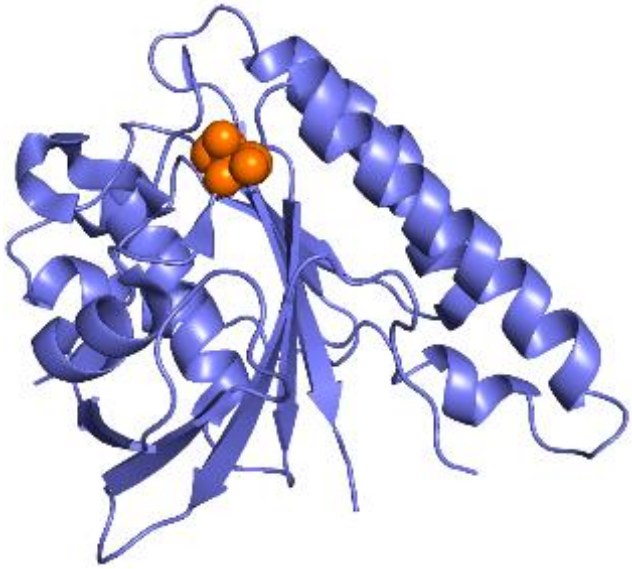

Figure 8. T138 equivalent position on *Pyrococcus furiosus* RadA

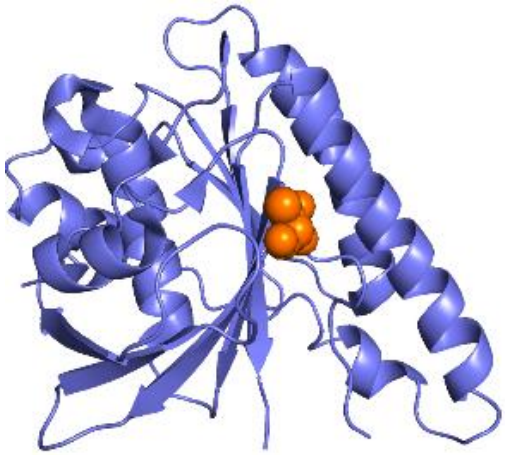

Figure 9. T477 equivalent position on *Pyrococcus furiosus* RadA

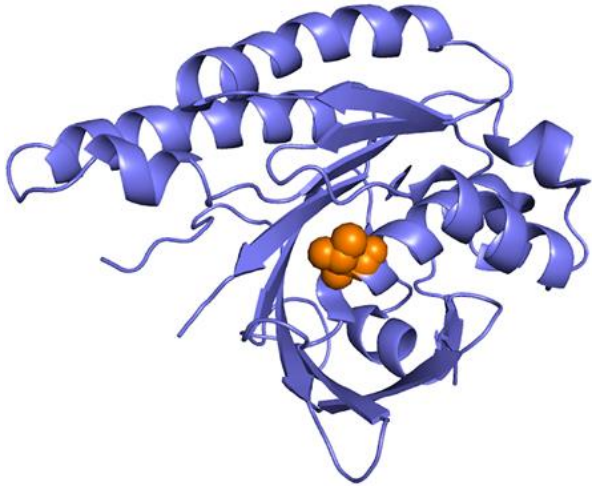

Figure 10. T325 equivalent position on *Pyrococcus furiosus* RadA

In addition, the projected splicing of T353 [Figure 3] did not occur even with the loop configuration and asparagine residue. Also, the mutant T408 was projected to show significant splicing, but the gel [Figure 5] shows that splicing did not occur even with its loop secondary structure and asparagine residue. The results showed that the WT control spliced under all conditions while the mutants did not splice under the addition of heat or ssDNA. Mutants T138 & T477 [Figure 7] also showed no ligated exteins even though they were projected to splice less effectively due to their secondary structure and -1 residues [Table 1].

DISCUSSION

The inteins inability to splice at any positions tested outside of the native context was surprising, particularly given that the *Pho* RadA intein is extraordinarily known for its ability to splice better when flanked by non-native extein (Topilina et al. 2015). Each mutant had the same +1 nucleophilic amino acid residue required for splicing in the WT and still no splicing was observed. The final residue of the N-extein (-1 residue) has been shown to influence splicing of the *Pho* RadA and other inteins. A study conducted previously with a non-native exteins had changed the residue in the -1 position (Oeemig et al. 2012), allowing us to predict splicing efficiency in Table 1. At least two of our tested inteins, T353 and T408, had a favorable residue present in the -1 position (Table 1).

There are a few plausible reasons as to why these mutants did not splice in their non-native settings. Firstly, the secondary structure of the surrounding protein body could be a hindrance on the inteins ability to splice even though it theoretically should not be when it comes to the loop or unstructured configuration. The T353 nucleophile in the +1 position was singled out for this experiment because of this theoretical thinking. This particular intein has a -1-asparagine residue with roughly a 95% predicted efficiency of splicing (Oeemig et al. 2012). As of right now there is not a clear explanation for why splicing results were not observed. Unfortunately, the results are fairly crude as they were not allowed to be conducted again due to the Covid-19 pandemic. If able the experiment would be conducted again to confirm the results presented in this thesis.

Nevertheless, the results of this experimentation raise exciting new questions concerning this topic and the factors that go into successful intein splicing. Original thought suggested that

only the flanking residues of the extein contributed to the splicing ability of an intein, but studies have shown that splicing inhibition can be mediated by interactions between the intein and remote residues within the extein (Topilina et al, 2015, Lennon et al, 2019). Closer analysis of the 3D structure should be observed for each mutant and compared to the *Pho* RadA WT to see if splicing is being inhibited in a more permanent way. Another plausible explanation is that remote residues in the flanking extein regions interact with the mutant inteins while they do not influence the WT intein. The resulting mutant proteins were soluble when expressed and could be purified which means that they must have a reasonable amount of stability due to the fact that they are not being degraded. Another possibility is that partial misfolding could explain the resulting inability to splice.

Overall, the results do not reflect upon the initial hypothesis, but they still raise new interesting questions. In the future lysine will be added to the -1 position to check if that changes the inactivity. This modification might work because lysine in the -1 setting is native in the native active site.

REFERENCES

1) Chan, Hon, et al. “Exploring Intein Inhibition by Platinum Compounds as an Antimicrobial Strategy.” Journal of Biological Chemistry, vol. 291, no. 43, 2016, pp. 22661–22670

2) Dong-Jiunn Jeffery Truong, Karin Kühner, Ralf Kühn, Stanislas Werfel, Stefan, Engelhardt, Wolfgang Wurst, Oskar Ortiz. “Development of an intein-mediated split– Cas9 system for gene therapy.” Nucleic Acids Research, vol. 43, Issue 13, 27 July 2015, Pages 6450–6458.

3) Eryilmaz, Ertan, et al. “Structural and Dynamical Features of Inteins and Implications on Protein Splicing.” Journal of Biological Chemistry, vol. 289, no. 28, 2014, pp. 19278–19278.

4) Lennon, Christopher W., and Marlene Belfort. “Inteins.” Current Biology, vol. 27, no. 6, 20 March 2017.

5) Lennon, Christopher W., et al. “Protein Splicing of a Recombinase Intein Induced by SsDNA and DNA Damage.” Genes & Development, vol. 30, no. 24, 2016, pp. 2663– 2668.

6) Lennon, Christopher W., Stanger, Matthew J., Belfort, Marlene. (2019). “Mechanism of Single-Stranded DNA Activation of Recombinase Intein Splicing.” *Biochemistry.* Vol 51. No 31. Pg 3335-3339.

7) Oeemig, Jesper S., et al. “NMR and Crystal Structures of the Pyrococcus Horikoshii RadA Intein Guide a Strategy for Engineering a Highly Efficient and Promiscuous Intein.” Journal of Molecular Biology, vol. 421, no. 1, 3 Aug. 2012, pp. 85–99.

8) Olga Novikova, Pradeepa Jayachandran, Danielle S. Kelley, Zachary Morton, Samantha Merwin, Natalya I. Topilina, Marlene Belfort, Intein Clustering Suggests Functional Importance in Different Domains of Life, *Molecular Biology and Evolution*, Volume 33, Issue 3, March 2016, Pg 783–799.

9) Topilina, Natalya I., et al. “Post-Translational Environmental Switch of RadA Activity by Extein–Intein Interactions in Protein Splicing.” Nucleic Acids Research, vol. 43, no. 13, 2015, pp. 6631–6648.

10) Wood, David W., and Julio A. Camarero. “Intein Applications: From Protein Purification and Labeling to Metabolic Control Methods.” Journal of Biological Chemistry, vol. 289, no. 21, 2014, pp. 14512–14519.

11) Zhang, Liyun, et al. “Cisplatin Inhibits Protein Splicing, Suggesting Inteins as Therapeutic Targets in Mycobacteria.” Journal of Biological Chemistry, vol. 286, no. 2, 2010, pp. 1277–1282.

12) Zhang, Liyun, et al. “Metal Ions Binding to RecA Inteins from Mycobacterium Tuberculosis.” Molecular BioSystems, vol. 5, no. 6, 2009, p. 644.