

Biochemical investigation of lariat debranching enzyme and spliceosomal RNA through the use of backbone branched RNA

by

Stephanie Nicole Mack

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

Carnegie Mellon University

2018

Abstract

Lariat introns are generated during pre-mRNA splicing. The 2'-OH of the branch point residue is linked to the 5'-terminus of the intron through a 2'-5'-phosphodiester bond. The 2'-5'-phosphodiester bond is not processed by traditional exo or endonucleases and is only cleaved by lariat debranching enzyme (Dbr1p). Introns are involved in a variety of regulatory processes in the cell and thus, Dbr1p is upstream from many critical cellular functions. Dbr1p is structurally very similar to other non-specific metallophosphoesterases and it is currently unknown why Dbr1p is specific while other enzymes in this superfamily are not.

Easy access to substrates for Dbr1p has previously been a major roadblock in conducting biochemical investigations of the enzyme. This lab has developed a robust method for synthesizing backbone branched RNA (bbRNA) that contain a 2'-5'-phosphodiester bond. A photolabile protecting group is employed during synthesis to create a branch point residue, from which a 2'-5'-phosphodiester bond can be formed. Improvements to the photodeprotection method are discussed. With ready access to easily modifiable bbRNA, the substrate requirements of Dbr1p are investigated. A variety of non-native bbRNA are incubated with Dbr1p and the change in catalytic activity is observed. Dbr1p is able to cleave a wide variety of backbone branched substrates but at a much slower rate than a traditional 2'-5'-phosphodiester bond. Non-cleavable analogues of bbRNA have been synthesized that contain a 2'-5'-triazole linkage, called click branched RNA (cbRNA). These RNAs interact with Dbr1p and are shown to be competitive inhibitors. The strength of inhibition is dictated by the size of the click branched inhibitor and kinetic analysis is performed using a fluorescence assay.

Dbr1p has been shown to be important in microRNA biogenesis through the mirtron pathway. Short oligomers are synthesized that are comprised of only the guide strand sequence of a miRNA. The 5'-terminus is linked to an internal 2'-OH to form a 2'-5'-phosphodiester bond

that is cleavable by Dbr1p. These lariat structures are used as siRNA and delivered into cells. Repression of mRNA and protein expression in human cancer cell lines is observed. These miRNA mimics have the potential to be used as a therapeutic, similar to how siRNA is currently being used.

Splicing occurs upstream of Dbr1p action. Lariat introns are formed during this process through the spliceosome and RNA is at the core of the spliceosomal catalytic machinery. With easy access to bbRNA, we are able to model the spliceosome immediately after the first catalytic step, branching. We set out to develop an assay to determine if the RNA at the core of the spliceosome is catalytic in a protein-free environment. Detailed analysis of the strength of RNA duplexed to bbRNA is performed using melting temperatures. As the spliceosomal RNA is not entirely duplexed, we created constructs that mimic these interactions. The melting temperatures are determined for bbRNA duplexes to see how the bbRNA structure might stabilize or destabilize a duplex interaction. Binding and cleavage assays are performed to see if spliceosomal RNAs can cause catalysis on bbRNA in a protein-free environment. This thesis focuses on understanding the processing of lariat introns upstream (splicing) and downstream (Dbr1p cleavage) of intron formation.

Acknowledgements

I would like to sincerely thank everyone who has been a part of my journey in graduate school. Many people have contributed to this experience through their encouragement, advice, and care.

First, I would like to thank my advisor, Dr. Subha R. Das for his support and guidance throughout my graduate career. I have learned so much during my time at Carnegie Mellon thanks to Das, all of which has pushed me to be a better scientist. I not only grew as a researcher, but also learned how to be a successful mentor based on his example. His expertise has been an invaluable resource. The skills I learned in his lab, both research and otherwise, will help me become a successful scientist in the future and for that I am very thankful.

I am grateful to the members of my graduate committee, Dr. Danith Ly and Dr. Newell Washburn. Their suggestions and feedback have made me think critically about my work and have caused me to be a better researcher. I would also like to thank Dr. Stefanie Sydlik who has kindly agreed to serve on my thesis committee. Additionally, I would like to thank Dr. Andrea Berman from the University of Pittsburgh, the external member of my Ph.D. committee, for her time and feedback.

I would like to thank all members of the Das lab, past and present for their comments and discussions that have made my experience in the lab very rewarding. I would particularly like to thank Dr. Sourav Dey who introduced me to graduate research and taught me how to be a thorough and thoughtful scientist. Also, I would like to thank Dr. Eduardo Paredes and Dr. Debashish Grahacharya for their initial research which has laid the groundwork for much of the research presented in this thesis. I have had tremendous help from Stephanie Wang, a fantastic undergraduate who has contributed to the protein-free splicing project. I would also like to thank Dr. Munira Fouz, Dr. John Pettersson, Sushil Lathwal, Chad Ratterman, and Megan Van Horn as

well as the past and present undergraduate students for their invaluable discussions and for making the Das lab a welcoming and friendly environment to walk into every day.

I would also like to thank all of the research collaborators I have had the pleasure of interacting with, and without whom this research would not be possible. Specifically, Dr. Andrew VanDemark and his lab members, currently Dr. Brian Graham and Grace Wang, for their efforts expressing and purifying Dbr1p. I would also like to thank Dr. Linda Peteanu for her contributions to the Dbr1p project. Dr. Girish Shukla and Dr. Jey S. Ebron have been fantastic collaborators on the mini-lariat project, and I thank them for their work on cellular studies with mini-lariat RNAs as well as their important insight. Also, a sincere thank you to Dr. Brigitte Schmidt who has been an invaluable resource for all things organic synthesis.

I am grateful to Dr. Leonard Vuocolo, Dr. Susan Graul, Dr. Marcel Bruchez, Carolyn Niederlander, and Marc McElhaney for teaching me how to teach. I am very appreciative of their guidance while I was a teaching assistant. I would like to acknowledge Dr. Rea Freeland, and the chemistry department staff including Valerie Bridge, Brenda Chambers, Patsey Haddock, Tim Sager, Sara Wainer, and Lorna Williams who have made my time in the department hassle free.

Finally, I would like to acknowledge the support of my family and friends. Without their constant love and support I would not be where I am today. I would first like to thank my parents who have always supported and encouraged me. They have taught me the value of hard work, nurtured my love of science, and pushed me to succeed in everything I do, and for that I am forever grateful. I would like to thank Alexander Boys for always motivating me, encouraging me, making me laugh when I'm stressed. I'm excited to share a lifetime of adventures with him. I would like to thank my friends in Pittsburgh for making my time in this city memorable, especially Chetali, Matt, Saransh, Steve, and Ross, among many others. I will always cherish my time at Carnegie Mellon because of the people that made it all possible.

Table of Contents

Abstract.....	ii
Acknowledgements.....	iv
Table of Contents.....	vi
List of Figures.....	viii
List of Tables.....	xi
List of Schemes.....	xiii
Chapter 1: Introduction to Splicing, Backbone Branched RNA and Lariat Debranching Enzyme.....	1
1.1. Dbr1p in Regulatory Processes and the Health Implications.....	3
1.2. Biochemistry of Lariat Debranching Enzyme.....	6
1.3. Analyzing the Dbr1p reaction <i>in vitro</i> with backbone branched RNA.....	9
1.4. Scope of Thesis.....	13
1.5. References.....	15
Chapter 2: Advancing backbone branched RNA synthesis through the improvement of the photodeprotection method.....	24
2.1. Introduction.....	24
2.2. Improving the Photodeprotection Method.....	29
2.3. Conclusion.....	36
2.4. Experimental.....	37
2.5. References.....	42
Chapter 3: Biochemical analysis of backbone branched RNA cleavage by lariat debranching enzymes.....	43
3.1. Introduction to Dbr1p.....	43
3.2. Requirements for <i>in vitro</i> Dbr1p activity.....	46
3.3. RNA requirements for the Dbr1p cleavage reaction.....	54
3.4. Conclusions.....	75
3.5. Experimental.....	77

3.6. References.....	83
Chapter 4: Inhibition of Lariat Debranching Enzyme with Triazole Branched RNA.....	86
4.1. Introduction.....	86
4.2. Synthesis of click branched RNAs.....	88
4.3. Inhibition of Dbr1p with cbRNA.....	91
4.4 Conclusions.....	96
4.5. Experimental.....	97
4.6. References.....	100
Chapter 5: Mini-Lariat RNAs for single-stranded RNA interference.....	102
5.1. Introduction.....	102
5.2. Synthesis and Purification of Mini-Lariats.....	107
5.3. RNA Interference Assays.....	119
5.4. Conclusions.....	124
5.5. Experimental.....	125
5.6. References.....	129
Chapter 6: Protein Free Splicing.....	132
6.1. Introduction.....	132
6.2. Synthesis of backbone branched RNA for Splicing Studies.....	136
6.3. Melting Temperature Studies.....	137
6.4. RNA Binding Assay.....	143
6.5. RNA Cleavage Assay.....	150
6.6. Conclusions.....	154
6.7. Experimental.....	156
6.8. References.....	164
Chapter 7: Conclusions and Future Directions.....	168
7.1. References.....	173

List of Figures

Chapter 1

Figure 1.1. Mechanism of pre-mRNA splicing.....	2
Figure 1.2. Overview of the splicing pathway.....	3
Figure 1.3. Alignment of Dbr1p amino acid sequences from a variety of organisms.....	7
Figure 1.4. Overlay of the active site of EhDbr1p from two crystal structures.....	9
Figure 1.5 Schematic of backbone branched RNA.....	10
Figure 1.6. Schematic representation of deoxyribozyme mediated lariat RNA synthesis.....	10
Figure 1.7. RNA synthesis cycle.....	12
Figure 1.8. Schematic of different branch point residues.....	12

Chapter 2

Figure 2.1. 6-nitroveratryloxymethyl and ortho-nitrobenzyl protecting groups.....	24
Figure 2.2. Diagram of the 2'-Method and 3'-Method for bbRNA synthesis.....	26
Figure 2.3. Detailed synthesis scheme of backbone branched RNAs.....	27
Figure 2.4. Results from photodeprotection of a 2'-NVOM protected DNA outside the synthesizer.....	32
Figure 2.5. Set up of UV LEDs and aluminum block around the synthesis column.....	33
Figure 2.6. Results from photodeprotection of a 2'-NVOM protected DNA inside the synthesizer.....	34

Chapter 3

Figure 3.1. Overlay of Dbr1p and Mre11 active site.....	43
Figure 3.2. Representative polyacrylamide gel and debranching plot.....	47
Figure 3.3. Mn ²⁺ titration with ScDbr1p and EhDbr1p.....	48
Figure 3.4. The effect of EDTA on EhDbr1p activity.....	49
Figure 3.5. The effect of preincubation of EhDbr1p with EDTA.....	50
Figure 3.6. EhDbr1p activity with added metal ions.....	51
Figure 3.7. The effect of reducing agents on the EhDbr1p cleavage rate.....	53
Figure 3.8. Structure of DNA/RNA hybrids.....	54
Figure 3.9. Observed cleavage rates of modified bbRNAs by Sc and EhDbr1p.....	55

Figure 3.10. Structure of EhDbr1p associated with branched RNA.....	57
Figure 3.11 Digestion of DNA/ RNA hybrids with high enzyme concentration.....	59
Figure 3.12. Digestion of RNA with a single 2'-5'-phosphodiester bond.....	59
Figure 3.13. Backbone branched RNA and duplexed bbRNA.....	61
Figure 3.14. Digestion of bbRNA-DNA hybrids by EhDbr1p-1.....	62
Figure 3.15. Digestion of bbRNA-DNA hybrids by EhDbr1p-2.....	64
Figure 3.16. Initial rate debranching assay of different sized bbRNAs.....	67
Figure 3.17. Initial rate debranching assay of bbRNA with hexynyl linker.....	69
Figure 3.18. Digestion of dual labeled bbRNAs with alternate branch points by EhDbr1p.....	69
Figure 3.19. Digestion of alternate branch point bbRNA.....	70
Figure 3.20. Proposed site of Dbr1p cleavage with pyrimidine branch point residues.....	71
Figure 3.21. Initial rate debranching assay of branch point adenosine and branch point uridine bbRNA.....	71
Figure 3.22. Kinetic model for substrate inhibition.....	72
Figure 3.23. Initial rate debranching assay of branch point cytidine by EhDbr1p.....	74

Chapter 4

Figure 4.1. Structure of bbRNA inhibitors for Dbr1p.....	87
Figure 4.2. Difference between the natural bbRNA substrate and the click branched inhibitor..	88
Figure 4.3. Purification gel of cbRNA4.....	89
Figure 4.4 Lineweaver-Burk plots of different types of inhibition.....	92
Figure 4.5. Competitive inhibition Lineweaver-Burk plots of the cbRNAs.....	93
Figure 4.6. Plots to determine inhibition constant of cbRNAs.....	94
Figure 4.7. Inhibitor strength of different cbRNAs.....	95

Chapter 5

Figure 5.1. Mirtron biogenesis pathway.....	103
Figure 5.2. Proposed mini-lariat action pathway.....	105
Figure 5.3. HPLC chromatogram of crude mini-lariat reaction.....	110
Figure 5.4. MALDI-TOF spectrum of mini-lariat RNA.....	111
Figure 5.5. HPLC chromatogram of crude mini-lariat reaction - CE removal.....	113
Figure 5.6. HPLC chromatogram of crude mini-lariat reaction - ammonium salts.....	114

Figure 5.7. Set up for cooling the mini-lariat reaction.....	115
Figure 5.8. HPLC chromatogram of crude mini-lariat reaction - cooling the reaction.....	116
Figure 5.9. HPLC chromatogram mini-lariat RNA with 2'-F modifications.....	117
Figure 5.10. MALDI-TOF spectrum of mini-lariat RNA with 2'-F modifications.....	117
Figure 5.11. Luciferase activity after treatment with mini-lariat RNAs.....	120
Figure 5.12. Quantitative real time PCR analysis of androgen receptor mRNA expression.....	121
Figure 5.13. Western blot of androgen receptor post treatment with mini-lariat RNAs.....	122

Chapter 6

Figure 6.1. Detailed overview of the splicing pathway.....	133
Figure 6.2. RNA configuration at the core of the spliceosome.....	134
Figure 6.3. 3-D representation of RNA in spliceosome core.....	134
Figure 6.4. Purification gel of bbRNA17.....	136
Figure 6.5. RNA hybrids for melting temperature studies.....	138
Figure 6.6. Representative melting curve and first derivative plot.....	139
Figure 6.7. First derivative plot of hybrid I.....	142
Figure 6.8. Schematic representation of bbRNA17, U2, and U6 RNA associated together.....	144
Figure 6.9. RNA binding assay.....	145
Figure 6.10. The effect of changing annealing procedures on RNA binding.....	146
Figure 6.11. The effect of pre-annealing on RNA binding.....	147
Figure 6.12. The effect of increasing MgCl ₂ concentration on RNA binding.....	148
Figure 6.13. The effect of a crowding agent on RNA binding.....	149
Figure 6.14. Cleavage reaction for in vitro splicing.....	151
Figure 6.15. Redesigned cleavage assay.....	153
Figure 6.16. Redesigned cleavage assay under a variety of conditions.....	153

List of Tables

Chapter 2

Table 2.1. Synthesis conditions for RNA and DNA, forward and reverse synthesis.....	27
Table 2.2. Solid phase synthesis conditions for DNA oligonucleotides.....	39

Chapter 3

Table 3.1. RNA used in 3.2.....	46
Table 3.2. RNA used in 3.3.1.....	55
Table 3.3. RNA used in 3.3.2.....	60
Table 3.4. RNA used in 3.3.3.....	61
Table 3.5. RNA used in 3.3.4.....	66
Table 3.6. RNA used in 3.3.5.....	68
Table 3.7. Calculated kinetic parameters for bbRNA9 and bbRNA12.....	73
Table 3.8. Sequences used in the biochemical studies of Dbr1p.....	77

Chapter 4

Table 4.1. Click branched RNA.....	90
Table 4.2. Relative amount of reagents used in CuAAC reaction.....	98

Chapter 5

Table 5.1. Mini-lariat synthesis conditions for steps 2-4.....	109
Table 5.2. Sequences used in the luciferase assay.....	119
Table 5.3. Concentration of reagents for CuAAC reaction.....	128

Chapter 6

Table 6.1. bbRNA synthesized by modified synthetic procedure.....	136
Table 6.2. Individual sequences for T_M experiments.....	137
Table 6.3. RNA hybrids and calculated melting point.....	140
Table 6.4. Sequences used in initial binding assay.....	143
Table 6.5. Sequences used in the cleavage experiments.....	151
Table 6.6. Sequences used in the redesigned splicing assay.....	152
Table 6.7. Sequences made using a modified 2'-Method.....	156
Table 6.8. RNA concentrations used in binding assay 1.....	158

Table 6.9. RNA concentrations used in binding assay 2.....	159
Table 6.10. RNA used in first cleavage test.....	160
Table 6.11. RNA used in second and third cleavage assays.....	162

List of Schemes

Chapter 2

Scheme 2.1. Synthesis of 2' and 3'-NVOM protected adenosine alcohol.....	25
Scheme 2.2. Synthesis of 2'-NVOM 3'-amidite monomer.....	25
Scheme 2.3. Outside synthesizer photodeprotection reaction scheme.....	31
Scheme 2.4. Inside synthesizer photodeprotection reaction scheme.....	34

Chapter 4

Scheme 4.1. Conjugation scheme for click branched RNA.....	89
--	----

Chapter 5

Scheme 5.1. Synthesis Scheme for Mini-Lariat RNAs.....	107
Scheme 5.2. Two different synthetic routes to mini-lariats.....	112
Scheme 5.3. Cyclization reaction to generate click-linked mini-lariats.....	118

Chapter 1 – Introduction to Splicing, Backbone Branched RNA and Lariat Debranching Enzyme

Over the years the perception that RNA is a minor macromolecule and doesn't play a large role in the cell has been drastically changing. All DNA is transcribed into RNA but only a fraction of that codes for protein, the remaining RNA untranslated². The non-coding RNA performs regulatory functions in the cell that range from modulating gene expression, to forming structures, to containing catalytic elements^{3,4}. Directly transcribed RNA contains both coding and non-coding RNA regions^{5,6}. In order for the message for the ribosome to be concise, non-coding RNA is excised from pre-messenger RNA (pre-mRNA) through a process called splicing⁷. The coding sequences in pre-mRNA are called exons and the non-coding sequences are called introns. Through splicing, the exons are ligated together while the introns are removed^{7,4}.

Splicing is a complicated process that involves a myriad of factors. There are small nuclear RNA (snRNA) which are required for splicing and make up the catalytic core of the complex, along with two metal ions that catalyze the reactions⁷. These snRNA are complexed with proteins and create small nuclear ribonucleoproteins (snRNPs) that move in and out of the splicing cycle at different points. The first catalytic step in the process involves cleaving the 5'-splice site (5'-SS) on pre-mRNA and separating the 5'-exon from the intron^{7,8}. The terminal 5'-OH of the intron then performs a metal-mediated S_N2 transesterification and reacts with the 2'-OH of the branch point residue and creates the 2'-5'-phosphodiester bond^{7,8,74}. The second catalytic step of splicing involves the 3'-terminus of the 5'-exon ligating with the 5'-terminus of the 3'-exon through a second S_N2 transesterification^{7,8}. The second ligation forms mature mRNA. These two steps are detailed in figure 1.1.

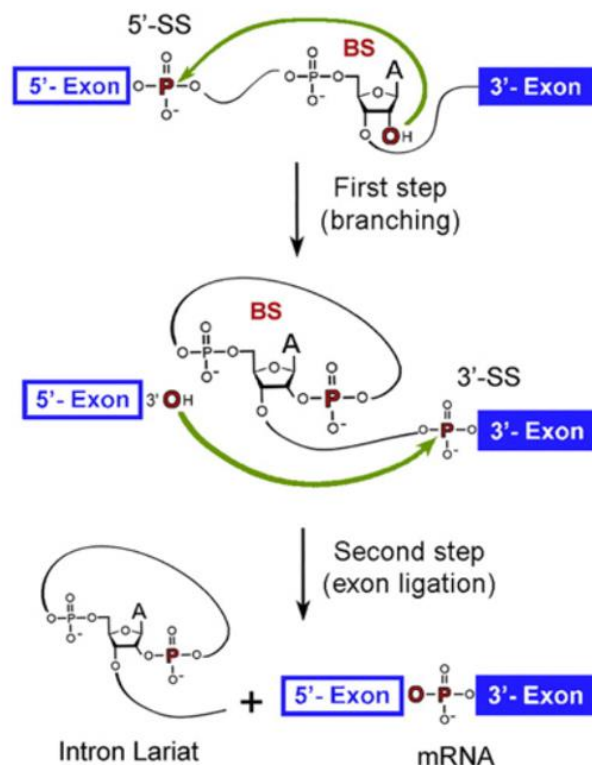


Figure 1.1. Mechanism of pre-mRNA splicing. The 5'-splice site (5'-SS) is attacked by the 2'-OH of the branch point residue, forming a 2'-5'-phosphodiester bond. The 3'-OH from the 5'-exon attacks the 3'-splice site (3'-SS). The splicing products are ligated exons and excised intron lariats. Figure from ref.⁷

The introns are removed as lariats, with the 5'-terminus linked to an internal 2'-OH group, forming a 2'-5'-phosphodiester bond. Once introns are excised, they are either broken down into their constituent monomers or perform a regulatory function in the cell^{7,9}. Introns are known to function in microRNA (miRNA) biogenesis, small nucleolar RNA (snoRNA) biogenesis, and transcription regulation among other functions⁹. Before an intron can go on to function in these different cellular pathways, the 2'-5'-phosphodiester bond needs to be broken. The 2'-5'-linkage is not cleavable by most cellular nucleases. Lariat debranching enzyme (Dbr1p) is a ubiquitous, evolutionarily conserved nuclease that specifically cleaves the 2'-5'-phosphodiester bond of lariat introns³⁸. The main focus of this thesis is the use of branched RNA that include a 2'-5'-linkage for biochemical interrogation of lariat debranching enzyme.

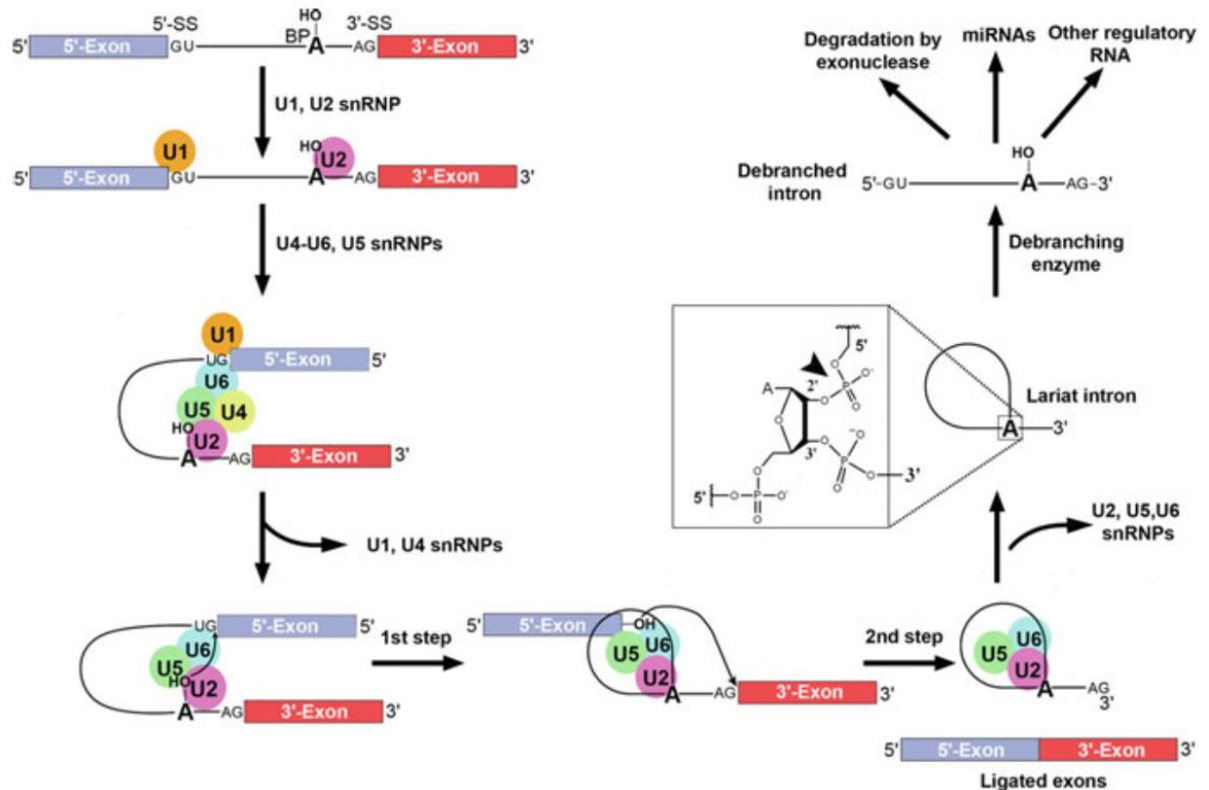


Figure 1.2. Overview of splicing pathway. The left column shows the pre-mRNA and different snRNP factors that enter and exit the pathway. The bottom row details the two catalytic steps, and the right column shows Dbr1p action. Figure from ref.¹⁰

1.1. Dbr1p in Regulatory Processes and the Health Implications

Because of Dbr1p's role in turning over introns, enzyme activity has been tied to a variety of different cellular processes. Lack of Dbr1p leads to an accumulation of lariat introns. Intron accumulation, in turn, prevents miRNA synthesis, snoRNA biogenesis, and nucleotide pool recycling, wreaking havoc on cells.

Splicing is a highly regulated pathway, with different factors entering and exiting the pathway in a tightly controlled manner^{7,11}. Dbr1p is associated with turnover of the splicing complexes after the two catalytic steps have occurred¹². After exon ligation, the intron remains associated with snRNPs¹². Intron disassociation and recycling of snRNPs to a pre-catalytic state is dependent on Dbr1p¹². With less Dbr1p to debranch introns, introns remain associated with snRNPs and prevent snRNP recycling¹². The pathway block causes subsequent splicing to occur only at sites that are highly complementary to

U2 snRNA branch recognition sequence, known as a strong branch site¹². If an intron contains a weak branch site, i.e. has low sequence homology to the U2 binding site, and the cells are deficient in Dbr1p, alternative splicing occurs, the weak branch site is not recognized and the adjacent exon is skipped, leading to alternative splicing¹². The increase in variant RNA that results from alternative splicing, results in an increase in improperly formed protein, and eventually leads to tumor growth in human cells¹².

In solid cancerous tumors, the environment is generally hypoxic. Under hypoxic conditions, Dbr1p expression is regulated by p53, a tumor suppressor, along with hypoxia inducible factor1 (HIF1)¹². Under hypoxic conditions Dbr1p is upregulated, indicating that it facilitates p53 tumor suppressor activity¹².

Dbr1p plays a role in miRNA biogenesis in a pathway that converges with the canonical miRNA pathway. MicroRNA are traditionally derived from specific miRNA transcripts, pri-miRNA, but a portion of miRNA are formed from introns, and are termed mirtrons^{13,14}. These will be discussed in greater detail in chapter 5, but mirtrons represent a significant portion of total miRNA¹⁵⁻¹⁷. Dbr1p has also been shown to play a role in snoRNA biogenesis, as many snoRNA are derived from introns¹⁸. Improper intron processing inhibits the biosynthesis of both of these important regulatory molecules.

Dbr1p has also been shown to be critical for retrotransposon function in *S. cerevisiae*¹⁹⁻²⁴. Retrotransposons are mobile genetic elements and retrotransposons function similar to retroviruses²⁵. A DNA element is transcribed into RNA, and then the RNA is reverse transcribed back into DNA, which can then imbed in the genome²⁵. While there is controversy in the literature over the exact role Dbr1p plays in Ty1 and Ty3 retrotransposition, Dbr1p depletion is linked with decreased transposon cDNA synthesis^{23,24}. Different groups have attributed the decreased amount of cDNA to the Ty1 mRNA being lariat shaped and containing a 2'-5'-phosphodiester bond^{21,22}. Others have

refuted that claim as they observed no evidence of a lariat in the Ty1 retrotransposon and the function that Dbr1p plays in this process remains to be determined^{19,20}.

Depleting cellular Dbr1p has been linked to inhibition of human immunodeficiency virus (HIV-1)^{26–28}. HIV-1 and other retroviruses are related to retrotransposons as they have a similar mechanism for replication and gene integration²⁶. It has been shown that Dbr1p is required for HIV-1 cDNA synthesis and that knocking down the enzyme decreases the total amount of nuclear cDNA²⁷. The 5'-end of HIV-1 RNA is thought to form a lariat structure that requires Dbr1p to cleave the 2'-5'-phosphodiester bond²⁸. Less Dbr1p means fewer opened lariats, and thus cDNA synthesis cannot proceed, making Dbr1p an interesting therapeutic target for HIV-1 treatment.

Humans that have bi-allelic Dbr1p mutants are more susceptible to brainstem viral encephalitis²⁹. Missense and nonsense mutations, are mostly localized to the metallophosphoesterase domain of the protein²⁹. The specific mutants found in five patients from three family lines are L13G, Y17H, I120T, and R197X²⁹. The exact cellular mechanism for increased viral susceptibility is unknown, but the mutations cause introns to accumulate²⁹. When cells derived from patients are infected with a virus like herpes simplex virus (HSV1), the amount of unprocessed introns increases further²⁹. The increased intron amounts may hinder the host cell's ability to detect an invading virus or cause misregulation of a pathway involved in virus recognition²⁹.

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that in some cases is brought on by aggregation of trans-activation response element DNA binding protein-43 (TDP-43) in the cytoplasm³⁰. In yeast cells overexpressing TDP-43, Dbr1p was found to be a toxicity modifier; less Dbr1p resulted in lower toxicity³⁰. The accumulated introns sequester TDP-43 in the cytoplasm and prevent protein aggregation and TDP-43 from interacting with cellular RNAs³⁰.

Overall, Dbr1p is associated with a variety of human diseases. Decreasing Dbr1p causes an increase in cellular abundance of lariat introns. Increasing the amount of introns usually is deleterious to proper cell function by inhibiting miRNA and snoRNA biogenesis and causing splicing machinery to stall in the post-catalytic complex^{12,13,18,31}. Decreased Dbr1p concentration has also been associated with tumor growth and inhibition of Ty1 retrotransposition^{12,21}. However, Dbr1p might be a therapeutic target for people with ALS or people infected with HIV-1^{28,30}. Further exploration of the biochemistry of Dbr1p is necessary to better understand the enzyme.

1.2. Biochemistry of Lariat Debranching Enzyme

Debranching activity was first seen in HeLa cell extracts in 1985 by Ruskin and Green and is ubiquitous in eukaryotic cells³². Deletion of the Dbr1p gene is lethal in higher organisms. *Saccharomyces cerevisiae* can survive without Dbr1p, but deletion of Dbr1p in *Schizosaccharomyces pombe* causes severe growth defects³³. The difference between these two forms of yeast is presumably because *S. pombe* has roughly 40x the number of introns than *S. cerevisiae*³³. Inhibiting Dbr1p causes introns to accumulate because no other enzyme can cleave the 2'-5'-phosphodiester bond³⁴. With introns accumulating in their lariat form, they cannot go on to perform regulatory functions and the cellular nucleotide pool is rapidly depleted, causing negative health effects, as described above³⁴.

Dbr1p is an evolutionarily conserved nuclease and debranching activity is seen in both prokaryotes and eukaryotes. Prokaryotes that do not have introns contain multicopy single-stranded DNA (msDNA), a DNA/ RNA hybrid linked together through a 2'-5'-phosphodiester bond, which is a substrate for Dbr1p^{35,76}. In eukaryotes, Dbr1p is required to cleave branched nucleic acid structures, like introns. Dbr1p is functional across species. In *S. cerevisiae* or *S. pombe* strains that are mutated to lack the DBR1 gene, human cDNA

for Dbr1p complement the phenotype of intron accumulation⁷⁵. The amino acid sequences of Dbr1p across different species is highly conserved, especially in the N-terminal portion of the protein (fig. 1.3).

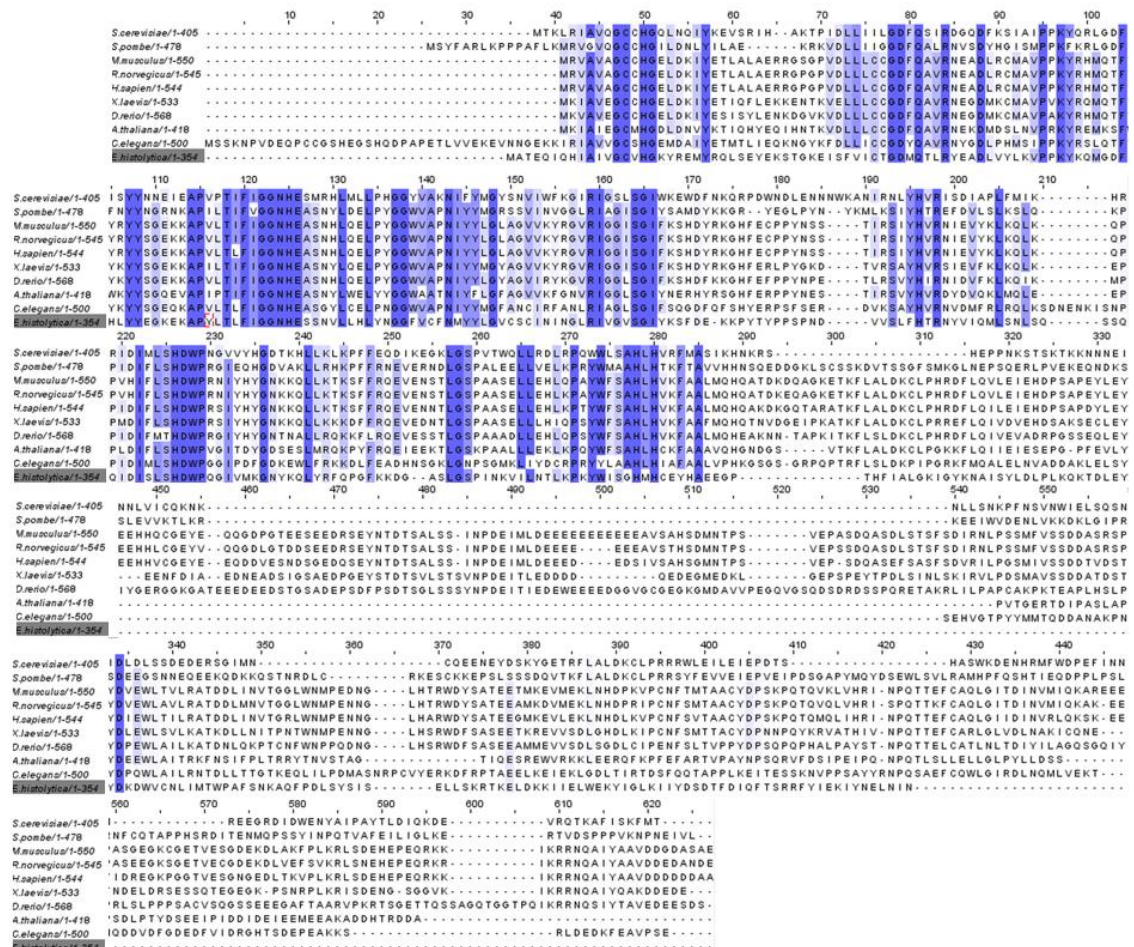


Fig. 1.3. Alignment of Dbr1p amino acid sequences from a variety of organisms laid out from N-terminus to C-terminus. Dbr1p amino acids from *S.cerevisiae*, *S.pombe*, *M.musculus*, *R. norvegicus*, *H.sapien*, *X.laavis*, *D.reio*, *A.thaliana*, *C.elegans*, and *E.histolytica*. Blue highlights denote sequence homology across species and the intensity of the color correlates with the strength of the conservation with darker being more conserved. The alignment was generated using ClustalW.

As a 2'-5'-phosphodiesterase, Dbr1p's main role in the cell is to cleave that bond in lariat introns.. Analysis of substrate requirements revealed that Dbr1p is active on substrates that contain phosphodiester bonds at the 2'-position as well as the 3'-position of the branch point residue^{32,35,36}. However, recently it was discovered that Dbr1p is able

to cleave bis-*p*-nitrophenylphosphate, a substrate quite unlike traditional backbone branched RNA³⁷.

Since the discovery of debranching activity, Dbr1p has been studied by a small number of groups. Dbr1p is a metallophosphoesterase that specifically cleaves 2'-5'-phosphodiesterbonds and does not interact with the vicinal 3'-5'-phosphodiester bond³². The enzyme shows remarkable sequence homology to Mre11³⁸. Mre11 is part of the metallophosphoesterase superfamily and is a DNA phosphodiesterase with 3'-5' exo and endonuclease activity³⁹. The active sites for Dbr1p and Mre11 are very similar, with one notable exception, an aspartic acid residue in Mre11 is a cysteine in Dbr1p^{38,39}.

The active site of Dbr1p not only holds the RNA substrate but also contains the catalytic metal ions. Metallophosphoesterases require two metal ions to complete catalysis^{8,39}. The specific ions that reside in the metal binding pocket of Dbr1p remains an unresolved issue in the literature. Biochemical analysis suggested that Dbr1p derived from *S. cerevisiae* was Mn²⁺-dependent³⁸. There are three published crystal structures of Dbr1p, all from *Entamoeba histolytica*⁴⁰⁻⁴². The first structure contained a Mn²⁺ ion in the beta pocket, and no metal in the alpha pocket⁴². The second structure observed Fe²⁺ in the beta pocket and Zn²⁺ in the alpha pocket⁴¹. The third structure determined Mn²⁺ and Zn²⁺ in the beta and alpha pockets respectively⁴⁰. Based on the variety of metal ions observed in the binding pockets of EhDbr1p crystal structures, the metal requirements for catalysis are flexible.

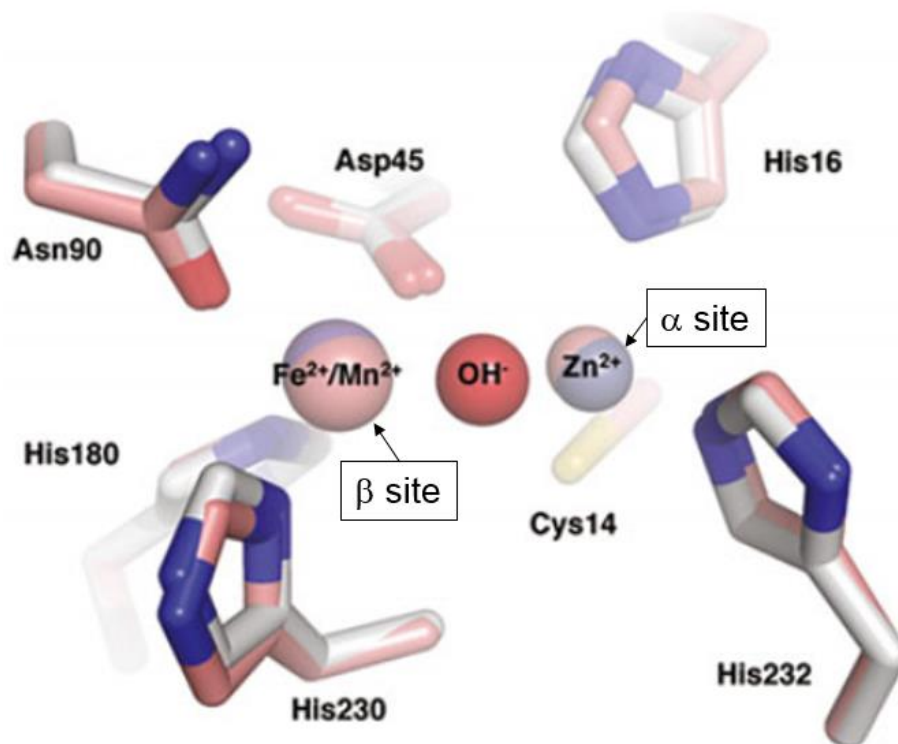


Figure 1.4. Overlay of the active site of EhDbr1p from two crystal structures. Similar coordination of Fe^{2+} and Mn^{2+} is observed. Figure adapted from ref⁴⁰.

1.3. Analyzing the Dbr1p reaction *in vitro* with backbone branched RNA

With the important role Dbr1p plays in the cell and the discrepancies in the literature about substrate and metal requirements, further investigation of Dbr1p is required. Lariat introns are difficult to work with, they need to be extracted from cells, are long, and there is little to no control over sequence. To overcome the obstacles of using lariat introns to study Dbr1p, small synthetic intron analogues have been synthesized. These are small RNAs that contain a 2'-5'-phosphodiester bond, like lariat introns. The RNAs are not looped together, but are Y shaped, containing a branch point residue with a 5'-arm, a 3'-arm, and a 2'-arm extending outward. The Y shaped RNAs, termed backbone branched RNAs (bbRNAs) are substrates of Dbr1p. Backbone branched RNA has been made in solution and by solid phase synthesis methods³⁵⁻⁵⁸. Additionally, a deoxyribozyme that creates bbRNA and lariat RNAs that contain a 2'-5'-phosphodiester bond has been reported^{62,67-}

⁷⁰. The 2'-hydroxyl of a branch point adenosine residue attacks a guanosine triphosphate, releasing pyrophosphate and ligating two RNA pieces together with a 2'-5'-phosphodiester bond⁶⁹. However, it is difficult to incorporate modifications or labels onto bbRNA when using a ribozyme.

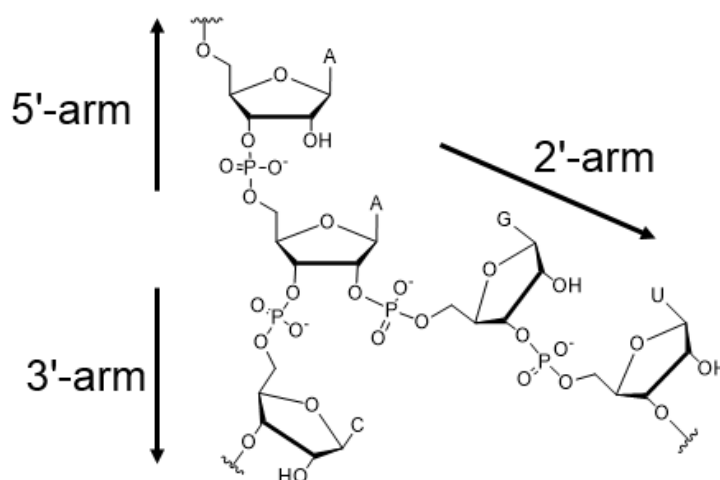


Figure 1.5. Schematic of backbone branched RNA centered around the branch point residue.

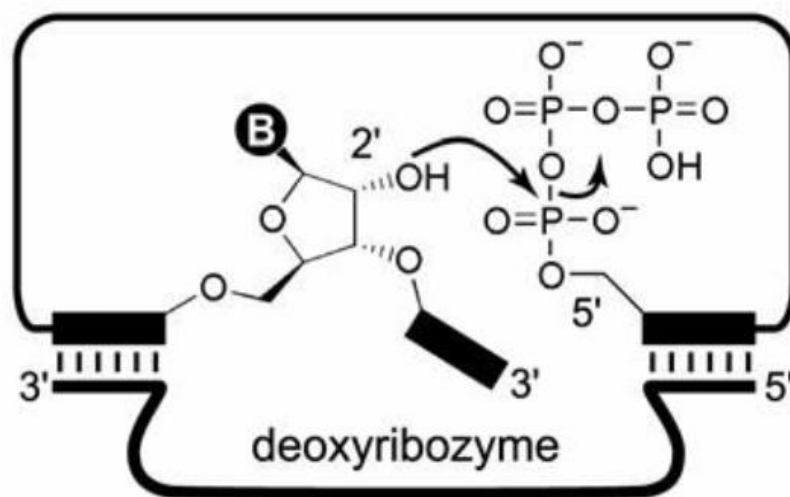


Figure 1.6. Schematic representation of deoxyribozyme lariat RNA synthesis. Figure from ref⁶⁸.

Other groups have explored solution and solid phase chemical synthesis to make bbRNA. Standard solid phase RNA synthesis is detailed in figure 1.7. RNA is grown from controlled pore glass beads in the 3'-to-5' direction using phosphoramidite monomers and

selective deprotection. To make backbone branched RNA, the standard procedure is slightly modified. RNA is grown using standard solid phase synthesis and then the branch point residue is coupled. Early work utilizing solid phase synthesis to create bbRNA uses a 2',3'-O-bisphosphoramidite branch point residue to crosslink two adjacent RNA strands together, creating a 3'-arm and a 2'-arm (fig. 1.8)^{51,54}. Traditional synthesis then continues to make the 5'-arm. The bisphosphoramidite technique generates bbRNA that have the same sequence in both the 2'-arm and the 3'-arm⁵⁴. To gain sequence control over every arm of the bbRNA, a protecting group is added at the 2'-position of the branch point ribose that is orthogonal to the acid, base, and HF labile protecting groups already employed in solid phase RNA synthesis. The Damha group uses a 2'-acetyl levulinyl ester group that is removed using hydrazine⁷¹. The Das lab has developed a similar method that used a 2'-photolabile protecting group that is removed using 365 nm light^{72,73}. Both methods for orthogonal 2'-protection give complete sequence control over each arm around the branch point residue.

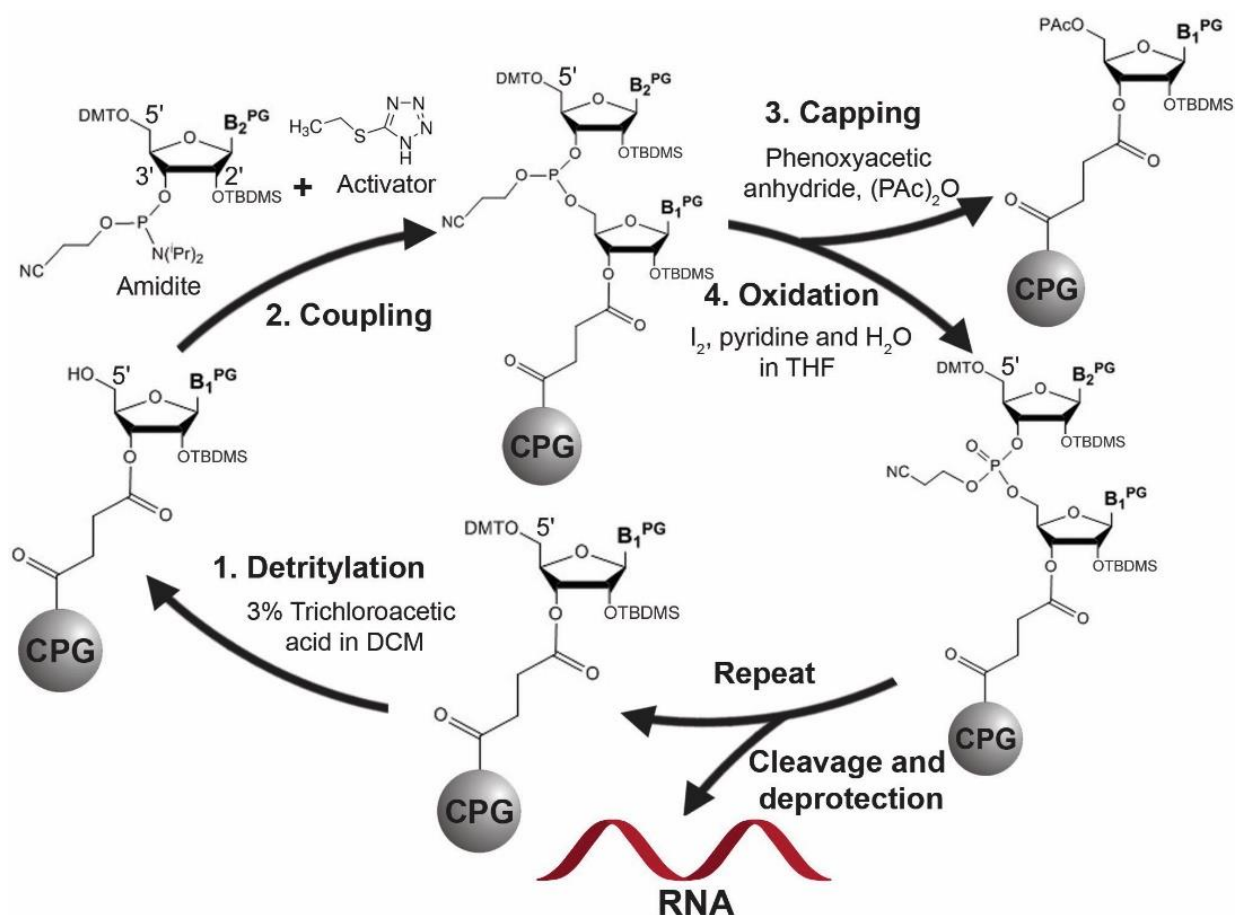


Figure 1.7. RNA synthesis cycle. RNA is linked to controlled pore glass (CPG). The 5'-protecting group, dimethoxytrityl (DMT), is removed first using 3% trichloroacetic acid, exposing the 5'-OH. The next RNA monomer is coupled using 5-ethylthio-1H-tetrazole. The newly formed phosphite linkage is oxidized into a phosphate using iodine. Uncoupled RNA is removed from the synthesis cycle by reacting the 5'-OH with phenoxyacetic anhydride. The cycle is repeated until the desired RNA is synthesized. The RNA is then cleaved from the bead and deprotected. Image from ref⁷³.

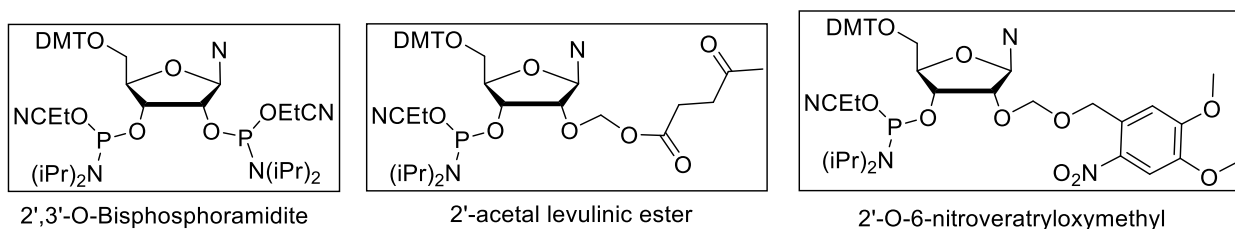


Figure 1.8. Schematic of different branch point residues to synthesize backbone branched RNA.

1.4. Scope of Thesis

There remains much work to be done to understand the full scope of Dbr1p's cellular activity. Dbr1p plays a critical role in many different cellular processes. Misregulation can have dire consequences but in certain circumstances, inhibiting Dbr1p has the potential to be greatly therapeutic. More information is needed about the extent of Dbr1p's debranching ability and kinetics in order to harness the Dbr1p pathway for potential therapeutic use. This thesis reports the biochemical interrogation of lariat debranching enzyme.

The major part of this thesis work focuses on the biochemistry of Dbr1p. In order to study the biochemistry, the synthesis method to make bbRNA is improved. The photolabile protecting group used in bbRNA synthesis previously was removed using a lamp, which was an overall inefficient method. The light source to remove the photolabile protecting group is upgraded to a UV-LED and kinetics of photodeprotection are explored (Chapter 2). With an improved bbRNA synthesis method, different Dbr1p substrates are made and the debranching capabilities of Dbr1p are further explored (Chapter 3). A multitude of non-traditional backbone branched RNAs are incubated with Dbr1p, including DNA/RNA hybrids, double stranded substrates, and bbRNA with non-canonical branch point residues. Further, the non-RNA requirements for optimal *in vitro* Dbr1p debranching activity are investigated, including Mn^{2+} , metal chelator EDTA, and other divalent metal ions.

With recent reports that decreasing Dbr1p activity is beneficial to HIV-1 and ALS diseased cells, inhibitors of Dbr1p are of therapeutic interest. A few inhibitors of Dbr1p are present in the literature, but are cleavable by the enzyme, just at a much slower rate than native 2'-5'-phosphodiester bonds. Using copper-catalyzed azide alkyne cycloaddition, backbone branched RNA that contain a 2'-5'-triazole linkage are made (Chapter 4). The

unnatural triazole linkage is uncleavable by Dbr1p. These click branched RNA (cbRNA) act as competitive inhibitors and the inhibition kinetics of different sized cbRNA are determined.

Taking inspiration from the mirtron pathway, small lariat shaped RNAs are synthesized on the solid phase, called mini-lariats (Chapter 5). Mini-lariats are short RNAs of the same sequence as the guide strand of a miRNA. The 5'-terminus of the mini-lariat is linked to an internal 2'-OH group, creating a 2'-5'-phosphodiester bond. Even with a loop size smaller than native introns, Dbr1p can cleave the 2'-5'-linkage and open up the mini-lariats, forming single stranded RNA. The efficiency of mini-lariats in the RNA interference pathway is determined in human prostate cancer cells.

Finally, backbone branched RNA are explored in the context of splicing. The spliceosome is upstream of Dbr1p, but its main function is to form lariat intron RNAs and subsequently remove the introns from pre-mRNA. The strength of RNA interaction in the spliceosome is explored through melting temperature analysis in a smaller model system that contains bbRNA and truncated snRNAs. Finally *in vitro* protein-free splicing assays are performed to look for both RNA association in binding assays, and RNA catalysis in cleavage assays. The work described herein will further the knowledge of Dbr1p activity in the context of *in vitro* debranching, inhibition, and RNA interference.

1.5. References

- (1) CRICK, F. Central Dogma of Molecular Biology. *Nature* **1970**, 227, 561–563.
- (2) Consortium, I. H. G. S. Finishing the Euchromatic Sequence of the Human Genome. *Nature* **2001**, 431, 931–945.
- (3) Cech, T. .; Steitz, J. . The Noncoding RNA Revolution - Trashing Old Rules to Forge New Ones. *Cell* **2014**, 157, 77–94.
- (4) Serganov, A.; Patel, D. J. Ribozymes, Riboswitches and beyond: Regulation of Gene Expression without Proteins. *Nat. Rev. Genet.* **2007**, 8, 776–790.
- (5) Sharp, P. A. Split Genes and RNA Splicing Nobel Lecture. *Cell* **1994**, 77.
- (6) Lee, Y.; Rio, D. C. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* **2015**, 84, 1–33.
- (7) Shi, Y. The Spliceosome: A Protein-Directed Metalloribozyme. *J. Mol. Biol.* **2017**, 429, 2640–2653.
- (8) Steitz, T. A.; Steitz, J. A. A General Two-Metal-Ion Mechanism for Catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, 90, 6498–6502.
- (9) Hesselberth, J. R. Lives That Introns Lead after Splicing. *Wiley Interdiscip. Rev. RNA* **2013**, 4, 677–691.
- (10) Dey, S. K.; Paredes, E.; Evans, M.; Das, S. R. From Nucleic Acids Sequences to Molecular Medicine. In *From Nucleic Acids Sequences to Molecular Medicine*; Erdmann, V.; J., B., Eds.; Springer: Berlin, Heidelberg, 2012; pp. 475–501.
- (11) Shi, Y. Mechanistic Insights into Precursor Messenger RNA Splicing by the Spliceosome. *Nat. Rev. Mol. Cell Biol.* **2017**, 18, 655–670.

- (12) Han, B.; Park, H. K.; Ching, T.; Panneerselvam, J.; Wang, H.; Shen, Y.; Zhang, J.; Li, L.; Che, R.; Garmire, L.; Fei, P. Human DBR1 Modulates the Recycling of SnRNPs to Affect Alternative RNA Splicing and Contributes to the Suppression of Cancer Development. *Oncogene* **2017**, 1–10.
- (13) Ruby, J. G.; Jan, C. H.; Bartel, D. P. Intronic MicroRNA Precursors That Bypass Drosha Processing. *Nat. Rev. Mol. Cell Biol.* **2007**, *448*, 83–86.
- (14) Berezikov, E.; Chung, W. J.; Willis, J.; Cuppen, E.; Lai, E. C. Mammalian Mirtron Genes. *Mol. Cell* **2007**, *28*, 328–336.
- (15) Chung, W.; Agius, P.; Westholm, J. O.; Chen, M.; Okamura, K.; Robine, N.; Leslie, C. S.; Lai, E. C. Computational and Experimental Identification of Mirtrons in *Drosophila Melanogaster* and *Caenorhabditis Elegans*. **2011**, 286–300.
- (16) Ladewig, E.; Okamura, K.; Flynt, A. S.; Westholm, J. O.; Lai, E. C. Discovery of Hundreds of Mirtrons in Mouse and Human Small RNA Data. *Genome Res.* **2012**, *22*, 1634–1645.
- (17) Wen, J.; Ladewig, E.; Shenker, S.; Mohammed, J.; Lai, E. C. Analysis of Nearly One Thousand Mammalian Mirtrons Reveals Novel Features of Dicer Substrates. *PLoS Comput. Biol.* **2015**, *11*.
- (18) Ooi, S. L.; Samarsky, D. a; Fournier, M. J.; Boeke, J. D. Intronic SnoRNA Biosynthesis in *Saccharomyces Cerevisiae* Depends on the Lariat-Debranching Enzyme: Intron Length Effects and Activity of a Precursor SnoRNA. *RNA* **1998**, *4*, 1096–1110.
- (19) Pratico, E. D.; Silverman, S. K. Ty1 Reverse Transcriptase Does Not Read through the Proposed 2',5'-Branched Retrotransposition Intermediate in Vitro. *RNA*

2007, 13, 1528–1536.

(20) Coombes, C. E.; Boeke, J. E. F. D. An Evaluation of Detection Methods for Large Lariat RNAs. *RNA* **2005**, 11, 323–331.

(21) Cheng, Z.; Menees, T. M. RNA Branching and Debranching in the Yeast Retrovirus-like Element Ty1. *Science* (80-.). **2004**, 303, 240–243.

(22) Salem, L. A.; Boucher, C. L.; Menees, T. M. Relationship between RNA Lariat Debranching and Ty1 Element Retrotransposition. *J. Virol.* **2003**, 77, 12795–12806.

(23) Karst, S. M.; Rütz, M. L.; Menees, T. M. The Yeast Retrotransposons Ty1 and Ty3 Require the RNA Lariat Debranching Enzyme, Dbr1p, for Efficient Accumulation of Reverse Transcripts. *Biochem. Biophys. Res. Commun.* **2000**, 268, 112–117.

(24) Lauermann, V.; Nam, K.; Trambly, J.; Boeke, J. D. Plus-Strand Strong-Stop DNA Synthesis in Retrotransposon Ty1. *J. Virol.* **1995**, 69, 7845–7850.

(25) Finnegan, D. J. Retrotransposons. *Curr. Biol.* **2012**, 22, 432–437.

(26) Ye, Y.; De Leon, J.; Yokoyama, N.; Naidu, Y.; Camerini, D. DBR1 SiRNA Inhibition of HIV-1 Replication. *Retrovirology* **2005**, 2, 63.

(27) Galvis, A. E.; Fisher, H. E.; Nitta, T.; Fan, H.; Camerini, D. Impairment of HIV-1 CDNA Synthesis by DBR1 Knockdown. *J. Virol.* **2014**, 88, 7054–7069.

(28) Galvis, A. E.; Fisher, H. E.; Fan, H.; Camerini, D. Conformational Changes in the 5' End of the HIV-1 Genome Dependent on the Debranching Enzyme DBR1 during Early Stages of Infection. *J. Virol.* **2017**, 91, 1–11.

(29) Zhang, S.-Y.; Clark, N. E.; Freije, C. A.; Hart, P. J.; Etzioni, A.; Casanova
Correspondence, J.-L.; Ciancanelli, M. J.; Biran, A.; Lafaille, F. G.; Tsumura, M.; Reino,

E. J. G.; Dobbs, K.; Hasek, M. Inborn Errors of RNA Lariat Metabolism in Humans with Brainstem Viral Infection. *Cell* **2018**, *172*, 952–965.

(30) Armakola, M.; Higgins, M. J.; Figley, M. D.; Barmada, S. J.; Scarborough, E. A.; Diaz, Z.; Fang, X.; Shorter, J.; Krogan, N. J.; Finkbeiner, S.; Farese, R. V.; Gitler, A. D. Inhibition of RNA Lariat Debranching Enzyme Suppresses TDP-43 Toxicity in ALS Disease Models. *Nat. Genet.* **2012**, *44*, 1302–1309.

(31) Okamura, K.; Hagen, J. W.; Duan, H.; Tyler, D. M.; Lai, E. C. The Mirtron Pathway Generates MicroRNA-Class Regulatory RNAs in Drosophila. *Cell* **2007**, *130*, 89–100.

(32) Ruskin, B.; Green, M. R. An RNA Processing Activity That Debranches RNA Lariats. *Science (80-.)*. **1985**.

(33) Nam, K.; Lee, G.; Trambly, J.; Devine, S. E.; Boeke, J. D. Severe Growth Defect in a Schizosaccharomyces Pombe Mutant Defective in Intron Lariat Degradation. *Mol. Cell. Biol.* **1997**, *17*, 809–818.

(34) Chapman, K. B.; Boeke, J. D. Isolation and Characterization of the Gene Encoding Yeast Debranching Enzyme. *Cell* **1991**, *65*, 483–492.

(35) Nam, K.; Hudson, R. H. E.; Chapman, K. B.; Ganeshan, K.; Damha, M. J.; Boeke, J. D. Yeast Lariat Debranching Enzyme: Substrate and Sequence Specificity. *J. Biol. Chem.* **1994**, *269*, 20613–20621.

(36) Ooi, S. L.; Dann, C.; Nam, K.; Leahy, D. J.; Damha, M. J.; Boeke, J. D. *RNA Lariat Debranching Enzyme*; Elsevier Masson SAS, 2001; Vol. 342.

(37) Schwer, B.; Khalid, F.; Shuman, S. Mechanistic Insights into the Manganese-Dependent Phosphodiesterase Activity of Yeast Dbr1 with Bis-p- Nitrophenylphosphate

and Branched RNA Substrates. *RNA* **2016**, 22, 1819–1827.

(38) Khalid, M. F.; Damha, M. J.; Shuman, S.; Schwer, B. Structure-Function Analysis of Yeast RNA Debranching Enzyme (Dbr1), a Manganese-Dependent Phosphodiesterase. *Nucleic Acids Res.* **2005**, 33, 6349–6360.

(39) Matange, N.; Podobnik, M.; Visweswariah, S. S. Metallophosphoesterases: Structural Fidelity with Functional Promiscuity. *Biochem. J.* **2015**, 467, 201–216.

(40) Ransey, E.; Paredes, E.; Dey, S. K.; Das, S. R.; Heroux, A.; Macbeth, M. R. Crystal Structure of the *Entamoeba Histolytica* RNA Lariat Debranching Enzyme EhDbr1 Reveals a Catalytic Zn²⁺/Mn²⁺ Heterobinucleation. *FEBS Lett.* **2017**, 591, 2003–2010.

(41) Clark, N. E.; Katolik, A.; Roberts, K. M.; Taylor, A. B.; Holloway, S. P.; Schuermann, J. P.; Montemayor, E. J.; Stevens, S. W.; Fitzpatrick, P. F.; Damha, M. J.; Hart, P. J. Metal Dependence and Branched RNA Cocrystal Structures of the RNA Lariat Debranching Enzyme Dbr1. *Proc. Natl. Acad. Sci.* **2016**, 113, 14727–14732.

(42) Montemayor, E. J.; Katolik, A.; Clark, N. E.; Taylor, A. B.; Schuermann, J. P.; Combs, D. J.; Johnsson, R.; Holloway, S. P.; Stevens, S. W.; Damha, M. J.; Hart, P. J. Structural Basis of Lariat RNA Recognition by the Intron Debranching Enzyme Dbr1. *Nucleic Acids Res.* **2014**, 42, 10845–10855.

(43) Damha, M. J.; Pon, R. T.; Ogilvie, K. K. Chemical Synthesis of Branched RNA: Novel Trinucleoside Diphosphates Containing Vicinal 2'-5' and 3'-5' Phosphodiester Linkages. *Tetrahedron Lett.* **1985**, 26, 4839–4842.

(44) Sekine, M.; Hata, T. Synthesis of Branched Ribonucleotides Related to the Mechanism of Splicing of Eukaryotic Messenger RNA. *J. Am. Chem. Soc.* **1985**, 107, 5813–5815.

- (45) Kierzek, R.; Kopp, D. W.; Edmonds, M.; Caruthers, M. H. Chemical Synthesis of Branched RNA. *Nucleic Acids Res.* **1986**, *14*, 4751–4764.
- (46) Sekine, M.; Heikkila, J.; Hata, T. A New Method for the Synthesis of Branched Oligoribonucleotides Using a Fully Protected Branched Triribonucleoside Diphosphate Unit. *Tetrahedron Lett.* **1987**, *28*, 5691–5694.
- (47) Fourrey, J. L.; Varenne, J.; Fontaine, C.; Guittet, E.; Yang, Z. W. A New Method for the Synthesis of Branched Ribonucleotides. *Tetrahedron Lett.* **1987**, *28*, 1769–1771.
- (48) Zhou, X.; Remaud, G.; Chattopadhyaya, J. New Regiospecific Synthesis of the “Branched” Tri-, Penta- & Hepta-Ribonucleic Acids Which Are Formed as the “Lariat” in the Pre-mRNA Processing Reactions[Splicing]. *Tetrahedron* **1988**, *44*, 6471–6489.
- (49) Damha, M. J.; Ogilvie, K. K. Synthesis and Spectroscopic Analysis of Branched RNA Fragments: Messenger RNA Splicing Intermediates. *J. Org. Chem.* **1988**, *53*, 3710–3722.
- (50) Balgobin, N.; Földesi, A.; Remand, G.; Chattopadhyaya, J. A New Regiospecific Synthesis of “Branched” Tetraribonucleotide and Its Three Analogues to Delineate the Chemospecific Role of the “Branch-Point” Adenine Nucleotide in Splicing. *Tetrahedron* **1988**, *44*, 6929–6939.
- (51) Damha, M. J.; Zabarylo, S. Automated Solid-Phase Synthesis of Branched Oligonucleotides. *Tetrahedron Lett.* **1989**, *30*, 6295–6298.
- (52) Sekine, M.; Heikkila, J.; Hata, T. Chemical Synthesis of Branched Oligoribonucleotides. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 588–601.
- (53) Sund, C.; Agback, P.; Chattopadhyaya, J. Synthesis of Tetrameric Cyclic Branched-RNA (Lariat) Modelling the Introns of Group II and Nuclear Pre-mRNA

Processing Reaction (Splicing). *Tetrahedron* **1991**, *47*, 9659–9674.

(54) Damha, M. J.; Ganeshan, K.; Hudson, R. H. E.; Zabarylo, S. V. Solid-Phase Synthesis of Branched Oligoribonucleotides Related to Messenger RNA Splicing Intermediates. *Nucleic Acids Res.* **1992**, *20*, 6565–6573.

(55) Sproat, B. S.; Beijer, B.; Grøtli, M.; Ryder, U.; Morand, K. L.; Lamond, A. I. Novel Solid-Phase Synthesis of Branched Oligoribonucleotides, Including a Substrate for the RNA Debranching Enzyme. *J. Chem. Soc. Perkin Trans. 1* **1994**, 419.

(56) Grøtli, M.; Sproat, B. S. A Universal Solid-Phase Synthesis of Branched Oligoribonucleotides. *J. Chem. Soc. Commun.* **1995**, 495–497.

(57) Ganeshan, K.; Tadey, T.; Nam, K.; Braich, R.; Purdy, W. C.; Boeke, J. D.; Damha, M. J. Novel Approaches to the Synthesis and Analysis of Branched RNA. *Nucleosides Nucleotides* **1995**, *14*, 1009–1013.

(58) Braich, R. S.; Damha, M. J. Regiospecific Solid-Phase Synthesis of Branched Oligonucleotides. Effect of Vicinal 2',5'- (or 2',3'-) and 3',5'-Phosphodiester Linkages on the Formation of Hairpin DNA. *Bioconjug. Chem.* **1997**, *8*, 370–377.

(59) Carriero, S.; Damha, M. J. Solid-Phase Synthesis of Branched Oligonucleotides. *Curr. Protoc. Nucleic Acid Chem.* **2002**, 4.14.1-4.14.32.

(60) Carriero, S.; Damha, M. J. Template-Mediated Synthesis of Lariat RNA and DNA. *J. Org. Chem.* **2003**, *68*, 8328–8338.

(61) Wang, Y.; Silverman, S. K. Characterization of Deoxyribozymes That Synthesize Branched RNA. *Biochemistry* **2003**, *42*, 15252–15263.

(62) Pratico, E. D.; Wang, Y.; Silverman, S. K. A Deoxyribozyme That Synthesizes 2',5'-Branched RNA with Any Branch-Site Nucleotide. *Nucleic Acids Res.* **2005**, *33*,

3503–3512.

(63) Mourani, R.; Damha, M. J. Synthesis, Characterization, and Biological Properties of Small Branched RNA Fragments Containing Chiral (R_p and S_p) 2',5'-Phosphorothioate Linkages. *Nucleosides, Nucleotides and Nucleic Acids* **2006**, *25*, 203–229.

(64) Lackey, J. G.; Sabatino, D.; Damha, M. J. Solid-Phase Synthesis and On-Column Deprotection of RNA from 2'- (and 3'-) O-Levulinated (Lv) Ribonucleoside Monomers. *Org. Lett.* **2007**, *9*, 789–792.

(65) Katolik, A.; Johnsson, R.; Montemayor, E.; Lackey, J. G.; Hart, P. J.; Damha, M. J. Regiospecific Solid-Phase Synthesis of Branched Oligoribonucleotides That Mimic Intronic Lariat RNA Intermediates. *J. Org. Chem.* **2014**, *79*, 963–975.

(66) Katolik, A.; Clark, N. E.; Tago, N.; Montemayor, E. J.; Hart, P. J.; Damha, M. J. Fluorescent Branched RNAs for High-Throughput Analysis of Dbr1 Enzyme Kinetics and Inhibition. *ACS Chem. Biol.* **2017**, *12*, 622–627.

(67) Wang, Y.; Silverman, S. K. Characterization of Deoxyribozymes That Synthesize Branched RNA. *J. Am. Chem. Soc.* **2003**, *125*, 15252–15263.

(68) Wang, Y.; Silverman, S. K. Efficient One-Step Synthesis of Biologically Related Lariat RNAs by a Deoxyribozyme. *Angew. Chemie - Int. Ed.* **2005**, *44*, 5863–5866.

(69) Wang, Y.; Silverman, S. K. A General Two-Step Strategy to Synthesize Lariat RNAs. *RNA* **2006**, *12*, 313–321.

(70) Javadi-Zarnaghi, F.; Höbartner, C. Functional Hallmarks of a Catalytic DNA That Makes Lariat RNA. *Chem. - A Eur. J.* **2016**, *22*, 3720–3728.

(71) Tago, N.; Katolik, A.; Clark, N. E.; Montemayor, E. J.; Seio, K.; Sekine, M.; Hart,

P. J.; Damha, M. J. Design, Synthesis, and Properties of Phosphoramidate 2',5'-Linked Branched RNA: Toward the Rational Design of Inhibitors of the RNA Lariat Debranching Enzyme. *J. Org. Chem.* **2015**, *80*, 10108–10118.

(72) Paredes, E. Triazole Linkages and Backbone Branches in Nucleic Acids for Biological and Extra-Biological Applications, Carnegie Mellon University, 2012.

(73) Dey, S. K. Biochemical and Single Molecule Studies of Backbone Branched RNAs and Larait Debranching Enzyme, Carnegie Mellon University, 2016.

(74) Shi, Y. Mechanistic Insights into Precursor Messenger RNA Splicing by the Spliceosome. *Mol. Cell Bio.* *18*, 655-70 (2017).

(75) Nam, J. *et al.* Human RNA lariat debranching enzyme cDNA complements the phenotypes of *Saccharomyces cerevisiae dbr1* and *Schizosaccharomyces pombe dbr1* mutants. *NAR.* **28**, 3666-73 (2000).

(76) Dhundale, A., *et al.* Structure of msDNA from *Myxococcus xanthus*: Evidence for a Long, Self-Annealing RNA Precursor for the Covalently Linked, Branched RNA. *Cell*, **51**, 1105-12 (1987).

Chapter 2 - Advancing backbone branched RNA synthesis through the improvement of the photodeprotection method

2.1 Introduction

As described in Chapter 1, multiple labs have worked to create backbone branched RNAs (bbRNAs) using various chemical and biochemical techniques. The Das lab previously developed a synthetic strategy to create backbone branched RNAs using a 2'-O-photolabile protecting group on the branching phosphoramidite^{1,2}. Initially, an *o*-nitrobenzyl (oNBn) protecting group was installed at the 2'-position of the branch point residue. This group has been widely used in the literature as a photolabile group for RNA photocaging, but has similar spectral overlap with, and can thus be damaging to, nucleobases³⁻⁵. Further optimization of the photolabile protecting group by previous lab members resulted in 6-nitroveratryloxymethyl (NVOM) as the photolabile group (Fig. 2.1). The absorption spectra is red-shifted to 365 nm, far enough away from the 260 nm absorbance of the nucleobases to not cause damage.

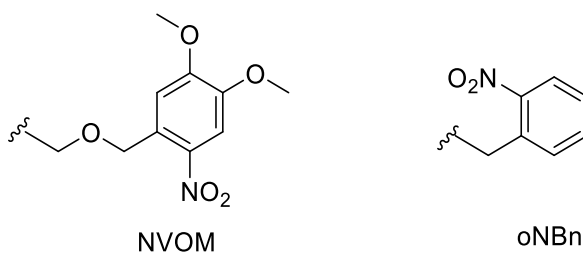
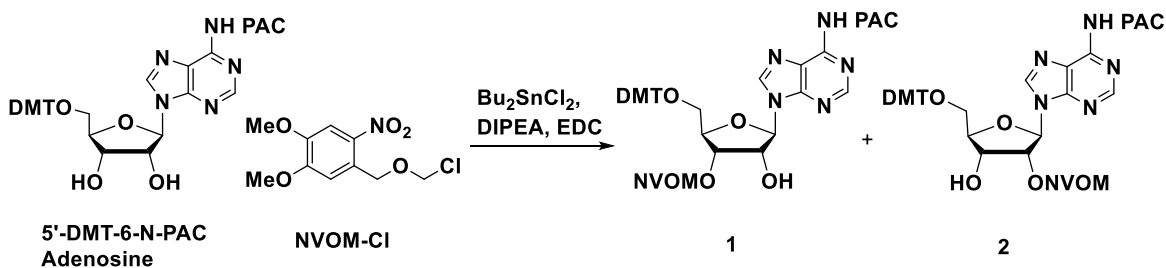


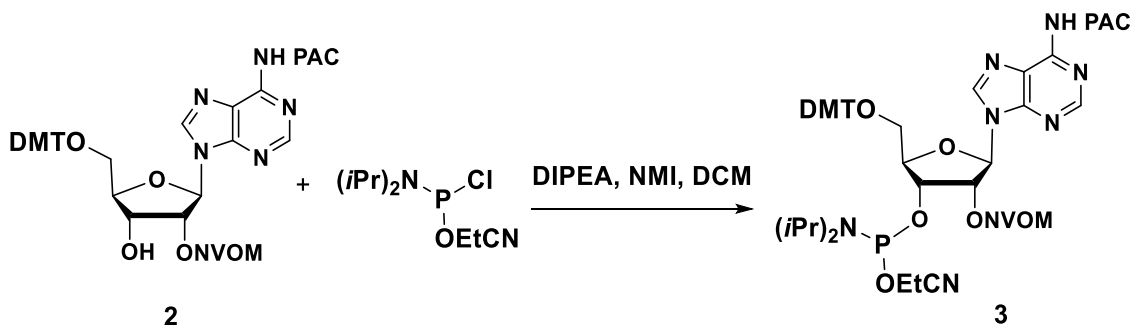
Figure 2.1. Structures of the 6-nitroveratryloxymethyl (NVOM) protecting group and ortho-nitrobenzyl (oNBn) protecting groups.

To make the protected nucleoside, the NVOM-Cl (prepared by Dr. Debashish Grahacharya) is attached to a commercially available phenoxyactyl (PAC)-protected nucleoside, such as adenosine, using a procedure outlined in scheme 2.1. The reaction yields both 2'-O- and 3'-O-protected adenosine (1 and 2 respectively) that are separated

by column chromatography. The purified nucleoside is subsequently reacted with a phosphoramidite chloride reagent to generate a phosphoramidite (scheme 2.2). After column chromatography, to separate the phosphoramidite from the starting material, the purified compound is then used in solid phase synthesis to construct backbone branched RNA.



Scheme 2.1. Synthesis of 2' or 3'-NVOM protected adenosine alcohol from commercially available adenosine diol.



Scheme 2.2. Synthesis of 2'-NVOM, 3'-amidite monomer. The reaction can be readily done with molecule 1 as well, resulting in a 2'-amidite, 3'-NVOM monomer.

It is possible to use either protected monomer when constructing backbone branched RNA (bbRNA), with slight modification to the synthetic procedure. The sequences on a bbRNA can be subdivided into three parts relative to the branch point residue, the 5'-arm, the 3'-arm, and the 2'-arm (fig. 1.4, 2.2). If using the 2'-O-protected branch point residue, the 3' and 5'-arms are synthesized first in the 3'-to-5'-direction using standard 5'-O-DMT, 3'-O-phosphoramidites. After removal of the 2'-O-NVOM group on the branch point residue, 'reverse' 3'-O-DMT, 5'-O-phosphoramidites are used to generate the

2'-arm. This procedure is termed the "2'-Method" as it uses the 2'-O-NVOM phosphoramidite and is shown in figure 2.2.

The 3'-O-NVOM-protected branch point phosphoramidite can also be used to make bbRNA by the alternate "3'-Method" (fig 2.2, right). First the 2'-arm is synthesized in the usual 3'-to-5'-direction and the 3'-O-NVOM protected branch point residue is included, and the 5'- arm synthesis continues. The 3'-O-photolabile group is removed, and the 3'- arm is made using reverse phosphoramidites. This "3'-Method" uses standard phosphoramidites to make the 2'-arm, and has the advantage of allowing for the incorporation of residues that contain modifications through readily available commercial phosphoramidites. Phosphoramidites that contain 2'-F or 2'-methoxy (OMe) modifications are not currently commercially available in the reverse form. The 3'-Method therefore allows these residues to be incorporated at any position along the 2'-arm of the backbone branched RNA. The ease of modification incorporation makes this method more versatile than others reported in the literature^{6,7}.

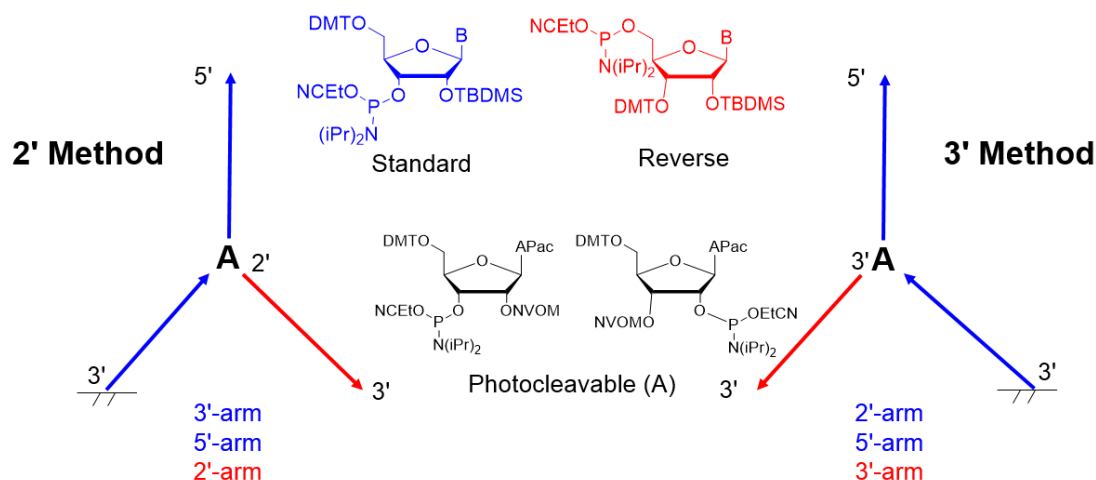


Figure 2.2. The 2'-Method and 3'-Method for bbRNA synthesis. The arm sequences in blue are synthesized first with standard phosphoramidites with the 2'-O-NVOM or 3'-O-NVOM photocleavable branch point residue. The red arm is synthesized last with 'reverse' phosphoramidites.

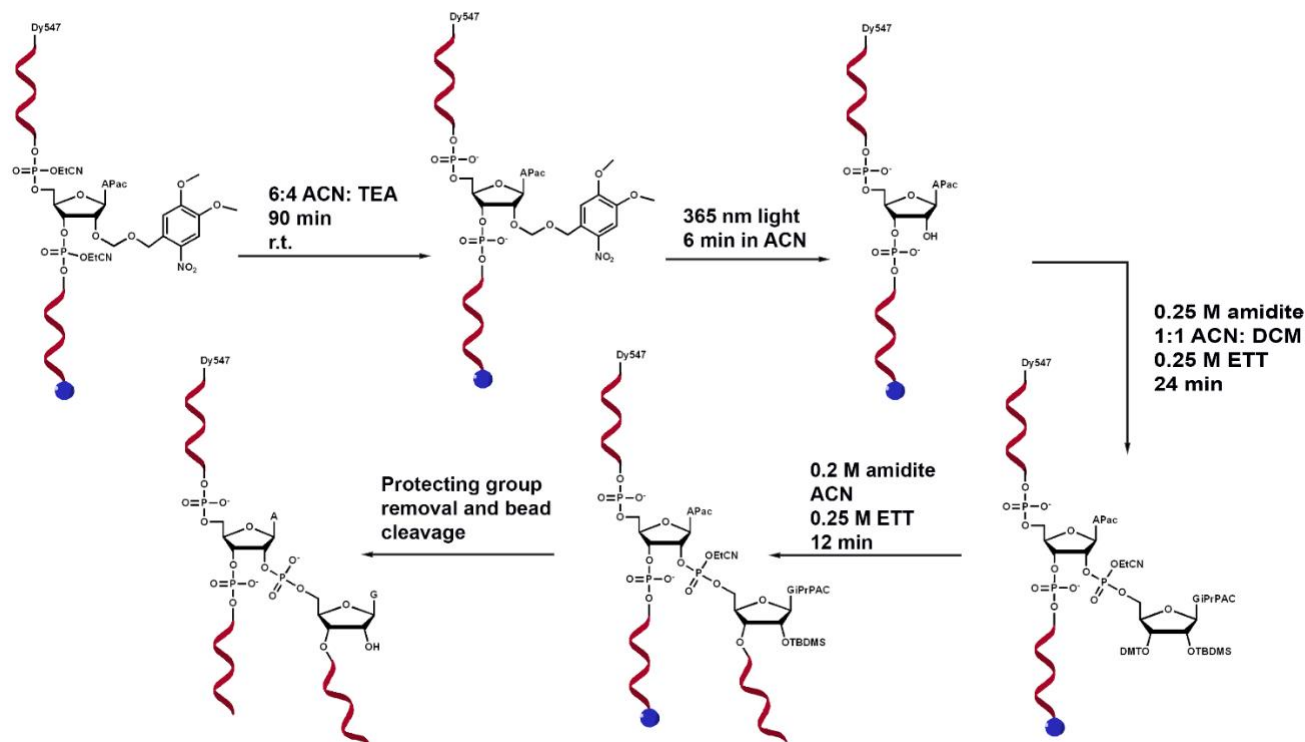


Figure 2.3. Detailed synthesis scheme of backbone branched RNAs.

Step	Reagent	Volume per injection (μL)	Reaction time per injection (s)	Number Injections	Total reaction time (s)
Deblock	3% Trichloroacetic acid	120	50	2	100
Coupling Forward	0.1 M amidite + 0.25 M 5-(Ethylthio)-tetrazole (ETT) activator	60 amidite + 60 activator	120 (for DNA amidites); 200 (for photoprotected amidite)	3	360(for DNA amidites); 600 (for photoprotected amidite)
Coupling Reverse	0.2 M amidite + 0.25 M ETT	60 amidite + 60 activator	360	2	720
Coupling First Residue in 2'-arm	0.25 M amidite + 0.25 M ETT	60 amidite + 60 activator	360	4	1440
Capping	Cap A: THF/Phenoxyacetic anhydride/ Pyridine. Cap B: 16% 1-methylimidazole in THF	60 Cap A + 60 Cap B	60	2	120
Oxidation	0.02 M I ₂ in THF/Pyridine/H ₂ O	120	50	2	100

Table 2.1. Table of synthesis conditions for RNA and DNA, forward and reverse synthesis.

The coupling details and times for both forward and reverse synthesis are outlined in table 2.1. After completion of forward synthesis, the terminus of the 5'-arm is capped using a phenoxyacetic group or through addition of any dye molecule. The capping step is necessary because it prevents unwanted extension of the 5'-arm during addition of reverse phosphoramidites. Next, the column is removed from the synthesizer and the cyanoethyl (CE) protecting groups on the phosphate backbone are selectively removed using a 6: 4 acetonitrile (ACN): triethylamine (TEA) solution⁸. The removal of the protecting group converts the phosphotriesters into less reactive phosphodiester to prevent 2'-to-3'-branch migration or strand cleavage once the NVOM group is removed. While the controlled pore glass (CPG) beads are outside the synthesizer, the photodeprotection step is done. Originally, the NVOM removal step was accomplished using a UV lamp. However, to achieve complete deprotection, the CPG beads were irradiated for 45 minutes. In this chapter, I investigated improving the photodeprotection method both inside and outside of the synthesizer with LED UV sources.

2.2 Improving the Photodeprotection Method

2.2.1. Practical considerations for photodeprotection

There are two possible ways to remove the NVOM moiety during bbRNA synthesis. The first is done outside the synthesizer, with the CPG beads moved into a glass vial and then illuminated. The second method keeps the column inside the synthesizer and then irradiates the beads. Both approaches have advantages and disadvantages and have been previously attempted by this lab while using a mercury UV lamp as an illumination source. The UV lamp is a 100 W long wave lamp (UVP-model B 100AP) and the light beam was focused with a glass lens. To improve upon the photodeprotection method, a UV LED was purchased (Prizmatix). The UV LED power source has three individually controllable fiber optic LED lines that illuminate at 365 nm and 100 mW power. The LEDs are more compact and powerful compared to the lamp, making the photodeprotection process more rapid.

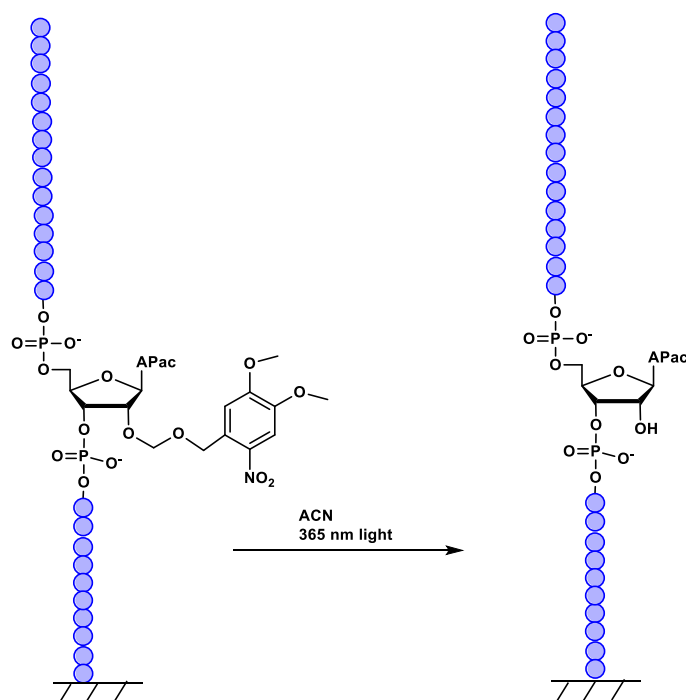
Determining which photodeprotection method to use, inside or outside synthesizer, depends on the desired end results. Keeping the CPG inside the synthesizer chamber and on the column maintains the moisture-free environment. To accomplish inside synthesizer deprotection, the LEDs either need to be shined through the chamber window or installed inside the airtight chamber. Even with a more powerful light source, the illumination through the chamber window is inefficient because the light needs to travel through glass and there is a significant distance between the chamber window and the column. Additionally, only one side of the synthesis column is illuminated, leaving the beads in the back of the column to be illuminated by scattering. We decided to illuminate directly inside the synthesizer. We placed the UV LEDs in the machine by drilling a hole in the side of the synthesizer big enough to insert three lines and adding a compression

fitting into the hole in order to properly seal the chamber and maintain a moisture free environment.

Photodeprotection performed outside of the synthesizer opens up the possibility of the moisture entrapment inside the beads, which would hinder further synthesis. However, any moisture can be removed by thoroughly drying the column after photodeprotection, re-purging the chamber with argon, and extensive washing of the column, once returned to the synthesizer, with synthesis grade acetonitrile. The advantage of removing the column, is that the beads are transferred into a glass vial, which has a broad surface area for uniform irradiation.

2.2.2. Photodeprotection of NVOM Modified DNA - Outside the Synthesizer

In order to determine photolysis efficiency of the UV LEDs, we synthesized a DNA of sequence 5'-t₁₅ A(2'-O-NVOM) t₁₀-3'. After synthesis, the cyanoethyl protecting groups are selectively removed. The CPG beads are washed thoroughly with acetonitrile to remove any lingering triethylamine from CE removal and are dried under house vacuum for 10 minutes. The beads are split into five aliquots (~50 nmol). Each aliquot is transferred to a 0.5 dram vial and 1 mL of ACN was added. The vial is clamped to a ring stand, and the UV LEDs are positioned underneath at a distance where the entire bottom of the vial is fully illuminated (~1 cm). Each aliquot is irradiated for a certain amount of time (0, 15, 45, 90, 180, and 600 seconds). After the allotted time, the beads are transferred into an empty column, washed with ACN, and dried under house vacuum. After drying, the CPG beads are transferred to a fresh 0.5 dram vial and the DNA was cleaved from the beads using 30% NH₄OH for 4 hours at room temperature. The beads are filtered away from the solution and the ammonia is removed by nitrogen gas flow. The DNA is lyophilized and resuspended in water before HPLC analysis.



Scheme 2.3. Outside synthesizer photodeprotection reaction scheme. After irradiation with 365 nm light, the NVOM group is removed, exposing a free hydroxyl.

Each reaction is run through a reverse phase HPLC. The exact conditions and column details are in the experimental section. The photolabile group is hydrophobic and increases the retention time of the oligonucleotide. By integrating the area under the observed peaks, the percent of protected and photodeprotected DNA is calculated, and thus a rate of photodeprotection is determined. Any DNA degradation from extended UV light exposure can also be observed on the chromatogram. As observed in fig 2.4, as time of irradiation increases, there is a decrease in the peak present at 0 second of irradiation. In conjunction, a new peak appears with a shorter retention time. The plot of fraction cleaved vs. time is fit to a biphasic exponential decay equation (Eq. 2.1), giving two rate constants, k_{fast} and k_{slow} . The fast fraction is rapid, and is believed to represent the portion of the population that is directly illuminated by the UV LED. The slow portion may be attributed to DNA that is being irradiated by scattered, less intense light in the interior of the column. The k_{fast} is 1.9 sec^{-1} and the k_{slow} is 0.011 sec^{-1} .

Equation 2.1. Biphasic exponential decay equation

$$y = A * e^{-(k_{fast})t} + C * e^{-(k_{slow})t}$$

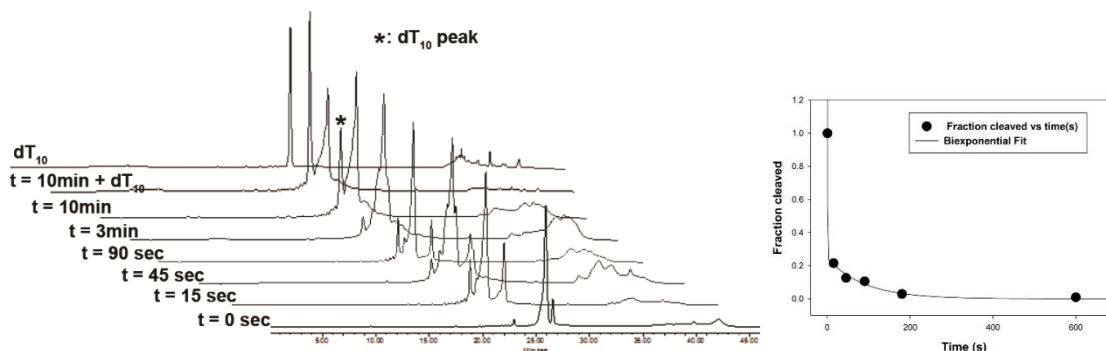


Figure 2.4. Results from photodeprotection of a 2'-O-NVOM protected DNA outside the synthesizer. Reverse phase HPLC chromatogram of deprotection reaction at different time points (left). There is a decrease in NVOM protected and an increase in NVOM removed DNA as irradiation time increases. At ten minutes of illumination, a t_{10} peak appears that corresponds to a cleavage product. The fraction of NVOM cleaved DNA is plotted over time (right) and fit to a biphasic exponential decay equation. The rate of reaction is determined from this fit.

These data establish a practical time for outside synthesizer photodeprotection to develop into a standard synthesis protocol. A six minute total irradiation time is used as the standard protocol because most of the NVOM on peak is removed by three minutes and the t_{10} cleavage has not begun. Halfway through, at three minutes, the vial containing the CPG beads is agitated. The shaking lets fresh CPG to come into direct contact with the UV light. These timings stop irradiation before degradation begins to occur (10 min) but ensures complete deprotection.

2.2.3. Photodeprotection of NVOM modified DNA - Inside the Synthesizer

The synthesizer itself requires modifications in order to place the UV LEDs inside. A hole is drilled into the side of the synthesizer and a compression fitting is installed. The

fitting is filled with a foam doughnut, which is sealed when not in use. When in use, the LEDs are fed through the opening and arranged around the synthesis column. An aluminum block is machined to fit around the standard synthesis block and hold the LEDs in place (fig 2.5). The aluminum block has three independently adjustable posts that each hold a LED line which are height and angle adjustable. The posts are arranged roughly 120° from each other to cover as much surface area of the CPG beads as possible. Since there are fewer CPG beads exposed on the surface than when the beads are transferred to a vial, arrangement of the LEDs is critical for complete deprotection of the NVOM group. The column itself needs be colorless, to minimize UV absorbance by the plastic.

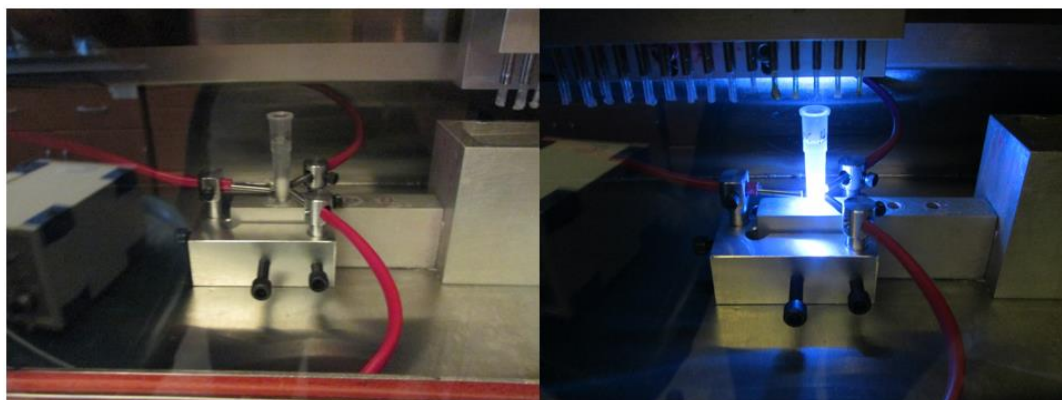
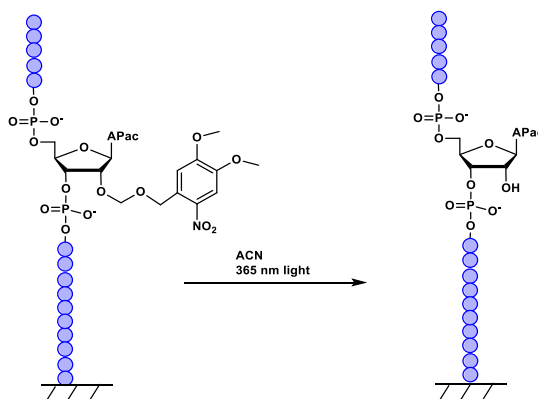


Figure 2.5. Set up of UV LEDs and aluminum block around the synthesis column. The LEDs are arranged ~120° from each other around the column and are as close as possible to the beads (left). ACN is injected into the column and the LEDs are turned on for the photodeprotection step (right).

Once the setup is complete, the same experiment described above is performed. A DNA of sequence 5' - t_5 A(2'-O-NVOM) t_{10} - 3' is synthesized at 1 μ mol scale and the CE groups are selectively removed using the ACN: TEA solution described above. The CE-deprotected beads are split into six fractions of ~150 nmol. Acetonitrile is injected into the synthesis column (200 μ L) and each fraction is irradiated for a given amount of time (0, 15, 45, 90, 180, 600 seconds). The DNA beads are moved into 0.5 dram vials and are incubated with 1:1 30% NH_4OH : methylamine for 12 min at 65 °C. The ammonia is

removed by nitrogen flow over the solution. The remaining solution is lyophilized, resuspended in water and analyzed by reverse phase HPLC.



Scheme 2.4. Inside synthesizer photodeprotection reaction scheme

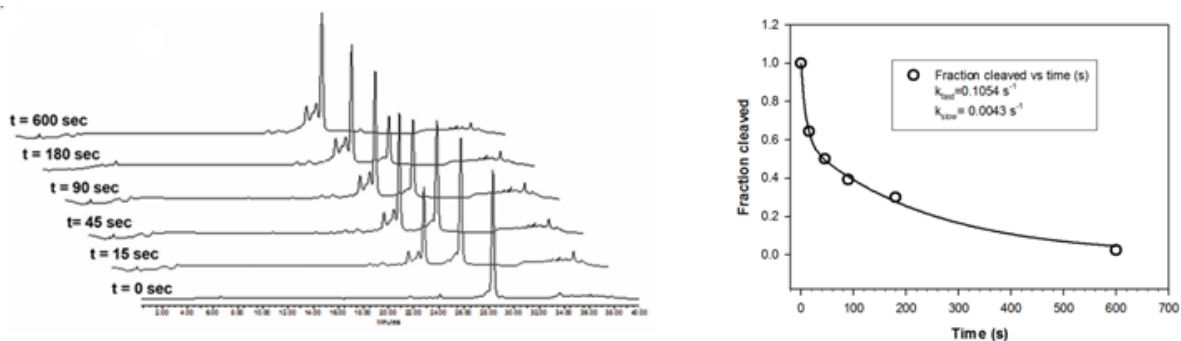


Figure 2.6. Reverse phase HPLC chromatogram of the inside synthesizer photodeprotection reaction after different UV irradiation times (left). Similar to the outside synthesizer reaction, after irradiation, a second peak appears that is 2'-NVOM deprotected. This peak appears over time as the protected peak disappears. The fraction cleaved is fit to a biphasic exponential decay equation to generate kinetic parameters (right).

The data was quantified as before and fit to a biphasic exponential decay (equation 2.1). The k_{fast} and k_{slow} , 0.1054 sec^{-1} and 0.0043 sec^{-1} respectively, are both slower than the outside of synthesizer method. The difference in deprotection rate between these two methods is expected as there is less surface illuminated inside the synthesizer. However, our data shows that there is complete deprotection at ten minutes and this is the time used henceforth as the standard protocol time for inside synthesizer photodeprotection.

Once the LEDs are removed, a hole remains in the side of the synthesizer. For typical operation not requiring photodeprotection, the hole is covered with a few layers of Teflon tape that are held in place by the compression nut. We have not observed any decrease in synthesis efficiency, indicating this is suitable to seal the synthesizer compartment.

2.3. Conclusion

In this chapter, I discussed the setup and kinetics for two different photodeprotection methods for oligonucleotide bound to solid support. The outside synthesizer method is rapid, complete within six minutes. However, this technique requires additional manipulation which may result in loss of product and exposes the CPG beads to the atmosphere, potentially hampering synthesis by trapping moisture in the beads. The inside synthesizer photodeprotection method solves the moisture problem, but requires modifying the synthesizer itself and careful setup.

The kinetics of deprotection for both methods are modeled with a biphasic exponential decay equation. The fit indicates that there is one fraction of oligonucleotide that has the protecting group rapidly removed and one that is cleaved much more slowly. The fast fraction is expected to be the oligonucleotide directly illuminated by the UV light. The slow fraction is illuminated by scattered light. When using the outside synthesizer method, the vial used during deprotection can be shaken, exposing a new fraction of beads to direct illumination. The inside synthesizer method does not allow for shaking, but the amount of CPG directly exposed to light can be increased by strategic placement of LED lines and loading the column with a small amount of CPG.

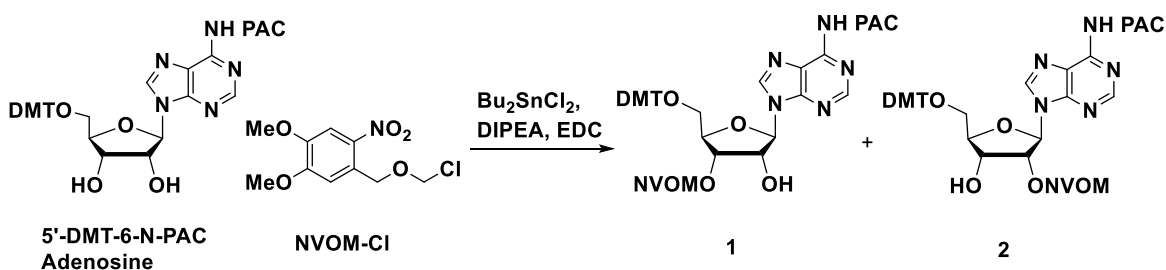
Both methods are efficient at removing the 2'-NVOM group from surface bound oligonucleotide and are an improvement over the previous method used in this lab. Both of these methods are implemented to make backbone branched RNAs as well as mini-lariat RNA, which is discussed in Chapter 5.

2.4. Experimental

The DNAs for photodeprotection were synthesized using standard DNA phosphoramidites and CPG solid supports purchased from Chemgenes. The starting material for the synthesis of 2'-NVOM phosphoramidite (5'-DMT-adenosine (N-PAC) nucleoside 2-cyanoethyl-N,N-diisopropyl-chloro-phosphoramidite) was also purchased from Chemgenes. Reagents for solid phase DNA synthesis, deblock, activator, ultramild CapA, CapB and oxidation reagents were purchased from Glen Research.

The UV LED source and leads were purchased from Prizmatix. The synthesizer is a MerMade 4 from Bioautomation. The HPLC is from Waters. All the dry solvents and other reagents were purchased from Sigma Aldrich or Fisher Scientific.

2.4.1. Synthesis of 5'-O-(Dimethoxytrityl)-2'-O-(6-Nitroveratryloxymethyl)-N6-phenoxyacetyl-adenosine

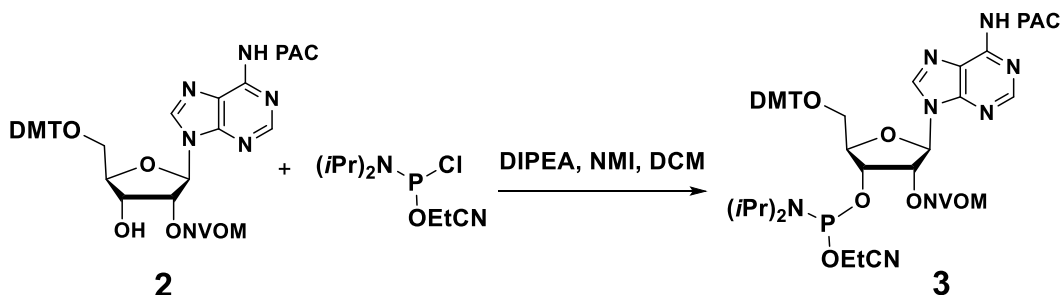


5'-DMT-6-N-PAC Adenosine (2 g) and NVOM-Cl (1.04 g) were dried under high vacuum overnight. The 5'-DMT-6-N-PAC was dried in a two neck round bottom flask, with one neck closed by a glass stopper and the other by a rubber septa. The NVOM-Cl was dried in a separate flask. 1,2-dichloroethane (EDC, 15 mL) was added to the round bottom flask followed by diisopropylethylamine (DIPEA, 2.5 mL). Bu_2SnCl_2 (1.04 g) powder was added to the reaction. The reaction was stirred at room temperature for 90 minutes. After 60 minutes, the NVOM-Cl was dissolved in 1.5 mL of EDC to create a slurry. Once the 90 minutes elapsed, the reaction mixture was placed in a 75 °C oil bath for 5 minutes. The

NVOM-Cl slurry was added to the reaction. After addition, the flask that contained the NVOM-Cl was washed with 2 mL of EDC and added to the reaction. The reaction was heated for 20 minutes at 80 °C.

Once the reaction is complete, the products **1** and **2** were separated on silica gel by column chromatography using 50 mL of each gradient solvent mixture. The solvent system was as follows: 50/ 50 ethylacetate (EA)/ hexanes, 60/40, 70/ 30, 80/ 20, 90/ 10, 100/ 0, 1% methanol in DCM, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%. All solutions have 2% triethylamine. Fractions that were determined to be the same by TLC are combined and dried down.

2.4.2. Synthesis of 5'-O-(Dimethoxytrityl)-2'-O-(6-Nitroveratryloxymethyl)-N6-phenoxyacetyl-adenosine-3'-N,N-diisopropyl(cyanoethyl) phosphoramidite



Compound **2** (240 mg) was dried in a flask overnight. Dichloromethane (3 mL) was added to compound **2** and the solution was stirred on ice for 30 minutes. DIPEA (240 μ L), 1-methylimidazole (NMI, 10 μ L), and N,N-diisopropyl(cyanoethyl) phosphoramidic chloride (88 μ L) was added to the reaction. The reaction was stirred on ice for 30 minutes and then slowly warmed to room temperature over 90 more minutes.

Once the reaction was complete, compound **3** was separated from compound **2** on silica gel by column chromatography using 100 mL of each gradient solvent with 2% triethylamine. The solvent system was as follows: 50/ 50 ethylacetate/ hexanes, 60/ 40,

70/ 30, 80/ 20. Collected fractions that were determined to be the same by TLC were combined and dried down.

2.4.3. Solid phase synthesis of DNA with 2'-NVOM residue

Solid phase synthesis was performed using a Mermade-4 automated synthesizer.

The conditions and reagents are outlined in the table 2.2.

Step	Reagent	Volume per injection (μL)	Reaction time per injection (s)	Number Injections	Total reaction time (s)
Deblock	3% Trichloroacetic acid	120	50	2	100
Coupling	0.1 M amidite + 0.25 M 5-(Ethylthio)-tetrazole activator	60 amidite + 60 activator	120 (for DNA amidites); 200 (for photoprotected amidite)	3	360 (for DNA amidites); 600 (for photoprotected amidite)
Capping	Cap A: THF/ Phenoxyacetic anhydride/ Pyridine. Cap B: 16% 1-methylimidazole in THF	60 Cap A + 60 Cap B	60	2	120
Oxidation	0.02 M I ₂ in THF/Pyridine/H ₂ O	120	50	2	100

Table 2.2. Solid phase synthesis conditions for DNA oligonucleotides (1 μmol scale).

Following synthesis, the cyanoethyl groups from the phosphate backbone were selectively removed using 10 mL of a 2:3 triethylamine: acetonitrile (ACN) mixture for 90 min. Every 15 minutes, ~2 mL of the solution was passed through the CPG beads using a syringe. The beads were then thoroughly washed with 10 mL of ACN, 10 mL of THF, and 10 mL ACN. After cyanoethyl removal, the 1x 1 μmol column was divided into smaller fractions as stated in the earlier section and illuminated.

2.4.4. Photodeprotection Method – Outside Synthesizer

1x 200 nmol of DNA of sequence 5'-t₁₅ A(2'NVOM) t₁₀-3' was split into six fractions of ~33 nmol each and moved into 0.5 dram glass vials. 1 mL of dry ACN was added to the vial. These were each irradiated with 365 nm light using the Prizmatix light source for an amount of time, 0, 15, 45, 90, 180, or 600 sec.

The beads were then moved back into a Bioautomation pipette style synthesis column to separate the beads from the ACN. The beads were dried over house vacuum and base deprotected with 1 mL of 30% NH_4OH per aliquot at ambient room temperature for 4 hrs while shaking. The solution was separated from the beads using a Pierce column. The ammonia was removed by nitrogen gas flow over the solution for 2 hrs and the remaining solution was lyophilized. Upon resuspension in water, each fraction was run through a reverse phase HPLC to determine fraction photo-protected and photo-deprotected.

The DNA was run through a Waters HPLC fitted with a XBridge BEH OST C_{18} column (4.6 x 150 mm) 5 μm . The temperature was 25 $^\circ\text{C}$ with a flow rate of 1 mL/ min. The solvents used were A: 0.1 M TEAA (pH 7); B: 80/ 20 ACN/ H_2O in 0.1 M TEAA buffer. The gradient ran from 2.5-25% solvent B in 30 min. Each peak was collected and lyophilized.

2.4.5. Photodeprotection Method - Inside Synthesizer

1 μmol of DNA of sequence 5'- $\text{t}_5\text{-A(2'-NVOM)-T}_{10}\text{-3'}$ was synthesized using the procedure detailed in table 2.2 and was cyanoethyl deprotected. The DNA was split into six fractions of ~160 nmol each. The three UV LED leads were run into the synthesizer and arranged around each column at 120 $^\circ$ angles from each other. The columns were washed 3x with 200 μL of ACN. Then, 200 μL of ACN was injected into the column and the UV power source was turned on for an amount of time, 15, 45, 90, 180, or 600 seconds. The light was turned off, the ACN drained from the column, and the beads were washed 3 more times with ACN before removal from the synthesizer. After this, the DNA was removed from the beads using 1 mL of a 1:1 30% NH_4OH : methylamine solution, heated at 65 $^\circ\text{C}$ for 12 minutes. The beads were separated from the solution using Pierce columns and the ammonia was removed using nitrogen gas flow for 2 hrs. The solution was subsequently lyophilized.

The DNA was run through a Waters HPLC fitted with a XBridge OST BEH C₁₈ column (4.6 x 50 mm) 2.5 μ m. The temperature was 25 °C and a flow rate of 1 mL/ min. The solvents used were A: 0.1 M TEAA (pH 7); B: 80/ 20 ACN/ H₂O in 0.1 M TEAA buffer. The gradient ran from 2.5-25% solvent B in 30 min. Each peak was collected and lyophilized.

2.5. References

1. Paredes, E. Triazole Linkages and Backbone Branches in Nucleic Acids for Biological and Extra-Biological Applications. (Carnegie Mellon University, 2012).
2. Dey, S. K. Biochemical and single molecule studies of backbone branched RNAs and lariat debranching enzyme. (Carnegie Mellon University, 2016).
3. Trzupek, J. D. & Sheppard, T. L. Photochemical Generation of Ribose Abasic Sites in RNA Oligonucleotides. *Org. Lett.* **7**, 1493–1496 (2005).
4. Chaulk, S. & MacMillan, A. M. Caged RNA: photo-control of a ribozyme reaction. *Nucleic Acids Res.* **26**, 3173–3178 (1998).
5. Das, S. R. & Piccirilli, J. A. General acid catalysis by the hepatitis delta virus ribozyme. *Nat. Chem. Biol.* **1**, 45–52 (2005).
6. Katolik, A. *et al.* Regiospecific solid-phase synthesis of branched oligoribonucleotides that mimic intronic lariat RNA intermediates. *J. Org. Chem.* **79**, 963–975 (2014).
7. Katolik, A. *et al.* Fluorescent Branched RNAs for High-Throughput Analysis of Dbr1 Enzyme Kinetics and Inhibition. *ACS Chem. Biol.* **12**, 622–627 (2017).
8. Braich, R. S. & Damha, M. J. Regiospecific solid-phase synthesis of branched oligonucleotides. Effect of vicinal 2',5'- (or 2',3'-) and 3',5'-phosphodiester linkages on the formation of hairpin DNA. *Bioconjug. Chem.* **8**, 370–377 (1997).

Chapter 3 –Biochemical analysis of backbone branched RNA cleavage by lariat debranching enzymes

3.1 Introduction to Dbr1p

As discussed earlier, lariat debranching enzyme plays a critical role in the cell, cleaving the 2'-5' phosphodiester bond in intron lariats. It cleaves this linkage with great specificity and has no effect on the vicinal 3'-5' bond in the branch point residue¹. However, the active site has remarkable sequence homology with Mre11, a nonspecific nuclease². Mre11 is a double-strand break repair enzyme and is a manganese-dependent metallophosphoesterase which has *in vitro* activity as a single-stranded DNA endonuclease, a double-stranded DNA 3' to 5' exonuclease, and a DNA hairpin opener³. The conserved residues of the Dbr1p active site differ from those of Mre11 in one residue - a cysteine in Dbr1p is an aspartate in Mre11 (fig. 3.1)². Given the similarity between the active sites of the two enzymes, the specificity of Dbr1p remains poorly understood. A detailed analysis of the substrate requirements may help in understanding the specificity of Dbr1p.

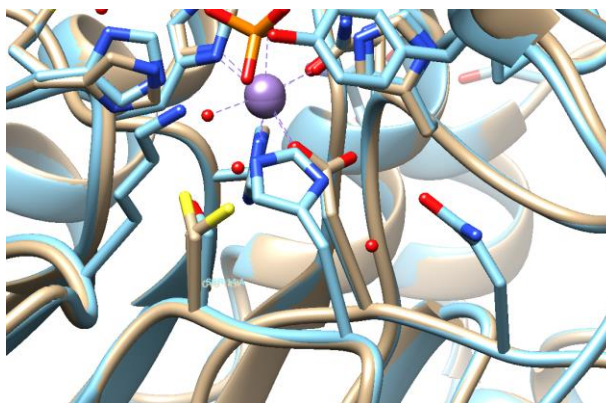


Figure 3.1. Overlay of Dbr1p active site (PDB: 4PEI) and Mre11 (PDB:3T1I). Mre11 is light blue and Dbr1p is tan. The metal ion (Mn^{2+}) is purple.

Dbr1p can work on a variety of substrates that contain a 2'-5' phosphodiester bond including lariat introns, multicopy single-stranded DNA (msDNA), and synthetically-derived branched RNA². Multicopy single-stranded DNA (msDNA) is a naturally occurring

nucleotide that has an RNA molecule linked to a DNA molecule through a 2'-5' phosphodiester bond⁴. Biochemical studies have been carried out on Dbr1p using synthetic backbone branched RNA (bbRNA), short RNA oligomers that contain a 2'-5' phosphodiester bond but are not looped together like introns. The short synthetic bbRNAs are readily cleaved by Dbr1p^{2,5}.

Analysis of the specific nucleotide requirements around the branch point of the substrate for proper activity of the enzyme has been done previously. A dinucleotide substrate linked only by a 2'-5' phosphodiester bond is not cleaved by the enzyme⁶. A nucleotide 5' of the branch point is not a requirement for specific debranching function⁷. A recent crystal structure of *Entamoeba histolytica* Dbr1p (EhDbr1p) bound to branched RNA supports the earlier biochemical observation that residues to the 5' of the branch point are not required for function, because when crystallized, that portion of the RNA is not observed to make any contacts with the enzyme and is entirely disordered in the solved structure⁸. Furthermore, it has been shown that the 3'-position of the branch point ribose needs to be connected to the 5'-phosphate of another nucleotide, even though the 3'-phosphate is not believed to participate in catalysis^{1,4,9}. The residue attached to the 3'-OH of the branching nucleotide is within hydrogen bonding distance of multiple residues on the enzyme, base stacks with the 2' branch, and has a variety of van der Waals interactions⁸.

Altering the identity of the branch point nucleobase has been investigated and it was found that cytidine, uridine, and guanosine branch point bbRNAs are all cleavable by debranching enzyme, but at a slower rate than adenosine^{6,10,11}. As seen in the crystal structure, the entire length of the branched RNA that protrudes from the 2'-5'-phosphodiester bond makes many contacts with the enzyme and makes the substrate cleavable by Dbr1p, while a 2'-5' linked phosphate will not be debranched^{6,8}. However, bbRNA with a deoxyribose in the first position of the branch can be debranched without a

significant effect on cleavage activity⁴. At the bridging phosphate, a pro-Rp sulfur abates activity, while a pro-Sp sulfur had no change in activity¹². In 2016, it was reported that lariat debranching enzyme can readily cleave bis-*p*-nitrophenylphosphate¹³. Based on the previous analysis of Dbr1p activity and substrate requirements, Dbr1p action on bis-*p*-nitrophenylphosphate is difficult to rationalize and outlines a need for a more thorough understanding of the substrate requirements, or lack thereof, for lariat debranching enzyme.

Another anomaly in the Dbr1p literature is the identity of the metal ion cofactor. Early studies of lariat debranching enzyme showed that the purified protein requires divalent metal ions to function properly^{1,6}. These studies, performed with human (hDbr1p) and *Saccharomyces cerevisiae* (ScDbr1p) debranching enzyme indicated that the divalent metal ion of choice was magnesium, the most abundant divalent metal ion in cells^{1,4,6,14}. There is a total of roughly 17-20 mM Mg²⁺ in mammalian cells, with 0.8 to 1.2 mM of that free¹⁴. Later research on ScDbr1p showed that Mn²⁺ is the preferred metal ion co-factor². Recent crystal structures of EhDbr1p tell a slightly different story. In the first crystal structure, published in 2014, the authors identify a Mn²⁺ ion in the β pocket with no ions observed in the α pocket¹⁵. The same group later crystallized EhDbr1p bound to bbRNA in 2016, and saw different metals in these sites. In the 2016 report, the authors detect a Zn²⁺ ion in the α pocket, and a Fe²⁺ ion in the β pocket⁸. In 2017 Ransey et. al. reported a EhDbr1p crystal structure with a Mn²⁺ ion in the α pocket and Zn²⁺ in the β pocket¹⁶.

Given the critical biological role of Dbr1p, and the literature discrepancies, investigating both substrate and metal requirements for the activity of Dbr1p is important. To gain a more thorough understanding of these requirements, both the RNA and non-RNA preferences of ScDbr1p and EhDbr1p are investigated and detailed in this chapter.

3.2. Requirements for *in vitro* Dbr1p Activity

Name	Sequence
bbRNA1	5'-GUA CUA A(2'-5'-GUA UGA) CAA GUU-3'

Table 3.1. Table of bbRNA used in section 3.2.

3.2.1 Titration of Mn^{2+} and the effect of EDTA on ScDbr1p and EhDbr1p

In vitro assays are performed using backbone branched RNA (bbRNA) that are labeled with a fluorophore at the 5'-terminus - either Dylight547 (Dy547) or Cyanine3 (Cy3). bbRNA1 is then incubated with EhDbr1p or ScDbr1p and metal ions. After a given amount of time, the reaction is stopped by adding two equivalents of 9:1 formamide and 0.1 M EDTA, 'stop' solution. The solution is loaded onto a denaturing polyacrylamide gel. The branched and debranched RNAs have different electrophoretic mobility shifts and the relative amount of each RNA piece is quantified based on the fluorescence after scanning the gel (fig. 3.2). The signal correlating to the amount of bbRNA is divided by the total signal from the two quantitated bands, the bbRNA and the stem RNA. This gives the fraction of bbRNA remaining at a given time point. The fraction of bbRNA remaining is plotted against the corresponding time (fig. 3.2. panel b).

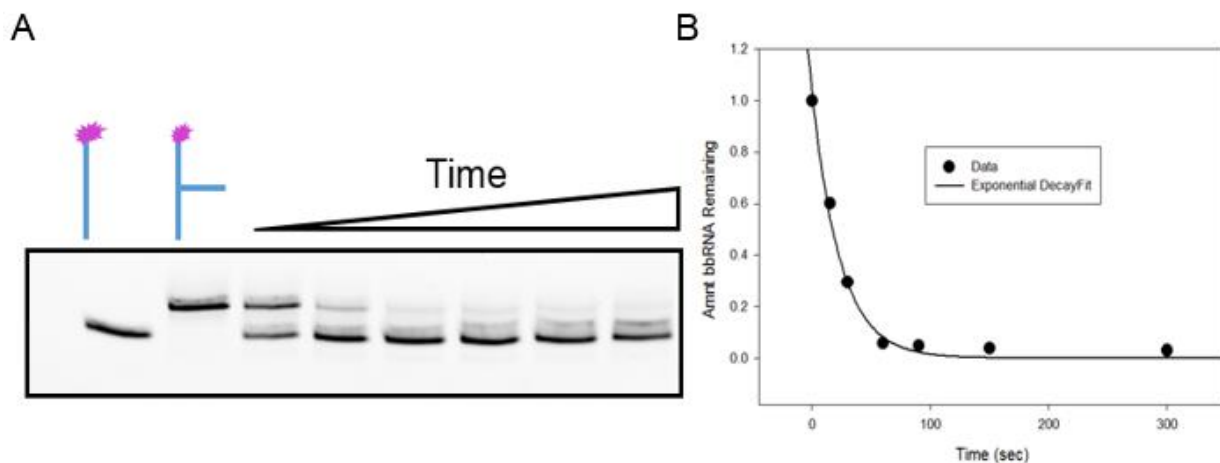


Figure 3.2. (A) Representative polyacrylamide gel of debranching reaction. The gel is imaged on a Typhoon FLA9000 scanned for Cy3 signal. The intensity of the gel band is quantitated using the corresponding software and amount of bbRNA is plotted over time (B). The data is fit to an exponential decay equation, where the time term gives the observed rate constant.

First, given the previous reports of Mn^{2+} as a critical cofactor for Dbr1p and the dearth of biochemical characterization of EhDbr1p, the optimal concentration of additional Mn^{2+} required for EhDbr1p activity *in vitro* is investigated alongside the more studied ScDbr1p^{2,8,13,15,16}. Reactions are performed with various concentrations of added Mn^{2+} , ranging from 0.5 mM to 10 mM. Further, the addition of metal chelator ethylenediaminetetraacetic acid (EDTA) without any Mn^{2+} added to the buffer is investigated. The metal ions or chelator are allowed to preincubate with the enzyme for three minutes before the addition of bbRNA1 to start the reaction. Over the course of five minutes, aliquots of the Dbr1p and bbRNA1 reaction are taken at 15, 30, 60, 90, 150, and 300 seconds and placed into stop solution. The relative amount of branched RNA to debranched RNA is calculated and plotted. The data is fit to an exponential decay equation $y = a * e^{-kt}$ and the k term generated from the fit is the observed rate constant (k_{obs}).

To determine the optimal Mn^{2+} concentration for *in vitro* reactions for both enzymes, the k_{obs} is plotted as a function of Mn^{2+} concentration (Fig. 3.3). Sc and Eh enzymes have similar overall profiles, with optimal activity peak between 1 and 4 mM, and then dropping as Mn^{2+} concentration increases. ScDbr1p has optimal Mn^{2+} at 4mM, consistent with previous reports². EhDbr1p debranching activity peaks around 0.5 mM Mn^{2+} with no significant difference between 0.5 and 1 mM. Therefore, for the remaining experiments, 1 mM Mn^{2+} was used for EhDbr1p assays.

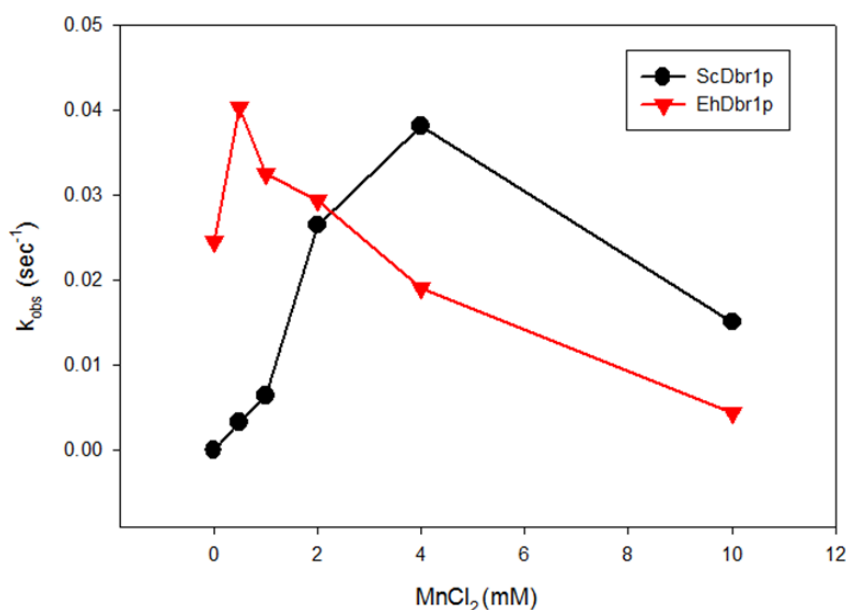


Figure 3.3. Mn^{2+} titration with ScDbr1p and EhDbr1p. Plotted points are observed rate constants generated from individual experiments.

3.2.2. The EDTA Effect

One striking difference between the two enzymes is the effect of EDTA. ScDbr1p activity is completely abolished, while EhDbr1p continues to cleave at EDTA concentrations 10^6 above enzyme. The effect of the metal chelator can be interpreted in a few ways. Since there is no additional metal added to the reaction, any metal present would be leftover in the protein from the expression process and has maintained

association throughout enzyme purification. These metal ions might be particularly hard to strip away. Another possibility to explain the EDTA effect is that the metal binding pockets of ScDbr1p are arranged differently than EhDbr1p. ScDbr1p might be more solvent accessible, and thus EDTA is able to extract the catalytic ions efficiently, rendering the enzyme inactive. Conversely, EhDbr1p might have a very tight affinity to its metal ions. A further option is that there might be metal ions around the surface of the ScDbr1p that are important for maintaining the surface of the protein²². These would be easily accessible to chelation by EDTA.

To investigate further the effect of EDTA on EhDbr1p, the enzyme is incubated for three minutes with increasing amounts of EDTA, under the hypothesis that given enough chelator, the active site metal ions might be removed. As seen in Fig 3.4, there is still minimal effect on the activity of EhDbr1p. The enzyme without EDTA is evaluated for debranching activity. The observed rate without added EDTA is used as a control to compare change in activity when EDTA is added to the reaction mixture.

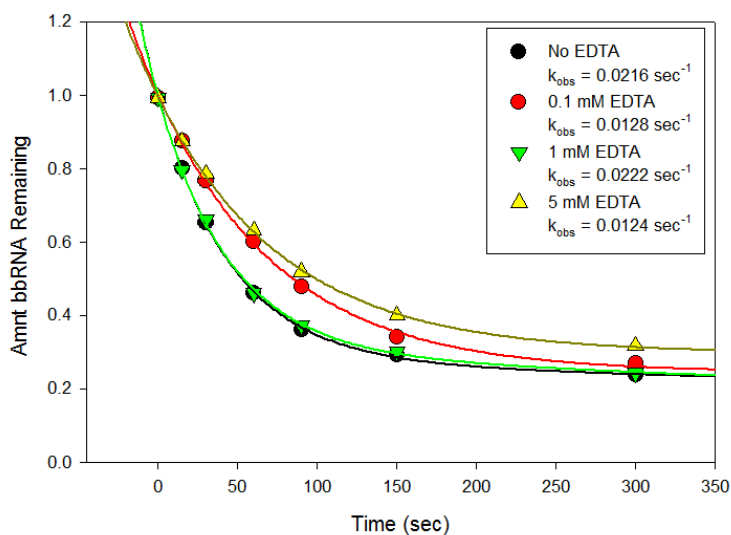


Figure 3.4. The effect of EDTA on EhDbr1p activity. The data is plotted as amount of bbRNA over time and fit to an exponential decay equation. The rate constants are detailed in the figure legend.

EhDbr1p is incubated with EDTA for an extended amount of time to see if more time is needed to remove metal ions and slow down EhDbr1p activity (Fig 3.5). EDTA is allowed to preincubate with EhDbr1p for the usual three minutes, or extended to 10 and 30 min. After preincubation, the reaction is started with addition of bbRNA1. At ten minutes of preincubation, the reaction rate was effectively the same as the three minute preincubation. However, after 30 min of preincubation, the rate decreases by 60%. The rate decrease suggests that long preincubation of EhDbr1p with EDTA will slow the rate. However, the rates are on the same order of magnitude, showing that this effect is still slight, and that EhDbr1p remains active with EDTA present in the reaction buffer.

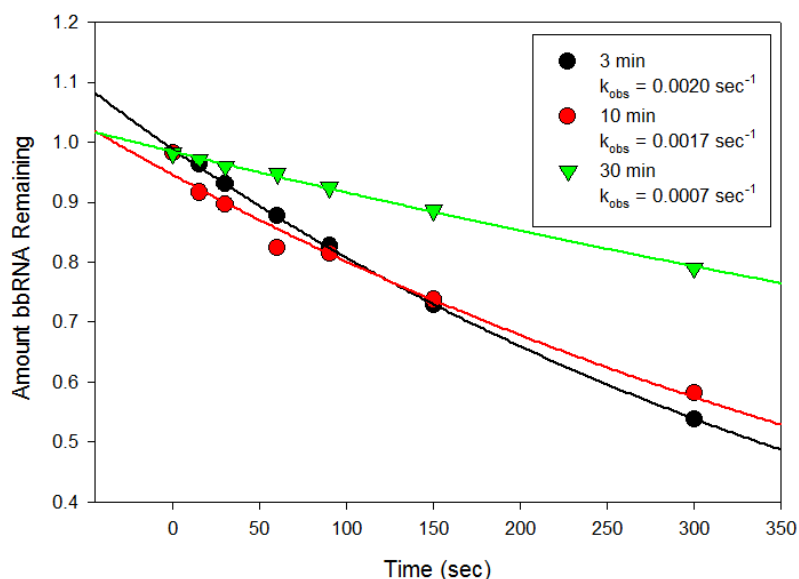


Figure 3.5. The effect of preincubation of EhDbr1p with EDTA. Data is plotted as amount of bbRNA vs. time. The data is fit to a single exponential decay equation and the observed rate constant is generated from the fit.

3.2.3 Other Divalent Ions with EhDbr1p

In 2005 the debranching ability of ScDbr1p with added Mg^{2+} , Mn^{2+} , Cu^{2+} , Zn^{2+} , Ca^{2+} , Ba^{2+} , or EDTA over a background of added Mn^{2+} was reported². The authors note that Cu^{2+} and Zn^{2+} halted activity, while addition of the remaining metal ions showed no

change in activity over background². In 2016, the second reported crystal structure of EhDbr1p was published and identified Fe²⁺ and Zn²⁺ in the metal-binding pockets⁸. The authors also investigated the activity of EhDbr1p with Mn²⁺, Zn²⁺, and Fe²⁺ *in vitro* to confirm what they were observing in the crystal structure⁸. However, they only investigated the metal ions predicted to be in the active sites and not the other transition metals explored in the 2005 report². To gain a better understanding of the catalytic requirements of EhDbr1p, the debranching activity with different metal ions, including both group II and transition metals, is investigated.

As an initial screen for activity, EhDbr1p is incubated with 1 mM of a given metal ion and aliquots are taken of the debranching reaction after 1 and 10 minutes. The results of this screen are consistent with previous reports (Fig. 3.6)². However, the activity cannot be entirely attributed to the given ion. Given that EhDbr1p is still active without added metal ion and even with added metal ion chelator, the metal ion screened may not significantly affect the metal ions in the active site. Therefore we can only say the screened metal either has no effect on or has an adverse effect on activity.

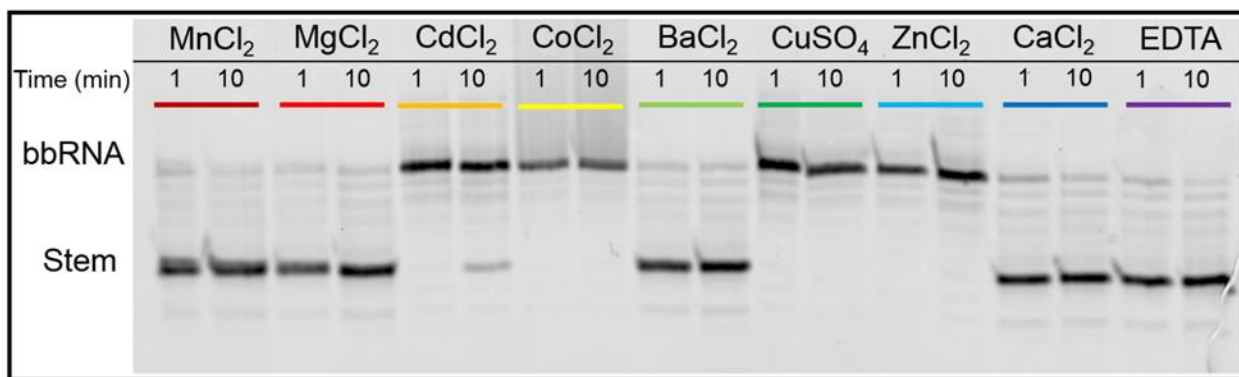


Figure 3.6. EhDbr1p activity with 1 mM of added metal ions. Debranching reaction after one and ten minutes are run on a 20% denaturing polyacrylamide gel. The gel is imaged on a Typhoon FLA9000 for Cy3 signal. Mn²⁺, Mg²⁺, Cd²⁺, Co²⁺, Ba²⁺, Cu²⁺, Zn²⁺, Ca²⁺, and EDTA are evaluated.

As seen in fig 3.6, Mn^{2+} , Mg^{2+} , Ba^{2+} , Ca^{2+} , and EDTA all permit activity at the given time points and have no negative effect on EhDbr1p activity. Cu^{2+} and Zn^{2+} both eliminated debranching activity. Cd^{2+} shows slight cleavage of the bbRNA after ten minutes, indicating that it hampers but does not eliminate activity of EhDbr1p. Finally, Co^{2+} precipitated out of the reaction after it interacts with the reducing agent in the buffer, so that experiment is inconclusive. These results support and expand previous investigations.

3.2.4 The Effect of a Reducing Agent on *in vitro* Dbr1p Activity

Reducing agents are frequently used in biochemical studies involving proteins because they stabilize free cysteines and reduce disulfide bonds. Including a reducing agent in a buffered solution helps prevent protein aggregation and facilitates proper function. Different reports on biochemical studies of the lariat debranching enzyme have used different reducing agents, including dithiothreitol (DTT) and tris (2-carboxyethyl) phosphine (TCEP) in the buffer. TCEP is known to be a more powerful reducing agent than DTT because it is more stable and irreversible¹⁷. In the initial reports of ScDbr1p, 2.5 mM DTT is used, and that amount of reducing agent has been used in the experiments described throughout this thesis^{2,4}. However, recent reports on the EhDbr1 use 0.5 mM TCEP in the buffer^{8,18}.

To compare our debranching reaction to the recent reports, the cleavage rate of EhDbr1p under different reducing conditions is investigated (Fig. 3.7). Equivalent amounts of reducing agent react at the same rate. Lowering the amount of TCEP reducing agent to 0.5 mM, as used in the published debranching assays with EhDbr1p, results in a slower reaction^{8,18}. The reduction in rate suggests that the assays published in the literature may not be optimized for maximal debranching rates.

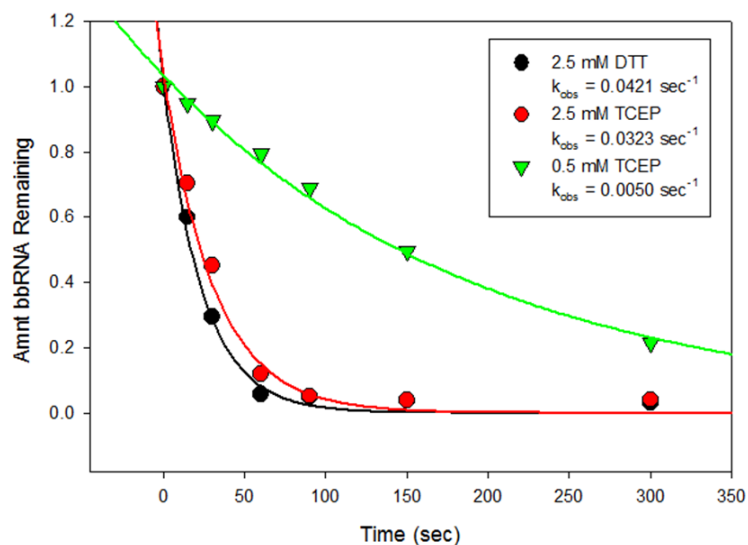


Figure 3.7. The effect of reducing agents, DTT and TCEP on the EhDbr1p cleavage rate. Data is plotted as amount of bbRNA vs. time. The data is fit to a single exponential decay equation and the observed rate constant is generated from the fit.

3.2.5. The Final Conditions

Based on the experiments outlined in the above sections, the optimal conditions for debranching buffer are 2.5 mM DTT as the reducing agent. The Mn^{2+} requirements for EhDbr1p and ScDbr1p are different, 1 mM and 4 mM Mn^{2+} respectively, and are used in the appropriate assays.

3.3. RNA requirements for the Dbr1p cleavage reaction

3.3.1. DNA/RNA Hybrids

With the advances we have made in the solid phase synthesis of bbRNAs, specific modifications and functional group substitutions can be readily incorporated to test their effects on Dbr1p cleavage activity. We explore substitutions of fluorine or methoxy at the 2'-position of the first residue in the branch. Further, DNA/ RNA hybrids, bbRNAs with a 2'-3'-linkage where the natural 2'-5'-linkage should be, as well as an entirely DNA backbone branched oligonucleotide are synthesized to evaluate in the debranching reaction. The bbRNA with a 2'-3' phosphodiester bond is obtained when the branch residues coupled during solid-phase synthesis are standard 5'-O-DMT, 3'-O-phosphoramidites (rather than reverse 3'-O-DMT, 5'-O-phosphoramidites). These sequences are detailed in table 3.2 and fig 3.8. All are found to be substrates for specific cleavage by both ScDbr1p and EhDbr1p. The cleavage rate of these bbRNAs with both ScDbr1p and EhDbr1p were measured and are reported in fig. 3.9.

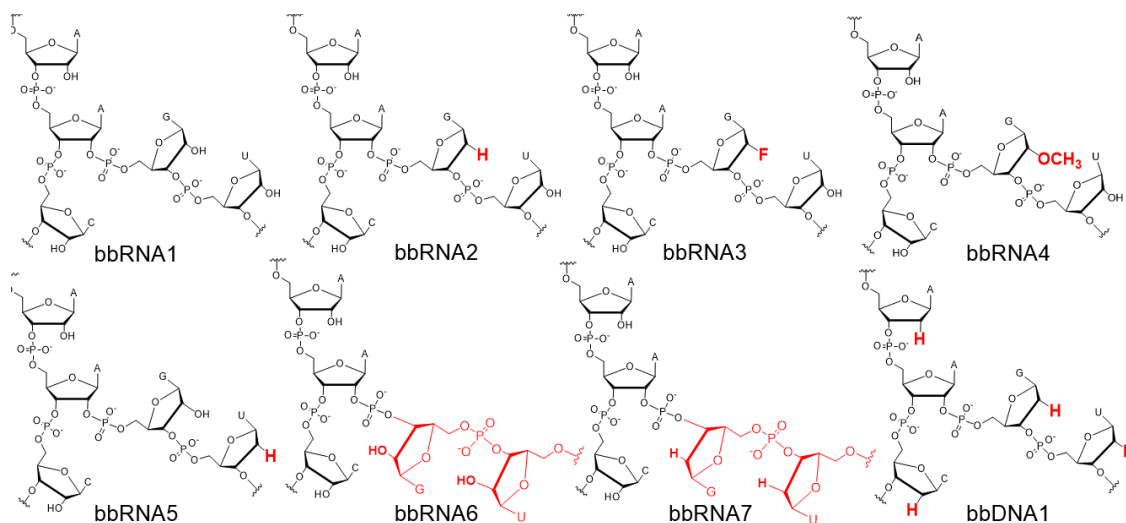


Figure 3.8. Structures of DNA/ RNA hybrids and 2' modified RNAs assayed for debranching activity. Red denotes a change from bbRNA1.

Name	Sequence
bbRNA1	5'-Cy3-GUA CUA A(2'-5'GUA UGA) CAA GUU-3'
bbRNA2	5'-Dy547-GUA CUA A(2'-5'gUA UGA) CAA GUU-3'
bbRNA3	5'-Dy547-GUA CUA A(2'-5'G(2'F)UA UGA) CAA GUU-3'
bbRNA4	5'-Dy547-GUA CUA A(2'-5'G(2'OMe)UA UGA) CAA GUU-3'
bbRNA5	5'-Dy547-GUA CUA A(2'-5'Gta tga) CAA GUU-3'
bbRNA6	5'-Dy547-GUA CUA A(2'-3'GUA UGA) CAA GUU-3'
bbRNA7	5'-Dy547-GUA CUA A(2'-3'gta tga) CAA GUU-3'
bbDNA1	5'-Dy547-gta cta a(2'-5'gta tga) caa gtt-3'

Table 3.2. Table of bbRNA used in section 3.3.1. Lower case letters are DNA and upper case letters are RNA.

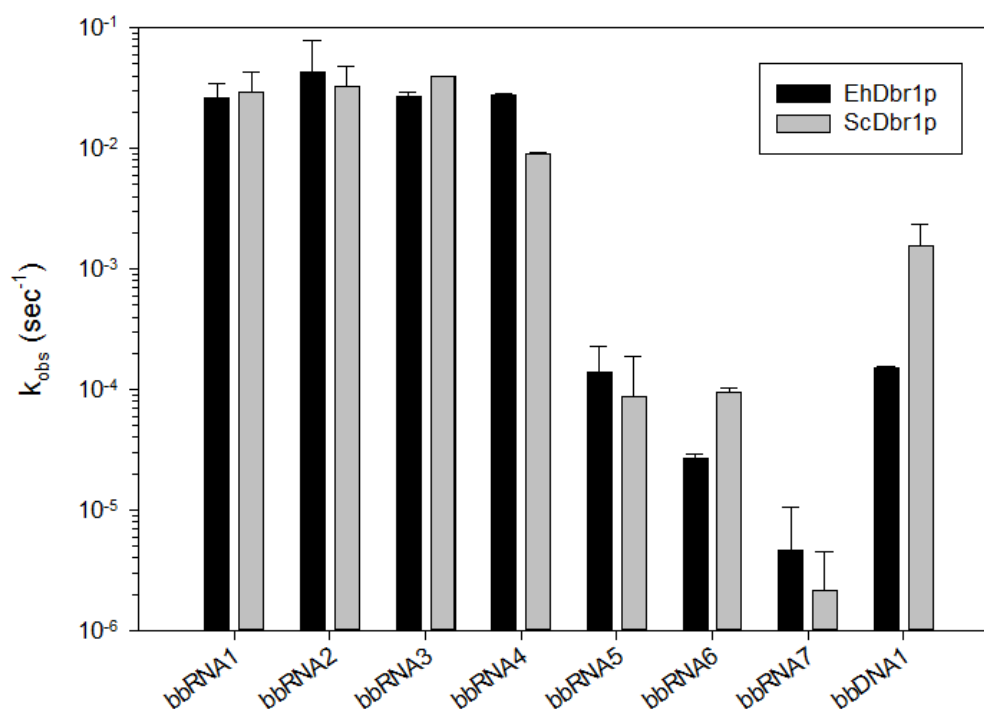


Figure 3.9. The observed cleavage rates of modified bbRNAs by Sc and Eh Dbr1p. The bbRNA substrates are detailed in table 3.2. The mean of two experiments is plotted and the error bars represent the standard deviation.

These results show that the replacing the 2'-OH, with a 2'-H (bbRNA2) or a 2'-F (bbRNA3) does not impact the debranching rate of either enzyme. Replacement with a

bulky 2'-OMe group (bbRNA4) does not impact the debranching enzyme from *E. histolytica*, but does slow down *S. cerevisiae* enzyme by three-fold. The observed rate change suggests differences in the active site between the two enzymes. The observation is supported by the recently published co-crystal structure of *E. histolytica* bound to backbone branched RNA, which shows no interactions with the 2'-OH of the first residue in the branch⁸. Such differences between the two enzymes' active sites are consistent with our observations of differences with EDTA addition, as described previously.

With the more structural changes to the backbone branched substrates, such as in DNA/RNA hybrids, a much greater reduction in cleavage rate is observed. A 100-fold reduction in cleavage activity is seen with bbRNA5. In bbRNA5, the branched oligonucleotide has 2'-deoxy modifications in the terminal five residues of the branch. In the co-crystal structure published by Damha et al., the 2'-hydroxyls of many residues of the branch, after the first residue, are within hydrogen-bonding distance of amino acid residues in the protein (Fig. 3.10)⁸. Based on the co-crystal structure, modifications to 2'-deoxy residues in the branch would likely lead to a change in binding affinity that affects the cleavage rate⁸. We observe the drop in observed activity for bbRNA5 but need further experiments to conclusively say the activity change is due to improper bonding than decreased catalysis.

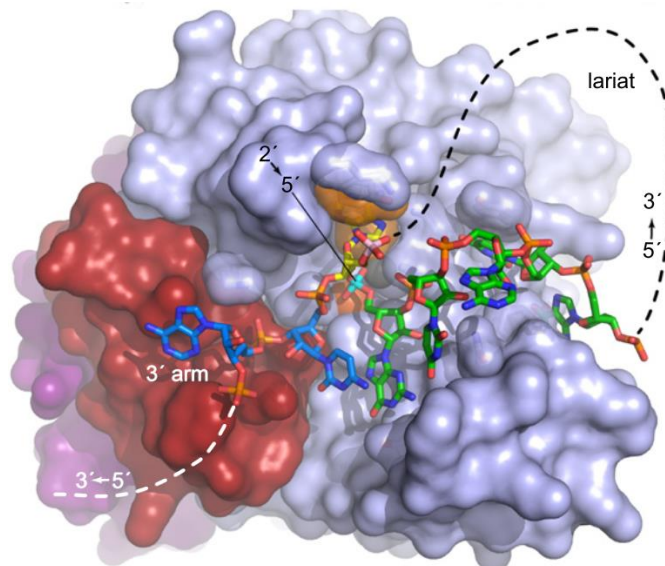


Figure 3.10. Structure of EhDbr1p associated with branched RNA from ref⁸. The branch point residue is in yellow, the 3'-arm is in blue, and the 2'-arm is in green. The 5'-arm was not observed.

The reduction in the observed cleavage rate constant with bbRNA5 is similar to the rate reduction seen when bbDNA1 is used. However, ScDbr1p cleaves bbDNA1 faster than EhDbr1p by 10-fold, further highlighting to differences between the enzymes. The specific differences between the enzymes, however, cannot be elucidated without access to a crystal structure of ScDbr1. Determining those differences through biochemical assays are outside the scope of this thesis.

Changing the directionality of the 2'-branch from the native 5'-3' to 3'-5' linkage causes a large drop in activity compared to the native substrate mimic. Considering the numerous contacts between the oligonucleotide and protein that will potentially be disrupted, and as the 2'-5'-phosphodiester bond that Dbr1p is specialized to cleave is replaced by a 2'-3'-linkage, an activity drop is to be expected. Even so, with all these significant structural differences, surprisingly, specific cleavage still occurs. When the entire backbone branched construct is RNA (bbRNA6), the drop in activity is

commensurate with the drop that occurs when changing the last five residues of the branch to DNA (bbRNA5). Again, there is a 10-fold difference between the two activities in reaction rate, with ScDbr1p reacting faster. Further altering the backbone branched oligonucleotide to contain both a 2'-3'-linkage and a DNA branch (bbRNA7) reduced the rate about 10-fold more, dropping off four orders of magnitude from bbRNA1. Additional studies will need to be done to determine if reduced binding affinity or decreasing catalytic rate constant primarily contribute to the decrease in observed rate. For such studies, assays that can distinguish between binding and catalytic reaction rate are required, including suitably labeled substrates.

As the unnatural substrates show a significant decrease in observed rate, a debranching reaction is done with an excess of enzyme over substrate concentrations. In previous studies, there is a large excess of substrate (100 nM) and a small amount of enzyme (1 nM). In those experiments the extent of enzymatic cleavage of the unnatural substrates could be determined. With greater enzyme concentrations (500 nM), and given sufficient time, all of the backbone branched oligonucleotides are cleaved (Fig. 3.11). EhDbr1p shows complete debranching and a second band that appears as a faster migrating band in the polyacrylamide gel. The identity of this band is not known, but might be a secondary cleavage of the stem, as it is faster running and thus smaller. ScDbr1p shows a similar pattern, with the exception of the 2'-3'-linked branched RNA and DNA. Both of these substrates do not go to completion in a ten minute incubation. The incomplete cleavage may be attributed to low binding affinity of the enzyme for this substrate analogue. The digestion of bbRNA6 and bbRNA7 with ScDbr1p also show pieces of oligonucleotide that migrate faster, during gel electrophoresis, than the stem. The cleavage product closest to the band identified as the stem is likely the same product that appears in the EhDbr1p digestion, based on its migration. The fastest migrating band

that appears only in ScDbr1p is still unknown, but might be from a second cleavage of the stem. However, its identity cannot be definitively established. Noteworthy, bbDNA1 is entirely debranched by both enzymes in the given time and shows no extraneous cleavage products; the only activity seen is on the 2'-5'-phosphodiester bond. It is unclear why this occurs. Does the RNA stem remain associated with the enzyme, causing the extra cleavages? Is linear RNA a substrate for Dbr1p while linear DNA is not? These are interesting questions to investigate in the future, with either computational modeling or carefully designed *in vitro* experiments.

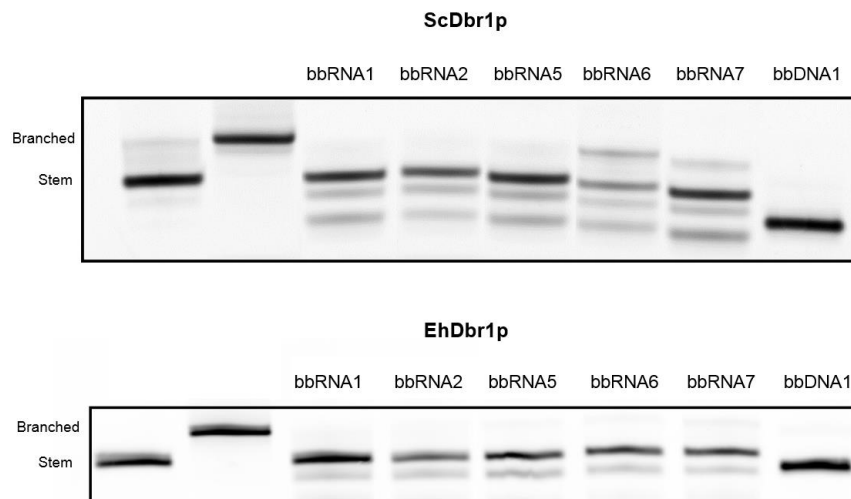


Figure 3.11. Denaturing polyacrylamide gel of cleavage of bbRNAs by ScDbr1 (top) and EhDbr1p (bottom) after ten minutes. These cleavage reactions included 500 nM enzyme, with 100 nM substrate in standard reaction buffer. The first two lanes are controls to mark the stem and bbRNA. The gel indicates that for all these substrates, cleavage proceeds to completion in 10 mins.

3.3.2. 2'-5' RNA

Since the discovery of lariat debranching enzyme and subsequent biochemical analyses, the minimal substrate requirement has been regarded as branched RNA with a 2'-5' linkage as well as a 3'-5' phosphodiester bond¹. We observe cleavage of 2'-3'-branched RNAs that suggests a much greater tolerance in the substrate. Therefore to further probe the Dbr1p substrate binding and cleavage requirement, we synthesized a

RNA that contained a single 2'-5'-linkage that was otherwise 'linear' (with typical 3'-5'-phosphodiester bonds), In comparison to bbRNA, is the top of the stem and the branch, and removes the bottom of the stem, leaving the branch point with a 3'-hydroxyl.

Name	Sequence
RNA1	5'-Dy547-GUA CUA A CAA GUU-3'
RNA2	5'-Dy547-GUA CUA A(2'-5'-GUA UGA)-3'

Table 3.3. List of sequences used in 3.3.2.

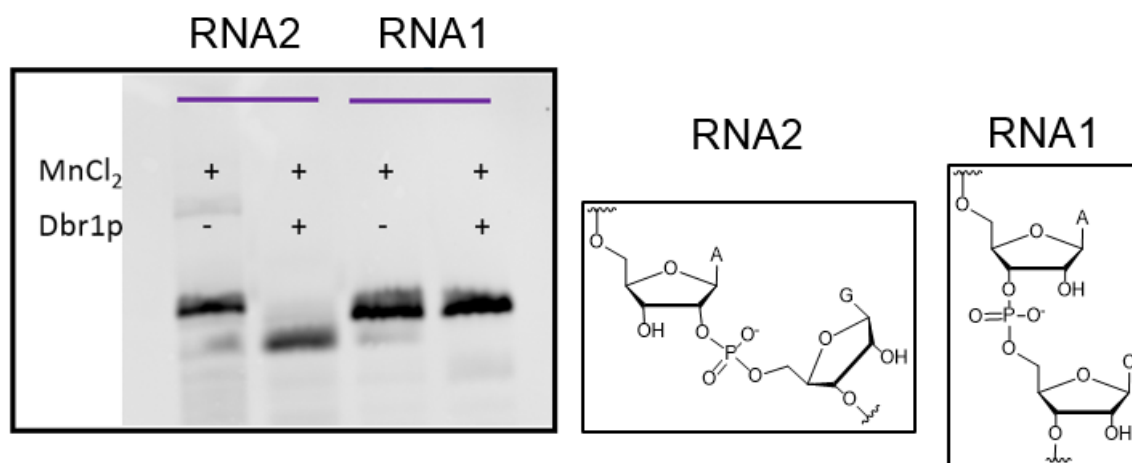


Figure 3.12. Denaturing polyacrylamide gel of EhDbr1p digestion of 13-mer RNAs either with a single 2'-5'-linkage (RNA2) or with all standard 3'-5'-linkages RNA (RNA1) (Left). After 90 minute incubation, EhDbr1p specifically cleaves the 2'-5' phosphodiester bond in the kinked RNA whereas no cleavage of the linear RNA is observed. Schematic representation of RNA2 and RNA1 are in the middle and right most panels, respectively.

The cleavage of linear RNA (RNA1) with Dbr1p was compared to RNA that has a single 2'-5' phosphodiester bond (RNA2) is shown in figure 3.12. After a 90 minute incubation, RNA1 remains intact while RNA2 is specifically cleaved. This shows that the 3'-5'-phosphate is non-essential for debranching enzyme activity and further shows the importance of having a branch off of the 2' position on Dbr1p activity.

3.3.3. Double stranded Substrates

To investigate the substrate requirements of Dbr1p further, we tested RNA that included double strands or duplexed RNA as substrates for EhDbr1p cleavage. The duplexed regions are RNA-DNA hybrids and the bbRNA used is larger than the one used

for the previously described studies. Its sequence is detailed in table 3.4. The larger bbRNA, a 29-mer, was used to ensure that the duplex regions remained intact under assay conditions. The 5'-arm is labeled with an Alexa 488 dye. The five duplexed structures used are detailed in fig. 3.13 and sequences in table 3.4.

Name	Sequence
bbRNA8	5'-Alexa488-ACU ACU GUA CUA A(2'-5'-GUA UGA) CAA GUU ACUU-3'
DNA1	5'- aagt aac ttg t tag tac agt agt – 3'
DNA2	5'- tca tac t tag tac agt agt – 3'
DNA3	5'-tca tac tag tac agt agt – 3'
DNA4	5'-tca tac cc tag tac agt agt – 3'
DNA5	5'-tca tac tt tag tac agt agt – 3'

Table 3.4. RNA sequences used in section 3.3.3.

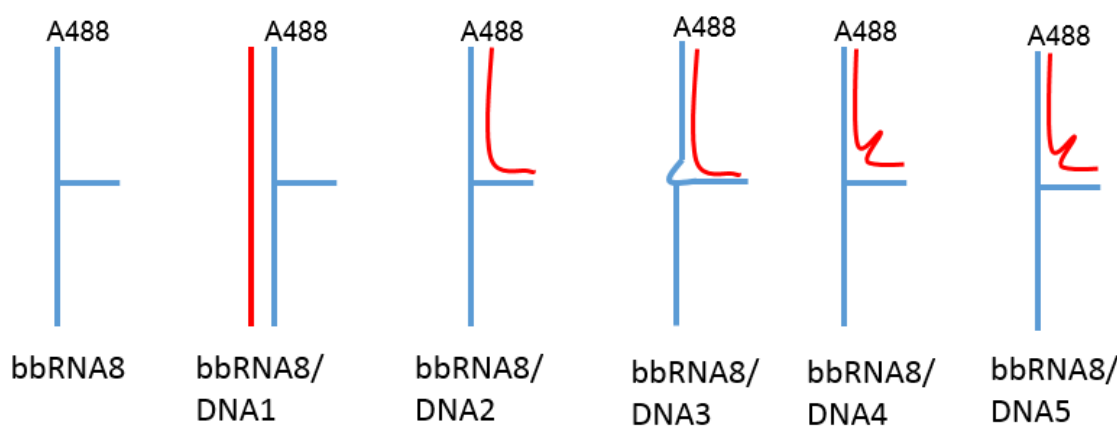


Figure 3.13. Backbone-branched RNA and duplexed bbRNA. The bbRNA is depicted in blue. The annealed DNA is depicted in red.

bbRNA8 is first hybridized with the appropriate DNA partner strand in PBS buffer by heating at 90 °C for one minute, then incubating at 65 °C for 15 min and then held at ambient room temperature (approximately 21 °C) for an hour. The melting temperature of the duplexes that hybridize to the top of the stem and branch are 48 °C and the duplexed

stem is 54 °C based on IDT calculations²¹. All of the duplexes have a T_M of at least ten degrees above the cleavage assay conditions. The bbRNA-DNA duplexes are all incubated with EhDbr1p in standard reaction buffer and cleavage over time is visualized by gel electrophoresis. When the stem is fully duplexed (bbRNA8/ DNA1), there is no cleavage of the RNA in five minutes. The bbRNA-DNA hybrids that contained a bulge near the branch point (bbRNA8/ DNA4, bbRNA8/ DNA5, bbRNA8/ DNA3), either on the RNA or on the DNA, are debranched at rates on the order of the single stranded control. However, bbRNA8/ DNA2 showed a five-fold decrease in activity relative to the single stranded RNA (Fig. 3.14). This indicates that the branch point residue being accessible is critical for debranching activity.

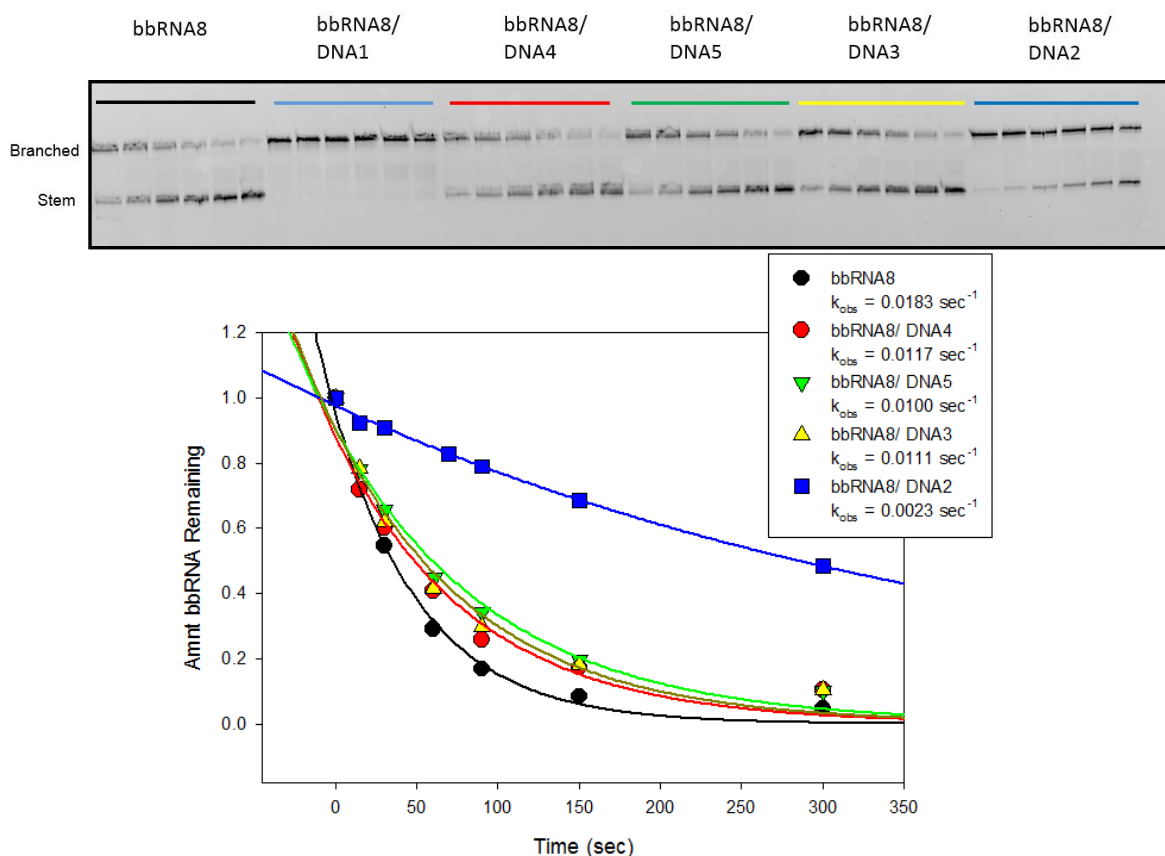


Figure 3.14. Digestion of bbRNA-DNA hybrids by EhDbr1p-1. Debranching reaction of duplexed substrates run on a 20% denaturing polyacrylamide gel (top). Plot of duplexed debranching reactions over time fit to a single exponential decay equation (bottom). The observed rate constant is generated from the fit.

The importance of the branch point residue is further explored later in this thesis, but the strong preference lariat debranching enzyme has for adenosine is a feature known throughout the literature. When debranching enzyme is mutated to decrease activity, the amount of introns with adenosine branch point increases¹⁹. We observe here that having the branch point in Watson-Crick base pairing to a complementary strand decreases the ability of Dbr1p to effectively cleave the 2'-5' linkage. Having duplex regions in other residues of the backbone branched RNA however, do not appear to affect the debranching rate.

To further confirm that the DNA was still duplexed during the debranching reaction DNA3, which bulges out the branch point A, is run at ambient room temperature (21 °C). The reaction is performed 27 °C below the calculated melting temperature of the DNA/RNA duplex. Changes in temperature have not previously been investigated with respect to the activity of the enzyme, so bbRNA8 is evaluated at both 37 °C and ambient temperature. The 37 °C reaction is done to calibrate this reaction to previous experiments and the 21 °C reaction was done to control for possible reaction rate reduction at the lower temperatures.

As seen in figure 3.15, decreasing temperature does decrease rate, but the stable duplex does not have any significant effect on the debranching rate. This confirms the previous experiments and supports the hypothesis that the branch point residue's accessibility is important for proper debranching activity.

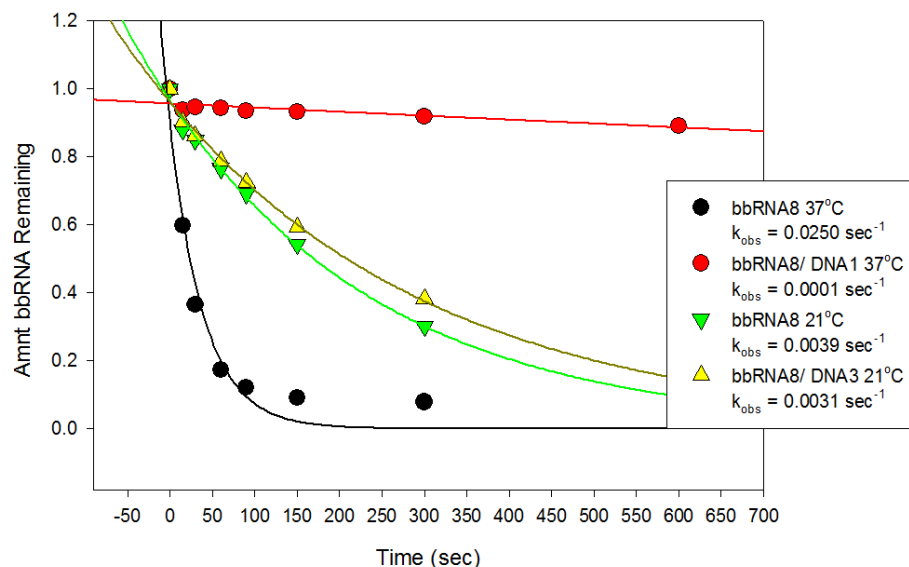


Figure 3.15. Digestion of bbRNA-DNA hybrids by EhDbr1p 2. The black and red plots are debranching reactions run at 37 °C and the green and yellow are run at 21 °C. The data points are generated from quantitation of a denaturing polyacrylamide gel, scanned using a Typhoon FLA9000 and band intensity was quantitated using the associated software. The lines are fits from a single exponential decay equation, which gives the observed rate.

To quantitate the debranching activity of the bbRNA8/DNA1 hybrid, the reaction time was increased from five minutes to ten minutes. The bands from the cleavage reaction are separated on a polyacrylamide gel and quantitated. The debranching reaction of the bbRNA8/ DNA1 is 250 times slower than the debranching of the bbRNA8 alone, at the same temperature confirming the importance of accessibility of the branch point adenosine for debranching. This also strongly suggests that the Dbr1p does not have a capability of melting or unwinding duplex RNA.

3.3.4. Different sized bbRNAs

Dbr1p's main role in the cell is to cleave the 2'-5' phosphodiester bond of introns removed during splicing. Therefore we sought to determine if Dbr1p had a preference for RNAs of a particular sequence length on a particular arm of the substrate. The Michaelis-Menten kinetic parameters of K_M and k_{cat} are determined for three different sized RNAs. Although multiple crystal structures of EhDbr1p have been elucidated, only one has the

protein bound to a bbRNA substrate⁸. Based on this structure, we reasoned that decreasing the length of the branch portion of the bbRNA will have the most impact on binding because it has the most contacts with the protein of all the arms of the bbRNA. No density for the 5'-portion of the stem was observed in the crystal structure and only a few contacts within hydrogen-bonding distance were seen between the 3' portion of the stem and EhDbr1p. It is also hypothesized, that the k_{cat} will not significantly change when altering the length of the bbRNA 'arms', as the catalytic site is not altered.

To investigate this, a Dbr1p cleavage assay based on Förster resonance energy transfer (FRET) is devised. Two fluorophores with sufficient spectral overlap for energy transfer are used, typically Cyanine3 (Cy3) conjugated to the 5'-terminus and a Cyanine5 moiety conjugated to the 3'-terminus of the 2'-arm. These dual labeled RNAs were synthesized using the 3' method described previously, with the Cy3 conjugated to the 5'-terminus through phosphoramidite chemistry and the Cy5 post synthetically attached by click chemistry. The Cy3 dye is excited in the fluorimeter, and the increase in Cy3 signal over time is observed as debranching enzyme works to cleave the 2'-5'-phosphodiester bond of the bbRNA. Cleavage of the bond by Dbr1p changes the distance between the two dyes, stopping energy transfer. The increase in Cy3 fluorescence signal is subsequently converted to a rate of reaction. These reactions are fit to Michaelis-Menten enzyme kinetics models. The Eadie-Hofstee linearization (Equation 3.1) is used to plot the data and determine the K_M and k_{cat} of EhDbr1p when acting on bbRNA of varying sizes.

Equation 3.1. Eadie-Hofstee Equation

$$v = -K_M * \frac{v}{[S]} + V_{max}$$

The standard bbRNA used in debranching reactions throughout this thesis is a 19mer, six residues in the 5'-arm, 3'-arm, and 2'-arm. Shortening the bbRNA by 3 nucleotides from the 5'-arm or 3'-arm is investigated. Therefore, we hypothesize that shortening the 3'-arm of the stem will have deleterious effects to the ability of Dbr1p to bind the RNA, but this effect might not be seen in the shortening of the 5'-arm as it does not make any contacts with the enzyme. However, this is not what was observed (fig. 3.16). The RNA used in this experiment is detailed in table 3.5.

Name	Sequence
bbRNA9	5'-Cy3-GUA CUA A(2'-5'-GUA UGA-t-Cy5) CAA GUU-3'
bbRNA10	5'-Cy3-GUA CUA A(2'-5'-GUA UGA-t-Cy5) CAA-3'
bbRNA11	5'-Cy3-CUA A(2'-5'-GUA UGA-t-Cy5) CAA GUU-3'

Table 3.5. List of RNA sequences used in section 3.3.4.

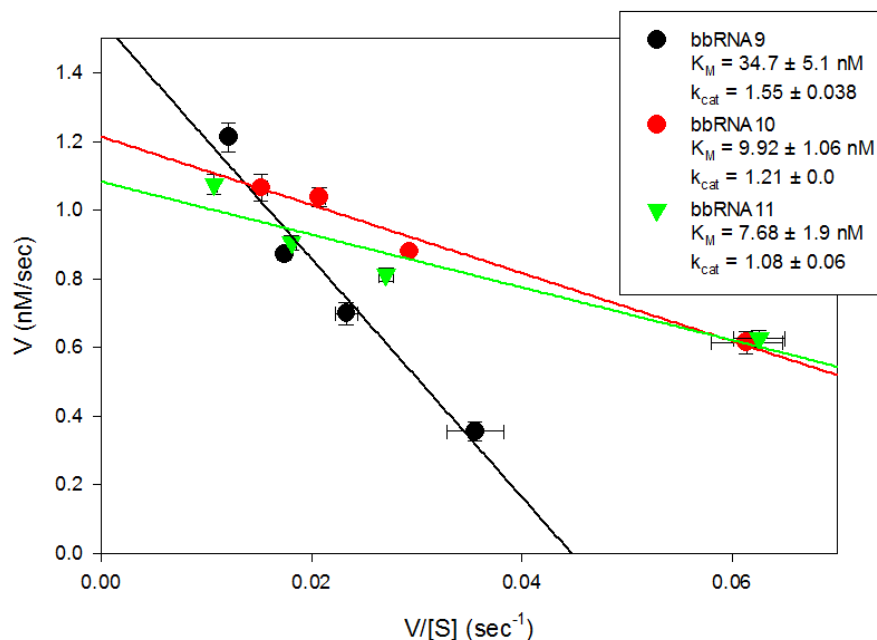


Figure 3.16. Initial rate debranching assay of different sized bbRNAs. Eadie-Hofstee plot of initial rate of debranching reaction with bbRNA9, bbRNA10, and bbRNA11. RNA concentrations range from 10 – 100 nM with EhDbr1p at 1 nM. Each point represents the mean of triplicate experiments and the standard deviation is represented by the error bars.

With these RNAs, the catalytic rate constant stays roughly the same for each of the tested RNAs. However the binding constants varied slightly with size. Shortening either the 5'-arm or the 3'-arm by three nucleotides improved binding ($K_M = 7.68$ nM) over the larger bbRNA9 ($K_M = 34.6$ nM).

3.3.5. Alternate Branch Point Residues

When introns are spliced, the branch point residue is almost exclusively adenosine. Introns exist that contain alternate nucleobases at the branch point, but are a smaller portion of the total intron pool¹⁹. Introns with cytidine branch points are poor substrates for Dbr1p¹⁰. However, preliminary data in the lab using ScDbr1p indicated that the enzyme is able to cleave bbRNA that contained uridine, cytidine, or guanosine as the branch point residues. With access to active EhDbr1p, we sought to examine these

alternate bbRNAs as substrates and their effects on the kinetics of the Dbr1p reaction. The alternate branch point bbRNAs are readily made using the same technique as described earlier. The only difference being the diol used to synthesize the photocleavable monomer was either cytidine or uridine instead of adenosine. The substrates are synthesized using the 3'-method described earlier and are dual labeled with Cy3 and Cy5 dyes so initial rates of the debranching reaction can be evaluated as well as both pieces of the RNA can be resolved on a polyacrylamide gel. The bbRNAs are detailed in table 3.6.

Name	Sequence
bbRNA9	5'-Cy3-GUA CUA A(2'-5'-GUA UGA- t -Cy5) CAA GUU-3'
bbRNA12	5'-Dy547-GUA CUA U(2'-5'-GUA UGA- t -Cy5) CAA GUU-3'
bbRNA13	5'-Dy547-GUA CUA C(2'-5'-GUA UGA- t -Cy5) CAA GUU-3'
bbRNA14	5'-Cy3-GUA CUA G(2'-5'-GUA UGA-(CH ₂) ₄ - t -Cy5) CAA GUU-3'
bbRNA15	5'-Dy547-GUA CUA A(2'-5'-GUA UGA-(CH ₂) ₄ - t -Cy5) CAA GUU-3'

Table 3.6. RNA used in section 3.3.5.

When labeling a guanosine diol with the photocleavable protecting group, only 2'-NOVM is a product of that reaction. Therefore, a hexynyl linker is required to attach a second dye to the branch. The hexynyl group is attached through phosphoramidite chemistry and the Cy5 dye is post-synthetically conjugated using CuAAC (bbRNA14). A cognate RNA is made that has an adenosine at the branch point as well as the hexynyl (bbRNA15) linker to test if the linker addition changes the interaction with Dbr1p. Both of the RNAs with the hexynyl linker are cleaved by Dbr1p over ten minutes (fig. 3.17). The initial rate assay for bbRNA15 did not match bbRNA9 and showed no change in rate over the same concentration range as bbRNA9, indicating that the hexynyl linker is negatively impacting interaction with Dbr1p (fig. 3.17). Therefore, bbRNA14 was not further

investigated. A debranching reaction was run to see if the branch point substrates are cleaved by EhDbr1p. After a ten minute incubation, the substrates were loaded on a 20% denaturing polyacrylamide gel (fig. 3.18).

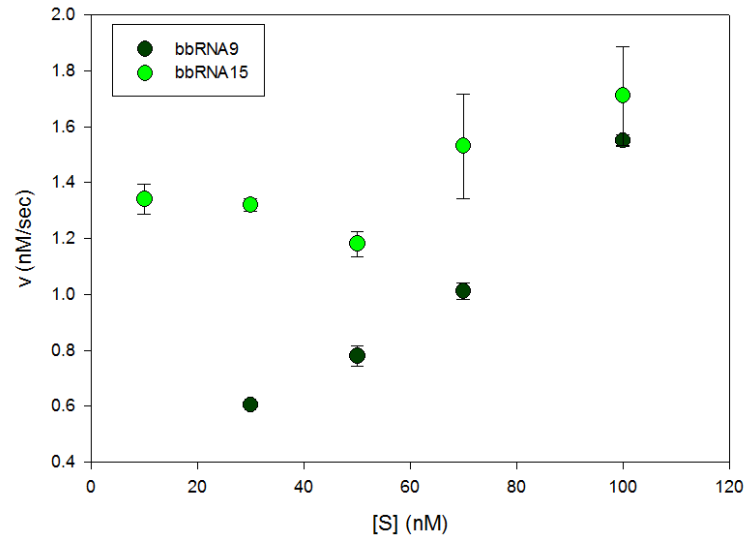


Figure 3.17. Initial rate debranching assay of bbRNA with hexynyl linker. bbRNA9 and bbRNA15 plotted as velocity against substrate concentration. bbRNA9 shows an increase in velocity as substrate concentration increases, but bbRNA15 does not. The plotted point is the mean of three experiments and error bars represent standard deviation.

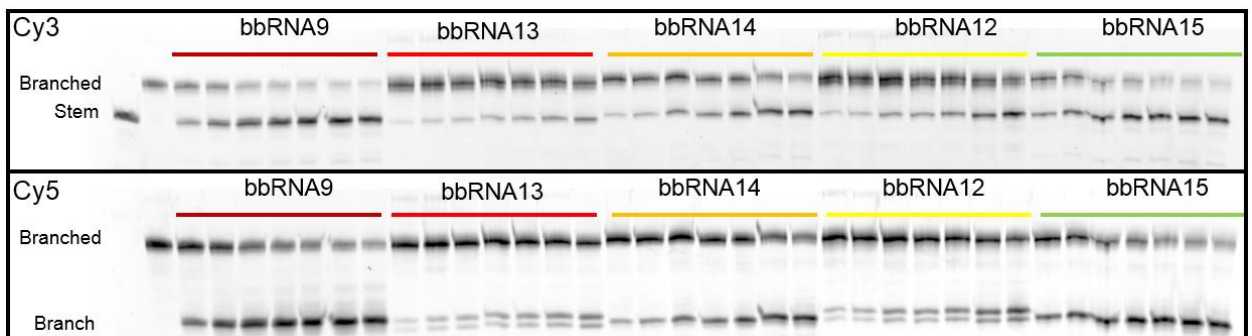


Figure 3.18. Images of 20% polyacrylamide gel of the EhDbr1p debranching of dual labelled bbRNAs with alternate branch points. The top box is a Cy3 scan, showing bbRNA and stem. The bottom box is a Cy5 scan of the debranching reaction and shows bbRNA and branch. Pyrimidines at the branch point (bbRNA12 and 13) show two bands in the Cy5 scan.

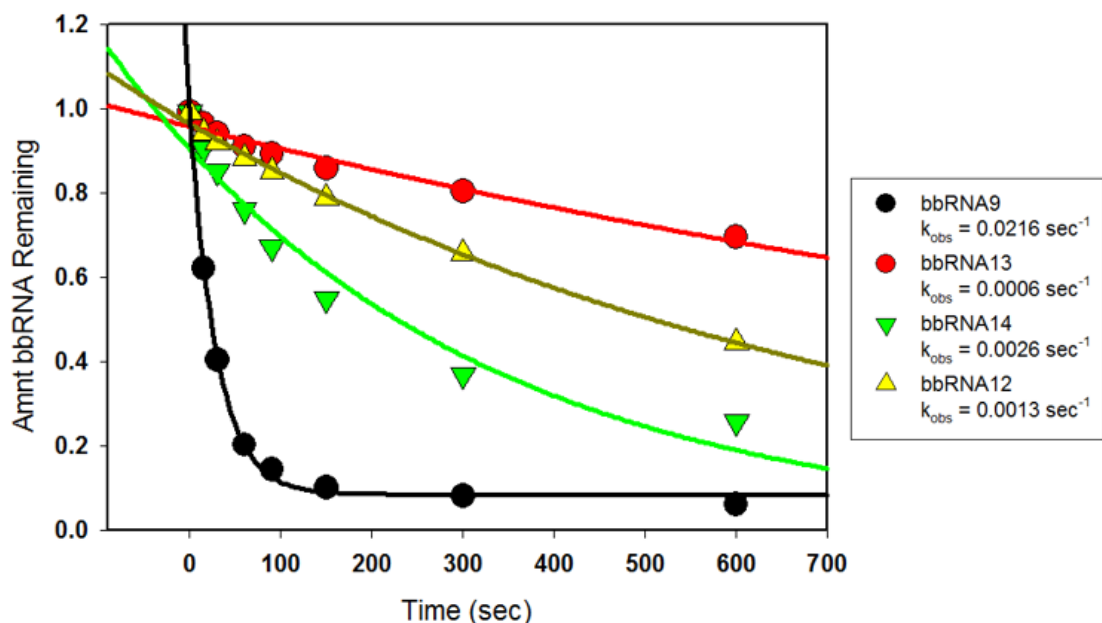


Figure 3.19. Digestion of alternate branch point RNA to get observed rates. The data points are generated from quantitation of a denaturing polyacrylamide gel (fig. 3.18), scanned using a Typhoon FLA9000 and band intensity was quantitated using the associated software. The lines are fits from a single exponential decay equation, which gives the observed rate.

As seen in Cy3 scan of figure 3.18, the stem appears over time and no other cleavage is observed. In the Cy5 scan, the branch accumulates over time, as expected. bbRNAs 12 and 13, branch points uridine and cytidine respectively, show two bands appearing over time in the Cy5 scan. All substrates contain the same Cy5, with the same method of attachment, so a dye effect does not explain this observation. There appears to be a difference in how binding and cleavage occurs when lariat debranching enzyme encounters a bbRNA with a pyrimidine at the branch point. Since there is only one cleavage product in the Cy3 channel, this indicates that Dbr1p is cleaving the 2'-5' phosphodiester bond first, separating the branch from the stem and then subsequently doing a second cleavage on the branch (fig. 3.20). The reaction rates are also slower for the bbRNA13 and bbRNA12 (fig. 3.19). How exactly this is occurring is unclear, but changing the branch point from a purine to a pyrimidine causes Dbr1p to react twice with the branch portion of the bbRNA.

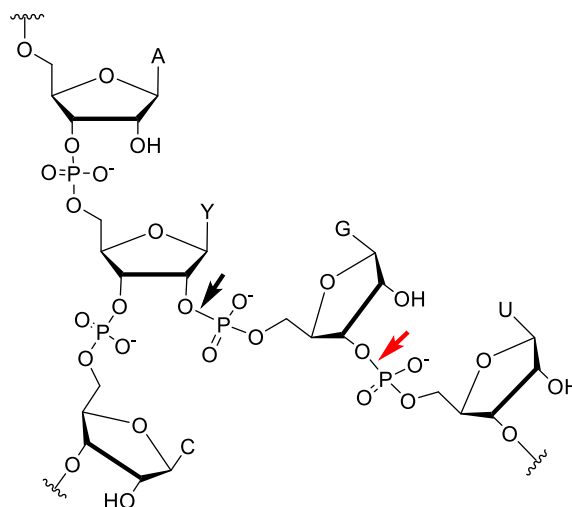


Figure 3.20. Schematic of proposed Dbr1p cleavage sites of bbRNA that contain a pyrimidine residue at the branch point. The black arrow denotes the known 2'-5'-phosphodiesterase activity of Dbr1p. The red arrow is the proposed second cleavage site that occurs subsequent to normal cleavage.

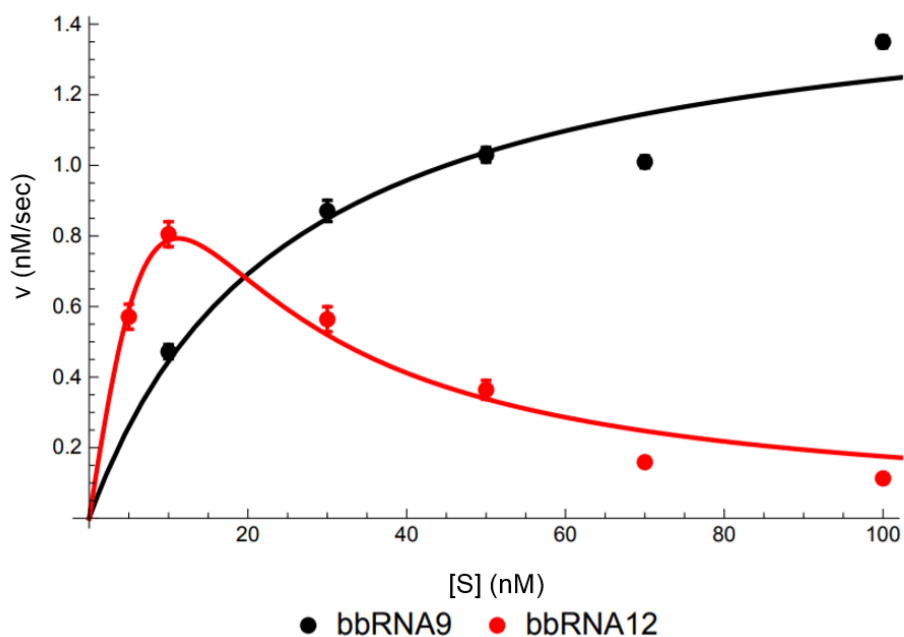


Figure 3.21. Initial rate debranching assay of branch point adenosine and branch point uridine bbRNA. Debranching rate as a function of substrate concentration increases for bbRNA9 (with branch point adenosine) and bbRNA12 (with branch point uridine) with EhDbr1p. Data for bbRNA9 cleavage is fit to the Michaelis-Menten equation while cleavage data for bbRNA12 is fit to the substrate inhibition equation. The data plotted is the mean initial velocity and error bars represent standard deviation from three experiments.

To explore the difference in cleavage kinetics of the alternate branch point bbRNA, initial rate reactions were performed. As seen in figure 3.21, in contrast to bbRNA9, the initial velocity of bbRNA12 first increases, and then decreases. The increase in activity occurs between 0 nM and 20 nM bbRNA12 added to EhDbr1p. After 20 nM, the velocity of the reaction decreases with increasing substrate. These data are fit a curve that models substrate inhibition, figure 3.19, equation 3.3. Substrate inhibition typically occurs when a second substrate molecule associates with the enzyme substrate complex and creates a catalytically dead complex. The inhibited complex usually occurs after a misdocking of the initial substrate molecule, leaving a second binding pocket open for a second substrate molecule. In this case, it is unclear exactly how a branch point A to U modification would cause such a dramatic misdocking or docking of a second substrate. The simplified kinetic expression for substrate inhibition is shown in equation 3.3 and the traditional Michaelis-Menten equation is shown in equation 3.2. The model for substrate inhibition is shown in figure 3.22.

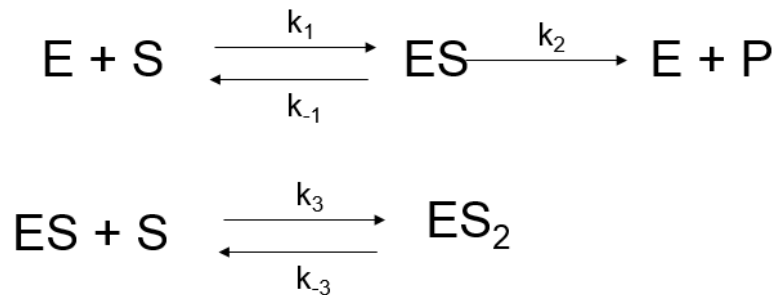


Figure 3.22. Kinetic model for substrate inhibition. Enzyme, substrate and product are denoted with the letters E, S, and P respectively. ES is the enzyme substrate complex and ES₂ is the inhibited complex.

Equation 3.2. Michaelis-Menten Equation

$$v = \frac{V_{max} * [S]}{K_M + [S]}$$

Equation 3.3. Substrate Inhibition

$$v = \frac{k_2 K' [S]}{K_M K' + K' [S] + [S]^2}$$

Equation 3.4. K' as the inhibition constant

$$K' = \frac{k_{-3}}{k_3}$$

When bbRNA12 is fit to a substrate inhibition curve, the K_M value is much larger than the K_M value bpA. The values are in table 3.7. The difference in K_M between these closely related substrates is too great to conclusively claim that substrate inhibition is the correct kinetic model. Other interactions might be occurring that are currently unknown and thus unaccounted for in the model.

RNA	K_M (nM)	k_{cat} (sec ⁻¹)	K' (nM)
bbRNA9	24.8	1.55	--
bbRNA12	4901	694	0.0255

Table 3.7. Calculated kinetic parameters for bbRNA9 and bbRNA12.

A similar increase and decrease in activity as substrate concentration increases is seen when the branch point residue is mutated to a cytidine (fig. 3.23). However this effect is much more dramatic. The significant decline in activity makes it difficult to accurately calculate initial rates because the change in Cy3 signal over time is minimal. To compensate for this, the enzyme concentration is increased from 1 nM to 5 nM. However, results from these experiments also could not be fit and are inconclusive.

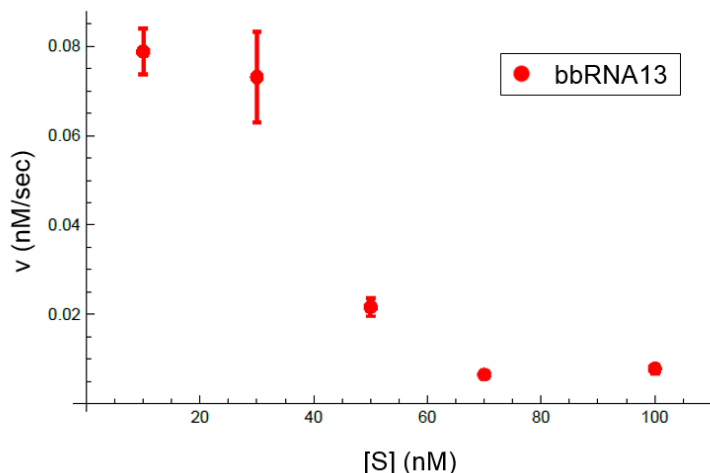


Figure 3.23. Initial rate debranching assay of branch point cytidine by EhDbr1p. Debranching initial velocity as substrate concentration increases for bbRNA with branch point cytidine substrates with 1 nM EhDbr1p. The mean of three experiments is plotted and error bars represent standard deviation.

The model for Dbr1p binding and cleavage of bbRNA with pyrimidine branch points needs to be examined as more information about these enzyme/RNA complexes becomes available. Should pyrimidine branch point bbRNAs be crystalized with an inactive form of enzyme, it may be possible to visualize how the RNA fits into the binding pocket. Alternately and complementarily, computational modeling may provide information on how the macromolecules will interact. These may provide additional or new insights for modelling the data for the kinetics of the cleavage reaction.

3.4. Conclusions

In this chapter, the versatility of lariat debranching enzyme with metal ions and substrates was explored. The Mn^{2+} requirements were investigated for lariat debranching enzyme from two different organisms, *S. cerevisiae* and *E. histolytica*. While both enzymes show a similar overall profile in the Mn^{2+} titration, a modest difference is observed between the two. EhDbr1p requires less Mn^{2+} (1 mM) added to the reaction mixture than ScDbr1p, which requires 4 mM to reach maximal rate. Further differences are seen between the enzymes when EDTA is added to the reaction buffer. EhDbr1p remains unaffected by the addition of the metal chelator while ScDbr1p loses activity. While the identity of the specific active site metal ions remains unresolved, the different effect of EDTA on the two enzymes indicates differences between the enzymes when it comes to the arrangement of their metal binding sites.

Besides manganese, other divalent metal ions were screened for any effects on EhDbr1p activity. The results are consistent with previous literature reports with ScDbr1p, and but expand those results. Previously, with ScDbr1p, 2.5 mM Mg^{2+} , Ba^{2+} or Ca^{2+} , added over a background of 4 mM Mn^{2+} did not cause a significant observable change in debranching enzyme activity². These divalent ions can be used with EhDbr1p, even without a Mn^{2+} background, and similar activity is retained. We also observe, and confirm, previous experiments that added EDTA does not negatively affect the debranching rate. However, Cu^{2+} , Zn^{2+} completely abolish Dbr1p activity and Cd^{2+} significantly slows down the reaction. Since these experiments, structures of EhDbr1p have been published, with differing metal ion occupancy in the active site: Mn^{2+} , Fe^{2+} and Zn^{2+} , and Mn^{2+} and Zn^{2+} . What the native metal ions are and whether they may be readily replaced and how that impacts activity remains to be resolved. As metalloenzyme active sites can be promiscuous without significant effects on *in vitro* activity, other direct methods to identify the active site metal ions, such as EPR spectroscopy, are required to address this.

The substrate requirements for Dbr1p activity were also investigated. Our results show that bbRNAs with 2'-3'-phosphodiester bonds are still cleavable by the enzyme. In these substrates the directionality of the branch is reversed from 5'-3' to 3'-5', potentially disrupting many enzyme to RNA contacts. In addition, we observe debranching enzyme cleavage of a single 2'-5'-phosphodiester bonds when there is no vicinal 3'-5'-phosphodiester bond on the ribose.

We show that modifications at the 2'-position of the first residue in the 2'-arm are well tolerated by both ScDbr1p and EhDbr1p. Duplexed branches and duplexed stems are debranched, however this occurs only when the branch point residue is accessible and not involved in Watson-Crick base pairing. Changing the length of the bbRNA substrates has minimal effect, and shortening the top or the bottom of the stem has only a modest increase in binding.

Changing the branch point nucleotide from a purine to a pyrimidine negatively affects the rate of cleavage at high substrate concentrations. The increase and subsequent decrease in activity observed during the initial rate assay fits well when modeled similar to substrate inhibition. Gel electrophoresis of the debranching products reveal an additional cleavage in the branch portion of the bbRNA after the initial cleavage of the 2'-5' phosphodiester bond.

As Dbr1p is a critical enzyme for cellular function, it is important to understand the limitations of debranching activity. Collectively, the experiments and results described here expand our insight into the range of Dbr1p cleavage activity and its requirements.

3.5. Experimental

Name	Sequence
bbRNA1	5'-Cy3-GUA CUA A(2'-5'-GUA UGA) CAA GUU-3'
bbRNA2	5'-Dy547-GUA CUA A(2'-5'-gUA UGA) CAA GUU – 3'
bbRNA3	5'-Dy547-GUA CUA A(2'-5'-G(2'F)UAUGA) CAA GUU – 3'
bbRNA4	5'-Dy547-GUA CUA A(2'-5'-G(2'OMe)UA UGA) CAA GUU – 3'
bbRNA5	5'-Dy547-GUA CUA A(2'-5'-Gta tga) CAA GUU – 3'
bbRNA6	5'-Dy547-GUA CUA A(2'-3'-GUA UGA) CAA GUU – 3'
bbRNA7	5'-Dy547-GUA CUA A(2'-3'-gta tga) CAA GUU – 3'
bbRNA8	5'-Alexa488-ACU ACU GUA CUA A(2'-5'-GUA UGA) CAA GUU ACUU-3'
bbRNA9	5'-Cy3-GUA CUA A(2'-5'-GUA UGA-t-Cy5) CAA GUU – 3'
bbRNA10	5'-Cy3-GUA CUA A(2'-5'-GUA UGA-t-Cy5) CAA – 3'
bbRNA11	5'-Cy3-CUA A(2'-5'-GUA UGA-t-Cy5) CAA GUU – 3'
bbRNA12	5'-Dy547-GUA CUA U(2'-5'-GUA UGA-t-Cy5) CAA GUU – 3'
bbRNA13	5'-Dy547-GUA CUA C(2'-5'-GUA UGA-t-Cy5) CAA GUU – 3'
bbRNA14	5'-Cy3-GUA CUA G(2'-5'-GUA UGA-(CH ₂) ₄ -t-Cy5) CAA GUU-3'
bbRNA15	5'-Dy547-GUA CUA A(2'-5'-GUA UGA-(CH ₂) ₄ -t-Cy5) CAA GUU-3'
bbDNA1	5'-Dy547-gta cta a(2'-5'-gta tga) caa gtt-3'
RNA1	5'-Cy3-GUA CUA A CAA GUU – 3'
RNA2	5'-Cy3-GUA CUA A(2'-5'-GUA UGA)-3'
DNA1	5'- aagt aac ttg t tag tac agt agt – 3'
DNA2	5'- tca tac t tag tac agt agt – 3'
DNA3	5'-tca tac tag tac agt agt – 3'
DNA4	5'-tca tac cc tag tac agt agt – 3'
DNA5	5'-tca tac tt tag tac agt agt – 3'

Table 3.8. Sequences used in the biochemical studies of Dbr1p. Uppercase letters denote RNA, lowercase DNA and t is a triazole linkage.

3.5.1. Gel analyzed debranching reaction – General

To perform a gel analyzed debranching reaction, the components of debranching buffer, 50 mM Tris.HCl pH 7.5, 25 mM NaCl, 2.5 mM DTT, 0.05% Triton X-100, and MnCl₂.

When using EhDbr1p, 1 mM MnCl₂ is used, and when using ScDbr1p, 4 mM MnCl₂ is

used. The RNA is at 100 nM and the enzyme is at 1 nM for reactions unless otherwise stated. To the debranching buffer, Dbr1p is added and is incubated at 37 °C for three minutes. The reaction is then started by addition of backbone branched RNA.

To evaluate the observed rate, 20 µL aliquots are removed from the reaction and added to 40 µL of stop solution (9:1 formamide: 0.1 M EDTA, trace bromophenyl blue). Time points are typically 15, 30, 60, 90, 150, 300 seconds. The aliquots are then run on a 20% denaturing polyacrylamide gel with 8 M urea and 1x TBE buffer. 15 µL of the aliquot is loaded into the wells and the gel was run in 0.5x TBE at 10 W for 2.5 hrs. This is sufficient to separate branched RNA from the debranched stem.

The gel is then visualized using a Typhoon FLA9000 (GE healthcare life sciences) using a fluorescent scanner stage. Most bbRNA are labeled with Dylight547 or Cy3 so the Cy3 settings were typically used, unless otherwise noted. The quantitation was done using associated software and the fraction of cleaved bbRNA was determined. The fraction cleaved was plotted against time and was fit to a single exponential decay fit using SigmaPlot.

3.5.2. Mn^{2+} titration

The Mn^{2+} titration was performed by using the general procedure outlined in 3.5.1. and changing the amount of MnCl_2 in the debranching buffer. Separate time course experiments were done to get an observed rate constant at each desired Mn^{2+} concentration. The MnCl_2 concentrations used were 0.5, 1, 2, 4, and 10 mM for both ScDbr1p and EhDbr1p. Following determination on individual rate constants, the observed rate was plotted against the concentration to determine optimal Mn^{2+} conditions for both enzymes.

3.5.3. EDTA reaction

The EDTA reactions were done using the general procedure outlined in 3.5.1. To test different amounts of added EDTA, instead of adding MnCl_2 to the reaction mixture, EDTA was added to the debranching buffer. Each concentration of EDTA was evaluated as stated above.

To test different incubation times of enzyme with EDTA, the general procedure was followed, using 1 mM EDTA instead of Mn^{2+} . Incubation time of enzyme with debranching buffer was increased to ten and thirty minutes as indicated. Workup of the reaction was performed as stated in 3.5.1.

3.5.4. Metal Screen

The metal screen was performed using the same procedure as outlined in 3.5.1 and a variety of metal ions were used as indicated, all at a 1 mM final concentration. The time points evaluated were one and ten minutes. The gel was imaged as detailed in 3.5.1 but not quantitated or evaluated for an observed rate constant.

3.5.5. Reducing Agent

Reducing agent was evaluated using the 3.5.1. procedure. Both 1,4-dithiothreitol (DTT) and tris(2-carboxyethyl)phosphine (TCEP) were evaluated. DTT was tested at 2.5 mM, the standard conditions, and TCEP was tested at 0.5 mM, the literature conditions, and 2.5 mM to directly compare to our assay. The gel was imaged and quantitated as detailed in 3.5.1 to compare observed rates.

3.5.6. DNA/ RNA hybrids

The DNA/ RNA hybrid bbRNAs were evaluated using both EhDbr1p and ScDbr1p. The metal ion used was Mn^{2+} at the concentration required by the individual enzymes. Each substrate was evaluated separately and at least twice. Time points varied with each hybrid. All of the substrates with the 2' position of the first residue in the 2'-arm were

evaluated at 15, 30, 60, 90, 150, 300, and 600 seconds. The bbRNA5 was evaluated at 2, 10, 20, 40, and 60 minutes. bbRNA6, bbRNA7 was evaluated at 5, 10, 30, 60, and 90 minutes. bbDNA1 was evaluated at 1, 2, 5, 10, and 30 minutes for ScDbr1p and 2, 10, 20, 40, 60 minutes for EhDbr1p. The individual k_{obs} were calculated and the mean of two runs was plotted in a bar graph. The error bars represent the standard deviation of the runs.

The high enzyme reactions were done using 500 nM enzyme and 100 nM RNA and let to incubate at 37 °C for 10 minutes. The reactions were then quenched and run on a polyacrylamide gel as described above.

3.5.7. 2'-5' Kinked RNA

The reactions were set up with debranching buffer and EhDbr1p. The RNAs evaluated are RNA1 and RNA2. The RNAs were incubated with the enzyme for 90 minutes and then the reaction was quenched using stop solution. The reactions were run on a 20% denaturing polyacrylamide gel and visualized using the Typhoon scanner as detailed in 3.5.1.

3.5.8. Duplexed Substrates

The first set of debranching reactions were done as described in the method outlined in 3.5.1. However, the duplexed backbone branched substrates were first hybridized together before addition to the debranching reaction. The hybridization was accomplished by making a 20 μ L reaction that has the final concentration of the backbone branched RNA be 1 μ M. The associated DNA strand was mixed with the bbRNA in water in a 1:1 ratio and heated at 90 °C for 3 minutes, 65 °C for 10 min, left at ambient room temperature (~21 °C) for 20 minutes and kept on ice until use. The hybridization was done immediately before debranching reaction. The RNA and DNA used is detailed in the table above. The analysis and quantitation was done as described.

The second round of duplexed reactions were performed in a similar fashion, with the hybridization reactions carried out identically. For the reaction performed at ambient room temperature, the preincubation and reaction was carried out on the bench top. Analysis and quantitation was done as previously described.

3.5.9. General Procedure for Debranching Reactions in the Fluorimeter

Dual labeled bbRNAs are used for studies performed in the fluorimeter to track the initial rate of cleavage reactions. These are done inside of a Fluoromax-2 fluorimeter using a rapid mixing device (Applied Photophysics RX2000) that allows for stop flow style mixing. Into one syringe, the RNA substrate is added in debranching buffer. In the second syringe, EhDbr1p is added and is also in debranching buffer. There is one small alteration to the debranching buffer, 0.005% Triton X-100 is used instead of 0.05% to prevent bubble formation in the lines of the rapid mixer which might result in poor reads in the fluorimeter.

The fluorimeter is set to run a time based data acquisition. The laser excites the Cy3 signal, 540 nm, and collect the emission signal at 560 nm. The slit width is 4 nm. Time points were collected every 0.05 seconds for 30 seconds. The data collection is started first, and then the rapid mixing device is used to push a fresh volume of liquid into the cuvette. The enzyme concentration stays constant at 1 nM, but the initial rates are calculated for five different substrate concentrations, 10, 30, 50, 70, and 100 nM bbRNA. Each concentration is run in triplicate.

The initial rates were calculated using the first five seconds of the reaction once the cuvette was filled with fresh solution from the rapid mixer. The rates are calculated from the slope of the fluorescence v. time plots. To convert initial rates from units of fluorescence to units of concentration, the donor dye (Cy3) signal is determined for a bbRNA (minimum fluorescence) and a stem (maximum fluorescence) for each concentration of substrate used. Taking the maximum fluorescence (I_0) and subtracting

from the minimum fluorescence (I_o) gives us the maximal possible signal change. The initial rate in terms of fluorescence (AU/sec) is multiplied by the concentration (nM) of substrate and then divided by the $I_e - I_o$ (AU) term to give the initial rate v_o (nM/sec).

These initial rates can then be plotted by substrate concentration to give the classic hyperbolic Michaelis-Menten plots. However, transforming the data and plotting v_o v. $v_o/[S]$ is the Eadie-Hofstee transformation and the kinetic parameters of k_{cat} and K_M can be calculated from a linear fit of the data. Since the data is collected in triplicate, the mean of the three runs is plotted and error bars represent the standard deviation of the three experiments.

3.5.10. Different Sized Substrates

The RNAs used in this experiment are detailed in the table above. The reactions were run and evaluated as stated in 3.5.9.

3.5.11. Alternate Branch Points

The RNAs used in the gel based assay are detailed above. The reactions were performed as described in 3.5.1 with the addition of a 600 second time point. The gel was scanned on the Typhoon FLA9000 at both the Cy3 and the Cy5 settings to visualize both dyes on the RNA.

The bpA and bpU sequences were run using an initial rate assay with the same procedure described in 3.5.9. The data was plotted as a standard Michaelis-Menten plot with $[S]$ v. v_o . The experiment was run in triplicate, with the means plotted and the error bars representing standard deviation of the measurements. The plots were generated in Mathematica using the non-linear model fit of the equations detailed in section 3.3.5.

3.6. References

1. Ruskin, B. & Green, M. R. An RNA Processing Activity that Debranches RNA Lariats. *Science* (80-.). (1985).
2. Khalid, M. F., Damha, M. J., Shuman, S. & Schwer, B. Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Res.* **33**, 6349–6360 (2005).
3. Hopfner, K. P. *et al.* Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell* **105**, 473–485 (2001).
4. Nam, K. *et al.* Yeast lariat debranching enzyme: Substrate and sequence specificity. *J. Biol. Chem.* **269**, 20613–20621 (1994).
5. Katolik, A. *et al.* Regiospecific solid-phase synthesis of branched oligoribonucleotides that mimic intronic lariat RNA intermediates. *J. Org. Chem.* **79**, 963–975 (2014).
6. Arenas, J. & Hurwitz, J. Purification of a RNA debranching activity from HeLa cells. *J. Biol. Chem.* **262**, 4274–4279 (1987).
7. Ganeshan, K. *et al.* Novel approaches to the synthesis and analysis of branched RNA. *Nucleosides Nucleotides* **14**, 1009–1013 (1995).
8. Clark, N. E. *et al.* Metal dependence and branched RNA cocrystal structures of the RNA lariat debranching enzyme Dbr1. *Proc. Natl. Acad. Sci.* **113**, 14727–14732 (2016).
9. Ooi, S. L. *et al.* *RNA lariat debranching enzyme. Methods in Enzymology* **342**, (Elsevier Masson SAS, 2001).
10. Jacquier, A. & Rosbash, M. RNA Splicing and Intron Turnover are Greatly Diminished by a Mutant Yeast Branch Point. *Proc Natl Acad Sci U S A* **83**, 5835–5839

(1986).

11. Pratico, E. D., Wang, Y. & Silverman, S. K. A deoxyribozyme that synthesizes 2',5'-branched RNA with any branch-site nucleotide. *Nucleic Acids Res.* **33**, 3503–3512 (2005).
12. Mourani, R. & Damha, M. J. Synthesis, characterization, and biological properties of small branched RNA fragments containing chiral (R_p and S_p) 2',5'-phosphorothioate linkages. *Nucleosides, Nucleotides and Nucleic Acids* **25**, 203–229 (2006).
13. Schwer, B., Khalid, F. & Shuman, S. Mechanistic insights into the manganese-dependent phosphodiesterase activity of yeast Dbr1 with bis-p- nitrophenylphosphate and branched RNA substrates. *RNA* **22**, 1819–1827 (2016).
14. Romani, A. M. P. Intracellular magnesium homeostasis. *Magnes. Cent. Nerv. Syst.* **512**, 13–58 (2011).
15. Montemayor, E. J. *et al.* Structural basis of lariat RNA recognition by the intron debranching enzyme Dbr1. *Nucleic Acids Res.* **42**, 10845–10855 (2014).
16. Ransey, E. *et al.* Crystal structure of the *Entamoeba histolytica* RNA lariat debranching enzyme EhDbr1 reveals a catalytic Zn^{2+}/Mn^{2+} heterobinucleation. *FEBS Lett.* **591**, 2003–2010 (2017).
17. Getz, E. B., Xiao, M., Chakrabarty, T., Cooke, R. & Selvin, P. R. A Comparison between the Sulfhydryl Reductants Tris(2-carboxyethyl)phosphine and Dithiothreitol for Use in Protein Biochemistry. *Anal. Biochem.* **273**, 73–80 (1999).
18. Katolik, A. *et al.* Fluorescent Branched RNAs for High-Throughput Analysis of Dbr1 Enzyme Kinetics and Inhibition. *ACS Chem. Biol.* **12**, 622–627 (2017).
19. Zhang, S.-Y. *et al.* Inborn Errors of RNA Lariat Metabolism in Humans with Brainstem Viral Infection. *Cell* **172**, 952–965 (2018).
20. Sproat, B. S. *et al.* Novel solid-phase synthesis of branched oligoribonucleotides, including a substrate for the RNA debranching enzyme. *J. Chem. Soc. Perkin Trans. 1*

419 (1994). doi:10.1039/p19940000419

21. Sugimoto, N. *et al.* Thermodynamic parameters to predict stability of RNA/ DNA hybrid duplexes. *Biochemistry*. **34**, 11211-6 (1995).

22. Ueda, E.K.M, *et al.* Current and prospective applications of metal ion-protein binding. *J. Chromatogr. A*. **988**, 1-23 (2003).

Chapter 4 – Inhibition of Lariat Debranching Enzyme with Triazole Branched RNA

4.1 Introduction

Proper function of lariat debranching enzyme is required for life in higher organisms. Dbr1p is involved upstream of a variety of different biochemical pathways because it is required for intron processing and intron processing is required for snoRNA and mirtron biogenesis¹⁻⁴. The most common result of Dbr1p inhibition is accumulation of introns^{1,5}. The importance of Dbr1p as a regulator of biochemical pathways, detailed in chapter 1, has prompted the exploration of therapeutic interventions. In ALS cells, Dbr1p has been shown to suppress the toxicity of TDP-43⁶. In Dbr1p has also been shown to play a role in HIV-1 replication⁷⁻⁹.

Knockdown of Dbr1p is used as a research tool to determine its role in specific pathways. Currently, the most used method for Dbr1p inhibition is knockdown of using short interfering RNA (siRNA). The positive health effects of inhibiting Dbr1p were elucidated using siRNA. However, it is of interest to generate chemical inhibitors of Dbr1p instead of relying on the RNA interference pathway, to develop possible therapeutics for those with ALS or HIV.

Recently, the Damha lab has synthesized inhibitors of Dbr1p^{10,11}. These inhibitors are backbone branched RNAs that contain modifications to the 2'-5'-phosphodiester bond or changes in the branch point ribose. In a 2017 paper, five inhibitors were synthesized (figure 4.1), and each inhibitor shown to inhibit Dbr1p activity¹⁰. The bbRNA that contained modified 2'-5'-linkages, phosphoramidate and phosphorothioate, are the most potent, but the inhibitors are still cleavable by Dbr1p¹⁰. All of the inhibitors described in the report are made using bbRNA synthesis methods.

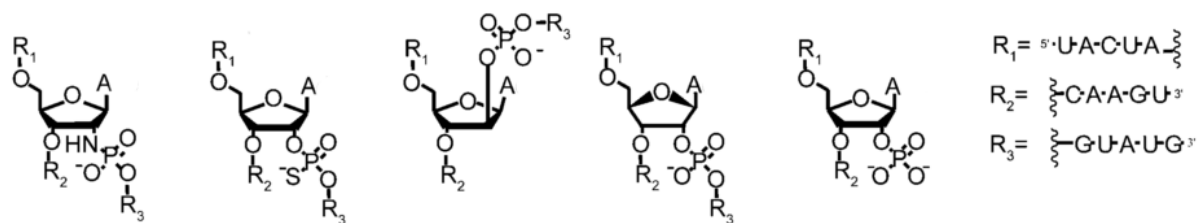


Figure 4.1. Structure of bbRNA made by Katolik et al. to inhibit Dbr1p. Figure adapted from ref.¹⁰.

We sought to make inhibitors of Dbr1p that were not cleavable by the enzyme and synthetically more readily accessible. Described in this chapter is the design and synthesis of branched RNAs that are made through the copper catalyzed azide alkyne cycloaddition (CuAAC or click) reaction. The resulting triazole linkage that connects the 2'-arm sequence is non-hydrolyzable. Four analogues are made and their inhibition kinetics are investigated.

4.2. Synthesis of click branched RNAs

The analogues described in this chapter are designed to be competitive inhibitors of Dbr1p. The overall structure is similar to bbRNA, but instead of containing a 2'-5'-phosphodiester bond, the inhibitors contain a 2'-5'-triazole linkage that is made through CuAAC or click chemistry (Fig. 4.2). These 'click-branched' RNA (cbRNA) are first synthesized as linear oligonucleotides and then conjugated via the click reaction in solution.

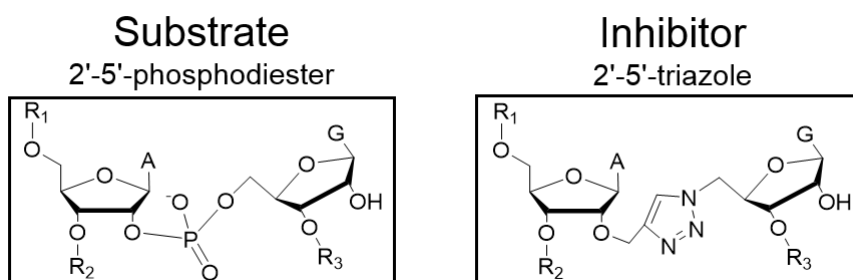
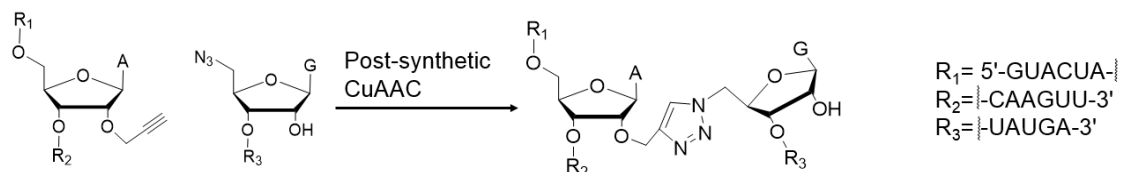


Figure 4.2. Difference between the natural bbRNA substrate and the click branched analogue.

The first linear oligonucleotide for the cbRNA contains the 5'-arm, the branch point nucleotide, and the 3'-arm. In making bbRNA through the methods described in earlier chapters, the branch point residue has the 2'-photolabile protecting group. For cbRNA, the branch point residue has a 2'-O-propargyl group. The 2'-O-propargyl phosphoramidite monomer is commercially available and is installed in the RNA strand using the same coupling conditions as regular RNA. The second oligonucleotide is the sequence corresponding to the 2'-arm sequence. The 2'-arm sequence is also made using standard RNA synthesis conditions with a 5'-azide group installed to replace the terminal 5'-OH¹².



Scheme 4.1. Conjugation scheme for click branched RNA. The stem and the branch are 'clicked' together using CuAAC to generate a 2'-5'-triazole linkage that connects the branch sequence.

The two linear sequences synthesized are used in a CuAAC reaction to generate the cbRNA. The RNA with the 2'-O-propargyl group is used as the limiting reagent, and the azide modified 2'-arm is used in 1.5x excess. Copper sulfate and the copper binding ligand tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) are added to the reaction along with a reducing agent, sodium ascorbate. The details of the CuAAC reaction are in the experimental section. An alternate method to make cbRNA utilizes pseudo-ligandless click chemistry, with ACN as a co-solvent^{13,14}. After three hours, the reaction is quenched using a 'stop' solution of 9: 1 formamide: 0.1 M EDTA and loaded onto a 20% denaturing polyacrylamide gel for purification. The branched RNA migrates slower than either of its constituent pieces (fig. 4.3). Following purification, the RNA is eluted, desalted, lyophilized, and resuspended in water before reactions with Dbr1p. Four different cbRNA were synthesized and are detailed in the table below.

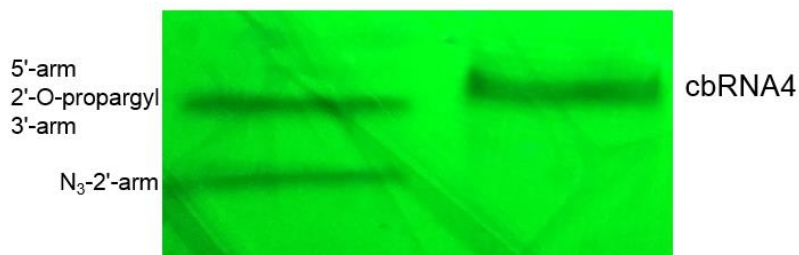


Figure 4.3. Purification gel of reaction to make cbRNA4. The left lane contains the RNA with the 2'-O-propargyl group and the RNA with the 5'-N₃ group. The right lane is the product of the CuAAC reaction. The RNA is run on a 20% denaturing polyacrylamide gel and visualized by UV-shadow.

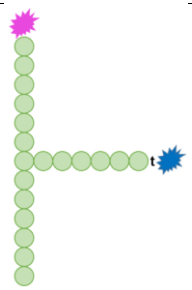
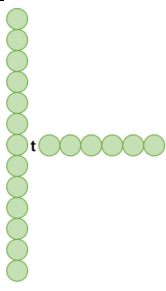
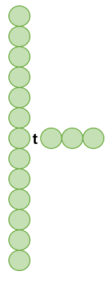
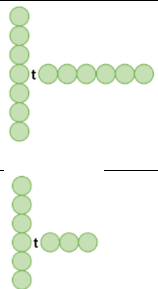

Name	Sequence	Schematic
bbRNA9	5'-Cy3-GUA CUA A(2'-5'-GUA UGA- t -Cy5) CAA GUU-3'	
cbRNA1	5'-GUA CUA A(2'-5'- t -GUA UGA) CAA GUU-3'	
cbRNA2	5'-GUA CUA A(2'-5'- t -GUA) CAA GUU-3'	
cbRNA3	5'-CUA A(2'-5'- t -GUA UGA) CAA-3'	
cbRNA4	5'-CUA A(2'-5'- t -GUA) CAA-3'	

Table 4.1. Click branched RNA. **t** represents the triazole linkage.

4.3. Inhibition of Dbr1p with cbRNA

The assays with cbRNA to inhibit the Dbr1p reaction are performed similarly to the initial rate assay described in chapter 3. Briefly, a bbRNA that is labeled with a pair of FRET-capable dyes (Cy3 and Cy5) is used to track the debranching reaction over time by monitoring the increase in Cy3 signal as the Cy5-containing piece of the RNA is cleaved by Dbr1p. This kinetic assay can be modified to study the effect of an inhibitor on debranching rate by including the cbRNA with the bbRNA substrate before the RNA is mixed with the enzyme in the cuvette.

The debranching action is evaluated at five different bbRNA concentrations (10, 30, 50, 70, 100 nM). At each bbRNA concentration, four different inhibitor concentrations are investigated (0, 10, 30, 50 nM) for a change in rate with enzyme concentration kept constant. A total of 20 different bbRNA/ cbRNA reactions are evaluated and each is run in triplicate. The experimental set up is the same as described in chapter 3, with bbRNA and cbRNA loaded into one line of the rapid mixing device and EhDbr1p loaded into the other line, both in debranching buffer. The fluorescence monitoring is started and fresh solution is pushed through the lines of the rapid mixing device into the cuvette which is located inside the fluorimeter. The rate of fluorescence increase over time is converted into a reaction velocity as before, and from this data Lineweaver-Burk plots are generated. A Lineweaver-Burk plot is a way to represent enzyme kinetics data that derives from linearizing the Michaelis-Menten equation. The linearization is accomplished by taking the reciprocal of the Michaelis-Menten equation, resulting in equation 4.1.

Equation 4.1. Lineweaver-Burk equation

$$\frac{1}{v_o} = \frac{K_M}{V_{max}} * \frac{1}{[S]} + \frac{1}{V_{max}}$$

Equation 4.1 has the form of $y = mx + b$ and the Michaelis-Menten parameters can be extracted from plotting $1/v_o$ v $1/[S]$. A linear regression line is fit to the experimental

data. Lineweaver-Burk plots are particularly useful for determining the method of inhibition, either competitive, non-competitive, or uncompetitive. When generating inhibition plots to determine method of inhibition, the initial rate for each concentration of inhibitor is fit to a linear equation. For a competitive inhibitor, as the concentration of inhibitor increases, the slope of the line increases and the y-intercept remains the same. For uncompetitive inhibition, with more inhibitor, the slope of the line stays the same, but the y-intercept increases. Finally, for non-competitive inhibition, both the slope and y-intercept will change as the inhibition amount increase (fig. 4.4).

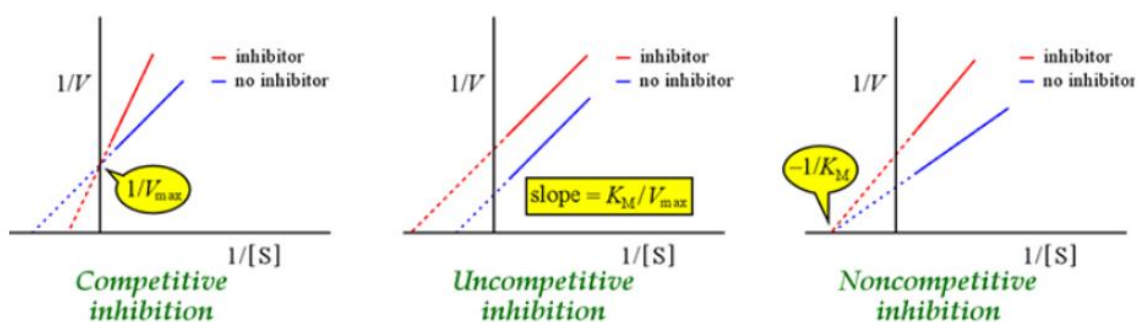


Figure 4.4. Lineweaver-Burk plots of different types of inhibition. Competitive inhibitors bind to the active site of the enzyme. Uncompetitive inhibitors bind to the enzyme substrate complex. Noncompetitive inhibitors bind equally well to the enzyme alone or the enzyme substrate complex. Image from ref¹⁵.

For each cbRNA described in table 4.1, the data is analyzed, plotted, and fit to linear lines. As seen in figure 4.5, as the amount of inhibitor tested increases, the slope of the line increases and the y-intercept stays essentially the same, which indicates that cbRNA are competitive inhibitors.

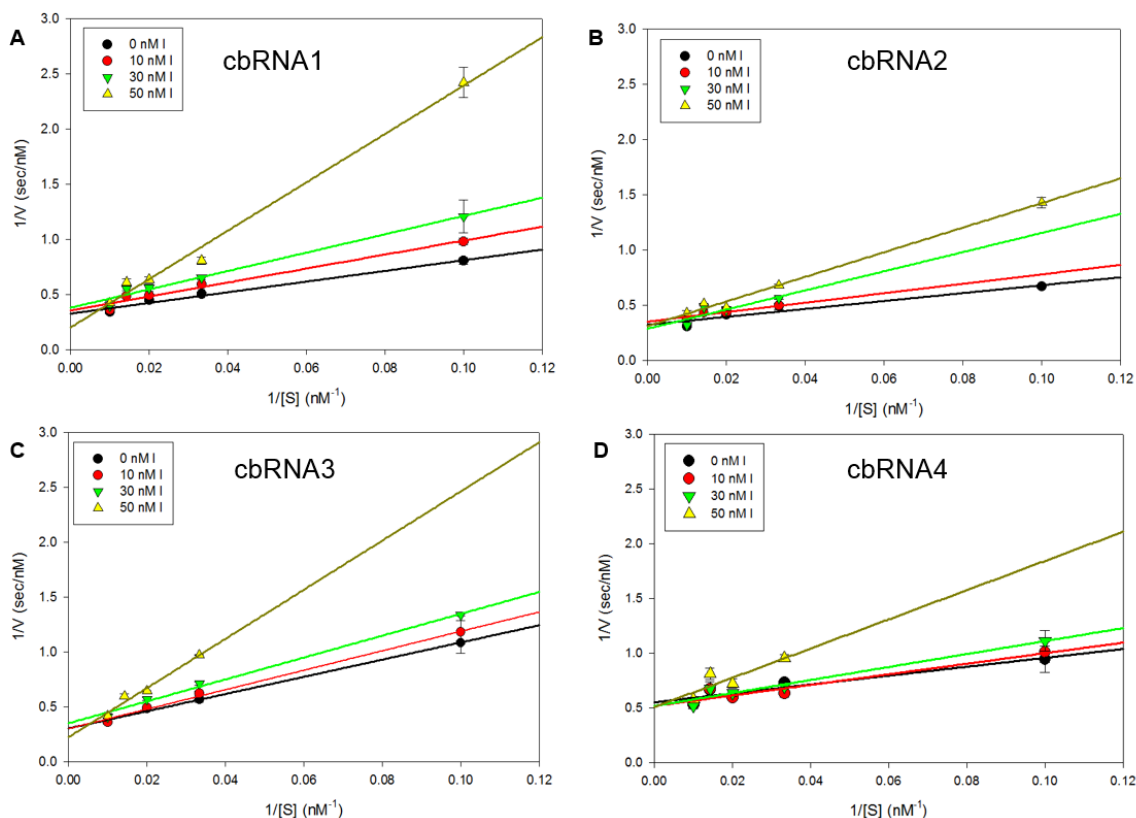


Figure 4.5. Lineweaver-Burk plots of the cbRNAs show competitive inhibition. Plots of the debranching kinetics of EhDBr1p (1 nM) in the presence of varying amounts of inhibitor and substrate. The initial rates were calculated from the initial rate assay and the inverse of the rate was plotted against inverse of concentration. The data was plotted as the mean of a triplicate run. Error bars represent standard deviation from the mean. Each data set was fit to a linear equation ($y = mx + b$). Black line is 0 nM cbRNA, red is 10 nM cbRNA, green is 30 nM cbRNA, dark yellow is 50 nM cbRNA.

In equation 4.2, K_i is the equilibrium constant of inhibition and defines the affinity of the enzyme for the particular inhibitor. The inhibition term is dependent on the concentration of inhibitor, and as this increases, the apparent K_M will increase (K_M^{app}). The apparent K_M term can be used to calculate the K_i term as shown in equation 4.3. Rearranging equation 4.3, equation 4.4 is generated. A plot of K_M^{app}/K_M vs. [Inhibitor] is fit to a linear equation and K_i can be calculated from the slope of the line. These plots for each inhibitor are outlined in figure 4.6.

Equation 4.2. Simplified competitive inhibition

$$v_o = \frac{V_{max}[S]}{[S] + K_M(1 + \frac{[I]}{K_I})}$$

Equation 4.3. Apparent K_M

$$K_M^{app} = K_M(1 + \frac{[I]}{K_I})$$

Equation 4.4. Inhibition equation

$$\frac{K_M^{app}}{K_M} = \frac{1}{K_I} * [I] + 1$$

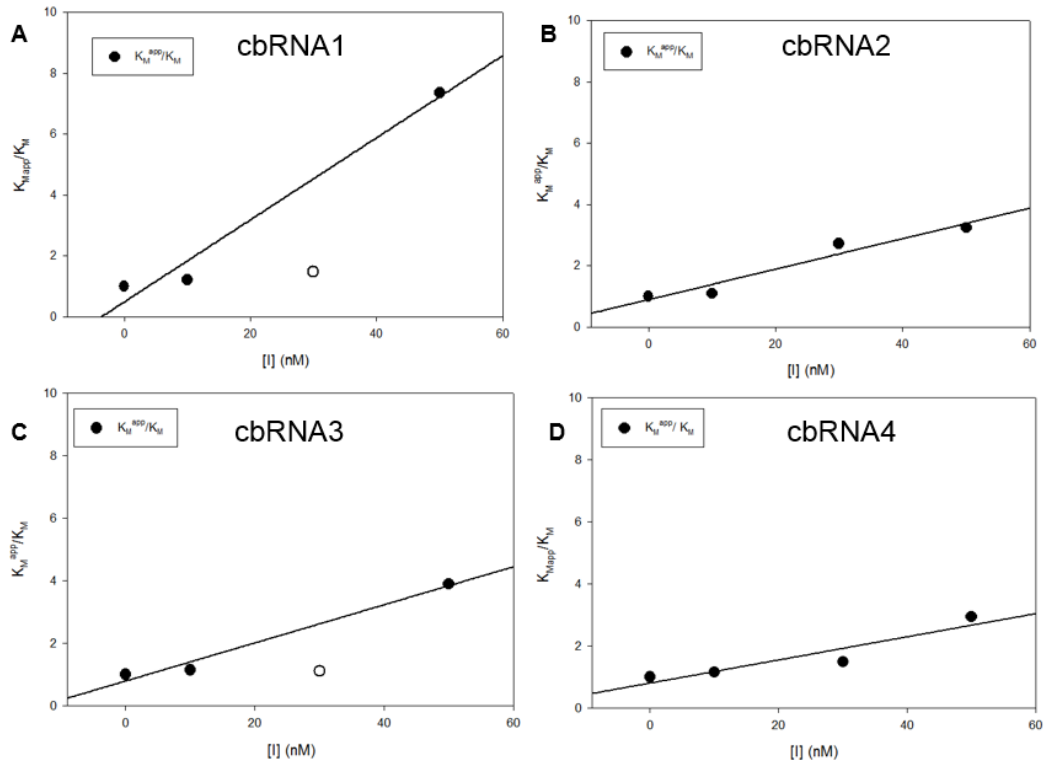


Figure 4.6. Plots to determine inhibition constant for each cbRNA. The apparent K_M , calculated from the fitted line in figure 4.5, is divided by the actual K_M and plotted against concentration of inhibitor. A line is fit to the data, and the inhibition constant is calculated from the fit and equation 4.4. Open circles on plots A and C were not included in the fit calculations.

The strength of the different inhibitors is compared based on the calculated inhibition constants. The relative strength of the inhibitors is detailed in figure 4.7. All of the inhibitors have low nanomolar inhibition constants but the constants vary slightly based on RNA size. The strongest inhibitor, cbRNA1, is the same size as the bbRNA substrate (bbRNA9) and is the largest of all the inhibitors. The weakest inhibitor is cbRNA4, the

smallest molecule. The next strongest inhibitor, cbRNA3, is a 13-mer RNA that contains the same size 2'-arm as cbRNA1. The third best inhibitor, cbRNA2, is a 16-mer. The fact that cbRNA3 is more potent than cbRNA2 suggests that a longer branch makes for stronger inhibitor, regardless of overall electrostatic interactions. The solved co-crystal structure of Dbr1p with bbRNA shows that the enzymes makes many contacts with the branch of the RNA and much fewer with the other arms¹⁶. Our observations of inhibitor strength correlate well with this data from the crystal structure.

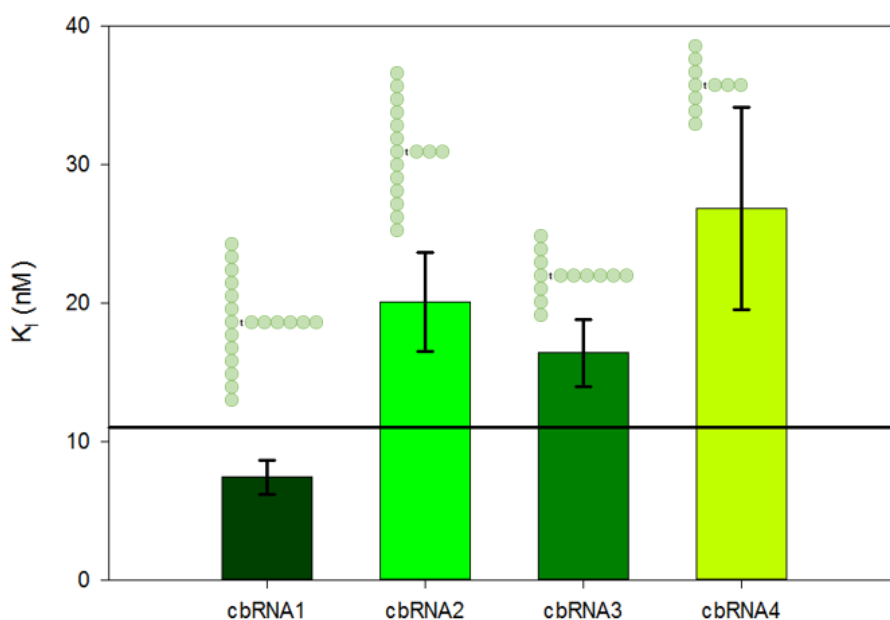


Figure 4.7. Bar chart of inhibition by cbRNAs. The horizontal reference line is the K_M of bbRNA9. The K_i comes from the slope of the fit line in figure 4.7 and the error bar represents the error in the generated fit.

4.4. Conclusions

In this chapter, click-branched RNA were designed and synthesized as inhibitors of Dbr1p and their mode of inhibition was established. The competitive inhibitors described here are easy to synthesize compared to other Dbr1p inhibitors described in the literature^{10,11}. The click branched RNAs are made by standard RNA synthesis and post synthetic conjugation using click chemistry. Other Dbr1p inhibitors are made through bbRNA synthesis and include modifications to the 2'-5'-phosphodiester bond. However the phosphorothioate and phosphoramidate linkages are still able to be hydrolyzed by Dbr1p, just much more slowly than the natural 2'-5'-phosphodiester bond¹⁰. The 2'-5'-triazole linkage described here is uncleavable by Dbr1p.

The size of the cbRNA determines the potency of Dbr1p inhibition. A 19-mer RNA is a stronger inhibitor than a 10-mer. However, the number of residues in the RNA arms is also important to consider. A longer 2'-arm makes for a more potent inhibitor presumably because it makes the most contacts with the enzyme. Click branched RNA are potent inhibitors of Dbr1p and the inhibition constant can be tuned based on the size of the RNA.

4.5. Experimental

4.5.1. Synthesis of cbRNA

Two pieces of RNA are synthesized independently to form one click branched RNA. The first piece, the stem, contains an internal 2'-O-propargyl group installed at a specific residue. A 2'-O-propargyl adenosine is used for these studies and is incorporated into the RNA using commercially available reagents (Chemgenes). The second piece, the branch, contains a 5'-azide. The azide transformation is accomplished using post synthetic modification on the solid support¹².

The 5'-OH of the branch is modified to a 5'-iodo. The CPG beads are treated with a 0.5 M methyltriphenoxyphosphonium iodide solution for 15 minutes inside the synthesizer. The CPG beads are then removed from the synthesizer and incubated with a saturated sodium azide solution in DMF for 90 minutes at 60 °C. The beads are then washed with 10 mL of DMF and 10 mL of ACN and dried on house vacuum before standard deprotection.

The stem strand contains the 5'-arm, 3'-arm, and branch point 2'-O-propargyl residue. The branch strand contains the 2'-arm. The stem and the branch strands are conjugated together using a CuAAC reaction. The reaction scale is 10 nmoles. The 2'-O-propargyl containing stem and the 5'-azide branch are mixed together with 1x PBS buffer, and water in one tube. A copper sulfate and THPTA mixture is made in a separate tube. Finally sodium ascorbate is kept in a third tube. Argon gas is blown on top of each solution for 3x 30 seconds. The required amount of sodium ascorbate is added to the RNA mixture, and is degassed for another 30 seconds. The copper sulfate and THPTA solution is added to the RNA reaction mixture and is degassed a fifth time for 30 seconds. The copper sulfate addition starts the reaction. The reaction is continued for three hours. The reaction is quenched with 9: 1 formamide: 0.1 M EDTA.

The reaction mixture was separated on a 20% denaturing polyacrylamide gel, run for 2 hours at 10 W in 0.5x TBE buffer. The cbRNA band is cut out and passively eluted for 24 hours using TE_{0.1} buffer. The elution solution is desalted using SepPak Classic C18 columns, lyophilized, and resuspended in water.

Reagent	Relative Amount
2'-O-propargyl RNA	1x
5'-N ₃ RNA	1.5x
CuSO ₄	10x
THPTA	50x
Sodium Ascorbate	50x
PBS Buffer	1x

Table 4.2. Relative amount of reagents used in the CuAAC reaction.

4.5.2. Inhibition Reaction

The inhibition reaction was performed inside of a fluorimeter using a rapid mixing device. Dual labeled bbRNA and inhibitor cbRNA are dissolved in 1x debranching buffer and mixed into one line of the rapid mixing device. The second line of the rapid mixing device is filled with EhDbr1p at 1 nM final concentration in debranching buffer. The debranching buffer contains 50 mM Tris.HCl pH 7.5, 25 mM NaCl, 2.5 mM DTT, 0.005% Triton X-100, and 1 mM MnCl₂.

The fluorimeter is set to a time based acquisition, monitoring change in fluorescence over time. The change in Cy3 signal is monitored. The laser is set to excitation of 540 nm and emission of 570 nm. The excitation and emission slit width is 4 nm. The data is collected at 0.05 second intervals.

The data acquisition is started and fresh solution is plunged into the cuvette. The increase in Cy3 signal is monitored. The change in fluorescence over time is then

converted into change in concentration over time, giving an initial rate of enzyme activity. Calculations are performed as described in the results section.

4.6. References

1. Ooi, S. L., Samarsky, D. a, Fournier, M. J. & Boeke, J. D. Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *RNA* **4**, 1096–1110 (1998).
2. Petfalski, E., Dandekar, T., Henry, Y. & Tollervey, D. Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.* **18**, 1181–9 (1998).
3. Ruby, J. G., Jan, C. H., Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nat. Rev. Mol. Cell Biol.* **448**, 83–86 (2007).
4. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in *Drosophila*. *Cell* **130**, 89–100 (2007).
5. Nam, K., Lee, G., Trambly, J., Devine, S. E. & Boeke, J. D. Severe growth defect in a *Schizosaccharomyces pombe* mutant defective in intron lariat degradation. *Mol. Cell. Biol.* **17**, 809–818 (1997).
6. Armakola, M. *et al.* Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nat. Genet.* **44**, 1302–1309 (2012).
7. Ye, Y., De Leon, J., Yokoyama, N., Naidu, Y. & Camerini, D. DBR1 siRNA inhibition of HIV-1 replication. *Retrovirology* **2**, 63 (2005).
8. Galvis, A. E., Fisher, H. E., Fan, H. & Camerini, D. Conformational Changes in the 5' End of the HIV-1 Genome Dependent on the Debranching Enzyme DBR1 during Early Stages of Infection. *J. Virol.* **91**, 1–11 (2017).

9. Galvis, A. E., Fisher, H. E., Nitta, T., Fan, H. & Camerini, D. Impairment of HIV-1 cDNA Synthesis by DBR1 Knockdown. *J. Virol.* **88**, 7054–7069 (2014).
10. Katolik, A. *et al.* Fluorescent Branched RNAs for High-Throughput Analysis of Dbr1 Enzyme Kinetics and Inhibition. *ACS Chem. Biol.* **12**, 622–627 (2017).
11. Tago, N. *et al.* Design, Synthesis, and Properties of Phosphoramidate 2',5'-Linked Branched RNA: Toward the Rational Design of Inhibitors of the RNA Lariat Debranching Enzyme. *J. Org. Chem.* **80**, 10108–10118 (2015).
12. Miller, G. P. & Kool, E. T. Versatile 5'-Functionalization of Oligonucleotides on Solid Support: Amines, Azides, Thiols, and Thioethers via Phosphorus Chemistry. (2004). doi:10.1021/JO035765E
13. Mack, S., Fouz, M. F., Dey, S. K. & Das, S. R. in *Current Protocols in Chemical Biology* 83–95 (John Wiley & Sons, Inc., 2016). doi:10.1002/cpch.1
14. Paredes, E. & Das, S. R. Optimization of acetonitrile co-solvent and copper stoichiometry for pseudo-ligandless click chemistry with nucleic acids. *Bioorg. Med. Chem. Lett.* **22**, 5313–5316 (2012).
15. Michaelis-Menten and Lineweaver-Burk Plots. (2013). at <<https://biochemaddict21.wordpress.com/2013/04/14/michaelis-menten-and-lineweaver-burk-plots/>>
16. Clark, N. E. *et al.* Metal dependence and branched RNA cocrystal structures of the RNA lariat debranching enzyme Dbr1. *Proc. Natl. Acad. Sci.* **113**, 14727–14732 (2016).

Chapter 5 – Mini Lariat RNAs for single-stranded RNA interference

5.1 Introduction

RNA interference (RNAi) is a biochemical pathway, innate in eukaryotes, that can protect the cells from invading RNA as well as regulate the expression of endogenous mRNAs^{1–4}. In the RNAi pathway, a primary microRNA (pri-miRNA) is transcribed from a specific microRNA gene and is then processed by Drosha into a pre-microRNA (pre-miRNA) that has a hairpin structure⁵. Following export of the pre-miRNA from the nucleus, the enzyme Dicer acts on the hairpin to remove the connecting loop and produce a RNA duplex⁵. One strand of this duplex, the guide strand, is loaded into the RNA induced silencing complex (RISC) that can effect either mRNA cleavage, repression of translation, or deadenylation of a mRNA transcript⁵. All three can repress protein synthesis⁵. The complementary strand of the RNA duplex, the passenger strand, is usually degraded, though there are reports of the passenger strand having a regulatory effect^{6–9}.

Besides originating from transcribed hairpins, miRNAs may also be generated from excised lariat introns^{10,11}. The 2'-5' phosphodiester linkage in excised lariat introns is cleaved by lariat debranching enzyme and the RNA folds into a hairpin, as a pre-miRNA, and enters into the traditional RNA interference pathway^{10–12}. These intron-derived miRNAs are termed mirtrons. As these pre-miRNA have a non-canonical origin, Drosha processing is not necessary for the biogenesis of mirtrons¹². Dicer processing is still required because introns vary widely in length and the hairpin formed after debranching needs to be trimmed into the miRNA duplex¹². The mirtron pathway is responsible for generating a large amount of miRNA. As reported in 2015, 478 mirtrons have been computationally identified, compared to 1228 canonical miRNAs in humans¹³.

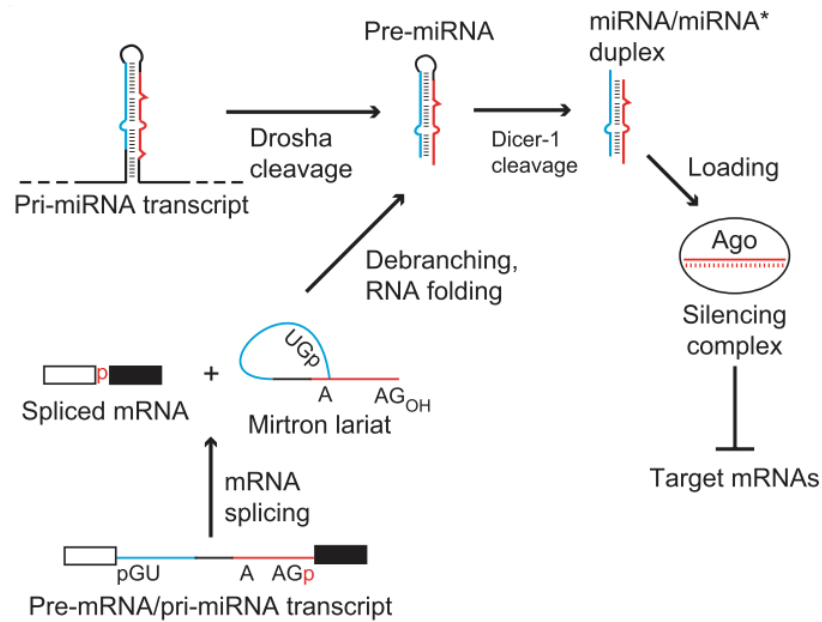


Figure 5.1. Mirtron biogenesis pathway and its convergence with the canonical miRNA pathway. Image from reference¹⁴.

Exogenous synthetic short interfering RNA (siRNA) can also be used to repress specific target mRNA¹. These siRNA are complementary duplexes that may be modified for stability, delivery and other desirable properties for their application. Like miRNA, siRNA targets and represses mRNA through the RNAi pathway¹⁵. These siRNA require two key features, a two-nucleotide 3'-overhang on either end of the duplex and a 5'-phosphate on the guide strand for optimal repression^{7,8,15,16}.

A 5'-phosphate is necessary for proper RNA loading into RISC¹⁷. Once an siRNA duplex is internalized, cellular kinases phosphorylate the 5'-ends of the duplex, removing the need to add a phosphate during synthesis⁶. Other modifications can be made to the siRNA to improve cellular stability. These include a phosphorothioates in the backbone or a 2'-hydroxyl substitution. The substitutions can include 2'-F, 2'-methoxy (OMe), or 2'-methoxyethyl (MOE) groups¹⁷. RNA association with Argonaute2 (Ago2) is not hampered by modifying the backbone or ribose substituents^{17,18}.

Single-stranded siRNA can also be effective in mRNA silencing and is advantageous because it lacks a passenger strand^{19–21}. Removal of the passenger strand cuts the siRNA production cost and eliminates the possibility of RNAi effects from the passenger strand^{33–35}. However, the use of single-stranded siRNA has its own set of challenges. Single-stranded RNAs are readily degraded by nucleases, so extensive chemical modifications are required for activity²⁰. A 5'-phosphate is required for proper loading into Ago2, but the kinase that phosphorylates duplex siRNA, Clp1, has impaired function with modified RNAs²⁰. A 5'-vinyl phosphonate modification has been used as a non-hydrolyzable phosphate mimic^{20,22}.

Here we design mini-lariat RNAs inspired by mirtron biogenesis, that combines the lariat structure with the utility of single-stranded RNAi. Mini-lariats are a single-strand of RNA the size of a mature miRNA guide strand (19-21 bases), have the looped shape of a lariat mirtron, and contain a 2'-5'-phosphodiester bond. These mini-lariats would be expected to be Drosha and Dicer independent. The mini-lariats would only require cleavage by Dbr1p before entering into RISC. Once cleaved by Dbr1p, the 5'-end of the RNA would have a terminal phosphate. The mini-lariat design reduces the cellular processing steps and the amount of RNA introduced into the cell. Mini-lariats are protected from 5'-exonucleases until acted upon by Dbr1p. Figure 5.2 shows how mini-lariats might enter into the RNAi pathway.

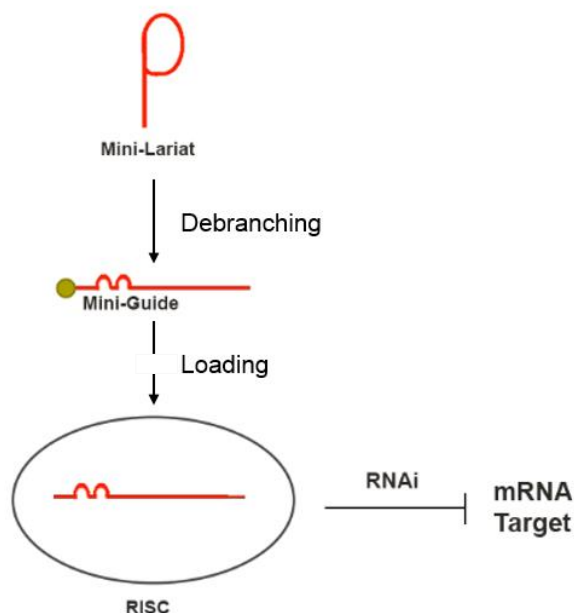


Figure 5.2. Overview of proposed mini-lariat action pathway post transfection. The phosphate of the 2'-5'-phosphodiester is in dark yellow.

Mini-lariats can be made by solid phase synthesis methods. The 'lariatization' or transformation from a linear RNA strand to the lariat structure is performed in the synthesizer. During standard 3'-to-5'-synthesis of RNA, a residue with a 2'-photolabile protecting group is coupled at a desired position. After completing the standard (linear) synthesis, the lariatization is accomplished in a three-step process while the oligonucleotide is still on solid support. First, the terminal 5'-hydroxyl is transformed into a reactive phosphoramidite for coupling. Second, the photolabile group is selectively removed to generate an internal 2'-hydroxyl. Finally, using standard phosphoramidite activating reagents, lariatization is completed by coupling the terminal 5'-phosphoramidite with the internal 2'-hydroxyl to generate the mini-lariat on solid support. We use standard RNA protocols for both release of the mini-lariat RNA from solid support and removal of all the protecting groups on the lariat oligonucleotide. Once these mini-lariats are generated and purified, they can be tested for RNAi activity.

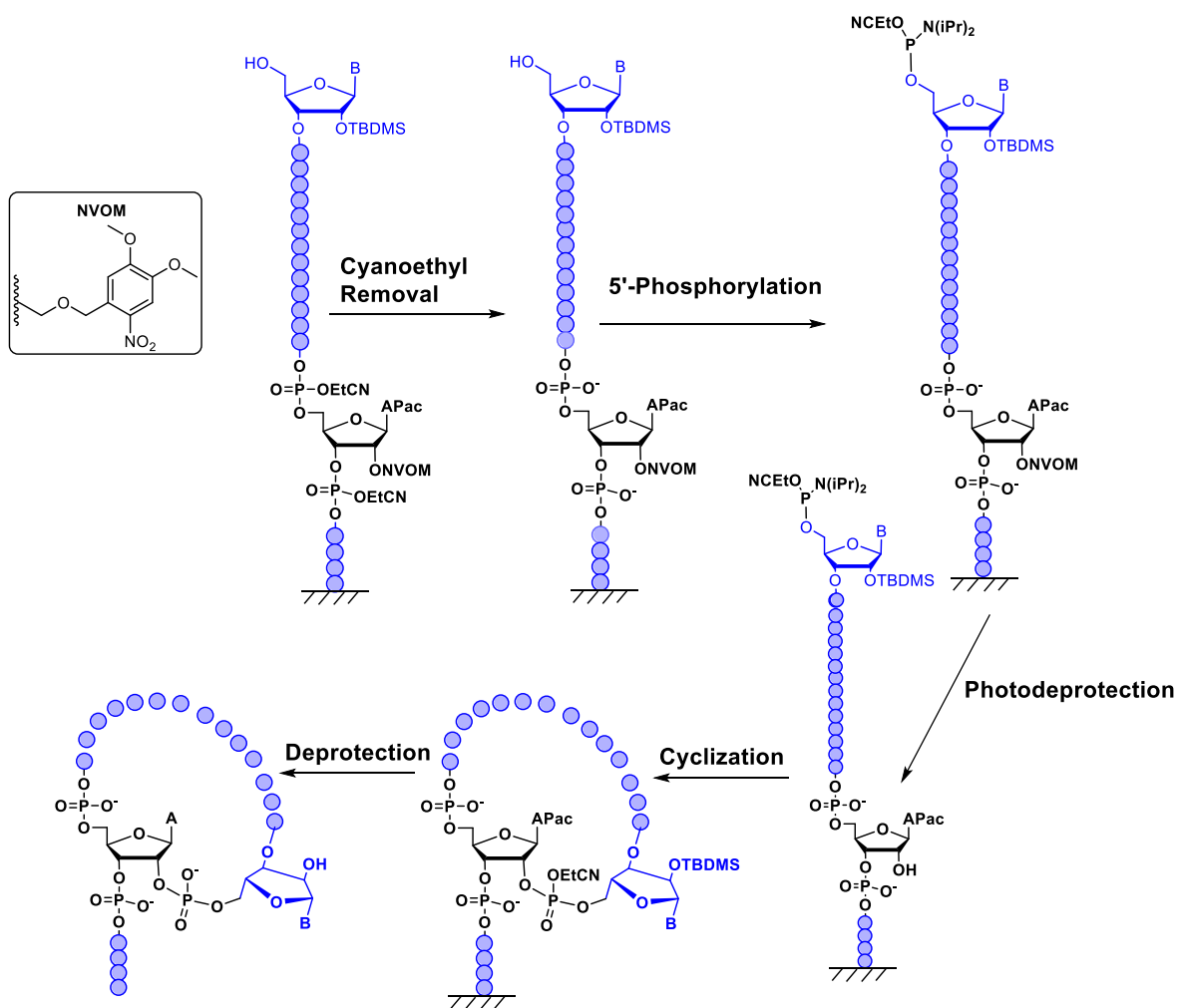
Preliminary methods to synthesize and test mini-lariat RNAs in RNAi were performed by Eduardo Paredes in the Das lab²³. This work showed that the mini-lariats can cause knockdown of a luciferase reporter by RNAi in *Drosophila melanogaster* S2 cells. My aim here is to further develop the method for mini-lariat synthesis and explore its efficacy in RNAi based knockdowns in human cells.

In this chapter, I describe the development and optimization of new protocols for mini-lariat synthesis. Subsequently, assays that demonstrate that mini-lariats work similarly to double stranded siRNA, are debranching dependent, and reduce both the endogenous protein and mRNA levels of the target, in a dose dependent manner.

5.2 Synthesis and Purification of Mini-Lariats

5.2.1. General Procedure

The mini-lariats are synthesized on the solid phase using a MerMade4 RNA synthesizer (Bioautomation). The procedure is customizable and can be performed on any RNA sequence desired as shown in scheme 5.1. However, Dbr1p has a preference for cleaving the 2'-5'-phosphodiester bond when an adenosine residue is at the branch point²⁴.



Scheme 5.1. Mini-Lariat RNA Synthesis. Conditions are detailed in section 5.5.1.

First, the RNA sequence is synthesized normally in the 3'-to-5' direction, with a 2'-O-NVOM protected residue incorporated at a desired position. The cyanoethyl (CE) protecting groups on the phosphate backbone are selectively removed using a 3:2 mixture of acetonitrile (ACN) and triethylamine (TEA) for 90 minutes at room temperature²⁵. The purpose of this step is to change the phosphotriesters into less reactive phosphodiester which prevents chain migration or cleavage at the branch point residue due to the 2'-hydroxyl that is free after photodeprotection. Second, the 5'-hydroxyl is coupled with N,N-diisopropylamino cyanoethylphosphoramidic chloride (PCI) to create a 5'-phosphoramidite. A vial on the synthesizer is loaded with the PCI reagent in dichloromethane (DCM) and another vial is filled with base (DIPEA) and catalyst (1-methylimidazole). Two consecutive and alternating 45 μ L injections of the prepared solutions, starting with the base causes layering of the solutions and the force of the injections causes mixing. After the final injection, the solution is left to couple for 10 minutes, with a vacuum pull every 2.5 minutes to bring fresh reagent into contact with the beads. After 10 minutes, the column is fully drained and the cycle is repeated, for a total coupling time of 90 minutes.

Third, after the phosphoramidate is installed, the 2'-O-NVOM photolabile group is removed to expose the 2'-OH. The 365 nm UV LED source is turned on and the column is irradiated for 10 minutes. The specifics of set up for in-synthesizer photodeprotection are as discussed in chapter 2. Vacuum is pulled every 2 minutes to drain the released products of the photodeprotection reaction. Then, the ACN is fully drained, the light turned off, and the column is washed with ACN six times to remove any lingering reagents.

The fourth and final step, cyclization, uses the standard 0.25 M 5-(ethylthio)-1H-tetrazole activator (ETT) activator already installed in the synthesizer. For cyclization, 65 μ L of the activator is injected into the column for 10 minutes with vacuum applied every

2.5 minutes. After ten minutes the column is drained, and 65 μ L of fresh ETT is introduced. Total lariatization time is 30 minutes. The newly formed phosphite linkage, from coupling of the 5'-phosphoramidite and the 2'-OH, is oxidized into a phosphate using standard oxidation conditions, 0.02M iodine in a THF/ pyridine/ water mixture. See experimental section 5.5.1 for details on lariatization reagents and conditions.

Step	Reagent	Injection Volume	Coupling Time Per Injection	# Injections	Total Coupling Time
5'-Phosphorylation	0.2 M N,N-diisopropylamino cyanoethylphosphoramidic chloride in DCM	90 μ L	10 min	9	90 min
	0.95 M DIPEA/ 13.25 mM 1-methylimidazole in DCM	90 μ L	10 min	9	90 min
Photodeprotection	365 nm light & ACN	200 μ L	10 min	1	10 min
Cyclization	0.25 M 5-Ethylthio-1H-tetrazole (ETT) in ACN	65 μ L	10 min	3	30 min

Table 5.1. Mini-lariat synthesis conditions for steps 2-4.

Following the lariatization process described above, the generated oligonucleotide sequence is base deprotected and cleaved from the beads using standard conditions: 1:1 30% NH_4OH : methylamine, for 12 minutes at 65 $^{\circ}\text{C}$. The 2'-O-TBDMS groups are subsequently removed using DMSO, TEA, and TEA.3HF for 2 hours at 65 $^{\circ}\text{C}$. The RNA is desalted with a reverse phase cartridge, lyophilized, and resuspended in water.

Post synthetic purification of mini-lariats is crucial as the cyclization is never quantitative. The products of the lariatization process are the desired mini-lariat, 5'-phosphorylated RNA, and 5'-hydroxyl RNA. To separate the three closely related products, reverse phase high performance liquid chromatography is used. The details are

in the experimental section. The buffers used are 0.1 M triethylamine acetic acid (TEAA) buffer as solvent A and 80: 20 ACN: H₂O in 0.1 M TEAA as solvent B. The purification is performed at 40 °C with a flow rate of 1 mL/min. The gradient is 4.8-16% of solvent B in 30 minutes.

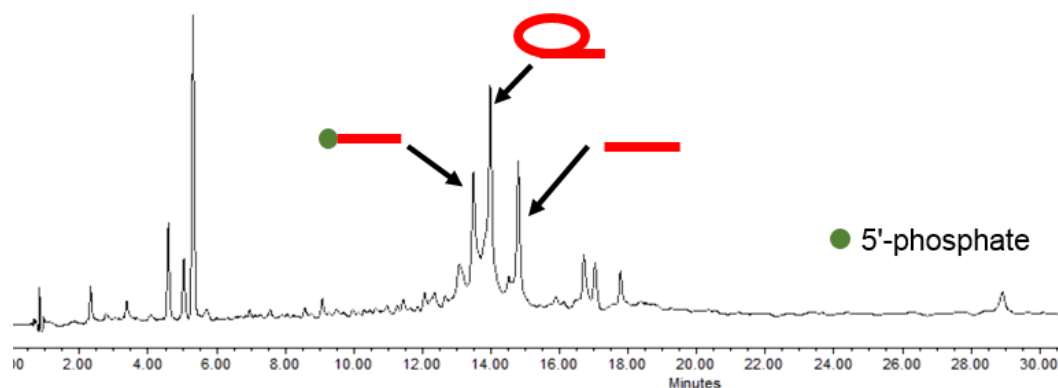


Figure 5.3. Reverse phase HPLC chromatogram of crude mini-lariat reaction. The three major products of the reaction are 5'-phosphorylated, mini-lariat, and 5'-OH that correspond to the indicated peaks.

Eluent corresponding to each peak in the HPLC (fig. 5.3) was collected and analyzed by matrix-assisted laser desorption/ ionization mass spectrometry (MALDI-TOF). The peak at ~5 minutes is from a truncation sequence (5'-A gc – 3') that is probably a cleavage product. Figure 5.4 shows the MALDI mass spectrum of the mini-lariat RNA following purification. From a 1 μ mol synthesis, 20 nmol of HPLC purified mini-lariat RNA was recovered.

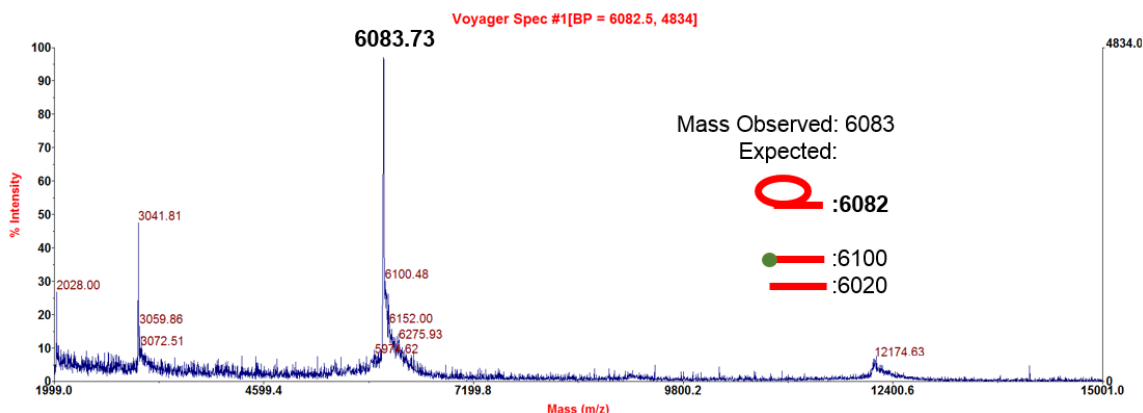
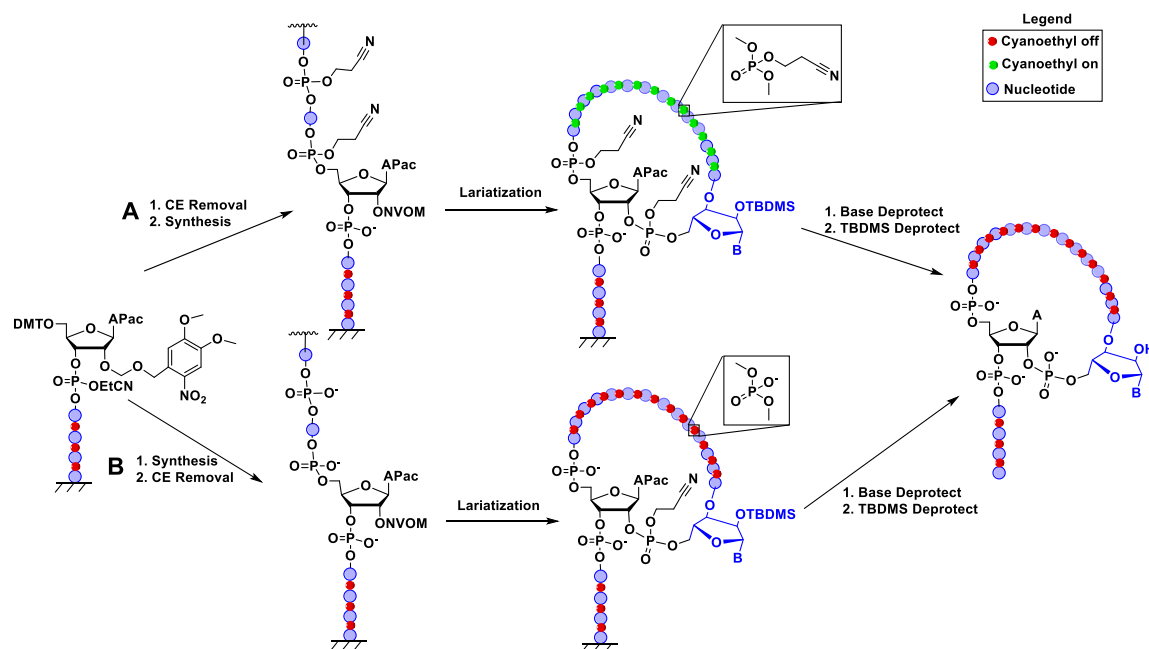


Figure 5.4. MALDI-TOF spectrum of mini-lariat RNA.

5.2.2. Exploring Synthetic Improvements – Charged or Uncharged Backbone

In the synthesis detailed in section 5.2.1 the phosphate backbone is negatively charged because of the complete removal of the CE groups. Alternately, only a few of the CE groups can be removed mid synthesis just after incorporation of the branch point residue. This results in a backbone for lariatization that has both negative charges where the CE groups have been removed, and CE groups in the backbone of residues 5'-to the branching residue. After incorporation of the branch point residue, the synthesis is paused, the column removed from the chamber, and the CE groups removed by incubation with 6:4 ACN: TEA for 90 minutes at room temperature. Following selective deprotection of the CE groups, the column is returned to the synthesizer and the addition of nucleotide residues can be continued. The selective deprotection changes the phosphotriester below the branch point residue into phosphodiester, preventing branch migration, but maintains the phosphotriesters and CE groups 5'-to the branch point. The two synthetic routes are detailed in scheme 5.2.



Scheme 5.2. Two different synthetic routes to mini-lariats. Route A details partial removal of the CE groups and route B is complete removal. Phosphate linkages with cyanoethyl groups off are represented with red circles. Phosphate linkages with cyanoethyl groups on are green circles.

Route A and route B detailed in scheme 5.2 both have potential advantages and disadvantages. In route A, partial removal of CE groups requires interruption of synthesis, but steric repulsion of the CE groups might aid intramolecular cyclization by enhancing bending of the strand. In route B, the removal of the CE groups is done after the entire strand is synthesized. However, revealing the negative charges on the backbone might hinder cyclization because of the repulsion of negative charges. For both routes, the crude reactions were analyzed through HPLC using the same gradient outlined in 5.2.1 (fig. 5.5).

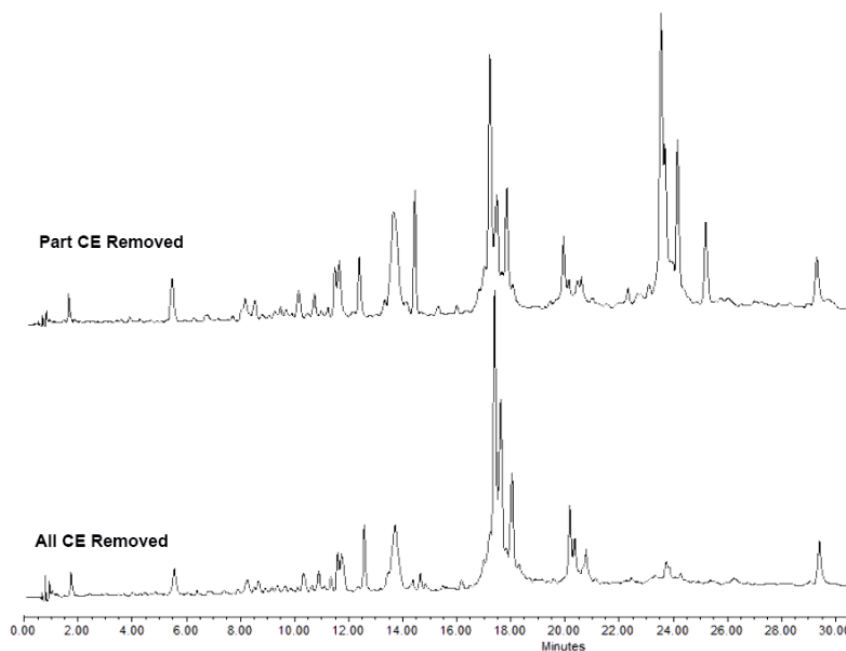


Figure 5.5. Reverse phase HPLC chromatogram of crude lariatization reactions. Route A, partial, or route B, complete, CE removal is compared.

As seen in figure 5.5, partial removal of the CE groups results in more peaks in the HPLC chromatogram. Complete removal of the cyanoethyl groups results in a cleaner synthesis. Based on this, complete removal of CE groups is adopted as the standard procedure for mini-lariat synthesis.

5.2.3. Exploring Synthetic Improvements – Ammonium Salts

The two major side products of the mini-lariat reaction are un-phosphorylated and un-cyclized RNA. To improve the cyclization step, tetraalkylammonium chloride is injected at the same time as the ETT activator. The quaternary ammonium ion is introduced to ion pair with the negatively charged phosphate backbone produced after CE removal. Tetrabutylammonium chloride is dissolved in ACN to a concentration of 0.1 M and 60 μL of the solution was co-injected with 65 μL of the activator. The solution is allowed to react for 30 minutes, the same time as the standard cyclization step. The RNA is deprotected according to standard protocols and the crude reaction is analyzed by HPLC. The

chromatogram shows no improvement compared to the standard synthesis (fig. 5.6). The addition of tetrabutylammonium chloride is determined to be ineffective at improving cyclization.

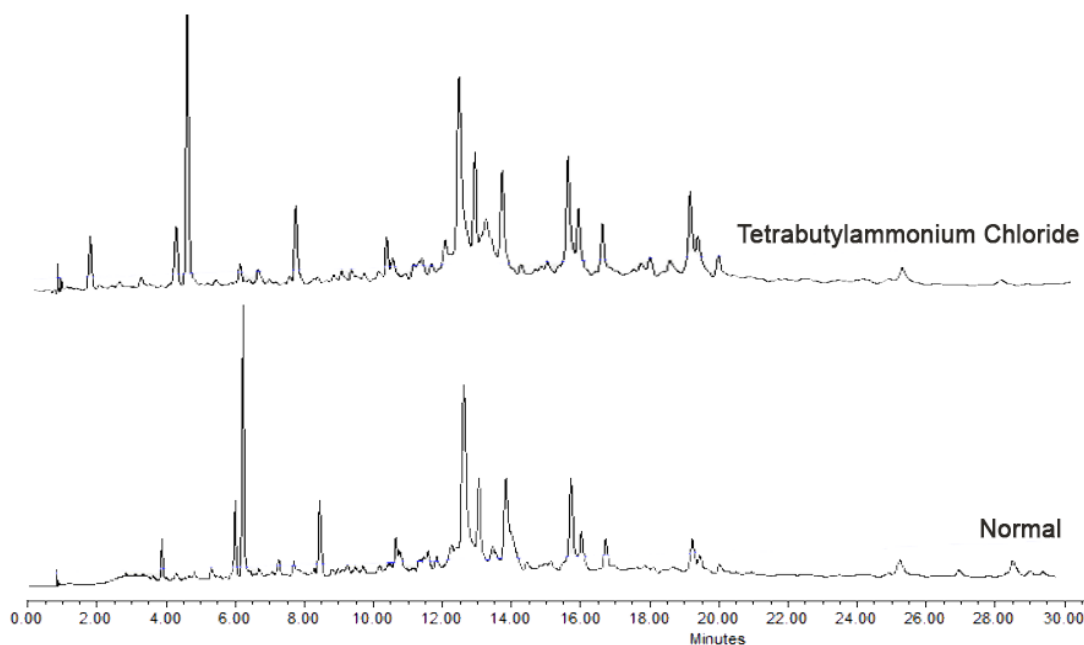


Figure 5.6. Reverse phase HPLC chromatogram of crude mini-lariat reaction with (top) and without (bottom) adding tetrabutylammonium chloride during the cyclization step. The column and gradient details are in section 5.5.2.

5.2.4. Exploring Synthetic Improvements – Cooling Down the Reaction

In an effort to improve the phosphorylation step, the reaction is cooled down to more closely mimic a phosphoramidite synthesis reaction performed in a flask. The phosphoramidite synthesis reaction flask is cooled in an ice bath and is slowly warmed to room temperature. However, when phosphorylating the 5'-terminus of a mini-lariat, the reaction is performed in the DNA synthesizer column at ambient room temperature. To cool down the reaction in the synthesizer, ice baths in beakers are placed under vials that contain phosphorylation reagent. Additionally, to cool the synthesis column, a Büchner flask filled with dry ice and capped with a rubber stopper. A clear tube is attached to the vacuum line, fitted to a drying tube, a valve, and then attached to a tube that has a smaller

inner diameter. The small tube is placed inside the synthesizer like the UV LEDs, through the side port and attached to the aluminum block. The cold carbon dioxide is blown through this tube, cooling the reaction down (fig. 5.7).

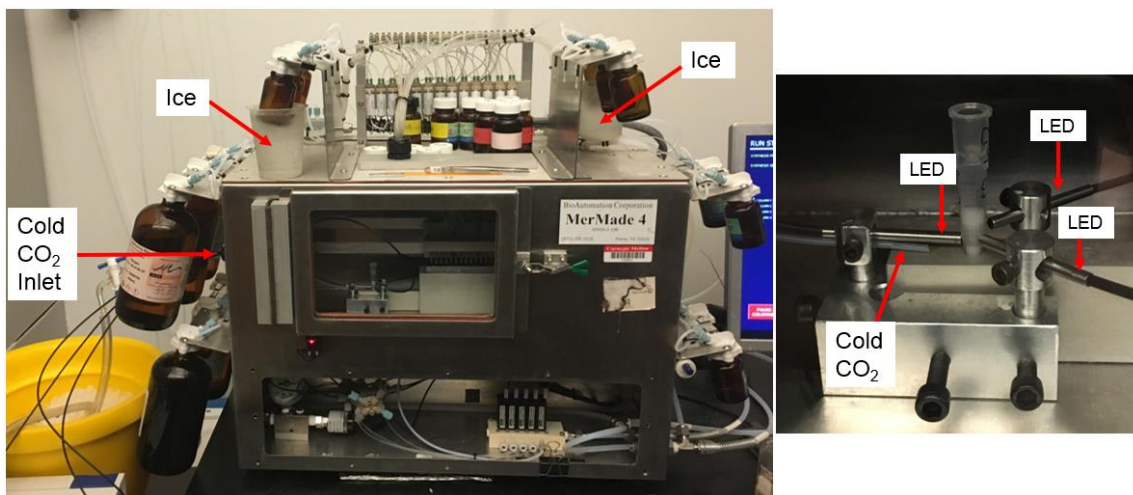


Figure 5.7. Set-up for cooling the mini-lariat reaction. Ice buckets are placed around the vials holding phosphorylation reagent (left). A tube is inserted through the LED hole that is attached to a flask full of dry ice. The cold CO₂ gas outlet is arranged underneath the UV LED to blow onto the CPG beads.

Once the cooling is set up, the phosphorylation step proceeds as outlined in section 5.2.1. The valve is closed once the phosphorylation step is complete and the ice buckets are removed. There was no change to the photodeprotection and cyclization steps. The RNA is deprotected like normal and the crude reaction is run on a reverse phase HPLC and compared to a normal, non-cooled reaction. There is no significant difference between the chromatograms of the two reactions and it is determined that cooling the reaction through the method described above was not effective in increasing the yield of the phosphorylation step.

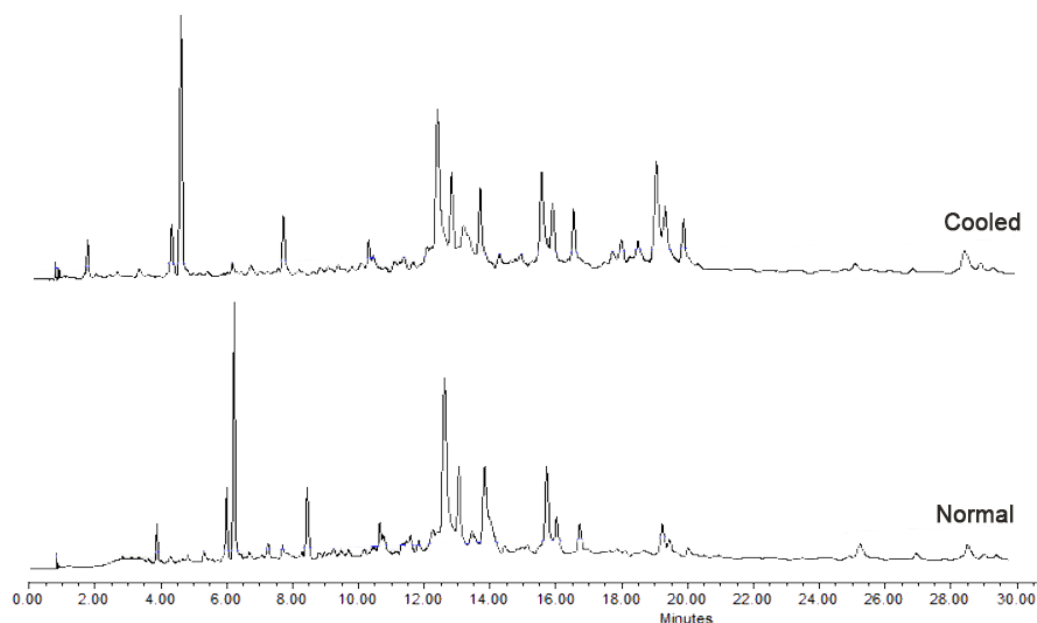


Figure 5.8. Reverse phase HPLC chromatogram of crude mini-lariat reaction when cooled (top) or room temperature (bottom) phosphorylation step. The column and gradient details are in section 5.5.2.

5.2.5. Ribose 2'-OH Modifications

Modifications to the 2'-position in the sugar backbone of siRNA confers stability against nucleases¹⁸. Modifications like 2'-F, 2'-OMe, and 2'-methoxyethyl (MOE) increase the half-life of siRNA, allowing greater gene knockdown to occur^{18,21,22}. Therefore, we made a batch of mini-lariats that contains three 2'-F modifications. These modifications were made using commercially available 2'-F phosphoramidites (Chemgenes) and incorporated using standard RNA synthesis conditions. The mini-lariats that contain the 2'-F modifications are synthesized and purified as described in 5.2.1. The HPLC trace of the crude reaction is shown in figure 5.9.

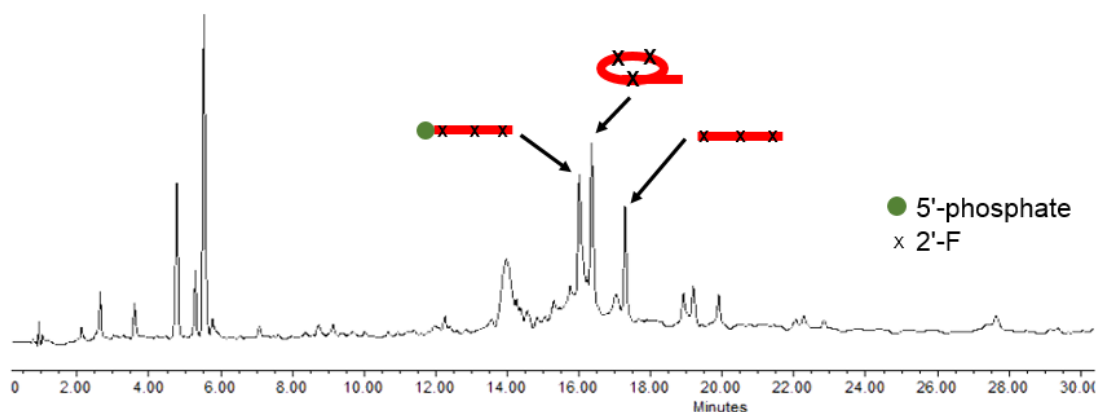


Figure 5.9. Reverse phase HPLC chromatogram of crude lariatization reaction with sequence that has three 2'-F modifications.

The retention time for the 2'-F modified RNA increases, compared to the non-fluorinated RNA but the three main products remain distinguishable. The MALDI trace is shown in figure 5.10. Considering the ease of 2'-F inclusion, suggests that other modifications such as 2'-OMe and 2'-MOE, for which phosphoramidites are commercially available, may be readily included into mini-lariat RNA.

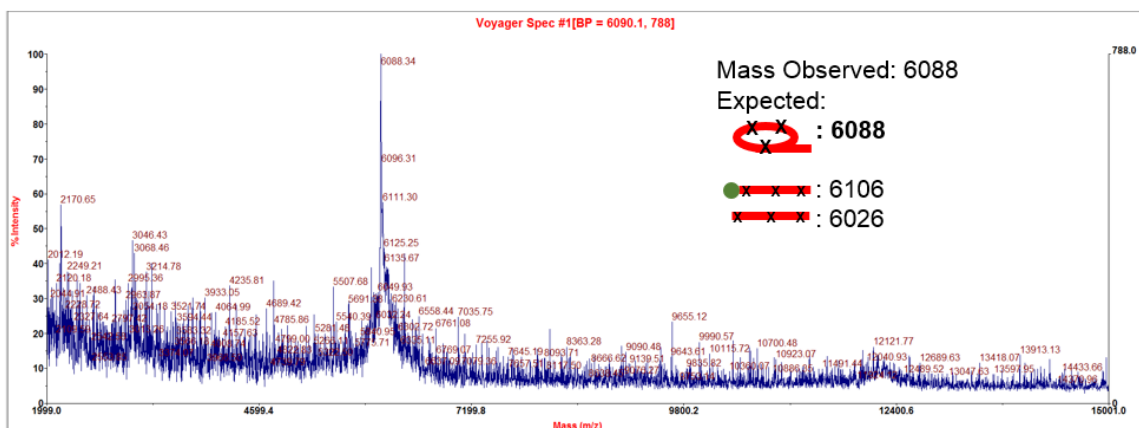
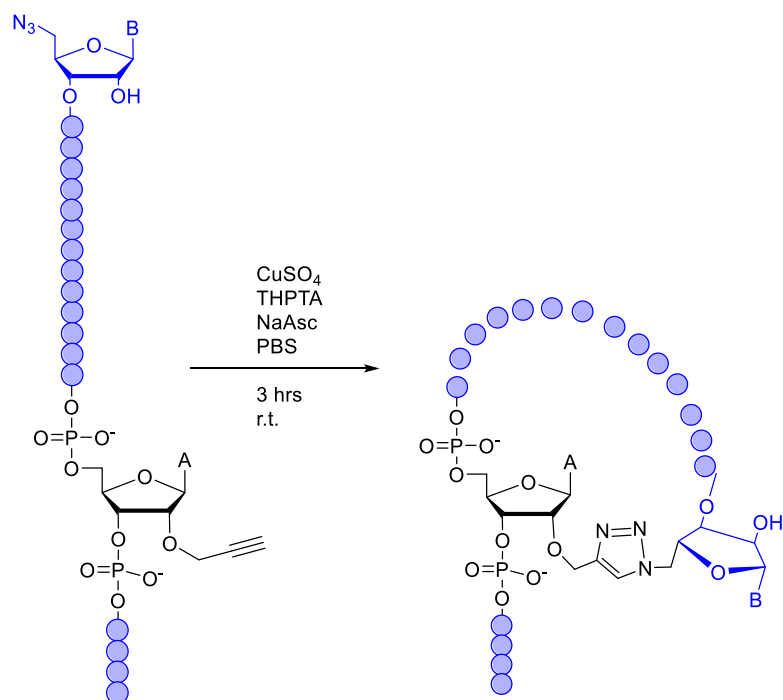


Figure 5.10. MALDI-TOF spectrum of fluorinated mini-lariat.

5.2.6. Synthesis of a Non-Cleavable Mini-Lariat

To show that any knockdown observed by a mini-lariat RNA is dependent on action by lariat debranching enzyme, a non-cleavable mini-lariat is required. The non-cleavable analogue is synthesized using a copper catalyzed azide alkyne cycloaddition (CuAAC or

'click' chemistry), reaction that substitutes the natural phosphate linkage with a triazole linkage, analogous to the inhibitors described in Chapter 4. The RNA is synthesized by incorporating a commercially available residue that has a 2'-O-propargyl at the desired branch point. An azide is installed at the 5'-terminus after strand synthesis, but before deprotection^{26,27}. Following, the RNA is cyclized using click chemistry. The scheme of this cyclization reaction is shown in scheme 5.3 and experimental details are described later in this chapter. Following synthesis, the RNA is purified by reverse phase HPLC.



Scheme 5.3. Cyclization reaction to generate click-linked mini-lariats.

5.3. RNA Interference Assays

The mini-lariat RNA synthesized as described above were tested in human prostate cancer cells by Jey S. Ebron from Dr. Girish Shukla's lab at Cleveland State University. Summarized here are the results of those studies that demonstrate the use and effectiveness of mini-lariat RNAs for RNAi and as a single-stranded miRNA mimic.

5.3.1. Luciferase Assay

The efficacy of mini-lariats is tested in human prostate cancer cells. As controls, a few different constructs re made and detailed in table 5.2.

Name	Sequence (5'-3')	Symbol
ss-644 (guide)	AGU GUG GCU UUC UUA GA gc	
ss-3F-644	AG <u>U</u> GUG GCU <u>UUC</u> <u>UUA</u> GA gc	
Passenger	UC UAA GAA AGC CAU AAU UU	
ss-5'p-644	p-AGU GUG GCU UUC UUA GA gc	
ss-5'p-3F-644	p-AG <u>U</u> GUG GCU <u>UUC</u> <u>UUA</u> GA gc	
C-ML-644	...AGU GUG GCU UUC UUA(2'-5'-t...) GA gc	
C-ML-644- linear	N ₃ -AGU GUG GCU UUC UUA(2'-O-propargyl) GA gc	
ML-644	...AGU GUG GCU UUC UUA(2'-5'-...) GA gc	
ML-3F-644	...AG <u>U</u> GUG GCU <u>UUC</u> <u>UUA</u> (2'-5'-...) GA gc	

Table 5.2. Sequences used in the luciferase assay. Legend for the sequences: uppercase is RNA, lowercase is DNA, red letter is 2'-F, t is a triazole linkage and a p at the 5' designates a 5'-phosphate. Legend for symbols: red is the guide strand, blue is passenger strand, green dot is 5'-phosphate, t is triazole, X is 2'-F

The miRNA sequence used in this study is that of miRNA644a. This miRNA is known to regulate the expression of androgen receptor (AR) mRNA²⁸. Androgen receptor has been shown to help cancerous cells survive, and thus inhibiting AR expression is therapeutically important²⁹. First a luciferase assay is performed in PC-3 cells to test that the mini-lariats are active RNAi agents in mammalian cells. PC-3 cells are human prostate cancer cells but lack endogenous AR. Therefore, the amount of knockdown seen with the introduced luciferase-AR construct can be accurately quantitated. Each of the RNAs listed

in table 5.2 was tested in PC-3 cells. The RNAs were transfected into the cells at 5 nM concentration using Lipofectamine3000 as a transfecting agent. After 48 hours the cells were evaluated for luciferase levels. The results of this experiment are shown in figure 5.11.

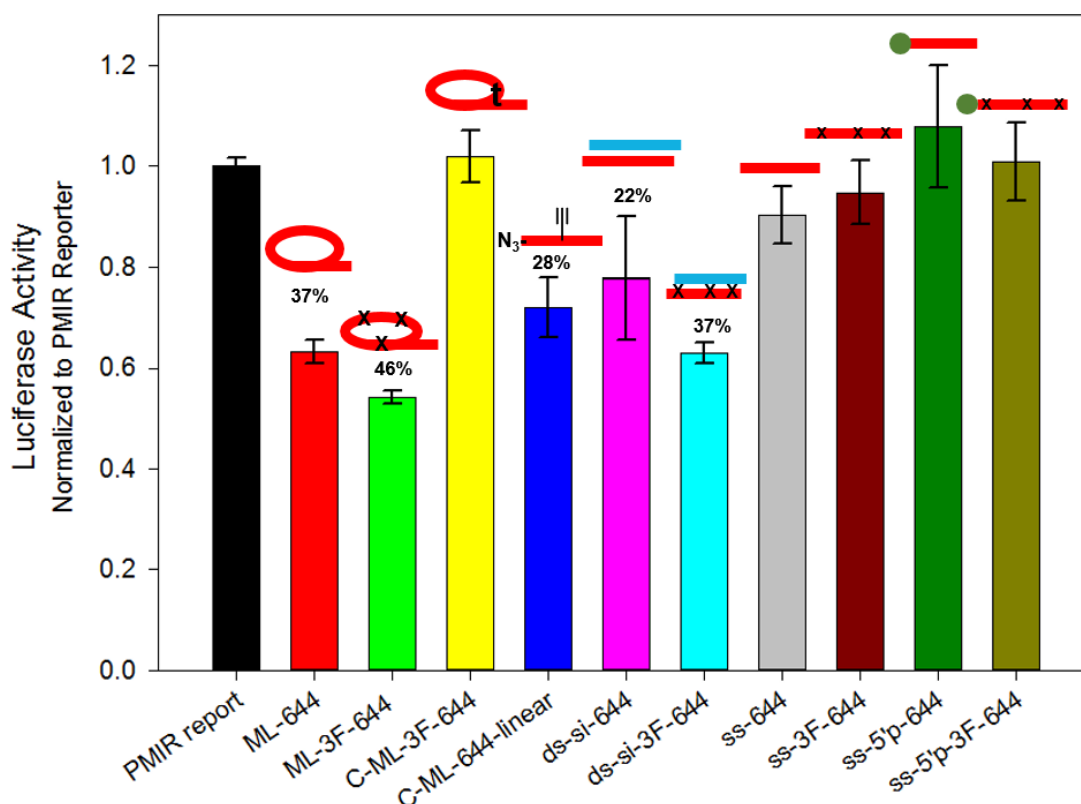


Figure 5.11. Plot of luciferase activity 24 hours after transfection with RNA constructs into PC-3 cells. 5 nM RNA is introduced and the luciferase activity is normalized to the PMIR reporter. Mean is plotted and the error bars represent standard deviation of three experiments. (Data from J.S. Ebron).

As seen in figure 5.11, standard duplex siRNA (ds-si-644) used as a control results in 22% knockdown of AR in 24 hrs, with the fluorinated RNA (ds-si-3F-644) inducing greater knockdown, likely due to its greater stability. Knockdowns with single-stranded (ss-644, ss-3F-644) and 5'-p single-stranded RNA (ss-5'p-644, ss-5'p-3F-644) are not appreciably different compared to the control duplex siRNA. The click linked lariat (C-ML-644) also does not show any knockdown. In contrast, the uncyclized version of the click linked lariat, with a 5'-azide and internal 2'-O-propargyl (C-ML-644-linear), shows

knockdown indicating that the mini-lariat will not knockdown mRNA expression unless debranched. The mini-lariat RNAs, both without and with 2'-F (ML-644, ML-3F-644) cause knockdown on the order of their double stranded counter parts.

5.3.2. Repression of Endogenous AR mRNA

Promising results from the reporter assay merited further investigation of the mini-lariats. LNCaP cells are used to investigate endogenous protein and mRNA levels post treatment with RNAi agents. LNCaP cells are prostate cancer cells that naturally express androgen receptor, unlike the PC-3 cells used in the luciferase assay. The cells are treated with 50 nM of active RNA. Meaning, the duplexes are transfected at 100 nM, since only half of the RNA is active. The constructs are introduced into the cells with Lipofectamine3000. After 24 hours, the RNA is extracted and probed for AR mRNA. The expression is normalized to 18S ribosomal RNA and is shown in figure 5.12.

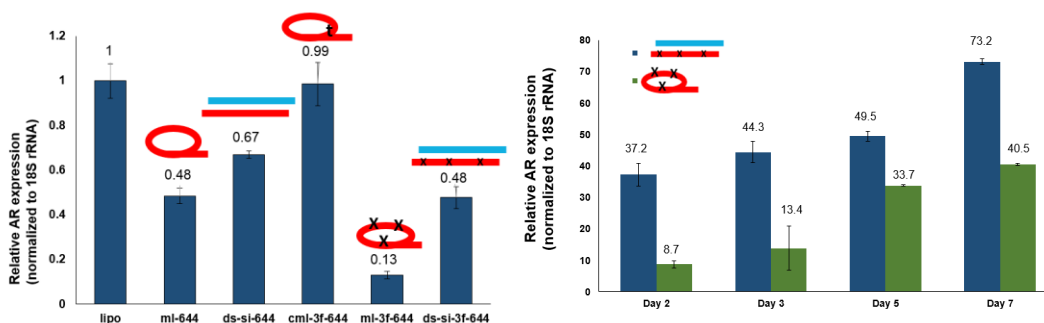


Figure 5.12. Quantitative real time PCR analysis of androgen receptor mRNA expression in LNCaP cells after transfection with RNAi agents. AR expression was normalized to 18S rRNA expression. Data are plotted as mean and error bars are standard deviation from three experiments. A variety of molecules are analyzed after 24 hours (Left). ds-si-3F-644 and ML-3F-644 RNAi activity was tested over one week (Right). Samples were analyzed at time points indicated on the graph. (Date from J.S. Ebron)

The mRNA levels recapitulate the pattern observed in the luciferase assay, with the mini-lariats (ML-644 & ML-3F-644) working better than the cognate siRNA (ds-si-644 & ds-si-3F-644). The click linked lariat (CML-3F-644) shows no appreciable silencing

effect. The data lends further support to a debranching dependent mechanism for mini-lariat action.

ML-3F-644 has the most effective in silencing AR mRNA. The length of silencing effect caused by ML-3F-644 is tested against ds-si-3F-644. On day 7, the mini-lariat RNA shows similar repression to the ds-si-3F-644 on day 2. The mini-lariats show a sustained knockdown effect over time and work better than traditional duplexed siRNA.

5.3.3 Repression of Endogenous AR Protein

Because ML-3F-644 is the strongest repressor of AR mRNA, the mini-lariat and its duplexed counterpart are further analyzed for a dose-dependent RNAi effect, this time looking at AR protein levels. The protein levels are analyzed by western blot and the house keeping gene, Hsp70, is unaffected by treatment with these RNAi agents. The concentrations used are 10, 20, 50, and 100 nM of RNA transfected into LNCaP cells. The results of the dose dependence test are shown in figure 5.13.

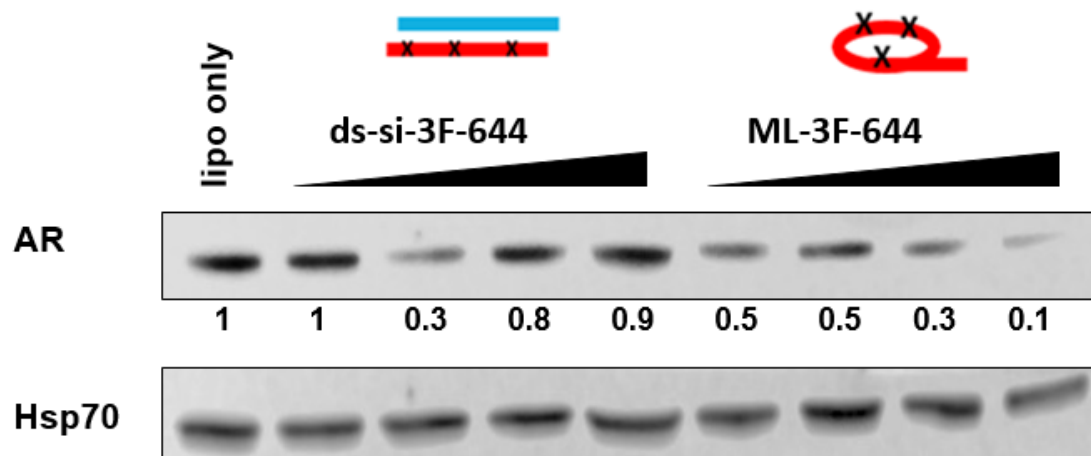


Figure 5.13. Western blot showing expression of androgen receptor and Hsp70 in LNCaP cells 48 hours after transfection with increasing amounts of ds-si-3F-644 or ML-3F-644. The RNAi active molecules were transfected at 10, 20, 50, and 100 nM concentration. (Data from J.S. Ebron)

At the lowest concentration, the mini-lariats perform better than the duplex strands. As the dosage of RNA increases, the ds-si-3F-644 construct shows a reversal of silencing activity, with almost no silencing seen at the highest doses of RNA. However, the mini-lariat constructs show increased AR protein suppression as the amount of RNA introduced to the cell increases, up to 90% knockdown at the highest RNA concentration. The exact reason why the ds-si-3F-644 construct reverses silencing is unclear. One possibility is that the passenger strand is getting loaded into RISC instead of the guide. The guide strand is selected by RISC through the thermodynamic interactions of the 5'-end. If the passenger strand is getting selected when there are higher concentrations of siRNA it is not targeting the androgen receptor, thus causing an increase in AR expression.

5.4. Conclusions

In this chapter I have shown that mini-lariat RNA can be synthesized on solid support. These mini-lariats are effective via the RNA interference pathway in knockdown of protein expression. The synthesis of the mini-lariat RNA is accomplished through the inclusion of the photo-protecting group at the 2'-position of a desired residue, similar to the synthesis of bbRNA. Conversion of the 5'-terminus into a phosphoramidite, followed by photo-deprotection, and cyclization can be performed entirely within the synthesizer with minimal manipulation. Completing the synthesis inside the RNA synthesizer has the benefit of minimizing exposure to moisture that is detrimental to synthesis. The synthetic scheme for lariatization allows incorporation of modified residues such as 2'-F residues. This may readily be extended to 2'-OMe or 2'-MOE substitution, as well as phosphorothioate linkages. These and other modifications to the RNA can increase the stability of the RNA and lend other useful properties towards increasing its effectiveness in cellular and biological applications¹⁸.

Mini-lariats are active in RNAi, knocking down endogenous protein and mRNA levels on the order of duplex siRNA (figs. 5.12 & 5.13). The mini-lariats show a dose dependent effect and maintain low levels of protein expression, while siRNA show an increase in protein levels after a certain concentration of introduced RNA. The mini-lariats have a more sustained effect in cells than traditional siRNA, with 60% knockdown of endogenous protein after seven days.

Mini-lariats have the added advantage of only introducing a single-strand of RNA into the cell. Whereas miRNAs are generated as a duplex, with both sides potentially having a cellular role, mini-lariats may be useful as a research tool to distinguish guide and passenger strand effects.

5.5. Experimental

5.5.1. Mini-Lariat Synthesis

First, a RNA molecule is synthesized using standard protocols and includes a single residue that has a 2'-NVOM protecting group at the 1 μ mol scale.

Step 1 – Cyanoethyl Removal

Following standard synthesis, the cyanoethyl protecting groups are selectively removed from the entire 1 μ mol column. A 10 mL syringe with a 6: 4 ACN: TEA solution is attached to the synthesis column. Every 20 minutes, ~2 mL of the solution is pushed through the column for a total of 90 min. Then the column is washed with 10 mL of ACN, then 10 mL of THF, then 10 mL of ACN and the column is dried under house vacuum for at least ten minutes. The 1 μ mol of material is split into three batches of ~330 nmol each, to ensure efficient photodeprotection while inside the synthesizer. After splitting the RNA into three separate columns, the columns are dried under house for ten minutes. One of the columns is loaded into the synthesizer for reaction.

Step 2 – 5'-Phosphorylation

The synthesizer is set up with two vials containing the reagents for phosphorylation. The first vial contained the diisopropylethylamine (DIPEA) base and the 1-methylimidazole catalyst, at a concentration of 0.95 M and 13.25 mM respectively, dissolved in DCM. The solution is at a final volume of 2 mL, with 335 μ L of DIPEA, and 7 μ L of 1-methylimidazole. The rest of the solution is DCM. The second vial contains the N,N-diisopropylamino cyanoethyl phosphoroamidic-Cl (PCI) phosphorylation reagent at 0.2 M concentration. This solution is also in 2 mL of DCM and required 90 μ L of PCI reagent.

Once the solutions are made and loaded onto the synthesizer, the sequence file is input. The remainder of mini-lariat synthesis is performed inside the synthesizer, so the

sequence file comprises steps 2-4. The file is as follows: U LL K ZQPQ ZQPQ ZQPQ ZQPQ ZQPQ ZQPQ ZQPQ ZQPQ A. The sequence is read by the synthesizer software in reverse order starting with letter A. The A step is a prewash of the column with ACN and a removal of the 5'-DMT group using the standard DMT removal solution of 3% trichloroacetic acid in DCM. Q represents 45 μ L injection of the DIPEA and 1-methylimidazole solution. P and Z both represent 45 μ L injection of phosphorylation reagent. Both Q and P have no wait time associated with them so the injections can immediately follow each other, while Z has a ten minute wait time and drain step.

Step 3 - Photodeprotection

The photodeprotection step is letter K in the sequence file. K is a ten minute ACN wash. Once the ACN is injected into the column, the UV LED is manually turned on from the control box located outside of the synthesizer irradiating the column with 365 nm light. Following photodeprotection, the column is washed with 200 μ L of ACN to remove any lingering product from the phosphorylation or deprotection steps (letter L in the sequence file).

Step 4 – Cyclization

Finally the cyclization step is represented by U. The MerMade software only allows injection of ETT activator that accompanies an injection from the amidite port. Because only the activator is necessary for cyclization, an empty bottle is loaded into the port associated with U. The 0.25 M ETT activator is injected and allowed to react with the column for ten minutes before draining and reinjection. This process repeats two more times for a total of thirty minutes. After coupling, oxidation occurs using the standard procedure used in linear synthesis. Finally the beads are washed three times with ACN.

Deprotection

After synthesis, the RNA is base deprotected using 1:1 30% NH_4OH : Methylamine for 12 minutes at 65 °C. The solution is removed from the CPG beads and the ammonia is blown off using a N_2 gas stream and the solution is lyophilized. The lyophilized RNA is then 2' deprotected using a DMSO, TEA, and TEA.3HF solution for two hours at 65 °C. This reaction is quenched, desalted, and lyophilized before purification.

5.5.2. Purification of Mini-Lariat RNAs

After synthesis, mini-lariat RNAs are purified using a reverse phase HPLC. The column is a Waters XBridge OST BEH C_{18} 2.5 μm (4.6 x 50 mm) column. Solvent A is 0.1 M TEAA and solvent B is 80/ 20 ACN/ H_2O in 0.1 M TEAA. The temperature is 40 °C and the flow rate is 1 mL/ min. The gradient is 4.8-16% B in 30 minutes.

5.5.3. Synthesis of Click Linked Lariat RNAs

The click linked lariats are first synthesized under standard conditions and incorporate a commercially available 2'-O-propargyl monomer at a desired position. The placement of the propargyl residue in the sequence matches the branch point in the phosphodiester linked mini-lariat RNA. After synthesis but while still on the solid phase, an azide is installed at the 5' position of the RNA. First, the 5'-OH is replaced with an iodo group by reaction with 0.5 M triphenylphosphonium iodine in DMF. The column is washed with DMF, reacted three times with 200 μL of the iodination reagent and rinsed with more DMF in the synthesizer. The column is removed from the chamber and attached to a syringe filled with a saturated solution of sodium azide in DMF. The solution is pushed through the column and put on a rotator inside of a 60 °C oven. Every 20 minutes, the syringe is removed from the oven and fresh solution is pushed into contact with the CPG beads. The total time for the azide attachment process is 90 minutes. Once the reaction

is complete, the column is washed with 10 mL of DMF, 10 mL of ACN, and dried under house vacuum. The RNA is then deprotected.

A copper catalyzed azide alkyne cycloaddition (CuAAC) is done to form the internal 2'-5'-triazole linkage. The RNA is reacted at a 5 nmol scale. The reagents and amounts used are detailed in the table below. The reaction was performed by mixing RNA, buffer, ACN, and water in a 0.65 mL microcentrifuge tube. A second tube contains 100 μ L of the sodium ascorbate solution, and a third tube has 100 μ L of the copper sulfate solution. Each individual tube is degassed with Ar three times for thirty seconds by blanketing the solution with the gas. The sodium ascorbate is added to the RNA solution and that is degassed a fourth time. The copper solution is added to the RNA solution and that is degassed a fifth time, starting the reaction. The reaction is run for three hours at room temperature. Once 1.5 hours has elapsed, more sodium ascorbate is added to the reaction mixture and the solution is degassed a sixth and final time. After the reaction is complete, the reacted RNA is separated from the salts and any unreacted RNA by reverse phase HPLC using the method detailed in 5.5.2.

Reagent	Final Concentration
RNA	50 μ M
CuSO ₄	10 mM
Sodium Ascorbate	10 mM
PBS buffer	1x
ACN	0.5%

Table 5.3. Final concentration of reagents for a 100 μ L CuAAC reaction

5.6. References

1. Fire, a *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
2. Zamore, P. D., Tuschl, T., Sharp, P. A. & Bartel, D. P. RNAi: Double-Stranded RNA Directs the ATP-Dependent Cleavage of mRNA at 21 to 23 Nucleotide Intervals. *Cell* **101**, 25–33 (2000).
3. Hutvágner, G. & Zamore, P. D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**, 2056–60 (2002).
4. Matzke, M. A. & Birchler, J. A. RNAi-mediated pathways in the nucleus. *Nat. Rev. Genet.* **6**, 24–35 (2005).
5. Okamura, K. & Lai, E. C. Endogenous small interfering RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **9**, 673–678 (2008).
6. Elbashir, S. M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 1–5 (2001).
7. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15**, 188–200 (2001).
8. Nykänen, A., Haley, B. & Zamore, P. D. ATP Requirements and Small Interfering RNA Structure in the RNA Interference Pathway. *Cell* **107**, 309–321 (2001).
9. Michler, T. *et al.* Blocking sense-strand activity improves potency, safety and specificity of anti-hepatitis B virus short hairpin RNA. *EMBO Mol. Med.* **8**, 1082–98 (2016).
10. Ruby, J. G., Jan, C. H. & Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**, 83–86 (2007).

11. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in *Drosophila*. *Cell* **130**, 89–100 (2007).
12. Westholm, J. O. & Lai, E. C. Mirtrons: microRNA biogenesis via splicing. *Biochimie* **93**, 1897–1904 (2011).
13. Wen, J., Ladewig, E., Shenker, S., Mohammed, J. & Lai, E. C. Analysis of Nearly One Thousand Mammalian Mirtrons Reveals Novel Features of Dicer Substrates. *PLoS Comput. Biol.* **11**, (2015).
14. Ruby, J. G., Jan, C. H. & Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**, 83–86 (2007).
15. Reynolds, A. *et al.* Rational siRNA design for RNA interference. *Nat. Biotechnol.* **22**, 326–330 (2004).
16. Ui-Tei, K. *et al.* Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res.* **32**, 936–948 (2004).
17. Schirle, N. T. *et al.* Structural Analysis of Human Argonaute-2 Bound to a Modified siRNA Guide. *J. Am. Chem. Soc.* **138**, 8694–8697 (2016).
18. Amarzguioui, M., Holen, T., Babaie, E. & Prydz, H. Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res.* **31**, 589–595 (2003).
19. Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R. & Tuschl, T. Single-Stranded Antisense siRNAs Guide Target RNA Cleavage in RNAi. *Cell* **110**, 563–574 (2002).
20. Prakash, T. P. *et al.* Identification of metabolically stable 5'-phosphate analogs that support single-stranded siRNA activity. *Nucleic Acids Res.* **43**, 2993–3011 (2015).

21. Lima, W. F. *et al.* Single-stranded siRNAs activate RNAi in animals. *Cell* **150**, 883–894 (2012).
22. Parmar, R. *et al.* 5'-(*E*)-Vinylphosphonate: A Stable Phosphate Mimic Can Improve the RNAi Activity of siRNA-GalNAc Conjugates. *ChemBioChem* **17**, 985–989 (2016).
23. Paredes, E. Triazole Linkages and Backbone Branches in Nucleic Acids for Biological and Extra-Biological Applications. (Carnegie Mellon University, 2012).
24. Nam, K. *et al.* Yeast lariat debranching enzyme: Substrate and sequence specificity. *J. Biol. Chem.* **269**, 20613–20621 (1994).
25. Carriero, S. & Damha, M. J. Solid-phase synthesis of branched oligonucleotides. *Curr. Protoc. Nucleic Acid Chem.* 4.14.1-4.14.32 (2002).
doi:10.1002/0471142700.nc0414s09
26. Miller, G. P. & Kool, E. T. Versatile 5'-Functionalization of Oligonucleotides on Solid Support: Amines, Azides, Thiols, and Thioethers via Phosphorus Chemistry. (2004). doi:10.1021/JO035765E
27. Miller, G. P. & Kool, E. T. A Simple Method for Electrophilic Functionalization of DNA. (2002). doi:10.1021/OL0264915
28. Ebron, J. S., Weyman, C. M. & Shukla, G. C. Targeting of Androgen Receptor Expression by Andro-miRs as Novel Adjunctive Therapeutics in Prostate Cancer. *J. Cancer Ther.* **4**, 47–58 (2013).
29. Narizhneva, N. V *et al.* Small molecule screening reveals a transcription-independent pro-survival function of androgen receptor in castration-resistant prostate cancer. *Cell Cycle* **8**, 4155–67 (2009).

Chapter 6 – Protein Free Splicing

6.1. Introduction

Pre-mRNA is processed into mature mRNA through splicing. The splicing process was discovered by Philip Sharp and Richard Roberts over 40 years ago¹. The spliceosome is a large megadalton complex that consists of both RNA and protein². The complex is highly dynamic, with different protein components associating, dissociating, and rearranging at each of the seven distinct steps in the splicing pathway². Key components of splicing are small nuclear ribonucleoproteins (snRNPs) that are smaller complexes of protein and RNA². The RNAs in snRNPs, small nuclear RNAs (snRNAs), are on average 150 nucleotides long and specifically recognize intron sequences in pre-mRNA through the 5'-splice site, 3'- splice site, and branch point sequence². An overview of the splicing pathway is shown in figure 6.1. The two main catalytic steps in splicing are freeing of the 5'-exon, by forming an intermediate RNA that contains both the lariat intron and the 3'-exon, and subsequently, ligating the 5'-exon with the 3'-exon. Both of these steps are S_N2 -transesterifications catalyzed by metals^{2,3}. The snRNAs are core components of each spliceosomal complex and are responsible for the catalytic activity of the spliceosome^{2,4-8}. The snRNPs as well as a myriad of other protein factors are important to properly position the different components together and to transition through the pathway. However, the catalytic core remains constant, containing the pre-mRNA, U2, U5, and U6 snRNPs along with at least 17 factors².

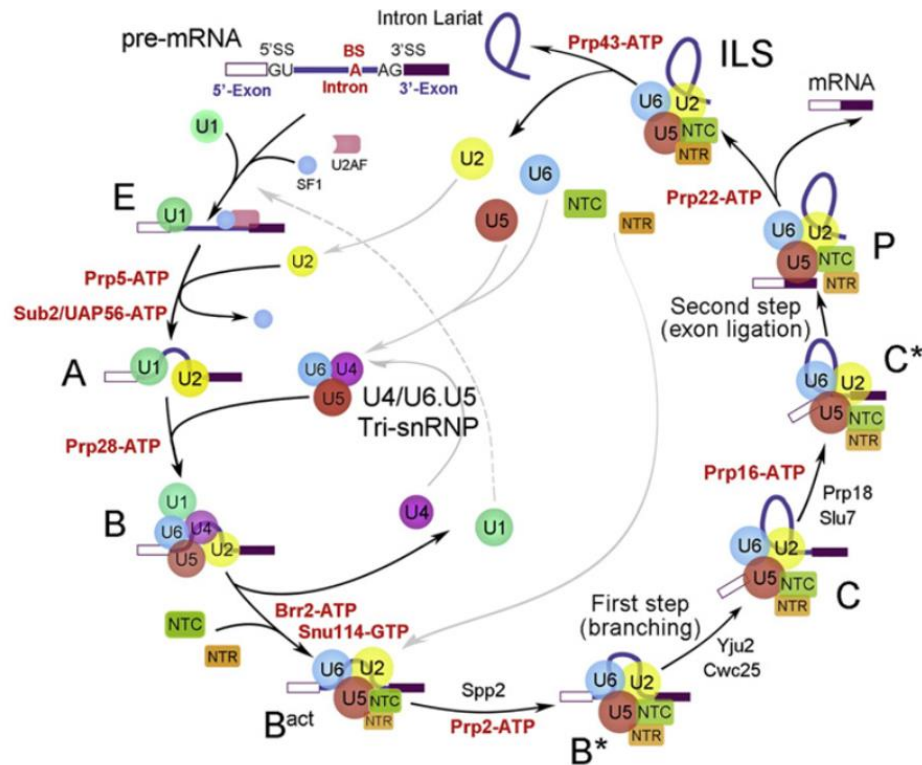


Figure 6.1. Detailed overview of the splicing pathway from ref⁹. snRNPs are colored circles. ATP dependent proteins are in red, select other protein factors are in black, and the NineTeen Complex (NTC) and NineTeen complex-Related (NTR) proteins are in colored boxes.

In 2016, two different groups used transmission electron cryomicroscopy (Cryo-EM) to determine the structure of the C complex in *S. cerevisiae*^{7,8}. Cryo-TEM structures of the spliceosome at other stages in the splicing cycle have also been elucidated^{10–15}. The C complex forms after the first catalytic step has occurred and the lariat has formed but is still ligated to the 3'-exon. Both groups observed very similar arrangement of snRNA around the intron lariat^{7,8}. The loop of the lariat was unable to be resolved, so the detected intron structure resembles backbone branched RNA (bbRNA) (fig. 6.2). U2 snRNA hybridizes to the 5' arm of the intron and bulges out the branch point residue. The 3'-arm makes two to three contacts with U2 snRNA. U6 snRNA hybridizes with the 2'-arm of the intron, leaving the first three residues in the arm unpaired. U2 and U6 snRNAs associate together as well, making a triplex near the branch point of the intron lariat as well as a second helix farther away from the intron interaction. U6 snRNA forms an internal stem

loop (ISL) that contains important catalytic residues. Loop I of U5 snRNA base pairs with the 5'-exon. These two RNAs do not make direct contacts with U2, U6, or the intron, but are positioned right in the active site.

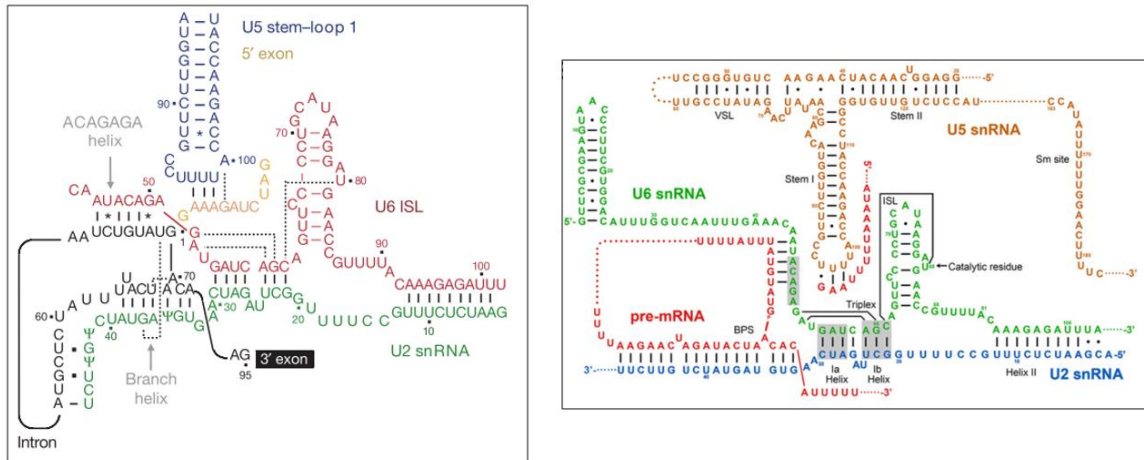


Figure 6.2. RNA configuration at the core of the spliceosome immediately after the first step of splicing (branching) elucidated by cryo-EM. Both independently acquired structures show similar interactions between the intron and U2 and U6 RNAs. The placement of U5 stem-loop1 and the 5'-exon are also shown. The left panel is from ref⁷ and the right panel is from ref⁸.

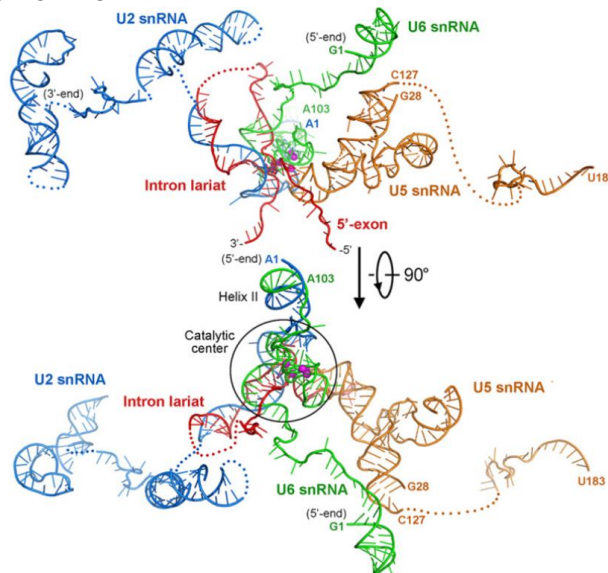


Figure 6.3. 3-D representation of RNA in spliceosome core from cryo-EM structure. Blue is U2 snRNA, green is U6 snRNA, gold is U5 snRNA, and red is the intron lariat. Disordered portions of RNA are represented by dotted lines. Fig from ref⁸.

The catalytic core of the spliceosome is composed of snRNAs and there are many similarities between the spliceosome core and group II introns^{3,4,6,16,17}. Group II introns are

self-splicing ribozymes that form a lariat with an adenosine residue at the branch point^{18–23}. The similarity between spliceosomal RNA and group II intron conserved sequences have prompted examinations into spliceosomal RNAs ability to act as a ribozyme^{5,6,24–27}. In one study, the authors used *in vitro* transcribed human U2 and U6 snRNAs and two linear oligomers, one that mimicked the 5'-exon and one that contained a short intervening (intron) sequence attached to the 3'-exon²⁴. After a 6 hour incubation at 45 °C, a spliced product containing both exons was observed²⁴. However, the product represented only 2% of the total input of radiolabeled RNA²⁴. Additionally, the pathway for the joining of the two RNA pieces appears to occur through a hydrolysis mechanism as opposed branching²⁴. A second study by the same group used a single RNA that contained both exons and a linking, intron-like sequence, more closely resembling a pre-mRNA²⁶. After a 15 hour incubation with U2 and U6, splicing was observed²⁶. Again, the splicing reaction represents a small fraction of the total input RNA²⁶.

With the synthetic access to bbRNA that represents the splicing intermediate, we set out to investigate protein-free splicing that starts with the C complex. The bbRNA is designed to have the same sequence and complementarity to U2 and U6 as observed in the cryo-TEM structures^{6,7}. The U2 and U6 sequences are from *S. cerevisiae* and are truncated to the portion of RNA observed in the cryo-TEM structures^{7,8}.

First, the melting temperature (T_M) of bbRNA hybridized to short RNAs that mimic spliceosomal contacts is investigated. Understanding the strength of specific RNA interactions will help develop an assay and give more information on the strength of interaction between snRNAs and the bbRNA. Assays that investigate the binding of U2 and U6 to bbRNA as well as the catalytic activity of the designed protein-free splicing system are performed and discussed in this chapter.

6.2. Synthesis of backbone branched RNA for Splicing Studies

Backbone branched RNA is synthesized as a lariat intron mimic. The general procedure is the same as detailed in previous chapters, but there are a few modifications to the procedure. The long 5' and 3'-arms make the coupling of the first residue in the 2'-arm difficult. To compensate, the amount of CPG used in each coupling reaction is reduced to ~100 nmol and the coupling time of the first residue is increased from 6 minutes to 8 minutes per injection, resulting in a 32 minute total coupling time. The remaining bases are coupled according to the standard bbRNA synthesis methods. Even with increasing the coupling time, the yields of this particular bbRNA are low, only recovering 16 nmol from a 1 μ mol synthesis. The purification gel is shown in figure 6.4.

Name	Sequence
bbRNA16	5'-AAC AGA UAC UA A(2'-5'-GUA UGU AUU GUU U) CAC AAU UUU-3'
bbRNA17	5'-Alexa488-NH-(CH ₂) ₆ -AAC AGA UAC UAA(2'-5' GUA UGU AUU GUU U) CAC AAU UUU – 3'
bbRNA18	5'-UUU AUA CUA A(2'-5'-GUA UGU AU UG-Dy547) CAC AAU UUU UCA GGU-t-Cy5- 3'

Table 6.1. bbRNA synthesized using the modified synthetic procedure. All three bbRNA are made using the 2'-Method. **t** represents a triazole linkage.

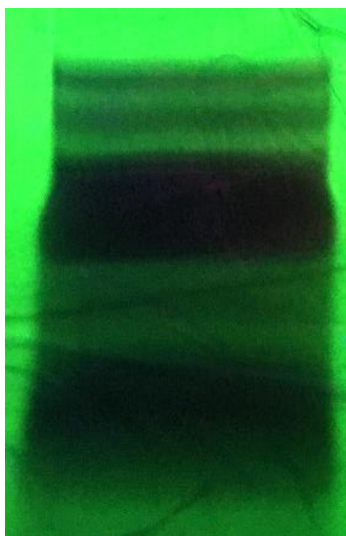


Figure 6.4. Purification gel of bbRNA17 used in in vitro spliceosome studies. The image shows a 20% denaturing polyacrylamide gel with RNA visualized by UV shadow. The top band is full length bbRNA17, the remaining bands are cleavage products and RNA where the synthesis of the 2'-arm failed.

6.3. Melting Temperature Studies

The RNA core of the spliceosome contains three pieces of RNA base paired together, the intron lariat, U2 snRNA, and U6 snRNA. To be able to quantitate some of the interactions between the three RNAs, a variety of RNAs are synthesized (table 6.2) and the melting temperature (T_M) of different hybrids is determined (table 6.3). The melting temperatures allow us to understand the amount of duplex destabilization that results from different RNA features seen in spliceosomal RNA, like a bulged base or a 2'-arm. Truncated versions of U2 and U6 and a large bbRNA are synthesized. The hybrids investigated are outlined in figure 6.5.

Hybrid I most closely resembles the arrangement of spliceosomal RNA. The branch point adenosine is bulged, the 3'-arm is not fully duplexed, and the 2'-arm has the first three nucleotides unpaired. However, to tease out the effect of these interactions on melting temperature, the remaining hybrids are also tested.

Name	Sequence
bbRNA16	5'-AAC AGA UAC UA A(2'-5'-GUA UGU AUU GUU U) CAC AAU UUU-3'
RNA3	5'-AAC AGA UAC UA A CAC AAU UUU-3'
RNA4	5'-AAC AGA UAC UA CAC AAU UUU-3'
RNA5	5'-AAA AUU GUG U UA GUA UCU GUU-3'
RNA6	5'-AAA AUU GUG UAG UAU CUG UU-3'
RNA7	5'-GUG UAG UAU CUG UU-3'
RNA8	5'-GUA UGU AUU GUU U-3'
RNA9	5'-AAA CAA UAC AUA C-3'
RNA10	5'-AAA CAA UAC A-3'
RNA11	5'-GUG Ψ AG UAU CUG UU-3'
RNA12	5'-GUG UAG UAU C Ψ G UU-3'
RNA13	5'-GUG Ψ AG UAU C Ψ G UU-3'

Table 6.2. Individual sequences for T_M experiments. Ψ is a pseudouridine residue.

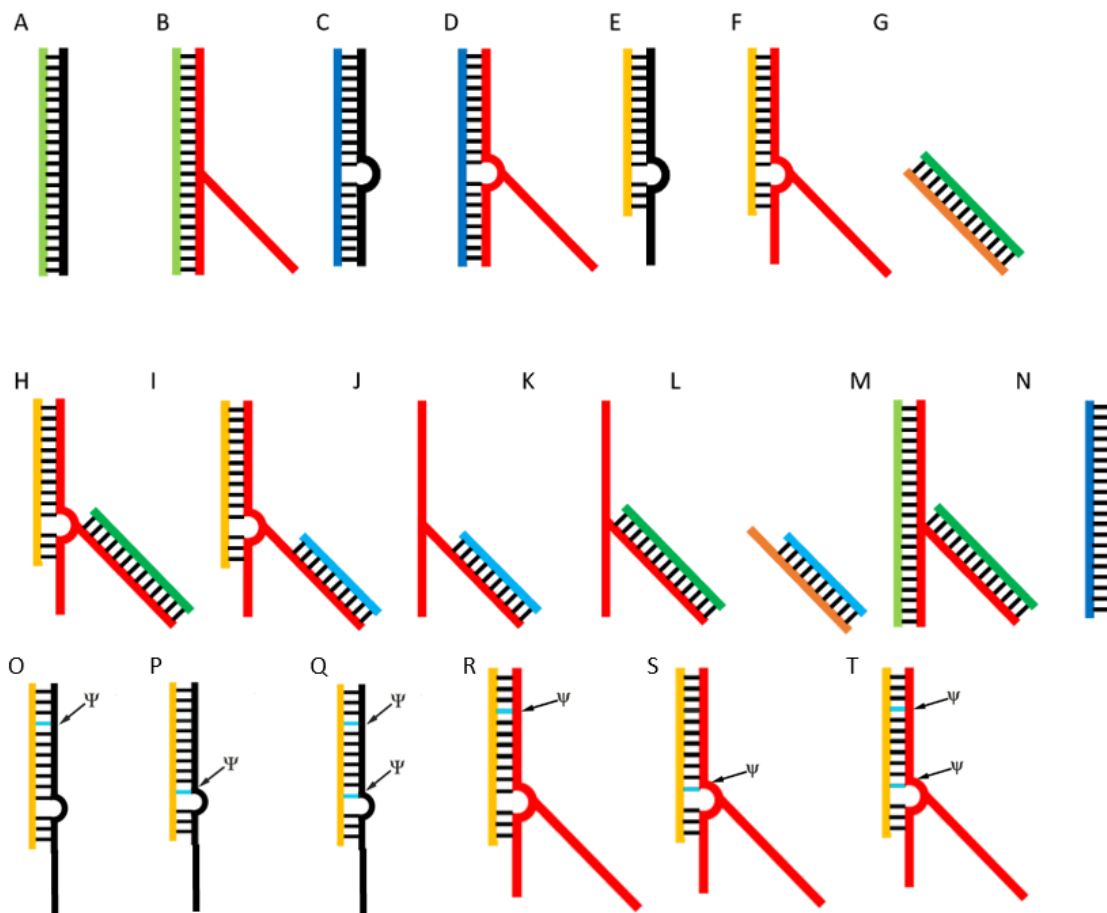


Figure 6.5. All of the RNA hybrids tested to determine melting temperatures of different duplexes. The RNAs strands and calculated T_M is in table 6.3. Hybrid I most closely resembles the spliceosome RNA interactions. Ψ represents a pseudouridine.

The melting point of the RNA hybrids from figure 6.5 is determined through UV-Vis spectrophotometry. 1 μ M of each RNA strand is mixed into a cuvette along with the T_M buffer (1 M NaCl, 100 mM sodium cacodylate, 0.5 mM EDTA). The solution is heated to 95 °C and cooled to 15 °C at a rate of 1 °C/ min, held at 15 °C for 5 minutes, and then heated to 95 °C at the same rate. The data is collected and the melting points are determined by calculating the first derivative and then the maximum of the data. Samples are run in triplicate, the mean is reported in table 6.3 and the error is the standard deviation. The majority of the UV-Vis experiments and first derivative calculations are

performed by undergraduate Stephanie Wang. Representative melting point and first derivative plots are in figure 6.6.

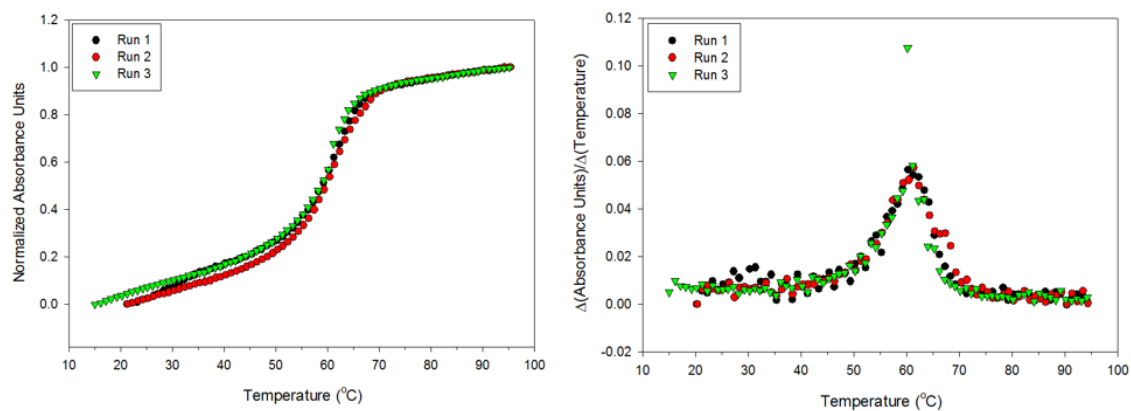


Figure 6.6. Representative melting curve (left) and first derivative plot (right) of hybrid R. The melting curve plots the normalized absorbance units by the temperature. The first derivative is generated for each line using SigmaPlot 12.5 and graphed against temperature. The maximum of each of the first derivative is determined and the mean of the three maximums determines the melting temperature.


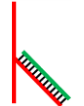

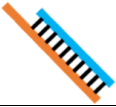

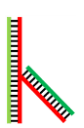





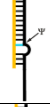



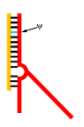

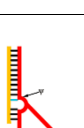

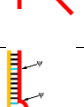
Hybrid Name	RNA sequences	T _M (°C)	Symbol	Hybrid Name	RNA sequences	T _M (°C)	Symbol
A	RNA5/ RNA3	71.2 ± 1.2		K	bbRNA16/ RNA9	53.4 ± 1.0	
B	RNA5/ bbRNA16	64.7 ± 1.4		L	RNA8/ RNA10	40.7 ± 1.1	
C	RNA6/ RNA3	63.5 ± 1.6		M	RNA5/ bbRNA16/ RNA9	55.6 ± 5.4 & 64.0 ± 0.87	
D	RNA6/ bbRNA16	60.1 ± 1.5		N	RNA6/ RNA4	71.2 ± 0.97	
E	RNA7/ RNA3	60.0 ± 1.2		O	RNA12/ RNA3	60.8 ± 1.6	
F	RNA7/ bbRNA16	57.0 ± 1.2		P	RNA11/ RNA3	60.8 ± 0.68	
G	RNA8/ RNA9	54.5 ± 1.5		Q	RNA13/ RNA3	63.2 ± 0.91	
H	RNA7/ bbRNA16/ RNA9	50.1 ± 3.3 & 53.2 ± 3.0		R	RNA12/ bbRNA16	60.6 ± 0.6	
I	RNA7/ bbRNA16/ RNA10	41.4 ± 5.2 & 55.4 ± 1.0		S	RNA11/ bbRNA16	59.7 ± 1.6	
J	bbRNA16/ RNA10	52.8 ± 3.1		T	RNA13/ bbRNA16	61.8 ± 3.2	

Table 6.3. Table of RNA hybrids used in the T_M studies, the calculated melting point, and a schematic representation of the hybrid. The experiments were run in triplicate and the melting temperature is the mean and the error is the standard deviation from the mean.

The calculated melting temperatures show that a bulge near the branch point or adding a 2'-arm decreases the stability of the duplexes. Comparing hybrid A to B, adding a 2'-arm destabilizes the duplexed 5' and 3'-arm by 6.8 °C. Comparing hybrid A to C, bulging out a residue destabilizes the duplex by 7.7 °C. Including both of these features does not compound the T_M decrease in an additive manner. There is only about a 3 °C further decrease in T_M from hybrid B or C and overall an 11.1 °C destabilization compared to hybrid A.

Hybridization of the 2'-arm is more complicated. When the 2'-arm of bbRNA16 is completely in a duplex (hybrid K), there is no significant change in T_M when compared to a duplex of the same size (hybrid G). So adding a 5' and 3'-arm does not impact the stability of the duplexed 2'-arm. However, hybrid J and L tell a different story. If the first three residues in the 2'-arm remain unpaired on bbRNA16 (hybrid J), there is a 12.1 °C increase in melting temperature compared to the similar duplex (hybrid L). It is unclear why there is such a large difference in T_M between these two hybrids, maybe the free residues in the 2'-arm of bbRNA16 base stack with residues in the 3'-arm, similar to how bbRNA base stacks with itself when complexed with Dbr1p and this interaction stabilizes of the duplex²⁹.

When all of the arms of bbRNA16 are duplexed, there are two transitions in the T_M graph and the first derivative plot shows two peaks (fig 6.7). The two peaks in the first derivative graph suggest that the melts are independent. Both maxima are reported in table 6.3. These melts were generated using a buffer that contains NaCl, sodium cacodylate, and EDTA. The melting temperature of duplexes is dependent on the ionic strength of the buffer. The bbRNA that are duplexed to two different RNA pieces can be analyzed in a buffer that contains divalent ions, like Mg^{2+} , which might separate the melting temperature differences to a greater extent than completed in these studies.

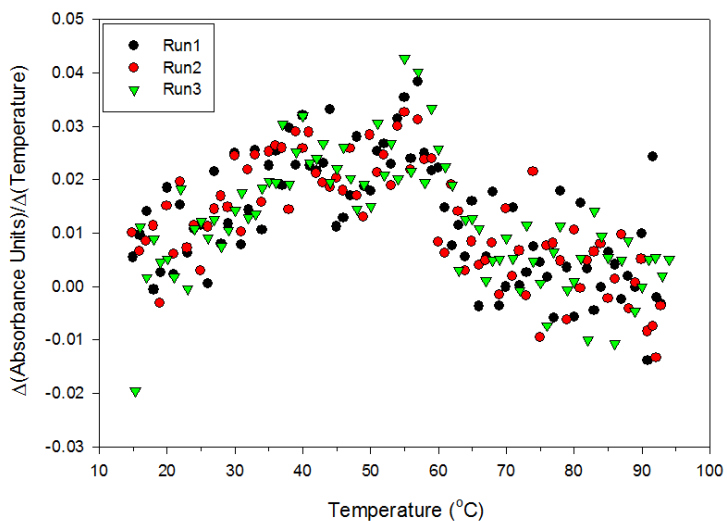


Figure 6.7. First derivative plot of hybrid I. There are two maxima in the plot, one roughly around 40 °C, and another around 55 °C. These represent independent melts of the two RNAs duplexed to bbRNA16.

Adding one pseudouridine does not change the T_M of a duplex as seen when comparing hybrid E to hybrids O and P. Two pseudouridine residues stabilize the duplex by roughly 3 °C. However, when hybridized to bbRNA16, one pseudouridine increases the T_M by about 3 °C. The increase in T_M here is important to note, because it is the opposite of the trend seen when duplexing the 5' and 3'-arms. Every instance when adding a 2'-arm to bulged and duplexed 5' and 3'-arms, the T_M decreases by 3 °C, (hybrids C and D, E and F). A second pseudouridine residue in the hybridizing strand does not increase the T_M a significant amount. The observed stabilization that pseudouridine affords may help the U2/ intron duplex form during the splicing process. Pseudouridine is known to play a role in maintaining loop structure in ribosomal RNA^{30,31}. However, the exact role the pseudouridine residue plays in the spliceosome remains to be determined.

6.4. RNA Binding Assay

6.4.1. Initial Binding Test

Before an *in vitro* cleavage assay with RNA components, it is imperative to ensure the components assemble together. To do this, the three RNA pieces that make base pairing interactions, U2, U6, and bbRNA, are annealed together under different conditions and analyzed on a non-denaturing polyacrylamide gel. If the RNAs associate together, the hybridized tri-RNA complex will migrate slower in the gel compared to each individual RNA as well as a complex comprised of two of the RNA components. To easily visualize the RNAs, each piece is labeled with a dye. The sequences are detailed in table 6.4. bbRNA17 and U2 are made using solid phase synthesis. U6 is *in vitro* transcribed and includes a 5'-N₃ (courtesy of Dr. Venkat Gopalan's lab at The Ohio State University). A dye is then conjugated onto the 5'-N₃ of U6 RNA using click chemistry. The U2 and U6 sequences used are derived from *S. cerevisiae* conserved sequences and designed to associate with the bbRNA. The specific interactions of the three designed RNAs is shown in figure 6.8.

Name	Sequence
bbRNA17	5'-Alexa488-NH-(CH ₂) ₆ -AAC AGA UAC UAA(2'-5' GUA UGU AUU GUU U) CAC AAU UUU – 3'
U2	5'-Alexa647-t-GGG AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU CUG UU-3'
U6	5'-Alexa594-t-GGG GAA ACA AUA CAG AGA UGA UCA GCA GUU CCC CUG CAU AAG GAU GAA CCG UUU UAC AAA GAG AU CCC-3'

Table 6.4. Sequences used in initial binding assay.

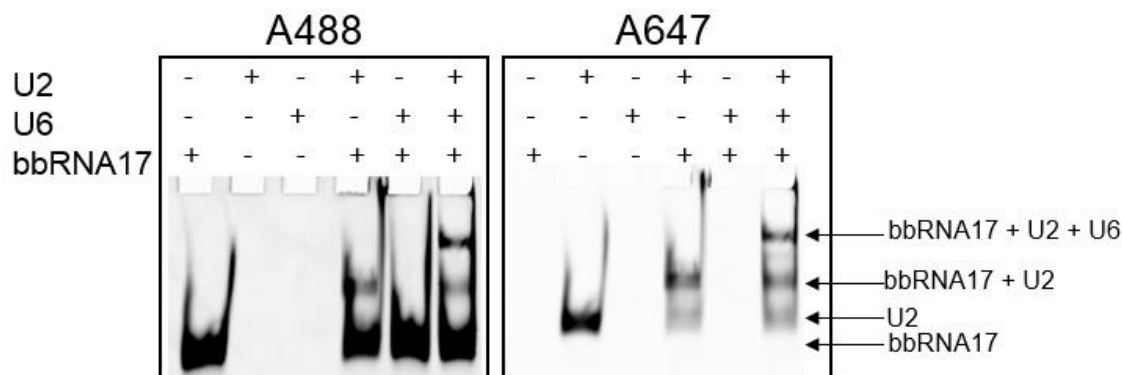


Figure 6.9. RNA binding assay. 12% non-denaturing polyacrylamide gel of RNA binding assay. The RNA was mixed with TMS buffer and incubated for 1 hour at 37 °C and analyzed on the gel. The gel is scanned twice, once per indicated dye. bbRNA17, U2 and U6 associate together, as seen in the last lane.

6.4.2. Improving the RNA Association – Annealing Procedure

The first test of RNA binding shows association between the three pieces, however not all of the RNA strands properly associated. To improve binding, different buffer and annealing conditions are investigated. First, the annealing procedure is changed. The melting temperatures of branched RNA duplexes determined in section 6.3 are higher (55 °C) than the annealing procedure used in section 6.4.1. To ensure proper assembly of U2, U6, and bbRNA17, the RNA strands are mixed together, heated above the melting temperature of the duplexes, and cooled. The RNA is heated to 85 °C for one minute, 65 °C for 10 minutes, and then kept at 37 °C for one hour. After the hour has elapsed, gel loading solution is added to the binding reaction and the RNA is run on a 10% non-denaturing polyacrylamide gel. There was no appreciable difference between the two annealing methods and no increase in the amount of properly assembled complex (fig. 6.10).

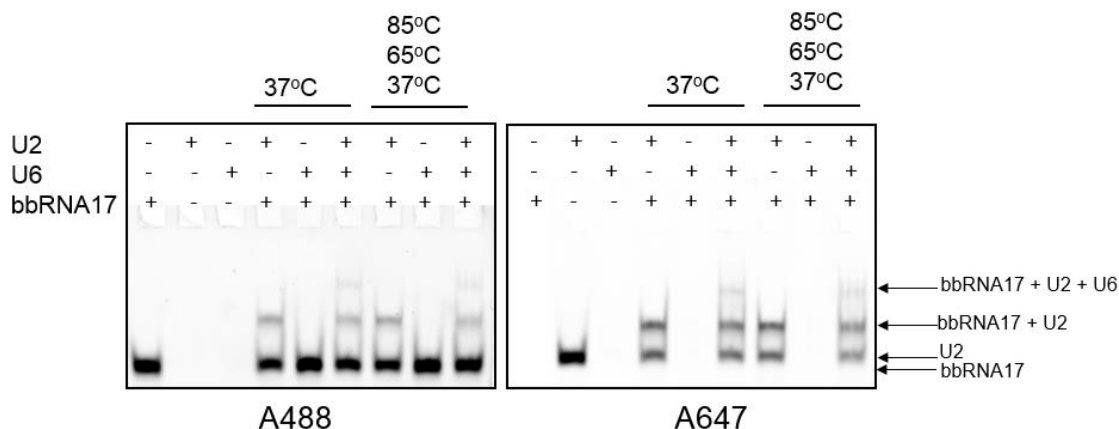


Figure 6.10. Difference in annealing procedures. The RNA is mixed, heated as indicated, and separated on a 10% non-denaturing polyacrylamide gel. The gel is run at 4 °C, and scanned using a Typhoon FLA9000 for Alexa488 and Alexa647 dyes.

6.4.3. Improving the RNA Association – Pre-Anealing

In another attempt to improve association of the U2, U6, bbRNA17 RNAs, U2 and U6 are annealed together before bbRNA17 is introduced. The pre-annealing is done to properly arrange the longer RNA pieces before adding bbRNA17. U2 and U6 are mixed with 89 mM Tris.HCl pH 7.5 and 100 mM NaCl, heated at 90 °C for one minute and then 65 °C for 10 minutes. 5 mM MgCl₂ is added to the U2 and U6 RNA. The mixture is then heated at 37 °C for five minutes. bbRNA17 is added and the mixture is heated at 37 °C for one hour. After one hour, gel loading solution is added and the binding reaction is run on a 10% non-denaturing polyacrylamide gel at 4 °C in TMS buffer. The gel is imaged on a Typhoon FLA9000 and scanned for Alexa488 and Alexa647. The bands observed in the Alexa488 scan are quantitated with the associated software (fig. 6.11). When the RNA is mixed together and annealed as in 6.4.2, 13.5% of bbRNA17 is in the three RNA complex, 30.2% is associated with U2 and 56.3% is free bbRNA17. When U2 and U6 are pre-annealed together, 18.1% of bbRNA17 is associated with U2+U6, 29.1% is with U2 and 52.7% is unassociated. Pre-annealing U2 and U6 increases the association of bbRNA17 with U2 and U6 by 5%.

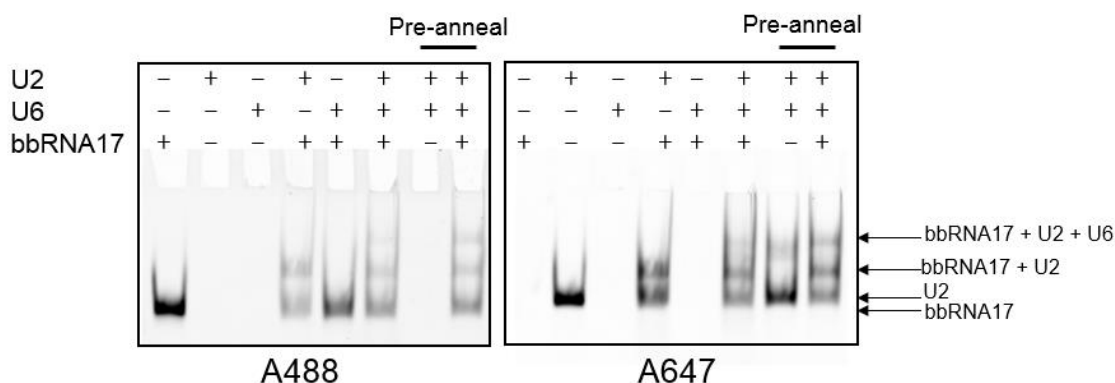


Figure 6.11. The effect of pre-annealing on RNA binding. The U2 and U6 RNA are mixed, heated, and bbRNA17 is added. The RNA are separated on a 10% non-denaturing polyacrylamide gel. The gel is run at 4 °C, and scanned using a Typhoon FLA9000 for Alexa488 and Alexa647 dyes.

6.4.4. Improving the RNA Association – More MgCl₂

High concentrations of divalent ions are used in *in vitro* experiments to fold large RNAs into secondary structures²². The amount of MgCl₂ is increased from 5 mM to 100 mM to force the RNA to fold and associate properly. U2, U6, and bbRNA17 are mixed into 1x TMS buffer. More MgCl₂ is added to one reaction to bring the total concentration up to 100 mM. Then the binding reaction is incubated at 37 °C for one hour. After one hour, gel loading solution is added to the tube, and the reaction is run on a 12% non-denaturing polyacrylamide gel in 1x TBM buffer at 4 °C. The gel is imaged on a Typhoon FLA9000 and scanned for Alexa488 signal. There is no appreciable difference between the two amounts of MgCl₂ (fig. 6.12).

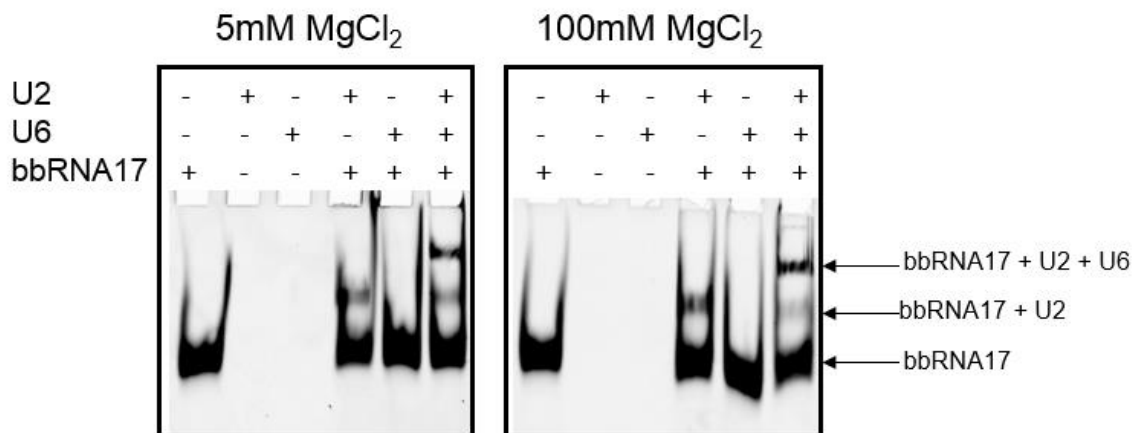


Figure 6.12. The effect of increasing the MgCl₂ concentration on RNA binding. bbRNA17, U2, and U6 are mixed together with 1x TMS buffer, and more MgCl₂ is added to one tube. The RNA are separated on a 12% non-denaturing polyacrylamide gel. The gel is run at 4 °C and scanned on a Typhoon FLA9000 for Alexa488 dyes.

6.4.5. Improving the RNA Association – Crowding Agent

The proteins of the spliceosome bring the small nuclear RNAs together². To bring the synthetic RNAs together in the *in vitro* splicing experiments, a crowding agent is added to force the RNA to interact and to mimic the crowded *in vivo* environment. In this case, 3% weight per volume (w: v) of 20,000 Da polyethylene glycol (PEG) is added to the binding reaction. Additionally, U2 and U6 are pre-annealed together, since that technique was shown to increase association of bbRNA17, U2, and U6. First, U2 and U6 are mixed together with 89 mM Tris.HCl pH 7.5 and 100 mM NaCl. The RNAs are heated at 90 °C for one minute and 65 °C for ten minutes. Then 100 mM MgCl₂ is added to the mixture and heated at 37 °C for five minutes. A solution of bbRNA17 in water, with or without PEG, is added to the U2 and U6 mixture. The binding reaction is incubated at 37 °C for one hour, gel loading solution is added and the mixture is separated on a 10% non-denaturing polyacrylamide gel. The gel is run at 4 °C and subsequently imaged at both the Alexa488 and Alexa647 channels (fig. 6.13). The gel bands are quantitated using the associated software and the Alexa488 image. The tri-RNA complex formed with PEG in solution is 14.7% of the total amount of bbRNA17. When PEG is not included in the binding reaction,

the top band represents 21.5% of the total amount of bbRNA17. These results indicate that adding PEG decreases the association of bbRNA17 with U2 and U6 by 6%.

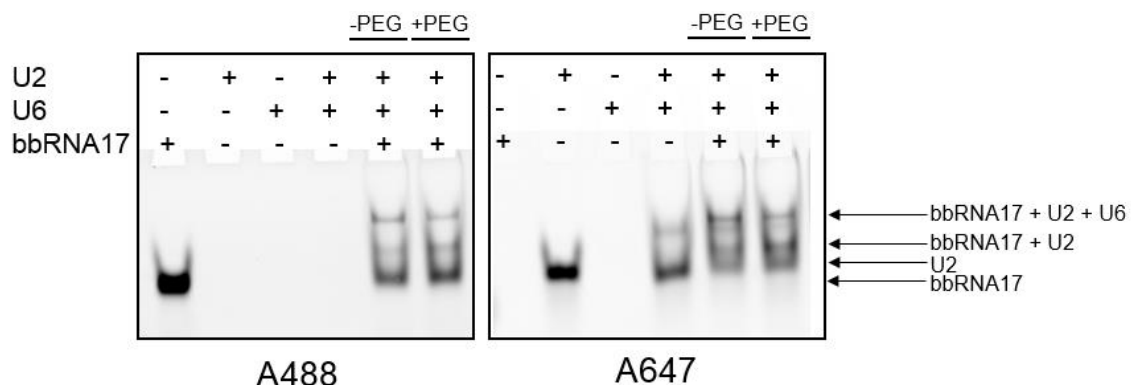


Figure 6.13. The effect of adding a crowding agent on RNA binding. U2 and U6 are preannealed and then incubated with bbRNA17 with or without adding 3% w: v PEG 20,000. The RNA are separated on a 10% non-denaturing polyacrylamide gel. The gel is run at 4 °C and scanned on a Typhoon FLA9000 for Alexa488 and Alexa647 dyes.

6.5. RNA Cleavage Assay

A cleavage reaction is performed to determine if bbRNA, U2, and U6 RNAs are catalytically active. As stated earlier, the RNAs used in the *in vitro* assays are modeled after the C-complex in the spliceosomal pathway. Aside from the direct base pairing interactions of U2 and U6 with the intron, the C-complex also contains U5 snRNA and the 5'-exon which are associated but not directly paired. Both catalytic steps of splicing are reversible and will result in different band patterns in a denaturing polyacrylamide gel³³. If the second step of splicing occurs in our *in vitro* experiments, the bbRNA will be smaller and migrate faster in the gel, as the 3'-exon will be removed. bbRNA18 contains dyes on the terminus of the 2' and 3'-arms which will be separated if the second splicing step occurs. The reverse of the first step of slicing can also occur, which results in cleavage of the 2'-5'-phosphodiester bond and ligation of the 5' end of the intron with the 3'-end of the 5'-exon. If the reverse of the first step occurs, the 2'-5'-phosphodiester bond will be cleaved, resulting in the 5'-arm, branch point residue, and 3'-arm of the bbRNA as one product and the 2'-arm ligated to the 5'-exon as the second product. The two pieces of RNA would be visualized separately on the gel as each would contain a different dye.

6.5.1. Initial Cleavage Test

The first test of *in vitro* RNA cleavage was performed before bbRNA17 was optimized for binding with U2 and U6 RNA. In the first experiment, U2u, U6u, U5, and 5'-exon RNA are mixed with bbRNA18 in PK buffer (60 mM Sodium phosphates, 3% w: v PEG 4,000) (table 6.5). The reaction is heated at 90 °C for 1 minute, heated at 65 °C for 15 minutes, and MgCl₂ is added to the reaction at 2.1 mM final concentration. The reaction is then left to shake for 16 hours at ambient room temperature. Stop solution (9: 1 formamide: 0.1 M EDTA) is added to quench the reaction and the mixture is loaded onto a 20% denaturing polyacrylamide gel. The gel is imaged on the Typhoon FLA9000 and

scanned three times, at the Cy3 channel, the Cy5 channel, and the fluorescein channel. Cleavage of bbRNA18 occurs only when U2u, U6u, U5, 5'-Exon are incubated with bbRNA18 in the presence of 2.1 mM MgCl₂ (fig. 6.14). If any of these components are missing, cleavage does not occur. However, there is not just one or two clear products from this cleavage reaction, as many bands appear on the gel. However, this is a clear indication that protein-free RNA cleavage can be catalyzed with these sequences.

Name	Sequence
U5	5'- AUG GUU CUU GCC UUU UAC CAG AAC CAU-3'
5'-Exon	5'-Fluorescein-CAG-3'
bbRNA18	5'-UUU AUA CUA A(2'-5'-GUA UGU AU UG-Dy547) CAC AAU UUU UCA GGU-t-Cy5- 3'
U2u	5'-GGG AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU CUG UU-3'
U6u	5'-GGG GAA ACA AUA CAG AGA UGA UCA GCA GUU CCC CUG CAU AAG GAU GAA CCG UUU UAC AAA GAG AU CCC-3'

Table 6.5. RNA sequences used in the cleavage experiment

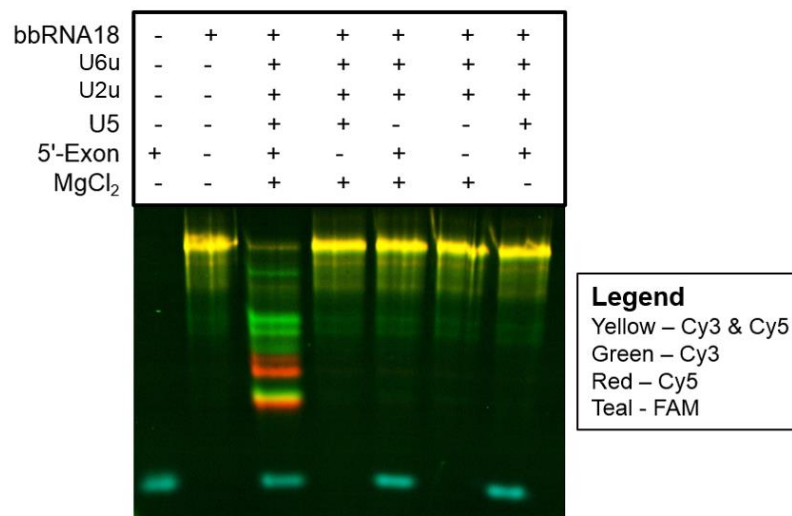


Figure 6.14. Cleavage reaction for *in vitro* splicing. After 16 hours of incubation, the cleavage reaction is run on a 20% denaturing polyacrylamide gel and scanned using a Typhoon FLA9000. False color gel generated from three separate gel images.

6.5.2. Cleavage Test Repeatability Issues

After the promising initial results, further experiments are performed to repeat the first cleavage assay. These included increasing the pH of the buffer to 8.5, the incubation time to 24 hours, the incubation temperature to 37 °C, the RNA concentration by 1.5x, the

MgCl₂ concentration by 2x, and the length of the 5'-Exon as well as investigating the order of RNA addition. All of these experiments resulting either in no observable cleavage or inconsistent results.

6.5.3. Redesigned bbRNA and Cleavage Tests

Given the inconsistency, bbRNA18 is redesigned to have more base pairing interactions with U2 and U6 RNA. bbRNA17 from section 6.4. is made and the binding assays are performed to show prove that the three RNAs are associating together (table 6.6). Once binding of U2, U6, and bbRNA17 is established, cleavage assays are once again performed.

Name	Sequence
U2	5'-Alexa647- t -GGG AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU CUG UU-3'
U6	5'-Alexa594- t -GGG GAA ACA AUA CAG AGA UGA UCA GCA GUU CCC CUG CAU AAG GAU GAA CCG UUU UAC AAA GAG AU CCC-3'
bbRNA17	5'-Alexa488-NH-(CH ₂) ₆ -AAC AGA UAC UAA(2'-5' GUA UGU AUU GUU U) CAC AAU UUU – 3'
U5	5'- AUG GUU CUU GCC UUU UAC CAG AAC CAU-3'
Exon2	5'-UCU AGA AAG-3'

Table 6.6. Sequences of RNA used in the redesigned splicing assay. **t** represents triazole link.

U2 and U6 are pre-annealed together in 89 mM Tris.HCl, pH 7.5 and 100 mM NaCl by heating to 90 °C for one minute and 65 °C for ten minutes. 100 mM MgCl₂ is added, the mixture is heated at 37 °C for five minutes, and finally, bbRNA17 is added. The reaction is left for 20 hours at 37 °C. After the reaction time elapsed, stop solution (9:1 formamide: 0.1 M EDTA) is added and the RNA is run on a 10% denaturing polyacrylamide gel and scanned for Alexa488 and Alexa647 signal. There appears to be no cleavage of bbRNA17 under these conditions (fig.6.15).

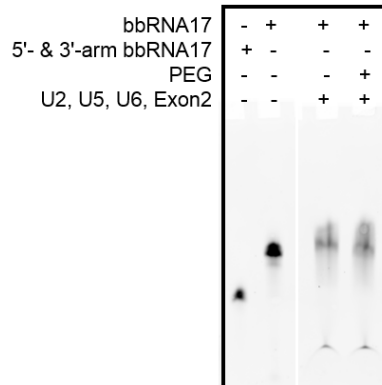


Figure 6.15. Cleavage of bbRNA17. U2, U6, U5, Exon2, and bbRNA17 are incubated at 37 °C for 20 hours with 100 mM MgCl₂. 3% w: v PEG 20,000 is added to one reaction. After 20 hours, the RNA is run on a 10% denaturing polyacrylamide gel and scanned on a Typhoon FLA9000 for Alexa488.

Further reactions are performed to determine cleavage conditions for the *in vitro* splicing assay. The conditions tested are as follows: changing the buffer system, adding 3% w: v PEG 20,000, doubling the amount of U6 in solution. In the first reaction, the Tris/NaCl buffer system is replaced with PK buffer (60 mM sodium phosphates, 3% w: v PEG 4,000) used in the first successful cleavage reaction. Each of these reactions are run at 37 °C and ambient room temperature (25 °C) for 20 hours. As seen in figure 6.16, none of the reactions, at either temperature, results in specific cleavage of bbRNA17.

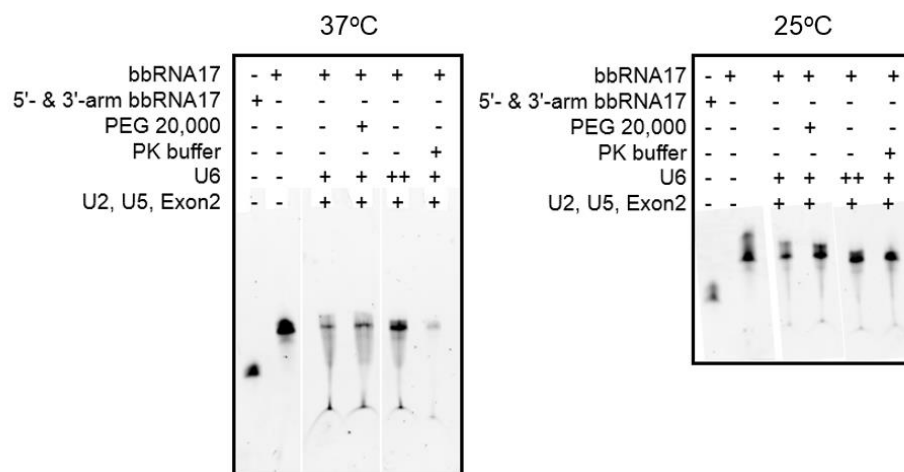


Figure 6.16. Cleavage of bbRNA17. U2, U6, U5, Exon2, and bbRNA17 are incubated at 25 °C or 37 °C for 20 hours with 100 mM MgCl₂ and buffer as indicated. After 20 hours, the RNA is run on a 10% denaturing polyacrylamide gel and scanned on a Typhoon FLA9000 for Alexa488.

6.6. Conclusions

This chapter describes different assays that investigate the stability and catalytic activity of the RNA core in the spliceosome. bbRNA is used as an intron mimic. The melting temperature of different RNA duplexes is determined. It was seen that adding a 2'-arm or bulging a residue have a similar destabilizing effect. However, the effect is not additive, and when both features are present there is only a minor destabilization of the duplex compared to when only one feature is present. Fully duplexing the 2'-arm of bbRNA16 does not change the T_M relative to a duplex that does not contain a 5' and 3'-arm. If the first three residues in the 2'-arm of bbRNA16 remain unhybridized, the melting temperature is much higher than a normal 10-mer duplex, indicating that the other residues in the bbRNA might be stabilizing the interaction. Adding a single pseudouridine residue to a duplex does not change the T_M but adding a pseudouridine to bbRNA that also contains a bulged 2'-arm stabilizes the interaction compared to the duplex with a bulge. The pseudouridine residue stabilization is opposite to the general trend that adding a 2'-arm destabilizes duplexes.

Truncated *S. cerevisiae* U2 and U6 snRNAs are able to hybridize with bbRNA and create a tri-RNA complex. U6 snRNA and bbRNA will only associate together when U2 is present. Higher $MgCl_2$ concentration aids in forming the tri-RNA complex. A PEG crowding agent doesn't stabilize the interaction. Pre-annealing U2 and U6 together slightly increased the amount of the U2/ U6/ bbRNA complex. Overall, no more than 20% of the total bbRNA is observed in the tri-RNA complex.

In the first attempt at a cleavage assay, incubating U2, U6, bbRNA18, U5, and 5'-exon together overnight shows cleavage of bbRNA. And further, if any of the individual RNAs are not present, cleavage was not observed. However, the cleavage was not consistently reproducible using bbRNA18 under a variety of different conditions. A longer

bbRNA (bbRNA17) was designed to increase association between the RNA pieces. However, under a few different conditions, bbRNA17 is not cleaved. Given the initial results from the first cleavage experiment and literature work indicating that the snRNAs can cleave RNA without associated spliceosomal proteins, further experimentation and careful assay condition design might prove ^{24,26}.

6.7. Experimental

6.7.1. Synthesis of bbRNA16, 17, and 18

Name	Sequence
bbRNA16	5'-AAC AGA UAC UA A(2'-5'-GUA UGU AUU GUU U) CAC AAU UUU-3'
bbRNA17	5'-Alexa488-NH-(CH ₂) ₆ -AAC AGA UAC UAA(2'-5' GUA UGU AUU GUU U) CAC AAU UUU – 3'
bbRNA18	5'-UUU AUA CUA A(2'-5'-GUA UGU AU UG-Dy547) CAC AAU UUU UCA GGU-t-Cy5– 3'

Table 6.7. bbRNA sequences made using a modified 2'-Method. t is a triazole linkage

The bbRNA used in this chapter, (bbRNA16, 17, and 18) were synthesized using the 2'-Method outlined in chapter 2. The 3'-arm and 5'-arm were synthesized in the forward direction. The CE groups were selectively removed using 6: 4 ACN: TEA for 90 min at room temperature. The beads were removed from the column and subdivided into eight fractions of ~150 nmol each. The subdivision is done in more fractions than typical bbRNA synthesis to decrease the amount of RNA reacting and hopefully increase yield. The RNA was transferred to a 0.5 dram glass vial and the 2'-O-NVOM group was removed by irradiation with 365 nm light for 6 minutes. The beads are transferred back into a synthesis column and placed into the synthesizer.

For reverse synthesis, the coupling time of the first base in the 2'-arm was extended. Each injection is typically 6 minutes of coupling for a total time of 24 minutes. For these RNAs, the time was increased to 8 minutes per injection for a total coupling time of 32 minutes. The remaining residues are coupled using the 2'-Method standard procedure (0.2 M amidite and 12 minute total coupling).

The RNA aliquots were combined into one vial and base deprotected using 1:1 40% NH₄OH: methylamine for 12 min at 65 °C. The ammonia was removed by nitrogen gas stream and the remaining solution was lyophilized. Next, the 2'-O-TBDMS groups were removed using DMSO, TEA, TEA.3HF. The RNA was quenched with RNA

quenching buffer (Glen Research) and desalted using Glen Pack desalting columns. The RNA was lyophilized and resuspended in water.

After deprotection, the RNA was gel purified using 20% denaturing (8 M Urea) polyacrylamide gel. The RNA bands were excised and passively eluted using TE_{0.1} buffer for 24 hours. After 24 hours, the RNA was desalted using Glen Pack desalting columns and lyophilized. The RNA was resuspended in water and analyzed by MALDI mass spectrometry.

After purification, bbRNA17 was conjugated with Alexa488-NHS ester. 1 mg of Alexa was dissolved in 60 μ L of dry DMSO. Unlabeled bbRNA17 was mixed with 0.1 M borate buffer (0.1 M sodium tetraborate decahydrate, pH 8.5) for one minute. 50 μ L of the dye solution was added to the RNA. The reaction was left to shake at room temperature overnight. bbRNA17 was purified using reverse phase HPLC.

6.7.2. Melting Temperature Studies

The melting temperature for each hybrid was determined using a Cary 300 BIO UV-Vis spectrophotometer. Each hybrid was run in triplicate. Solutions of 1 mL were made that contained each oligonucleotide at 1 μ M concentration, and a buffer of 1 M NaCl, 100 mM sodium cacodylate, and 0.5 mM EDTA. These solutions were transferred to quartz cuvettes and run on the UV-Vis. The spectral bandwidth was set to 3 nm. The temperature run ramped from 95 °C to 15 °C at 1 °C per minute. The temperature was held at 15 °C for five minutes. The temperature was then ramped from 15 °C to 95 °C at a rate of 1 °C per minute. The data was analyzed and the average melting point was calculated using the first derivative using SigmaPlot 12.5.

6.7.3. Binding Assay

The first binding assay was performed using U2 and U6 RNA that was post synthetically labeled with Alexa647 and Alexa594 respectively using click chemistry. The reaction was performed like the other click reactions described in previous chapters. The RNAs were then gel purified, eluted, desalted, lyophilized, and resuspended in water.

U2, U6, and bbRNA17 were incubated with TMS buffer (5 mM MgCl_2 , 100 mM NaCl, 89 mM Tris.HCl pH 7.5) for 1 hour at 37 °C (Table 6.8). Each piece of RNA was at 50 nM. The reaction had a total volume of 10 μL . After 1 hour, 10 μL of gel loading solution (40% glycerol) was added to the reaction and the reaction was run on a 12% non-denaturing polyacrylamide gel in 1x TBM buffer (5 mM MgCl_2 , 200 mM boric acid, 89 mM Tris.HCl pH 7.5). The gel was run at 4 °C. The gel was imaged using a Typhoon FLA9000 and scanned at Alexa488 and Alexa647 wavelengths.

Reaction	bbSplice	U2	U6
1	50 nM	--	--
2	--	50 nM	--
3	--	--	50 nM
4	50 nM	50 nM	--
5	50 nM	--	50 nM
6	50 nM	50 nM	50 nM

Table 6.8. RNA concentrations used in binding assay 1.

The second set of binding reactions, U2, U6, and bbRNA17 (50 nM each) were incubated in TMS buffer. One reaction was incubated at 37 °C for one hour, as in the first reaction. The second reaction was incubated at 85 °C for one minute, 65 °C for 10 minutes, and then kept at 37 °C for one hour. After the 37 °C incubation, 10 μL of 40% glycerol gel loading solution was added to the 10 μL reaction. The binding reactions were run on a 10% native polyacrylamide gel in 1x TBM buffer at 4 °C. The gel was then imaged on a Typhoon FLA9000 and scanned at Alexa488 and Alexa647 wavelengths.

Reaction	bbSplice	U2	U6
1	50 nM	50 nM	--
2	50 nM	--	50 nM
3	50 nM	50 nM	50 nM

Table 6.9. RNA concentrations used in binding assay 2. Each reaction is made in duplicate, with one annealed at 37 °C and the other with the altered procedure.

In the third set of binding reactions, U2 and U6 (50 nM each) were incubated with 100 mM NaCl and 89 mM Tris.HCl pH 7.5 and heated at 90 °C for one minute, at 65 °C for ten minutes. Then, 5 mM MgCl₂ is added to the reaction and it is kept at 37 °C for five minutes. While U2 and U6 were heating, bbRNA17 (50 nM) and water were heated at 37 °C for ten minutes. The preheating is done to have both solutions at the same temperature when mixing together. The bbRNA17 solution was added to the U2 and U6 mixture and left at 37 °C for one hour. After one hour, 10 µL of 40% glycerol gel loading solution was added. The binding reactions were run on a 10% native polyacrylamide gel in 1x TBM buffer at 4 °C. The gel was then imaged on a Typhoon FLA9000 and scanned at Alexa488 and Alexa647 wavelengths. The bands in the gel was quantitated using the associated software from the image taken at Alexa488 wavelength.

In the fourth set of binding reactions, the amount of MgCl₂ was increased from 5 mM to 100 mM. In two tubes U2, U6, and bbRNA17 (each at 50 nM) are mixed with 100 mM NaCl and 89 mM Tris.HCl pH 7.5. In one tube, 1 µL of 100 mM MgCl₂ was added to make a final concentration of 5 mM. In a second tube, 2 µL of 1 M MgCl₂ was added to make a final concentration of 100 mM. Both tubes were incubated at 37 °C for one hour. After one hour, 40% glycerol was added to the solution (1:1 v:v) and the reactions were loaded into a 12% non-denaturing polyacrylamide gel. The gel was run at 4 °C in 1x TBM buffer. The gel was imaged on a Typhoon FLA9000 and scanned at Alexa488 and Alexa647 wavelengths.

In the fifth set of binding reactions, 3% w: v PEG 20,000 was added as a crowding agent. U2 and U6 RNA (50 nM each) were pre-annealed together by mixing them with 100 mM NaCl and 89 mM Tris.HCl pH 7.5. U2 and U6 were then heated at 90 °C for one minute, and 65 °C for ten minutes. 100 mM MgCl₂ was added to the reaction and then the mixture was heated at 37 °C for five minutes. While U2 and U6 were annealing bbRNA17 (50 nM) was mixed in another tube with water, and PEG. The bbRNA17 solution was incubated at 37 °C for ten minutes. The bbRNA17 solution was added to the U2 and U6 mixture and incubated at 37 °C for one hour. After the incubation time was complete, 40% glycerol was added to the solution at 1:1 v: v. The binding reaction was run on a 10% non-denaturing polyacrylamide gel run at 4 °C in 1x TBM buffer. The gel was imaged using a Typhoon FLA9000 and scanned at Alexa488 and Alexa647 wavelengths. The gel bands were quantitated using the associated software at the Alexa488 channel.

6.7.4. Cleavage Assays

The first cleavage test was performed using the RNA outlined in table 6.10 at the concentration indicated in that table.

Name	Sequence	Concentration
bbRNA18	5'-UUU AUA CUA A(2'-5'-GUA UGU AU UG-Dy547) CAC AAU UUU UCA GGU-t-Cy5- 3'	50 nM
5'-Exon	5'-Fluorescein-CAG-3'	100 nM
U5	5'- AUG GUU CUU GCC UUU UAC CAG AAC CAU-3'	100 nM
U2u	5'-GGG AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU CUG UU-3'	100 nM
U6u	5'-GGG GAA ACA AUA CAG AGA UGA UCA GCA GUU CCC CUG CAU AAG GAU GAA CCG UUU UAC AAA GAG AU CCC-3'	100 nM

Table 6.10. RNA used in first cleavage test and final concentration of RNA used in the reaction.

All of the RNAs, bbRNA18, 5'-Exon, U5, U2u, and U6u were mixed together in PK buffer (60 mM sodium phosphates, 3% w:v PEG 4,000, pH 7.0). The reaction was heated at 90 °C for one minute and heated at 65 °C for 15 minutes. The tube was spun down

using a tabletop centrifuge and MgCl_2 was added to a final concentration of 2.1 mM. The reaction was left shaking at ambient room temperature for 16 hours. Gel loading solution (9: 1 formamide: 0.1 M EDTA) was added 1: 1 v: v and the reactions were run on a 20% denaturing (8 M urea) polyacrylamide gel. The gel was run at 0.5x TBE buffer at ambient temperature. The gel was imaged on the Typhoon FLA9000 and scanned at Cy3, Cy5, and fluorescein wavelengths. The individual gel images were false colored and overlaid to generate image 6.14.

After investigation of RNA binding conditions a second cleavage assay was run. bbRNA17 was used along with U2 and U6 RNAs that were labeled with Alexa dyes (table 6.9). For this reaction, U2 (37.5 nM) and U6 (37.5 nM) were preannealed and thus incubated with 89 mM Tris.HCl pH 7.5 and 100 mM NaCl. This solution was heated at 90 °C for one minute and 65 °C for ten minutes. MgCl_2 (100 mM) was added and the reaction was heated at 37 °C for five minutes. While U2 and U6 were annealing, bbRNA17 (25 nM), U5 (37.5 nM), Exon2 (37.5 nM), water and PEG (3% w: v PEG 20,000) were mixed and heated at 37 °C for ten minutes. The solution containing bbRNA17 was added to the U2 and U6 solution. The RNA was left to incubate for 20 hrs at 37 °C. After 20 hours, 1:1 v:v gel loading solution was added (9:1 formamide: 0.1 M EDTA) and the RNA was separated on a 10% denaturing polyacrylamide gel. The gel was run in 0.5x TBE buffer at ambient temperature and imaged on the Typhoon FLA9000. The gel was scanned at Alexa488, and Alexa647 wavelengths.

Name	Sequence	Concentration
bbRNA17	5'-Alexa488-NH-(CH ₂) ₆ -AAC AGA UAC UAA(2'-5' GUA UGU AUU GUU U) CAC AAU UUU – 3'	25 nM
Exon2	5'-UCU AGA AAG-3'	37.5 nM
U5	5'- AUG GUU CUU GCC UUU UAC CAG AAC CAU-3'	37.5 nM
U2	5'-Alexa647-t-GGG AUC UCU UUG CCU UUU GGC UUA GAU CAA GUG UAG UAU CUG UU-3'	37.5 nM
U6	5'-Alexa594-t-GGG GAA ACA AUA CAG AGA UGA UCA GCA GUU CCC CUG CAU AAG GAU GAA CCG UUU UAC AAA GAG AU CCC-3'	37.5 nM

Table 6.11. RNA used in the second and third cleavage assays.

A third cleavage assay three experimental conditions were tried. In tube A, PEG 20,000 was used as a crowding agent. In tube B, double the amount of U6 was used. In tube C, PK buffer was used instead of TMS buffer. The RNAs and concentrations used are in table 6.11. In tube A, U2 (37.5 nM) and U6 (37.5 nM) were mixed with 100 mM NaCl and 89 mM Tris.HCl pH 7.5. In tube B, U2 (37.5 nM) and U6 (75 nM) were mixed with 100 mM NaCl and 89 mM Tris.HCl pH 7.5. In tube C, U2 (37.5 mM) and U6 (37.5 mM) were mixed with PK buffer (60 mM sodium phosphates, 3% w:v PEG 4,000, pH 7.0). All three tubes were heated at 90 °C for one minute, and 65 °C for ten minutes. The tubes were spun down using a tabletop centrifuge and MgCl₂ (100 mM) was added. Then, the tubes were placed at 37 °C for five minutes. While tubes A, B, and C were annealing, three more solutions are made. In tubes D and E, bbRNA17, Exon2, U5, and water were mixed. In tube F, bbRNA17, Exon2, U5, water, and 3% w:v PEG 20,000 were mixed.

These solutions were heated at 37 °C for ten minutes. The bbRNA17 solutions were added to the U2 and U6 solutions. Tube D was added to tube A; E to B; F to C. Then, the reactions were allowed to incubate at 37 °C for one hour. After the hour elapsed, the reactions were split in half, 10 µL of each solution remained at 37 °C for 20 hours, and 10 µL was incubated at ambient room temperature for 20 hours. Once the reaction time elapsed, gel loading solution (9:1 formamide: 0.1 M EDTA) 1:1 v: v with the reactions. The RNA was run on a 10% denaturing (8 M urea) polyacrylamide gel in 0.5x TBE at ambient

temperature. The gel was then scanned using the Typhoon FLA9000 at wavelengths for Alexa488 and Alexa647.

6.8. References

1. Sharp, Phillip, A. Split Genes and RNA Splicing. *Nobel Lect.* (1993).
2. Shi, Y. The Spliceosome: A Protein-Directed Metalloribozyme. *J. Mol. Biol.* **429**, 2640–2653 (2017).
3. Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6498–502 (1993).
4. Madhani, H. D. & Guthrie, C. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71**, 803–817 (1992).
5. Guthrie, C. & Collins, C. A. The question remains: Is the spliceosome a ribozyme? *Nat. Struct. Biol.* **7**, 850–854 (2000).
6. Valadkhan, S. Role of the snRNAs in spliceosomal active site. *RNA Biol.* **7**, 345–53 (2010).
7. Galej, W. P. *et al.* Cryo-EM structure of the spliceosome immediately after branching. *Nature* **537**, 197–201 (2016).
8. Yan, C. *et al.* Structure of a yeast catalytic step I spliceosome at 3 . 6-angstrom resolution. *Science (80-.).* **349**, 1182–1191 (2015).
9. Shi, Y. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat. Rev. Mol. Cell Biol.* **18**, 655–670 (2017).
10. Golas, M. M. *et al.* 3D Cryo-EM Structure of an Active Step I Spliceosome and Localization of Its Catalytic Core. *Mol. Cell* **40**, 927–938 (2010).
11. Deckert, J. *et al.* Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological

conditions. *Mol. Cell. Biol.* **26**, 5528–43 (2006).

12. Bertram, K. *et al.* Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell* **170**, 701–713.e11 (2017).
13. Azubel, M., Wolf, S. G., Sperling, J. & Sperling, R. Three-Dimensional Structure of the Native Spliceosome by Cryo-Electron Microscopy. *Mol. Cell* **15**, 833–839 (2004).
14. Bertram, K. *et al.* Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **542**, 318–323 (2017).
15. Nguyen, T. H. D. *et al.* Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* **530**, 298–302 (2016).
16. Burke, J. E., Sashital, D. G., Zuo, X., Wang, Y.-X. & Butcher, S. E. Structure of the yeast U2/U6 snRNA complex. *RNA* **18**, 673–83 (2012).
17. Valadkhan, S. & Manley, J. L. A tertiary interaction detected in a human U2-U6 snRNA complex assembled in vitro resembles a genetically proven interaction in yeast. *RNA* **6**, 206–19 (2000).
18. Jarrell, K. A., Peebles, C. L., Dietrich, R. C., Romiti, S. L. & Perlman, P. S. Group II intron self-splicing. Alternative reaction conditions yield novel products. *J. Biol. Chem.* **263**, 3432–3439 (1988).
19. Padgett, R. a, Podar, M., Boulanger, S. C. & Perlman, P. S. The stereochemical course of group II intron self-splicing. *Science* **266**, 1685–1688 (1994).
20. Griffiin, E. A., Qin, Z., Michels, W. J. & Pyle, A. M. Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2'-hydroxyl groups. *Chem. Biol.* **2**, 761–770 (1995).

21. Michels Jr., W. J. & Pyle, A. M. Conversion of a Group II Intron into a New Multiple-Turnover Ribozyme That Selectively Cleaves Oligonucleotides: Elucidation of Reaction Mechanism and Structure/ Function Relationships. *Biochemistry* **34**, 2965–2977 (1995).
22. Chin, K. & Pyle, a M. Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection. *RNA* **1**, 391–406 (1995).
23. Qin, P. Z. & Pyle, A. M. Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain I is an independent folding unit with a requirement for specific Mg²⁺ ions in the tertiary structure. *Biochemistry* **36**, 4718–4730 (1997).
24. Valadkhan, S., Mohammadi, A., Jaladat, Y. & Geisler, S. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11901–6 (2009).
25. Guo, Z., Karunatilaka, K. S. & Rueda, D. Single-molecule analysis of protein-free U2–U6 snRNAs. *Nat. Struct. Mol. Biol.* **16**, 1154–1159 (2009).
26. Jaladat, Y., Zhang, B., Mohammadi, A. & Valadkhan, S. Splicing of an intervening sequence by protein-free human snRNAs. *RNA Biol.* **8**, 372–377 (2011).
27. Warnasooriya, C. & Rueda, D. Single-molecule fluorescence-based studies on the dynamics, assembly and catalytic mechanism of the spliceosome. *Biochem. Soc. Trans.* **42**, 1211–8 (2014).
28. Wong, C. C. L. & Shi, Y. Structure of a yeast spliceosome at. 1–16 (2015).
29. Clark, N. E. *et al.* Metal dependence and branched RNA cocrystal structures of the RNA lariat debranching enzyme Dbr1. *Proc. Natl. Acad. Sci.* **113**, 14727–14732

(2016).

30. Desaulniers, J.-P. *et al.* Pseudouridines in rRNA helix 69 play a role in loop stacking interactions. *Org. Biomol. Chem.* **6**, 3892 (2008).

31. Sakakibara, Y. & Chow, C. S. Role of Pseudouridine in Structural Rearrangements of Helix 69 During Bacterial Ribosome Assembly. *ACS Chem. Biol.* **7**, 871–878 (2012).

32. Shu, Y. *et al.* Fabrication of 14 different RNA nanoparticles for specific tumor targeting without accumulation in normal organs. *Rna* **19**, 767–777 (2013).

33. Tseng, C.-K. & Cheng, S.-C. Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* **320**, 1782–4 (2008).

Chapter 7 – Conclusions and Future Directions

During the process of splicing, coding sequences known as exons, are ligated together while non-coding sequences, introns, are removed from pre-messenger RNA (pre-mRNA) to make messenger RNA (mRNA)^{12,13}. The cellular machinery that performs the process of splicing is known as the spliceosome and is comprised of a variety of factors as well as small nuclear RNA (snRNA)^{12,13}. The snRNA make up the catalytic core of the spliceosome and carry out two metal catalyzed S_N2 transesterifications on pre-mRNA to form lariat introns and ligated exons¹². After splicing is complete, the introns are processed by lariat debranching enzyme (Dbr1p) which cleaves the 2'-5'-phosphodiester bond and linearizes the RNA. Introns then go on to participate in a variety of regulatory processes in the cell including snoRNA and mirtron biogenesis^{10,11}. Dbr1p is a unique phosphodiesterase, cleaving only the 2'-5'-phosphodiester bond, unlike other metallophosphoesterases that are less specific¹⁴. Understanding the substrate and non-substrate requirements for catalysis will give us more insight on how Dbr1p functions.

In order to study Dbr1p catalysis *in vitro*, substrate analogues are needed. The Das lab has previously developed a method to synthesize backbone branched RNA that contains a 2'-5'-phosphodiester bond. Solid phase synthesis is used and a branch point residue is made that contains a photolabile protecting group on the 2'-OH. Previously, the protecting group was removed using a UV lamp, which was an inefficient method. Described here is an improved method for photodeprotection using UV-LEDs. The kinetics of photodeprotection are determined for two different experimental setups, one where the 2'-photolabile group is removed while the CPG beads remain inside the synthesizer, and one where the beads are removed and deprotection occurs in a glass vial. The protecting group is removed in ten or six minutes, respectively. The inside synthesizer deprotection

method is used when making mini-lariat RNAs while the outside synthesizer method is used to make backbone branched RNAs (bbRNAs) for *in vitro* Dbr1p biochemical studies.

The optimal buffer conditions for *in vitro* debranching assays are determined. Dbr1p derived from different organisms has different metal requirements for optimal catalysis. The enzyme is still active after incubation with Mg^{2+} , Ba^{2+} , and Ca^{2+} . Activity is hampered or abolished after incubation with Cd^{2+} , Cu^{2+} , and Zn^{2+} . Dbr1p from *E. histolytica* is active after incubation with EDTA, while enzyme from *S. cerevisiae* is not. Also, the amount of reducing agent can change the speed of reaction.

The diversity of RNA cleavable by Dbr1p *in vitro* is much broader than initially anticipated. Backbone branched RNA that contain modifications like 2'-F and 2'-OMe at the first residue in the 2'-arm are rapidly cleavable by Dbr1p. Further, bbRNA that contain an unnatural 2'-3'-phosphodiester bond are cleavable, just at a much slower rate than the natural substrate. Additionally, bbDNA is cleaved. Contrary to previous work, cleavage of the 2'-5'-phosphodiester bond was accomplished when there is no 3'-5'-phosphodiester bond on the branch point residue, a feature previously thought to be required for catalysis¹. Changing the canonical branch point residue from adenosine to uridine creates a substrate still cleavable by Dbr1p but under a different mechanism of action, putatively thought to be substrate inhibition. Further work needs to be done to determine the exact mechanism as a second cleavage is observed after analysis of debranching using gel electrophoresis.

Inhibition of Dbr1p causes an increase in lariat introns in the cell and is linked with a variety of negative health consequences^{2,3}. However, in a few specific instances, HIV and ALS, inhibition of Dbr1p has the potential to be a useful therapeutic⁴⁻⁶. Previous efforts in the literature to make Dbr1p inhibitors utilize complicated synthetic pathways^{7,8}. We have developed analogues of bbRNA that have a 2'-5'-triazole linkage instead of a 2'-5'-

phosphodiester bond. These act as competitive inhibitors and are easier to synthesize. The strength of inhibition can be tuned depending on the size of each arm of the click branched RNA.

Introns play an important role in mirtron biogenesis^{9,10}. The microRNA generated from this pathway then go on to regulate important cellular processes. Taking inspiration from mirtrons, mini-lariat RNAs are synthesized. The mini-lariat RNAs are synthesized on the solid phase, using the inside synthesizer photodeprotection technique established in this thesis. These miRNA mimics are active in the RNA interference pathway in a Dbr1p dependent manner, bypassing Drosha and Dicer processing. Mini-lariats that mimic miR644a, a regulator of the androgen receptor in prostate cancer cells, work on the order of traditional short interfering RNA. Mini-lariats are shown to suppress both mRNA and protein expression. The mRNA repression with mini-lariats was stronger over the course of a week than siRNA. Further, mini-lariat RNAs show increased repression as dose increases unlike siRNA, where the protein repression reverses after a certain concentration.

Finally, this thesis explored the interactions of the spliceosomal RNA core. Truncated snRNAs are used and their strength of interaction with a bbRNA intron mimic is studied using melting experiments. Many different RNA hybrids are investigated and it was observed that adding a 2'-arm or adding a bulged out residue in a duplex confer a similar amount of destabilization. Combining these two features only decrease the thermodynamic stability a small amount compared to only one. Pseudouridines residues can have a stabilizing effect. bbRNAs can associate with U2 and U6 snRNAs. The ribozyme capabilities of the snRNAs was also investigated and cleavage was observed, but not in a consistent repeatable manner.

The catalytic activity of the snRNA bears further investigation. Given the fact that cleavage has been observed in protein-free *in vitro* experiments, different reaction conditions can be tested to develop a repeatable cleavage assay. As the snRNAs described in this thesis are truncated, full length RNAs can be transcribed and tested in the cleavage assay. While the truncated RNAs were carefully designed using two different cryo-EM structures as a guide, eliminating the spliceosomal proteins as well as shortening the RNA may have had compounding negative effects on the repeatability of the experiment. Aside from using full length snRNAs, alternate buffer conditions, more divalent metal ions, longer incubation times, and higher incubation temperatures can all be explored in order to develop a robust, repeatable cleavage assay.

Dbr1p can also be further investigated. It was observed and reported in this thesis that Dbr1p is able to cleave a broader spectrum of RNA than initially believed. However, exactly how this occurs has yet to be determined. The specific interaction between RNA and enzyme can be elucidated through co-crystal structures using a catalytically dead enzyme and modified RNA. In collaboration with Professor Andrew VanDemark at the University of Pittsburgh, we are currently investigating different enzyme mutants. A variety of mutants have been made and are designed to disrupt either the metal binding pockets or specific residues involved in RNA recognition. Out of these mutants, we expect to find one that is catalytically dead and that can be used for crystallization. It would be particularly interesting to see how bbRNA that contain a 2'-3'-phosphodiester bond fit into the active site of Dbr1p. Further, we have synthesized click branched RNA (cbRNA) that can be used in crystallization experiments with active enzyme. Since it is established that cbRNAs are competitive inhibitors, these RNAs should fit into the active site in a similar manner to bbRNA.

This thesis has also laid the groundwork for deeper investigation into mini-lariat RNAs. The synthesis yields are low and have the potential to be improved. The loop size can be reduced and coupling times of both the phosphorylation step and the cyclization step can be extended. The activator can be changed from ETT (5-ethylthio-1H-tetrazole) to 1H-tetrazole or 5-benzylthio-1H-tetrazole. Aside from synthesis improvements, mini-lariats raise some interesting biological questions that can be explored. Since Dbr1p is found in both the cytosol and the nucleus, it is believed that mini-lariats are debranched in the cytosol¹¹. A fluorescence microscopy experiment can be done to confirm this hypothesis. Dye-labeled mini-lariats can be synthesized and incubated with cells. The cells can then be fixed to microscope slides and Dbr1p can be fluorescently probed with antibodies. Co-localization of Dbr1p with the dye labeled mini-lariats will show where in the cell debranching of mini-lariats occurs. Additionally, it is observed that mini-lariats sustain repression of protein expression at high concentrations of mini-lariats, while siRNA of miR44a do not continue to repress past a certain concentration. Detailed mechanistic studies in cells can be carried out to better understand the difference between these two constructs.

The work presented in this thesis advances the understanding of Dbr1p biochemistry and uses Dbr1p as a therapeutic agent, both as a target (inhibitors) and as a prodrug processor (mini-lariats).

7.1. References

1. Ruskin, B. & Green, M. R. An RNA Processing Activity that Debranches RNA Lariats. *Science* (80-.). (1985).
2. Zhang, S.-Y. *et al.* Inborn Errors of RNA Lariat Metabolism in Humans with Brainstem Viral Infection. *Cell* **172**, 952–965 (2018).
3. Han, B. *et al.* Human DBR1 modulates the recycling of snRNPs to affect alternative RNA splicing and contributes to the suppression of cancer development. *Oncogene* 1–10 (2017). doi:10.1038/onc.2017.150
4. Galvis, A. E., Fisher, H. E., Fan, H. & Camerini, D. Conformational Changes in the 5' End of the HIV-1 Genome Dependent on the Debranching Enzyme DBR1 during Early Stages of Infection. *J. Virol.* **91**, 1–11 (2017).
5. Galvis, A. E., Fisher, H. E., Nitta, T., Fan, H. & Camerini, D. Impairment of HIV-1 cDNA Synthesis by DBR1 Knockdown. *J. Virol.* **88**, 7054–7069 (2014).
6. Armakola, M. *et al.* Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nat. Genet.* **44**, 1302–1309 (2012).
7. Katolik, A. *et al.* Fluorescent Branched RNAs for High-Throughput Analysis of Dbr1 Enzyme Kinetics and Inhibition. *ACS Chem. Biol.* **12**, 622–627 (2017).
8. Katolik, A. *et al.* Regiospecific solid-phase synthesis of branched oligoribonucleotides that mimic intronic lariat RNA intermediates. *J. Org. Chem.* **79**, 963–975 (2014).
9. Ruby, J. G., Jan, C. H., Bartel, D. P. Intronic microRNA precursors that bypass Drosha processing. *Nat. Rev. Mol. Cell Biol.* **448**, 83–86 (2007).
10. Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M. & Lai, E. C. The Mirtron

- Pathway Generates microRNA-Class Regulatory RNAs in *Drosophila*. *Cell* **130**, 89–100 (2007).
11. Kataoka, N., Dobashi, I., Hagiwara, M. & Ohno, M. hDbr1 is a nucleocytoplasmic shuttling protein with a protein phosphatase-like motif essential for debranching activity. *Sci. Rep.* **3**, 1090 (2013).
 12. Shi, Y. The Spliceosome: A Protein-Directed Metalloribozyme. *J. Mol. Biol.* **429**, 2640-53 (2017).
 13. Shi, Y. Mechanistic Insights into Precursor Messenger RNA Splicing by the Spliceosome. *Mol. Cell Bio.* **18**, 655-70 (2017).
 14. Matange, N. et al. Metallophosphoesterases: Structural Fidelity with Functional Promiscuity. *Biochem. J.* **467**, 201-16 (2015).