RNA BINDING PROPERTIES OF A TRANSLATIONAL ACTIVATOR THAT ALSO
FUNCTIONS IN GROUP I INTRON SPLICING

by

BEN KASPAR

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science

Thesis Adviser: Dr. Mark Caprara

Department of Biochemistry
CASE WESTERN RESERVE UNIVERSITY

August 2008
We hereby approve the thesis/dissertation of

Ben Kaspar
Candidate for the Master's degree*

(Signed)
Dr. James Bruzik
(chair of the committee)
Dr. Alan Tartakoff
Dr. Pieter deHaseth
Dr. Piet de Boer
Dr. Mark Caprara
(Advisor)
(Date) July 9, 2008

We also certify that written approval as been obtained for any proprietary material contained therein
Table of Contents

List of Figures .................................................................................................................. 5
Preface .............................................................................................................................. 5
Acknowledgements .......................................................................................................... 6
Abstract ............................................................................................................................ 7

Chapter I: Introduction ................................................................................................. 8

Introduction ..................................................................................................................... 9

The role of Pet54p in translation of mitochondrial encoded COX3 mRNA ................. 16

The role of Pet54p in splicing of the group I intron aI5β .............................................. 21

Group I introns ................................................................................................................ 21

PET54 plays a role in splicing of a single group I intron .............................................. 22

Coupling, the “RNP code and the evolution of a dual function protein ...................... 28

Biochemical characterization of Pet54p ........................................................................ 32

Chapter II: A shared RNA binding site in the Pet54 protein is required for translational activation and splicing in the yeast mitochondria ........................................... 34

Abstract ........................................................................................................................ 35

Introduction ..................................................................................................................... 36

Results ............................................................................................................................ 38

Cloning PET54, the 5’ UTL of COX3 and aI5β containing pre-RNA into expression vectors ...................................................................................................................... 38

Pet54p specifically associates with aI5β RNA and facilitates the exon ligation step of splicing in vitro ........................................................................................................... 39

Overlapping binding sites in Pet54p for COX3 5’ UTL and COX1 aI5β containing RNAs ............................................................................................................................. 43

Binding of COX3 5’ UTL and COX1 aI5β pre-RNA to Pet54p ...................................... 44
Mapping the Pet54p binding site in COX3 5’ UTL and COX1 aI5β containing RNAs

Footprinting the Pet54p binding sites on the aI5β pre-RNA and COX3 5’ UTL

Discussion

Translation activation of COX3

Splicing of the aI5β intron

Evolutionary implications

Materials and Methods

Molecular cloning and mutagenesis

Protein expression and purification

RNA transcription purification

uv cross-linking

RNA splicing assays

Delineation of Pet54p binding sites

RNA binding

RNA footprinting

Chapter III: Pet54p interacts with the aI5β intron and COX3 5’ UTL through an unusual RRM motif.

Abstract

Introduction

Results

Binding stoichiometry of the Pet54p protein

Pet54p contains a putative RRM domain with homology to the human Fox-1 RRM
A semi-conserved residue in the Pet54p RRM does not affect the RNA binding properties of Pet54p

Iodosobenzoic acid cleavage of Pet54p

Discussion

A hypothesis regarding the potential role of the RRM in group I intron splicing

Materials and Methods

Molecular cloning and mutagenesis

Generation of a recombinant PET54 mutant

Yeast transformation

uv cross-linking

Binding experiments

Assay of respiratory function

o-Iodosobenzoic acid cleavage

Chapter IV: Summary and Future Directions

Summary

A shared RNA binding site in the Pet54 protein is required for translational activation and group I intron splicing in yeast mitochondria

Pet54p interacts with the aI5β intron and COX3 5’ UTL through an unusual RRM motif

Future Directions

Distinguishing between two models of RNA recognition and evolution

Determining the contribution of the RRM domain to Pet54p affinity, specificity and catalysis of the aI5β intron

Bibliography
List of Figures

**Figure 1-1:** Mechanism of group I intron splicing______________________________25

**Figure 1-2:** Models of group I intron splicing________________________________26

**Figure 1-3:** A hypothesized mechanism for the function of Pet54p in splicing of the aI5β intron and translation of COX3____________________________________31

**Figure 2-1:** Schematic of Pet54p ligands_______________________________________41

**Figure 2-2:** Pet54p binds directly to the aI5β pre-RNA and facilitates exon ligation_______________________________________________________________42

**Figure 2-3:** Specific binding of Pet54p to the COX3 5’ UTL and aI5β pre-RNA_____________________________________________________________46

**Figure 2-4:** Pet54p binds with high affinity to subdomains within the COX3 5’ UTL and the aI5β pre-RNA__________________________________________48

**Figure 2-5:** Mapping of the Pet54p binding site on the COX3 5’ UTL and COX1 aI5β containing RNA__________________________________________49

**Figure 2-6:** Visualization of the Pet54p-binding sites by RNase footprinting____52

**Figure 2-7:** Model for Pet54p function in translational activation___________58

**Figure 3-1:** Binding stoichiometry of Pet54p___________________________________71

**Figure 3-2:** Residues conserved between the Pet54p RRM and the human Fox-1 RRM.________________________________________________________72

**Figure 3-3:** Residues conserved between the Pet54p and Fox-1 RRMs mapped on the NMR structure of Fox-1____________________________________74

**Figure 3-4:** Analysis of Pet54p F241A and F244A mutants______________________78

**Figure 3-5:** o-iodosobenzoic acid cleavage____________________________________79

**Figure 3-6:** Two possible identities for the cross-linked pet54p cleavage fragment mapped onto the Fox-1 RRM structure________________________81

**Figure 3-7:** Two hypotheses regarding the role of Pet54p in splicing of the aI5β intron________________________________________________________85
Preface

The work presented in this thesis dissertation deals with the biochemical characterization of the dual function protein Pet54p from *Saccharomyces cerevisiae*. The entire work was not performed by me alone. The contributions of other lab members to each chapter are outlined below.

The splicing experiments shown in Figure 2-2, B and C were performed by Abby L. Bifano. Dr. Mark Caprara performed the footprinting assays shown in figure 2-6 and well as the RT-PCR experiment shown in Figure 3-4, B. In the conception and interpretation of all experiments herein, I was also greatly assisted by Dr. Caprara.
Acknowledgements

I would like to thank my advisor, Dr. Mark Caprara, for his support throughout my graduate career. This work would not have been possible without his guidance. He has always exhibited great patience.

I would also like to thank the members of my thesis committee, Dr. Alan Tartakoff, Dr. Pieter deHaseth, Dr. James Bruzik and Dr. Piet de Boer. Their advice has been invaluable in helping me find the right direction in this challenging field.

I also owe a debt of gratitude to the past and present members of my lab, Kristina Brady, Maureen Downing, Piyali Chatterjee, Abby Bifano, and Ed Turk. It has been a great experience working in the lab and I have learned an enormous amount from each member.
RNA binding properties of a mitochondrial translation factor from *Saccharomyces cerevisiae* that is also involved in group I intron splicing.

Abstract

by

BEN KASPAR

In *Saccharomyces cerevisiae*, the protein product of the nuclear gene PET54 is required for both translation of mitochondrial COX3 mRNA and splicing of the mitochondrial encoded aI5β intron. Here, we define the protein as a component of two distinct ribonucleoproteins (RNPs). We show that Pet54p binds to both the aI5β intron and COX3 5’UTL with high specificity and affinity and facilitates exon ligation. We also demonstrate that Pet54p binds both substrates with the same or overlapping binding sites through recognition of similar sequences. In a second study, the stoichiometry of Pet54p binding to its natural ligands is determined to be 1:1. Lastly, the specific RNA crosslink formed by Pet54p and its natural substrates was mapped to the putative RNA Recognition Motif (RRM) of the protein. These results suggest that Pet54p recognizes its RNA substrates with an unusually high level of specificity for a single RRM containing protein.
Chapter I:

Introduction
Introduction:

Ribonucleoproteins (RNPs) are complexes of RNA and protein with essential roles in virtually all aspects of RNA metabolism. The importance of RNPs in cellular biology is dramatically demonstrated by disorders of RNP function such as spinal muscular atrophy and retinitis pigmentosa (Meister et al., 2002; Mordes et al., 2007). Our research seeks to address the question of how proteins recognize and interact with their RNA substrates in order to form RNPs. The number of components involved in the function of large RNPs, such as the spliceosome and ribosome, make this question difficult to address in these complexes. For this reason, studies of simple RNPs, consisting of one or two protein factors and a single RNA, can reveal important principles that may not be easily studied in more complex RNPs. Because of the coupled nature of RNA processing events in the eukaryotic nucleus, such simple RNPs are difficult to define in higher eukaryotes. However, within the mitochondria of Saccharomyces cerevisiae, such RNPs are abundant, functioning in processes as diverse as splicing, tRNA maturation and degradation.

Pet54p is a nuclear encoded protein that functions in both splicing and translation within the mitochondria of Saccharomyces cerevisiae. Because both of these processes frequently rely on RNA/Protein interactions, we hypothesized that Pet54p exerts its functions through RNP formation. Consistent with this hypothesis, we felt that biochemical analysis of Pet54p in the presence of its putative RNA substrates would generate significant contributions to our primary objective of elucidating the mechanisms of RNP assembly and add to our current understanding of mitochondrial biology in Saccharomyces cerevisiae.
In the last decade, the study of RNP complexes has resulted in a number of important breakthroughs. For example, structural and biochemical studies of the prokaryotic ribosome showed conclusively that RNA was responsible for catalysis of peptide bond formation and demonstrated in unprecedented detail the structural basis of codon recognition and GTP hydrolysis in the A site and stabilization of tRNA binding in the P site of the RNP (Schuwirth et al., 2005; Selmer et al., 2006). Yet crystallization experiments cannot answer the question of how such a complex assemblage of RNA and protein is brought together to produce a functional enzyme. Genetic studies and advances in affinity purification methods in \textit{Saccharomyces cerevisiae} revealed that the eukaryotic ribosome undergoes an ordered assembly process requiring at least 150 non-ribosomal proteins (Zemp and Kutay, 2007). Some of these proteins bind to specific ribosomal RNA sequences to facilitate formation of appropriately folded RNA structures. Others are RNA chaperones. Chaperones are factors defined by their ability to prevent the formation of misfolded RNAs or to unfold those RNAs trapped in non-productive conformations (Schroeder et al., 2004). RNA helicases also comprise a fraction of the proteins required for proper assembly. These proteins are thought to use ATP hydrolysis to facilitate structural changes in RNAs and displace nucleic acid bound proteins (ibid).

The presence of so many RNA-interacting proteins underscores the importance of appropriate ribosomal RNA (rRNA) folding throughout the ribosomal assembly pathway (Herschlag, 1995; Lorsch, 2002). By reducing misfolded intermediates, these factors play a vital role in decreasing the energy of activation required for ribosome formation (ibid). Currently, very little is known about how these specific RNA binding proteins, chaperones and helicases interact with their rRNA substrates and each other. In
particular, determining the specificity and function of individual factors has been an ongoing challenge.

Recently, *in vitro* biochemical studies of small spliceosomal RNAs revealed that the spliceosome, like the ribosome, likely possesses a catalytic domain composed of RNA (Valadkahn and Manley, 2001). Yet the two transesterfication reactions that define the splicing reaction cannot occur on a biologically relevant timescale without proteins. Indeed, each of the five small nuclear snRNAs (snRNAs) involved in the reaction are associated with proteins. These small nuclear ribonuclear proteins (snRNPs) are responsible for the complex series of RNA-RNA and RNA-protein rearrangements necessary for splicing to take place. However, snRNPs do not act alone. In fact, they represent only a small fraction of the more than 200 proteins implicated in spliceosomal assembly. Among these are the hnRNP proteins, which bind RNA through a conserved fold known as the RNA Recognition Motif (RRM). Frequently, these proteins behave much like ribosomal chaperone proteins by preventing the formation of potentially inhibitory RNA structures (Krecic and Swanson, 1999). Another group of RRM containing proteins, known as SR proteins, have also been shown to bind RNA. These factors have been implicated in a variety of functions from RNA annealing to protein recruitment (Krainer et al., 1990; Valcarcel and Gebauer, 1997).

Individual spliceosomal components have been studied in a number of ways. Depletion and rescue experiments using nuclear extracts elucidated the involvement of protein factors in splicing of specific mRNAs (Wang et al., 1998). Cross-linking studies have been used to identify which RNAs are bound by spliceosomal proteins (Sawa and Shumura, 1992). Furthermore, the rescue of protein deletion phenotypes through
mutations in snRNAs or pre-mRNA transcripts has suggested specific functions for some spliceosomal factors (Siatecka and Reyes, 1999). While these experiments are powerful, they say very little about the spliceosome as a dynamic complex. Much like ribosomal assembly, splicing requires multiple rearrangements of RNA and protein components and is not amendable to study in vitro. Thus, while progress has been made in assigning functions to individual spliceosomal components, the dynamics of the interactions between these components remain largely unknown.

One approach to the study of complex RNPs such as the spliceosome and ribosome has been to step backward and explore systems that began with RNA alone and evolved, in some tractable way, to rely on proteins for efficient function. The best examples of such systems are formally autocatalytic RNAs that have grown dependent on protein co-factors for appropriate folding and catalysis (Lambowitz et al., 1999; Smith et al., 2007). While this may seem at first like a foray into the biological past, these simple RNPs are far from vestigial in many organisms. In the mitochondria of many lower eukaryotes, RNPs, sometimes consisting of a single RNA and protein, are vital components of the cellular machinery.

Much of our current understanding of RNA/Protein interactions is informed by mobile genetic elements that parasitize the mitochondrial genomes of fungi. Particularly well studied are the group I introns: stretches of selfish DNA frequently found disrupting the open reading frames (ORFs) of mitochondrial genes. The RNAs encoded by group I introns possess the potential to be excised from host mRNA, thereby mitigating the deleterious effect of the selfish gene on host ORF expression. How they realize this potential, however, can differ dramatically from intron to intron. Some are capable of
autocatalysis, while others have evolved to rely on protein factors encoded by open reading frames within them. In some cases, these parasitic elements may even develop a dependence on factors donated by the nucleus of the host organism. As in ribosomal assembly and spliceosome function, these factors include RNA helicases and RNA chaperones (ibid; Caprara et al., 1996; Ho and Waring., 1999).

In many cases, group I introns and their cognate splicing factors are amenable to exploration in vitro. For example, the RNA helicase CYT-19 is a member of the DEAD box family of RNA helicases involved in splicing of the Neurospora group I intron LSU (Mohr et al., 2002). In eukaryotes, ribosome biogenesis requires at least nine helicases from this same family while mRNA splicing requires at least four, making it difficult to define the role of individual DEAD box helicases in these processes (Bleichert and Baserga, 2007). In contrast, CYT-19 requires a single additional factor, CYT-18, to facilitate splicing of its target group I intron (Mohr et al., 2001). Thus all the necessary components for splicing can be studied in vitro by generating recombinant CYT-18 and CYT-19 protein, and transcribing the RNA of the Neurospora LSU intron. In this way it was determined that CYT-19 disrupts non-native RNA structures subsequent to CYT-18 binding, thereby lowering the energy of activation required for the splicing reaction. Furthermore, CYT-19 did not appear to bind specifically to any particular RNA sequence or structure (Mohr et al., 2002).

Such studies allow researchers to refine their questions about the function of similar proteins in more complex RNPs. For example, it would be difficult to explore the functional specificity of a particular helicase in the context of the spliceosome because of the number of RNA and protein factors involved as well as the dynamic nature of the
RNP. Yet with a single group I intron RNA and one additional co-factor, it is possible to
determine if the helicase acts on particular sequences, structures, or even steps in the
splicing pathway. Thus, the comparative minimalism of group I intron systems presents
the potential for well controlled experiments, and what has arisen from these experiments
is a deeper understanding of the ways in which proteins can interact with RNA to
accomplish biological functions.

While the study of fungal group I intron RNPs has improved our knowledge of
RNA/protein interactions, many of the RNPs that form in fungal mitochondria remain
poorly understood. Early mutagenesis experiments in *Saccharomyces cerevisiae* revealed
that most of the factors responsible for appropriate mitochondrial function are encoded in
the nucleus (Costanzo et al., 1986). Many of these were characterized as components of
the respiratory complex. However, mitochondrial transcription experiments revealed that
a subset of nuclear mutants eliminated expression of the respiratory complex proteins
encoded within the mitochondria (Towpik, 2005). Northern blots using respiratory RNA
specific probes indicated that mRNA levels were largely the same as wild type in these
mutant backgrounds. Thus, it became apparent that the genes responsible for regulation
of mitochondrial respiratory expression operate largely at the level of mRNA processing
(ibid).

Further northern blot experiments showed that many of these putative RNA
processing genes encoded splicing factors for mitochondrial introns (Krieke et al., 1986).
Others prevented translation even in the presence of fully processed mRNAs (Cabral and
Schatz, 1978; Costanzo et al., 1986). These genes appeared to function as a bridge
between RNA processing and translation. Remarkably, it was subsequently demonstrated
that each of the seven RNAs encoding polypeptide chains synthesized in the mitochondria rely on specific translation factors (Towpik, 2005). These genes require sequences within the 5’ untranslated leader sequences (UTLs) of their target RNAs to facilitate translation.

Because the proteins involved in mitochondrial RNA translation require specific mRNA sequences for function, it is parsimonious to suggest that these factors form RNP complexes to facilitate translation. Yet the genetic studies suggest that the RNPs potentially formed by these factors would be quite different from the large RNPs that facilitate RNA processing in the nucleus. Rather than acting on a broad group of substrates through shared machinery as the spliceosome and ribosome do, these mutually exclusive factors appear to act only on their target RNAs to achieve the same function. In this way, the yeast mitochondrial translation factors appear to resemble the nuclear encoded proteins involved in group I splicing, which are generally specific for their target RNAs and accomplish the same function without shared machinery.

Unfortunately, it has been very difficult to prove that the proteins involved in translation of mitochondrial mRNAs act as RNPs. Solubility studies have indicated that most of these factors are integral membrane proteins associated with the inner mitochondrial membrane. For this reason, mitochondrial translation has not been studied through cellular extracts in the manner of the spliceosome, nor have its individual components been reconstituted from recombinant proteins to form functional complexes in the manner of group I intron studies. Thus, direct association of RNA with translation factors remains to be demonstrated.

Of the known mitochondrial translation factors, only one is not an integral
membrane protein. Remarkably, this protein, known as Pet54p, functions both in the translation of the COX3 protein as well as the splicing of the α5β group I intron. Through the study of this unique protein, some of the current experimental challenges in the study of yeast mitochondrial translation factors might be overcome. Particularly, it may be possible to determine the types of interactions that take place between this translation factor and its native RNA substrate, the COX3 5' UTL. Such information would be an important step toward a mechanistic understanding of translation in the yeast mitochondria. Furthermore, because Pet54p is also a specific group I intron splicing factor, it may be characterized through binding and splicing assays similar to those described for CYT-19. Such studies are critical for determining the principals of RNA/protein interactions that result in the structural changes necessary for catalysis in both simple and complex RNPs. Lastly, comparison of Pet54p's functions in α5β splicing and translation of the COX3 UTL may provide information that is relevant to the evolution of multiple functions among individual RNA processing factors, such as the spliceosomal SR and hnRNP proteins.

The role of the PET54 gene in translation of the mitochondrial encoded COX3 mRNA

The PET54 gene was originally identified in a genetic screen for mutations in the nuclear genome of *Saccharomyces cerevisiae* that resulted in loss of mitochondrial respiration (Costanzo et al., 1986). As a direct consequence of this inability to respire, strains carrying these mutations can no longer grow on non-fermentable media. Thus, the PET designation derives from the characteristically “petite” appearance of these mutants.
when grown under low glucose conditions (Tzagoloff and Diekmann, 1990). Yeast strains possessing this phenotype contain deleterious mutations primarily in genes critical for oxidative metabolism, for example, subunits within the respiratory complex or proteins essential for functional expression of this complex (Towpik, 2005).

Additional analysis revealed that the PET54 strain lacked detectable cytochrome c oxidase activity (Costanzo et al., 1986). Subsequent experiments involving radiolabeling of mitochondrial translation products showed that a major component of this complex, the cytochrome c oxidase subunit 3 (COX3) protein, was not present in the PET54 strain. The absence of this protein explains the respiratory phenotype. One of three proteins contributed to the multi-subunit cytochrome c oxidase complex by the mitochondrial genome, COX3 is an essential component of the holoenzyme active site (Herrmann and Funes, 2005). This inability to accumulate COX3 occurred despite the observation that the COX3 mRNA was present at wild type levels within the mitochondria, suggesting that the absence of COX3 protein was the result of defects in the translation of COX3 mRNA (Costanzo et al., 1986).

A phenotype similar to that observed in the PET54 strain had already been reported in another respiration deficient mutant, PET494 (Cabral and Schatz, 1978). In this case, the absence of COX3 translation could be suppressed by mitochondrial genome fusions that replaced the 608 bp 5’ Untranslated Leader Sequence (UTL) of the COX3 mRNA with the 5’ UTLs of other transcripts (Costanzo and Fox, 1986). These same suppressors were found to restore respiratory function in the PET54 mutant strain suggesting that the two proteins may act together to facilitate COX3 translation. A third strain, PET122, was later implicated in translation of the COX3 UTL through a similar
screen (Fox et al., 1988).

These studies demonstrated that three nuclear encoded protein factors, PET54, PET494, and PET122, facilitate mitochondrial translation of COX3. Furthermore, suppressor mutations established a role for the 5’UTL of COX3 in this process. To tie these two observations together, Costanzo and Fox (1988) employed a creative genetic approach. Their work took advantage of the observation that other mitochondrial mRNAs also required unique translation factors. Knowing that their previously identified suppressors of COX3 translation factor mutations contained fusions of 5’UTLs from other mitochondrial transcripts, they sought to determine if the fusion of the COX3 UTL with another yeast mRNA would result in effective translation of that mRNA in the absence of its wild type co-factors. To this end, a strain was generated that contained a deletion of the Cob translational activator CBS1. This deletion eliminates mitochondrial respiration by abolishing translation of cytochrome b (Rodel and Fox, 1987). This enabled the authors to search for suppressor mutations that contained the COX3 5’UTL at the 5’end of the COB gene.

This suppressor screen was successful. The authors identified a strain capable of respiration in the absence of CBS1 and demonstrated that this strain contained a mitochondrial gene fusion containing the first two-thirds of the COX3 5’UTL fused to the cytochrome b transcript. Deletion of PET54, PET122 or PET494 all resulted in loss of suppression. These experiments demonstrated that the COX3 translation factors and sequences or structures within the COX3 5’UTL are required to facilitate translation.

Thus, a role for PET54 was defined genetically. Working either independently, or in conjunction with co-factors PET122, and PET494, the protein facilitates translation of
COX3 in a manner that also requires the COX3 5’ UTL. What is missing from this analysis is any evidence for direct interaction between Pet54p and its associated cofactors, or the COX3 5’ UTL.

In yeast studies, there is a trajectory of understanding often followed in the characterization of a biological process. This path traditionally begins with a genetic approach leading to determination of genes essential for some function that can be selected for or against. Further definition requires tools, offered by molecular biology, that allow the researcher to explore the physical interactions of proteins encoded by the essential genes.

For PET54, and its functionally associated co-factors PET122 and PET494, the first work to address the physical properties of these proteins explored interactions between the translation factors and the inner mitochondrial membrane (McMullin and Fox, 1993). As an initial approach, this made good sense for two reasons. First, it is experimentally feasible. Mitochondria can be purified and associations between proteins and membranes can be distinguished by treatment with alkaline carbonate, which solubilizes peripheral membrane proteins, but leaves integral membrane proteins in the membrane. Secondly, it begins to address an important biological question. COX3 is an integral membrane protein, thus, it must be inserted appropriately in the membrane to function properly. Furthermore, as a component of a complex holoenzyme, COX3 must be translated as part of an intricate, temporally controlled assembly process along with a number of other nuclear and mitochondrial encoded co-factors (Fontanesi, 2006). For these reasons, it is likely that the COX3 mRNA must be localized to the inner mitochondrial membrane, where COX3 is inserted and assembled, before it can be
inserted and assembled.

Working within this framework, researchers hypothesized that the three translation factors might play a role in localizing the COX3 mRNA to the appropriate region of the IMM to facilitate insertion and assembly (McMullin and Fox 1993). Such localization could conceivably occur through the interaction of one or more of these factors with the inner mitochondrial membrane. To test this, experiments were conducted to isolate mitochondrial membranes and probe them for the presence of the translation factors. Both Pet122p and Pet494p were found solely in the fraction associated with the inner mitochondrial membrane. Furthermore, this association could not be disrupted by the addition of alkaline carbonate, which, suggests that Pet122p and Pet494p are integral membrane proteins (McMullin and Fox, 1993).

The pattern of localization for Pet54p however, was unique. The protein was present in both the IMM pool as well as a soluble fraction. This suggested that the protein was peripherally associated with the membrane. Consistent with this localization pattern, Pet54p could be extracted with alkaline carbonate. These results supported the hypothesis that COX3 translation factors localize the COX3 mRNA to the IMM.

Co-Immunoprecipitation and two hybrid experiments have also been performed and these demonstrated specific interactions between the three COX3 specific translation factors. These experiments also showed an interaction between Pet122p and the mitochondrial ribosome (Naithani et al., 2004).

The experiments above indicate that, despite numerous challenges, progress has been made in understanding the role of specific translation factors in the yeast mitochondria. The current model suggests that these factors interact with one and other,
the inner mitochondrial membrane, the COX3 mRNA, and the ribosome. Yet this model is profoundly static. There exists no mechanistic or biochemical data to enhance our knowledge of the significance of these interactions. There are, of course, numerous technical obstacles that must be overcome if this subject is to be probed in more detail. The elements of this machinery cannot be reconstituted and studied *in vivo*. In addition, the proteins themselves are highly unique, providing little evidence of homologous functional domains for mutagenesis studies. Yet the scientific questions at the heart of this matter are important and fascinating.

**The role of PET54 in splicing of the group I intron a15β:**

a. Group I introns

While the mitochondria of all organisms require efficient mRNA translation, the organelles of fungi and plants must also contend with the transcription products of numerous parasitic genes. For instance, the mitochondrial genome of *S. cerevisiae* has been invaded by a number of mobile genetic elements encoding catalytic RNAs. When transcribed, these chimeric pre-mRNA transcripts must have their intervening sequences, called introns, removed through a process called splicing. Splicing ensures that the genetic disruption caused by mobile genetic element insertion is not translated by the ribosome. Because these elements typically disrupt essential genes, splicing is critical to the survival of the host and, by extension, the parasitic gene itself. Mitochondrial introns fall into two groups based upon their mode of catalysis as well as conserved structural features. Group I introns splice via two sequential transesterification reactions the first of
which requires an exogenous guanosine molecule (Figure 1-1). The result of these reactions is ligation of the RNA regions, called exons, upstream and downstream from the intron thus restoring the RNA transcript to its pre-invasion state. Group II introns also utilize two transesterfication reactions to accomplish splicing; however, the mechanism of splicing requires no exogenous guanosine, relying instead upon a bulged adenosine within the intron to initiate the reaction.

Within each class of intron, facilitation of splicing can occur in numerous ways as outlined in Figure 1-2. In the simplest case, the introns can self-splice. This occurs when the intron is capable of folding into the appropriate conformation for autocatalysis independent of protein co-factors. In more complex examples, introns may grow to be dependent on a variety of proteins to facilitate proper folding and catalysis. In either case, these elements must be excised efficiently as any harmful decrease in functional host mRNA would cause rapid elimination of these selfish genes. Indeed, the remarkable persistence of these elements suggests that most introns cause little or no detriment to their host organisms. This persistence is not, however, due completely to the mobile genetic elements themselves as host contributed proteins often play a major role in efficient splicing.

b. The PET54 gene plays a role in splicing of a single group I intron

When the PET54 deletion phenotype was initially characterized, it was observed that some strains presented defects in the translation of cytotochrome oxidase subunit I (COX1) as well as COX3 (Kloeckener-Gruissem et al., 1987; Valencik et al., 1989). Interestingly, northern blots revealed an accumulation of intron containing precursor
RNAs and a corresponding decrease in mature COX1 mRNA in these strains. This surprising result suggested that mutations in PET54 could also cause splicing defects within the COX1 gene (Valencik et al., 1989). The phenotype of the splicing deficiency closely mirrors that of the COX3 translation defect in that respiration is lost and the COX1 protein fails to accumulate (ibid).

In the strain of yeast employed in this experiment, the COX1 mRNA contained eight introns. Of these introns, two were “optional,” meaning, they did not occur in all laboratory strains (Kloeckener-Gruissem et al., 1987). The size of the precursor that formed in the PET54 deletion strain suggested that one of these two optional introns was retained in the COX1 mRNA. Northern blots subsequently revealed that the retained intron was α15β, a group I intron between exons 5 and 6 of the COX1 ORF (Valencik et al., 1989).

To understand how the protein encoded by PET54 might facilitate group I intron splicing, it is useful to examine group I intron splicing in more detail, starting with autocatalysis and building in complexity to multi-protein systems.

Self-splicing group I introns require no proteins at all to be excised from transcripts. These introns are capable of folding into the appropriate conformation to form a ribozyme active site without protein co-factors. It is important to note that the difference between folding and catalysis in a group I intron is minimal. The two sequential transesterification reactions that result in ligation of exons follow directly from the formation of a complex three-dimensional RNA structure (Woodson, 2005). At its core, this structure consists of two domains referred to as P4-P6 and P3-P9 (Cech and Herschlag, 1996; Hougland et al., 2005). This nomenclature reflects the importance of
paired (P) regions in group I intron structure. The domains are essentially groups of RNA helices that must fold and stack in a defined conformation to facilitate catalysis. In addition, most introns possess peripheral structures that facilitate folding and add additional stability to the core structure (ibid; Woodson, 2005). Self-splicing group I introns contain all the necessary information for active ribozyme formation in their RNA. Thus, the study of self-splicing introns is essential to our understanding of group I intron containing RNPs.

Experimental data suggests that the P4-P6 domain folds rapidly while formation of P3-P9 occurs more slowly (Hougland et al., 2005; Woodson, 2005). This indicates that the RNA forms stable intermediates while folding. These stable intermediates resolve rapidly in some introns, while others may remain trapped in inactive conformations for long periods of time. If the intron remains trapped in a conformation long enough to prevent splicing on a biologically relevant timescale, this deficiency will result in loss of gene expression. One can envision an equilibrium between inactive and active conformations. In efficiently self-splicing introns, this equilibrium may be shifted toward the active conformations by the presence of peripheral RNA structures. Those introns that cannot undergo efficient self-catalysis independently generally rely on protein co-factors.

Because the properties of group I intron-splicing co-factors are so diverse, the function of PET54 is difficult to predict from genetic data. It is implicated in splicing of a single group I intron, but this specificity is not unique. Many introns encode highly specific splicing factors, called maturases within open reading frames that are spliced out along with the catalytic RNA (Solem et al., 2002; Longo et al., 2005; Szczepanek 2006). These proteins exist as a component of the mobile genetic element and often recognize
Figure 1-1. Mechanism of group I intron splicing. (A) Prior to endogenous guanosine binding, the 5’ exon pairs with the intron to form the P1 helix. Subsequently, an endogenous guanosine binds within the G-binding site (B) leading to nucleophilic attack at the 5’ splice site (C). This first transesterification reaction results in a conformational change freeing the guanosine from the G binding site (D). This opens the G-site for occupation by the ΩG (E) allowing the second transesterification reaction to take place and facilitating formation of the ligated exon (F). Adapted from Hougland et al, 2005.
Figure 1-2. Models of group I intron splicing. The ways in which group I introns are spliced within host organisms are detailed above. Within the green box, autocatalytic splicing is depicted. This is hypothesized to be the original state of all group I introns. Over evolutionary time, introns may remain autocatalytic (blue box) or become dependent on proteins for splicing. Often, as shown in the tan box, these proteins are encoded within the introns themselves. Such proteins, known as maturases are frequently specific for the intron in which they are encoded. In other cases, introns may evolve to require factors contributed by the host nucleus (pink box). Nuclear encoded factors may act as general co-factors. However, some of these factors act specifically on individual introns.
idiosyncratic features of their cognate introns. Remarkably, even among the known nuclear encoded splicing factors, high substrate specificity for a single intron is not an unusual trait. Yeast nuclear proteins such as CBP2 have also been shown to recognize and splice single introns (Shaw et al., 1996; Shaw and Levin 1997).

What is unique about PET54 is its dual function in translation and splicing. The optional nature of the intron targeted by PET54 suggests that splicing is a secondary adaptation. This raises interesting questions as to how the role of the protein in splicing differs from translation. The αl5β intron itself is also unusual in its protein requirements. At least five factors have been shown genetically to be involved in splicing of this intron, a surprisingly large number considering that most group I introns appear to require one or two factors for efficient splicing. This is intriguing as it suggests that protein/protein interactions may also be important in facilitating αl5β splicing.

The complexity of the intron as an RNP further broadens the possible roles for PET54 in splicing. Conceivably, the protein may be acting by binding directly to the RNA. It could do this though transient interactions, in the manner of a chaperone, by preventing misfolding events along the pathway to a catalytic structure, or it may act by binding tightly to the RNA in a way that stabilizes the active conformation of the intron. Alternatively PET54 may facilitate splicing though protein/protein contacts, for instance, by recruiting other splicing factors to the RNP complex. It is also possible that both types of interactions may be critical for efficient splicing if, for example PET54 functions to stabilize an RNA structure through direct binding and recruitment of additional proteins involved in other aspects of the splicing pathway.
Coupling, the “RNP code” and the evolution of a dual function protein

The ascendancy of complex regulatory networks in RNA metabolism is a fascinating and important area of molecular evolution. Indeed, it is one of the great wellsprings of eukaryotic diversity. It has been suggested that the fate of nuclear mRNAs during post transcriptional processing is dictated by an “RNP code,” meaning, it is the aggregate of mRNA and proteins bound that determine the outcome of the transcribed RNA (Singh and Valcarcel, 2005). Frequently, small RNA binding proteins, such as those in the SR and hnRNP families, play important roles in defining this code (Valcarcel and Gebauer, 1997). As an mRNA goes through various stages of processing, the protein complement bound to it changes: new factors bind, others leave, and some remain throughout. This ensures that each processing event is “coupled” to the preceding one through the presence of shared factors (ibid).

It seems increasingly likely that the yeast mitochondria have an RNP code of their own. Just as eukaryotic mRNAs require unique protein components to recruit the spliceosomal machinery, mitochondrial mRNAs require specific factors, unique to each transcript, to facilitate translation. How this RNP code facilitates recruitment of the ribosomal machinery is unknown.

Remarkably, many of the factors responsible for nuclear mRNA regulatory networks share similarities with Pet54p. They are often small proteins implicated in RNA binding as well as protein/protein interactions. Furthermore, they usually act as components of larger RNP complexes, ensuring specificity. Even more compelling are the functional similarities between these proteins. Like the hnRNP and SR proteins described above, Pet54p appears to function in multiple steps throughout the lifetime of
an mRNA. Because one of these steps is clearly a secondary function, Pet54p could provide a rare and unique window into the evolution of RNA processing components.

If PET54 exerts its function in splicing directly through protein/protein interactions, it does not appear to do so in complex with the translation factors encoded by PET122 and PET494. Though some evidence suggests that group I intron splicing is a co-translational process (Sandegren and Sjoberg, 2007), deletion of PET122 and PET494 does not affect splicing of a15β. Given that a15β splicing requires at least five distinct protein co-factors, it is certainly possible that PET54 recruits a second set of proteins involved in splicing. Such a mechanism would be similar to that observed with SR and hnRNP proteins which will recruit different sets of factors depending on the surrounding molecular context. The recruitment of different sets of proteins may also reflect the dual pattern of Pet54p localization. Data suggests that separate pools of the protein may be involved in splicing and translation, as one fraction of the protein is membrane associated while the other appears to be free in the cytosol (McMullin and Fox 1993). Figure 1-3 shows one hypothesis for the distinct localization pattern of Pet54p. In this scenario, the cytosolic fraction of Pet54p participates in a15β splicing through interactions with the intron as well as additional cytosolic co-factors. The membrane bound fraction, acting on the COX3 5’UTL, interacts with Pet122p and Pet494p to facilitate translation.

If the assumption is made that Pet54p, like the hnRNP and SR proteins, interacts directly with RNA, there are two evolutionary pathways by which the dual functions of the protein may have arisen. The first possibility is that the protein has evolved a binding surface to accommodate the a15β intron and that this binding surface is separate from the one employed in translation. Precedent for such an evolutionary pathway may be seen in
the *Neurospora crassa* tryosyl-tRNA synthetase CYT-18. This protein functions specifically to aminoacylate tRNA\textsubscript{tyr} and as a general splicing factor of group I introns. Crystal structures revealed that CYT18 facilitates splicing though an extended binding surface distinct from that employed in aminoacylation of tRNA (Paukstelis et al, 2008). The maturase I-\textit{AniI} also facilitates splicing though a binding surface independent of that employed in its primary function as a DNA cleaving enzyme (Chatterjee et al., 2003).

It is also possible that the separate roles of PET54 in translation and splicing are accomplished through recognition of both ligands by the same RNA binding surface. This would be a highly novel result as such a mechanism has never been convincingly demonstrated for a group I intron splicing factor. While not common in group I intron splicing factors, a mechanism by which multiple functions are accomplished through the same RNA binding specificity would be consistent with observations of nuclear RNA processing components. Many RNA binding proteins involved in these functions accomplish distinct tasks while utilizing the same RNA binding domain. The SR and hnRNP families of spliceosomal proteins, for example, can have varied effects on splicing, import and degradation while employing the same binding site to interact with RNA (Valcarcel and Gebauer, 1997; Sanford et al, 2005).

While little is known as to how PET54 stimulates translation or splicing, different regions of the protein appear to be critical for each process. Research conducted by Valencik and McEwen (1991) suggests that the roles of PET54 in translation and splicing are genetically separable. The authors made a series of insertion mutations encoding four amino acids throughout the PET54 coding sequence. These mutations were transformed into a PET54 deletion strain with variant of the COX1 gene containing the aI5\textbeta intron.
Figure 1-3. A hypothesized mechanism for the function of Pet54p in splicing of the α15β intron and translation of COX3. A membrane associated fraction of Pet54p interacts with Pet494p, Pet122p, and the COX3 5′UTL to facilitate translation of the COX3 protein. Splicing of the α15β intron is accomplished by a soluble cytoplasmic fraction of Pet54p through associations with a separate set of proteins.
The various PET54 mutants were subsequently analyzed for the accumulation of COX1 and COX3 proteins as a measure of splicing and translation respectively. The authors observed that certain mutations singly affected splicing or translation. This would seem to indicate that some regions or domains within the protein are critical for both splicing and translation while other regions function specifically for a single task. Thus there are areas of overlapping function and areas of independent utility.

The results of the mutagenesis experiment cannot be used to rule out either hypothesis regarding PET54's potential role as an RNA binding protein. They do, however, suggest that the functions are not completely overlapping. On the surface, this would seem to favor the hypothesis that PET54, like CYT-18, accomplishes separate tasks by utilizing different RNA binding surfaces. However, unlike CYT-18, PET54 appears to engage in protein/protein interactions with other translation factors. Thus, one must also consider the hypothesis that PET54 employs separate domains to recruit different sets of proteins for facilitate splicing and translation. In this case, the protein could employ the same binding site to interact with both RNAs.

**Biochemical characterization of Pet54p**

Our own work has dealt primarily with the biochemical characterization of Pet54p *in vitro*. In chapter II, we explored the question of RNA specificity, demonstrating that Pet54p bound with high affinity to each of its natural ligands, but not to truncated RNA mutants. We also demonstrated that Pet54p specifically facilitates the second step of α15β splicing, a function unique among known group I intron splicing factors. We then showed through cross-linking competition experiments that Pet54p recognizes both α15β and the
COX3 UTL through a similar or overlapping binding site. We also characterized the RNA sequences in aI5B and the COX3 UTL recognized by Pet54p.

In chapter III, we began to characterize the RNA binding region of the protein. We demonstrate that a specific RNA/protein cross link, formed in the presence of the COX3 UTL or aI5β intron, maps to that of the putative RNA recognition motif of Pet54p.
Chapter II

A shared RNA binding site in the Pet54 protein is required for translational activation and group I intron splicing in yeast mitochondria

Benjamin J. Kaspar, Abby L. Bifano and Mark G. Caprara*
Abstract:

The Pet54p protein is an archetypical example of a dual functioning (“moonlighting”) protein: it is required for translational activation of the COX3 mRNA and splicing of the α15β group I intron in the COX1 pre-mRNA in *S. cerevisiae* mitochondria (mt). Genetic and biochemical analyses in yeast are consistent with Pet54p forming a complex with other translational activators that, in an unknown way, associates with the 5’ untranslated leader (UTL) of COX3 mRNA. Likewise, genetic analysis suggests that Pet54p along with another distinct set of proteins facilitates splicing of the α15β intron, but the function of Pet54 is, also, obscure. In particular, it remains unknown whether Pet54p is a primary RNA binding protein that specifically recognizes the 5’ UTL and intron RNAs or whether its functional specificity is governed in other ways. Using recombinant protein, we show that Pet54p binds with high specificity and affinity to the α15β intron and facilitates exon ligation *in vitro*. In addition, Pet54p binds with similar affinity to the COX3 5’ UTL RNA. Competition experiments show that the COX3 5’UTL and α15β intron RNAs bind to the same or overlapping surface on Pet54p. Delineation of the Pet54p binding sites by RNA deletions and RNase footprinting show that Pet54p binds across a similar length sequence in both RNAs. Alignment of the sequences shows significant (56%) similarity and overlap between the binding sites. Given that its role in splicing is likely an acquired function, these data support a model in which Pet54p’s splicing function may have resulted from a fortuitous association with the α15β intron. This association may have led to the selection of Pet54p variants that increased the efficiency of α15β splicing and provided a possible means to co-regulate COX1 and COX3 expression.
Introduction:

Group I introns are mobile genetic elements that have colonized diverse sets of organisms including bacteria, bacteriophage and unicellular eukaryotes (Haugen et al., 2005). As one example, the mitochondria (mt) of Saccharomyces cerevisiae (S.c.) contain up to nine group I introns that are found within the cytochrome oxidase (COX) 1, apocytochrome b (COB) and large ribosomal RNA genes (Foury et al., 1998). While some group I introns can self splice in vitro, most are believed to require proteins for splicing in vivo. In fact, genetic analysis in fungi has shown that a number of nuclear encoded mt proteins are required for group I intron splicing. In most cases, individual proteins show high substrate specificity and facilitate splicing of only one intron. Interestingly, in a few specific cases, these proteins have additional mt functions, such as aminoacylation or translational activation, and splicing function likely reflects the secondary adaptation of each protein (Lambowitz et al., 1999). Thus, investigations of such group I intron co-factors provides a rich opportunity to understand details of molecular evolution that may be pertinent in the selection of other multifunctional “moonlighting” proteins in more complex organisms, such as pre-mRNA splicing co-factors which are secondarily involved in mRNA export and translational regulation (Valcarcel and Gebauer, 1997; Sanford et al., 2005).

The Pet54p protein is an intriguing example of a nuclear encoded moonlighting protein in yeast mt. Disruption of the PET54 gene affects both the processing of the a15β group I intron in the COX1 pre-mRNA and the translation of the COX3 mRNA (Valencik et al., 1989; Costanzo et al., 1986). Despite extensive investigations, the mechanisms of Pet54p function in these two unrelated processes remains elusive. In
particular, it is unknown whether Pet54p is a primary RNA binding protein that specifically recognizes both RNAs or whether its functional specificity is governed in other ways.

Pet54p is a small, ~35 kDa protein that contains an RNA recognition motif (RRM) near its C-terminus. Genetic and biochemical analyses have established that Pet54p interacts with two other factors, Pet122 and Pet494, and this complex functionally interacts with the 5’ untranslated leader (UTL) of COX3 mRNA. The multi-part complex is essential for COX3 translation, since disruption of any component severely reduces COX3 protein production (Brown et al., 1994; Fox, 1996). How the complex assembles on the 5’ UTL and promotes translation is not known, but, once bound, it may serve to facilitate ribosome recognition of the AUG start codon (ibid; Haffter et al., 1991; Chacinska and Boguta 2000). The presence of an RRM in Pet54p is suggestive that the protein binds to the 5’ UTL of COX3, but no direct evidence exists to show this is the case. However, deletions within the COX3 5’ UTL exacerbate the respiratory-deficient (mt impaired) phenotype of a mis-sense mutation just downstream of the RRM motif in Pet54p. These observations are thought to reflect a functional interaction between Pet54p and RNA (Brown et al., 1994).

The role of Pet54p in promoting a1β splicing is even less clear. Like all group I introns, the a1β intron must fold into an active conformation to catalyze the self-splicing reaction (Houghland et al., 2005; Woodson 2005). Genetic analyses have shown that a1β splicing efficiency is reduced when one of five mt targeted nuclear genes, including PET54, is disrupted (Lambowitz et al., 1999; Valencik et al., 1989). An attractive hypothesis is that Pet54p along with the other proteins function to facilitate folding of the intron. Protein involvement in group I intron folding has been demonstrated in vitro for
the \textit{N. crassa} mt tyrosyl tRNA synthetase, \textit{S. cerevisiae} mt CBP2 and \textit{A. nidulans} I-\textit{AnI} proteins (Caprara et al., 1996; Weeks and Cech 1995; Ho and Waring 1999; Caprara et al., 2007). The disruption of the PET54 gene does not appear to affect the splicing of any other yeast mt group I intron suggesting that the protein may interact directly with the \textit{aI5\beta} intron or one of its other co-factors (Valencik et al., 1989).

As a starting point to understand how Pet54p functions as both an activator of COX3 translation and as an essential co-factor in the splicing of \textit{aI5\beta} intron, we have analyzed its interaction with both RNA ligands using recombinant Pet54p purified from bacteria. UV cross-linking experiments provide evidence that Pet54p specifically recognizes both the COX3 5’ UTL and the COX1 pre-mRNA containing the \textit{aI5\beta} intron. \textit{In vitro} experiments are consistent with Pet54p facilitating the exon ligation step of \textit{aI5\beta} splicing. Competition experiments show that both COX3 5’ UTL and \textit{aI5\beta} containing pre-RNA bind to the same or overlapping sites in Pet54p. RNA mapping and mutational experiments demonstrate that Pet54p binds to regions in each RNA that share sequence similarity. Collectively, these observations are consistent with Pet54p using analogous interactions to bind both RNAs and support a model in which a fortuitous association with the \textit{aI5\beta} intron provided the opportunity by which Pet54p adapted to function in splicing.

\textbf{Results:}

Cloning PET54, the 5’ UTL of COX3 and \textit{aI5\beta} containing pre-RNA into expression vectors.
The PET54 gene was amplified by PCR from yeast genomic DNA and cloned into a bacterial expression vector. The construct contained a short His6-tag fused to the N-terminus of Pet54p that was used to purify the protein using affinity chromatography (see Materials and Methods). The 613 nucleotide (nt) 5’ UTL of COX3 along with the upstream tRNAval gene and the aI5β intron with flanking exon sequences were amplified by PCR and cloned into a vector downstream of a T7 RNA polymerase promoter for use in generating RNAs via in vitro transcription (see Materials and Methods). The COX3 5’ UTL is numbered relative to the first base (+1) of the mRNA coding region: 5’ end of the leader is designated –613 and the last base is –1 (7). The 1604 nt aI5β pre-RNA includes a 12 nt 5’ exon, the catalytic core, an open reading frame encoding a degenerate homing endonuclease, and the 3’ splice site (SS) followed by an 18 nt 3’ exon (Figure 2-1). Notably, the aI5β intron contains a non-canonical 162 nt AU rich insertion sequence that interrupts the normally 3 nt J3/4 sequence between the 5’ strands of P3 and P4 (Figure 2-1). To facilitate in vitro transcription yield, we also created a shorter RNA that contained the 5’ exon and catalytic core followed by a short sequence extension (aI5β 3’TRE pre-RNA). Equilibrium binding experiments demonstrated that Pet54p has the same affinity for the truncated and full-length aI5β pre-RNAs (data not shown), and we used this RNA for some of the experiments described below.

**Pet54p specifically associates with aI5β RNA and facilitates the exon ligation step of splicing in vitro**

To test if recombinant Pet54p was capable of binding the aI5β pre-RNA in the
absence of other splicing co-factors, we performed uv cross-linking experiments. Internally labeled, 4-thiouridine containing RNAs were incubated with Pet54p, exposed to uv light, the RNA degraded by RNases and the products separated by SDS-PAGE. As shown in Figure 2-2A, Pet54p was cross-linked with the αI5β pre-RNA in this assay. Control experiments showed that the crosslink was dependent on 4-thiouridine substitution and uv irradiation (data not shown). Some group I intron co-factors recognize conserved structural features in group I introns and thus binds to multiple introns indiscriminately. To test the specificity of the Pet54p interaction, unlabeled group I intron competitor RNAs were included in the cross-linking reactions. Figure 2-2A shows that inclusion of full-length αI5β pre-RNA significantly decreased cross-linking efficiency. In contrast, the presence of the mt COB group I intron RNA from *A. nidulans* did not decrease cross-linking efficiency. Like the αI5β intron, the COB intron transcript is AU rich (65%) and folding studies suggest that almost all of the conserved secondary structure elements are formed in this intron. These data suggest that cross-linking is reporting a specific interaction between Pet54p and the αI5β pre-RNA and, furthermore, Pet54p does not recognize conserved secondary structural features in group I introns.

We have recently developed an *in vitro* splicing assay with the αI5β intron and the Mrs1p protein (Bifano and Caprara, in preparation). The αI5β intron does not self-splice *in vitro* under near physiological concentrations of divalent cations (7 mM Mg2+). However, binding of recombinant Mrs1p protein efficiently facilitated the first step of αI5β splicing *in vitro* (Bifano and Caprara, in preparation; Figure 2-2B). A plot of the fraction of pre-RNA spliced versus time showed that the first step occurred with a rate of 0.02 (±0.005) min-1. However, the fraction of RNA that underwent the second step was
Figure 2-1. Schematic of Pet54p RNA ligands. (A) Secondary structure model of COX1 and aI5β intron RNA. The noncanonical sequence inserted (‘insertion sequence’) between P3 and P4 is shaded in gray. (B) Schematic of the COX3 5’ UTL. Gray shading represents a region of the UTL that was previously deleted resulting in loss of translational activation (23; see text). Note that the 5’ UTL is numbered relative to the start of the coding region (+1).
Figure 2-2. Pet54p binds directly to the aI5β pre-RNA and facilitates exon ligation. (A) UV cross-linking. Internally $^{32}$P labeled 4-thiouridine containing aI5β pre-RNA (20nM) was incubated with Pet54p in the presence of increasing concentrations (50-750nM) competitor RNA. After irradiation at 312 nm, the RNA was digested with ribonucleases and the reactions electrophoresed through an SDS-PAGE gel. Radiolabeling of Pet54p indicates a cross-link with the RNA. aI5β FL, full-length aI5β pre-RNA; COB, *A. nidulans* mt COB group I intron pre-RNA (B) Time course of Pet54p facilitated exon ligation. 5′-$^{32}$P-labeled aI5β pre-RNA was incubated with 1µM Mrs1p with or without 250nM Pet54p. (C) Plots of the total fraction RNA spliced and fraction of ligated exon (LE) produced. The data are averages from two independent experiments and the error bars represent standard deviation. The data were fit to the first-order equation: Fraction RNA species = $A \left(1 - e^{-kt}\right)$ where $A$ is the amplitude of RNA processed and $k$ represents the pseudo-first-order rate constant, $k_{obs}$. A third experiment with a different preparation of RNA also yielded a 1.5 fold increase in exon ligation. Titration experiments confirmed that the concentration of Pet54p was saturating for these experiments (see also figure 2-4).
only 0.04 (±0.009) after 3 hrs and may reflect that the 3’SS was not properly positioned in the αlβ catalytic core or the intron did not undergo the appropriate conformational changes required for the exon ligation. Addition of Pet54p slightly decreased the overall fraction of RNA that underwent splicing but increased the amplitude of ligated exons to 0.06 (± 0.002), a modest but reproducible increase of 1.5 fold (Figure 2-2B). The rate of ligation was not changed significantly by the addition of Pet54p (0.028 ±0.007 vs. 0.022 ±0.006 min⁻¹). Control experiments showed that Pet54p did not facilitate splicing in the absence of Mrs1p (not shown). Thus, in the minimal αlβ in vitro splicing reaction, Pet54p appears to function primarily to increase the efficiency of exon ligation.

**Overlapping binding sites in Pet54p for COX3 5’ UTL and COX1 αlβ containing RNAs**

We performed uv cross-linking experiments (see above) to assess if Pet54p also directly interacts with the COX3 5’ UTL. Indeed, as with the αlβ pre-RNA, Pet54p formed an efficient crosslink with 4-thiouridine containing COX3 5’ UTL RNA (Figure 2-3A). To assess if the Pet54p used a single site for binding to the COX3 5’UTL RNA and COX1 αlβ containing RNAs, unlabeled competitor UTL and group I intron RNAs were included in the reaction mixture. Figure 2-3A shows that inclusion of unlabeled COX3 5’UTL, full length or truncated (αlβ 3’TR) αlβ RNAs abolished cross-linking of 4-thiouridine substituted COX3 5’UTL. In contrast, inclusion of the *A. nidulans* COB group I intron had no effect on cross-linking efficiency. Identical results were found for 4-thiouridine substituted αlβ 3’TR RNA (Figure 2-3B). Collectively, these observations provide evidence that Pet54p interacts specifically with both COX3 5’ UTL and αlβ
pre-RNA. In addition, the ability of αI5β RNA to compete for binding to COX3 5’UTL and vice versa suggest that Pet54p contains a single or overlapping binding sites for each RNA.

**Binding of COX3 5’ UTL and COX1 αI5β pre-RNA to Pet54p**

To directly measure the binding affinity of Pet54p to its respective ligands and begin to delineate the RNA binding sites, we carried out equilibrium binding assays on full length and truncated transcripts. For the αI5β 3’TR RNA (493 nts), Pet54p bound with a $K_d^{app}$ of 18.3 (± 6) nM while the *A. nidulans* COB intron showed little detectable binding (Figure 2-4A). Pet54p binds to the COX3 5’ UTL with a $K_d^{app}$ of 20.8 (± 8) nM, an affinity equal to that of the αI5β RNA (Figure 2-4B). The finding that Pet54p binds with similar affinity to the intron and UTL RNAs raises the possibility that Pet54p recognizes similar binding sites in both RNAs.

We next split the αI5β 3’TR RNA into two fragments to assess whether Pet54p bound within the αI5β group I intron catalytic core or to a region peripheral to the core. In this regard, a 233 nt RNA (P1+INS) was synthesized that contained the 5’ exon through to the end of the idiosyncratic sequence that is inserted between the 5’ strands of P3 and P4 helices (Figures 2-1 and 2-4A). Pet54p bound this RNA bound with near wild-type affinity ($K_d^{app} = 18.2$ (± 9) nM) providing evidence that the Pet54p recognizes sequences and/or structures within P1 through the insertion sequence (Figure 2-4A). In contrast, a 306 nt RNA (P3-3E) that contained the catalytic core of the intron (with the insertion sequence deleted), from P3 through the 3’ exon, was not appreciably bound by Pet54p (Figure 2-4A). These findings demonstrate that Pet54p does not bind to the intron.
catalytic core, but, instead, primarily recognizes the 5’ region of the aI5β RNA.

In a parallel analysis, the COX3 5’ UTL was split into approximately equal sized fragments containing the 5’ or 3’ halves of the RNA. Pet54p bound with ~ 2-fold decrease in affinity \(K_d^{app} = 31.9 (\pm 3) \text{ nM}\) to an RNA fragment composed of the first 302 nts of the UTL (from A –613 to G -311; COX3UPST, Figure 2-4B). In contrast, Pet54p did not bind an RNA fragment containing the remaining UTL sequence from A – 312 to U-1 (COX3DNST; Figure 2-4B). Thus, as in the case for the aI5β pre-RNA, Pet54p shows high affinity and selective binding to a subdomain within the COX3 5’ UTL.

**Mapping the Pet54 binding site in COX3 5’ UTL and COX1 aI5β containing RNAs**

The truncation experiments described above demonstrated that Pet54p bound to 5’ upstream regions in both COX3 5’ UTL and COX1 aI5β RNAs. To more precisely define the smallest 5’ fragment that can be bound by Pet54p, we employed a “boundary assay” that has been used to map protein-binding sites in rRNA (e.g. Ryan and Draper., 1989). In these experiments, either 5’ end-labeled phosphorothioate (NαS) substituted RNAs were treated with iodine that cleaves at each phosphorothioate linkage or unsubstituted RNA was treated with RNase T1 that cleaves after guanosine. In both cases, cleavage generates a number of fragments that are then incubated with protein. RNA fragments capable of binding protein are captured on a nitrocellulose filter, eluted and then separated on a denaturing gel. Comparison of the bound and unselected fragments identifies the shortest RNA capable of binding protein (Ryan and Draper., 1989).
Figure 2-3. Specific binding of Pet54p to the COX3 5’ UTL and aI5β pre-RNA. Internally \(^{32}\)P-radiolabeled 4-thiouridine containing 5’ UTL (A) or aI5β pre-RNA (B) was incubated with Pet54p in the absence or presence of increasing amounts of unlabeled competitor RNAs (250 and 500nM). aI5β FL, full-length aI5β pre-RNA; aI5β 3’TTR, aI5β pre-RNA truncated after the P9.1 structure (see Figure 1); UTL, COX3 5’ UTL; COB, \(A.\ nidulans\) mt COB group I intron pre-RNA.
For 5’ end-labeled αS substituted aI5β 3’TR RNA, Pet54p strongly bound RNAs that contained intron sequence up to A60, within the non-canonical insertion sequence (Figure 2-5A). Furthermore, Pet54p also bound fragments terminating at A50, but the reduced intensity of these bands relative to the other bound fragments is consistent with the protein having a reduced affinity for these RNAs. Taken together these results show that virtually the entire intron catalytic core was not required for Pet54p binding and that the minimal 5’ terminal fragment bound by Pet54p is 62 nts. This includes 5’ exon (12 nts), the remainder of P1 helix, the 5’ strand of P3 and some sequences in the insertion sequence (50 nts total) of the aI5β RNA (see the schematics in Figures 2-1 and 2-6A).

The COX3 5’ UTL RNA is very A-rich, and fragments generated from iodine cleavage of αS substituted material were not sufficiently resolved by gel electrophoresis to accurately assign the identity of fragments bound to Pet54p. We therefore used CaS substituted RNA as well as T1 digested material. Experiments with cleaved CaS UTL RNAs showed that Pet54p bound 5’ end-labeled fragments that contained sequence up to C –410 while assays with the T1 digested material showed that Pet54p could also bind a shorter fragment terminating at G- 476 (Figure 2-5B). Interestingly, COX3 5’UTL RNA terminated just upstream of G –476, at C –480, was poorly bound by Pet54p (see asterisk, Figure 2-5). This suggests that the 137 nt fragment defined by the first base of the UTL (A –613) and terminating at G-476 contains the Pet54p binding site.

**Footprinting the Pet54p binding sites on the aI5β pre-RNA and COX3 5’ UTL**

To identify the Pet54p binding sites within the 5’ ends of both RNAs, RNA
Figure 2-4. Pet54p binds with high affinity to subdomains within the COX3 5’ UTL and the a15β pre-RNA. In each experiment, trace amounts of internally $^{32}$P-radiolabeled RNA was incubated with increasing amounts of Pet54p (0-500nM) and complexes filtered onto nitrocellulose and the fraction of RNA bound calculated by the fraction of radioactivity retained on the filters vs. the total filtered. The data were fit to an equilibrium binding equation (see materials and methods section). (A) Binding of a15β 3’TR RNA and 5’ and 3’ truncations. As shown in the schematic, P1+INS contains the 5’ exon through the end of the insertion sequence; P3-3E, contains the intron core with the insertion sequence deleted. The open diamond symbols represent binding of the A. nidulans COB group I intron. In the schematic, the 5’ exon is black, the 5’ strand of P3 is gray and the insertion sequence is striped. (B) Binding of the COX3 5’ UTL and 5’ and 3’ truncations. As shown in the schematic, UPST contains the first 302 nt of the UTL; DNST contains the UTL sequence from A-312 to U-1. COX3 coding sequence (+1 and beyond) is shown for orientation but was not included in the RNAs tested. The open diamond symbols represent the binding of A. nidulans COB group I intron.
Figure 2-5. Mapping the Pet54p binding site on COX3 5’ UTL and COX1 al5β containing RNA. (A) Boundary experiment for the al5β 3’TR RNA. For 5’ end-labeled, AαS substituted al5β 3’TR, Pet54p tightly bound RNAs fragments up to A60 and with lesser affinity to fragments terminated up to A50. IN, input; BD, Pet54p bound RNA. (B) Boundary experiment COX3 5’ UTL RNA. For 5’ end-labeled UTL RNA digested with RNase T1, Pet54p tightly bound RNAs fragments terminated at G -476. For 5’ end-labeled CaS substituted COX3 5’ UTL RNA, Pet54p tightly bound RNAs fragments up to C -480 (asterisk) OH, hydrolysis ladder.
footprinting experiments were performed. The accessibility of the RNA ligands’ phosphate backbones in the presence and absence of Pet54p was probed by partial ribonuclease digestion. The RNA was digested with the double stranded specific RNase V1 while single stranded regions were probed with RNase 1, and the results are shown in Figure 2-6. Interestingly, in both cases, Pet54p binding did not result in protection of any V1 cleavage sites, and, importantly, only a subset of RNase 1 cleavage sites were protected suggesting that Pet54p primarily binds single-stranded regions in both RNAs.

For the a15β 3’TR pre-RNA, there were four regions of protection from RNase 1 digestion: within the P1 stem-loop, both in exon and intron sequences (from residues –g3 to A9), in the junction between P1 and P3 (J1/3; C22-A27) and in two segments of the insertion sequence (from A52 to U53 and U67 to A69; Figure 2-6A). Interestingly, positions –g3 to U3 are also cleaved with V1 consistent with these positions being base-paired in the P1 helix (Figure 2-6A). The cleavage by both V1 and 1 RNases suggests that the paired and unpaired states of this helix are in dynamic equilibrium. The lack of protection from RNase V1 cleavage by protein binding may reflect that Pet54p binding results in stabilization of the P1 helix while protection of positions in the P1 loop (A4 – A9) may reflect direct binding by the protein. Interestingly, in the insertion sequence, the two regions protected from RNase 1 cleavage are both surrounded by regions cleaved by V1 suggesting that they are in loops or bulges within adjacent helices (Figure 2-6A).

Importantly, the protected regions lie within and adjacent to the sequences found to be critical for Pet54p binding in the boundary assay (Figure 2-5A). In this regard, the first two protected regions are within the minimal Pet54p bound 5’ fragment that terminated at A50 (shown in Figure 2-6A and see above). The third and fourth regions
are just upstream and downstream of the high affinity boundary, defined by the fragment ending at A60 (Figures 2-5A and 2-6A). Collectively, these results demonstrate that Pet54p binding sites are within single stranded regions close to the 5’ end of the α15β RNA.

The COX3 5’UTL also showed multiple regions of protection from RNase 1 when bound to Pet54p (Figure 2-6B). Two of these regions (A -537 to A -529 and C -523 to A -512) are separated by only 5 nts, while the third region (A -590 to A -582) is further upstream. Poor V1 cleavage efficiency in this region of the RNA makes it difficult to judge the amount of secondary structure, however, given that RNase 1 cleaves only at selective positions suggests that the UTL contains significant secondary and/or tertiary structure. Presumably the protected regions are close in the folded structure of the UTL.

The lengths of RNA containing the binding sites for α15β not including the P1 paired nts –3 – 3) and COX3 5’ UTL identified by RNase footprinting are of similar size (66 versus 78 nts, respectively) and the sequences are extremely AU rich (89% and 96%, respectively) suggesting that Pet54p may recognize related sequences in both RNAs. Indeed, alignment of these sequences show 56% identity (40/71), with the 5’ and 3’ boundaries of the protected fragments overlapping (Figure 2-6C). Although the 5’ most protected residues are not precisely aligned, both contain an AAUA sequence (Figure 2-6C). The 3’ most protected residues align at a UAA (Figure 2-6C). Interestingly, the protected sequence from C22 to A27 in α15β RNA shares significant identity (67%) with the aligned COX UTL sequence. However, the lack of RNase 1 digestion in this region of the UTL prevents making strong conclusions about whether Pet54p binds to this region of the COX3 5’ UTL. Thus, while RNase footprinting analysis is restricted by the limited
Figure 2-6. Visualization of the Pet54p-binding sites by RNase footprinting. 5’ end-labeled aI5β 3’TR RNA (A) or COX3 5’ UTL (B) were incubated with buffer (-), RNase VI or I in the absence (first lane of each set) or presence of saturating Pet54p (1.25 or 2.5 µM, second and third lanes of each set). Cleavage products were separated by denaturing gel electrophoresis and the data quantified as described in the Materials and Methods section. Only those sites that showed a 2-fold or higher reduction in phosphorimager counts under both protein concentrations were identified as ‘protected’ (gray bars). The asterisk in (A) refers to bands of site-specific degradation caused by incubation with Pet54p. Below each gel image is a schematic of the relevant sequence protected from RNase1 digestion. Gray shading represents protection and the red shaded residues demarcate the 5’ binding boundary as defined in Figure 5. In (A), a secondary structure of the first 21 nt defining the P1 helix is also shown. Note that residues G-3 to U3 are also cleaved by the double-stranded specific VI RNase, but Pet54p binding does not protect these residues from VI cleavage. (C) Alignment of the RNase 1 protected sequence. The sequence from positions A4 to A69 of aI5β intron and positions A-590 to U-512 of the COX3 5’ UTL were aligned using the EBOSS program (32). Gray shaded residues are protected from RNase 1 digestion in the presence of Pet54p [see (A) and (B) above].
number of regions that show cleavage, these observations are consistent with Pet54p recognizing related sequences and/or structures within similar spatial constraints in both of its RNA ligands. These observations are reminiscent of the *E. coli* S7 and S15 ribosomal proteins that interact with different subdomains of 16S rRNA as well as their own mRNAs. In these cases, the RNA binding sites are spread over a rather long, differently folded sequences that contain multiple, short (2-5 nts) stretches of strong identity (Robert and Brakier-Gingras 2001; Nikulin et al., 2000 Philippe et al., 1995).

**DISCUSSION:**

Pet54p cooperates with two distinct sets of proteins to activate COX3 translation and facilitate excision of the aI5β intron from the COX1 pre-mRNA. However, prior to this analysis, it was not known whether Pet54p was a specific RNA binding protein, and if so, whether it recognized both of its ligands by a similar mechanism. In this study we show that recombinant Pet54p bound with high specificity and affinity to both the COX3 5’ UTL and aI5β pre-RNA in the absence of other co-factors demonstrating that it functions, in part, by interacting directly to both RNAs. Competition experiments were consistent with Pet54p using overlapping binding sites to interact with both RNAs. In support of this, we identified Pet54p binding sites within the COX3 5’ UTL and aI5β pre-RNA by truncation experiments and RNase footprinting and showed that the sites are similar in spatial context and sequence. Collectively, these results suggest that Pet54p binds both RNAs in an analogous manner and, as discussed below, provides insight into the individual functions of Pet54p and its evolution as a splicing co-factor.
Translation activation of COX3

The protein products of the PET54, PET122 and PET494 genes are required for translation of the COX3 mRNA, and our results are consistent with Pet54p interacting directly with the 5’ UTL to mediate translational activation. Synthetic lethal and genetic suppression studies were consistent with a model in which Pet54p and Pet122p functionally interacts with the 5’ UTL of COX3, and mutational analysis has shown that an internal deletion between U –605 to A –406 of the 5’ UTL abolishes translational activation (Brown et al., 1994; Costanzo and Fox 1993; Wiesenberger et al., 1995). Here we have shown that Pet54p physically interacts with the 5’ UTL of COX3, and furthermore, it specifically recognizes a 137 nt fragment (residues A –613 to G -476) that overlaps with the 199 nt region identified genetically to be important for COX3 translation. Furthermore, mutational analysis has also shown that deletion of COX3 5’UTL sequence from G –607 to G –527 severely diminish COX3 mRNA levels and protein production (Wiesenberger et al., 1995). The AU-rich segment (A -590 to U -512) that contains Pet54p binding sites, as determined by RNase footprinting, overlaps this region. These results are consistent with Pet54p binding to the upstream region of the 5’ UTL and playing an important role in helping to assemble the activation complex onto an appropriate location within the COX3 mRNA to facilitate translation (8; Figure 2-7).

Splicing of the aI5β intron

Genetic analysis has shown that the protein products of MRS1, PET54, MSS18, MSS116 and SUV3 affect splicing of the aI5β intron, but whether all of these proteins act directly to facilitate splicing and what roles they play are unknown (Lambowitz et al.,
We have shown that Mrs1p binding is sufficient to facilitate α15β splicing in vitro, but the exon ligation step is inefficient (Figure 2-2B; Bifano and Caprara, in preparation). Here we show that Pet54p binds with high specificity to the α15β pre-RNA, and it associates primarily with the very 5’ end of the intron. Binding of Pet54p does not facilitate general catalytic activity of the α15β intron in vitro, but, instead, increases the efficiency of the exon ligation step. The increase in the fraction of molecules that undergo the second step is a modest ~1.5 fold suggesting that other factors may collaborate with Pet54p to assist in this step. While the mechanism of action is still under investigation, the finding that Pet54p binds to the 5’ end of α15β raises the possibility that it may assist in stabilizing the interaction of the 3’ exon with the intron’s internal guide sequence (P10 pairing) that serves to position the 3’ SS for ligation after the first step in splicing (Figure 2-7; 12). On the other hand, results from the RNase mapping experiments suggests that binding of Pet54p also stabilizes pairing of the 5’ exon to the internal guide sequence (P1 pairing). Thus, Pet54p may assist in splicing by preventing pre-mature dissociation of the 5’ exon after the first step in splicing.

Previously it was shown by northern analysis that α15β containing COX1 pre-mRNAs accumulated in PET54 disruption strains suggesting that the protein also is important in the first splicing step (Valencik et al., 1989). Presumably our in vitro conditions mask this contribution of Pet54p in splicing. However, given that Pet54p binding appears to stabilize the P1 pairing, it may act by promoting docking of the 5’ exon into the intron’s catalytic core. Regardless of the specific mechanism, the observation that Pet54p interacts near (or at) the splice sites makes it unique among characterized group I intron co-factors, which primarily bind to and stabilize the active
structure of the intron core (Lambowitz et al., 1999; Caprara et al., 1996; Ho and Waring 1999; Caprara et al., 2007).

**Evolutionary implications**

Our findings show that Pet54p uses the same or overlapping binding sites to interact with the COX3 5’ UTL and the α15β intron and a part of the recognition involves related features that span ~60 nts in both RNAs. Given that splicing is likely a secondary adaptation of the protein (see Introduction), these observations support a model in which Pet54p became involved in α15β processing by virtue of a fortuitous interaction with sequences and/or structures in the intron that resembled its native recognition site in COX3 5’ UTL. A specific function for Pet54p may have been selected over time if changes in the protein increased the efficiency of splicing. It would be essential that these changes did not impair Pet54p translational activation function; thus, in this model, “effector domains” of Pet54p are predicted to be separate. In support of this hypothesis, mutations that affect one or the other function have been discovered providing evidence that both functions of Pet54p are genetically separable (Valencik and McEwen 1991). Together with our data, this latter observation suggests that the splicing function of Pet54p may include both RNA binding and effector domains (Figure 2-7).

Group I intron co-factors either use pre-existing nucleic acid binding sites to facilitate splicing or “develop” new binding sites from non-functional surfaces. For example, a conserved RNA binding domain in the yeast mt leucyl-tRNA synthetase is critical for its role in aminoacylation, as well as promoting splicing of bI4 intron (Chatterjee et al., 2003). In contrast, the homing endonuclease I-Anil from *Aspergillus*
and the mt tyrosyl tRNA synthetase from *Neurospora* both have developed independent group I intron RNA binding sites (ibid; Bolduc et al., 2003; Paukstelis et al., 2005). The Pet54p protein appears to have traveled a unique evolutionary path. On the one hand, it uses a pre-existing RNA binding site for binding to the aI5β intron, but another, yet to be defined surface, is primarily important for splicing and not translational activation. Thus, in other cases, the evolution of coordinated gene expression at the RNA level may similarly involve the development of protein co-factors that use a single recognition mechanism for two or more substrates, with binding regulating different steps in the expression of individual genes. Continuing studies of the structure and functions of Pet54p should provide new insights into how RNA binding proteins, in general, may be exploited for other functions *in vivo*.

**MATERIALS AND METHODS:**

**Molecular cloning and mutagenesis**

The aI5β intron containing COX1 gene was amplified from mt genomic DNA isolated from the BY4741 yeast strain using the primers AI5Top (5'-GCGGAATTC TAATACGACTCACTATAGGGATCCACGATATTAATTTAATAAGTGTCG TGC) and AI5Bot (5'-GGAAGCTTCAACTTTATTATTAGTTAATTAAGTGTCG TGC) to make paI5FL. The downstream truncated aI5β intron construct (paI5Prox) was made using AI5Top primer and the A15 Prox (5'-GGAAGCTTCAACTTTATTATTAGTTAATTAAGTGTCG TGC), which hybridized just downstream of the catalytic core of aI5β (see Fig. 1-1). The COX3 5' UTL was also amplified from mt genomic DNA using the primers T7COX3UTL (5'-
Figure 2-7. Model for Pet54p function in translational activation (A) and splicing (B). A common RNA-binding domain is used to interact with both RNAs, whereas different effector domains are required for each function. (A) The model for COX3 translational activation is adapted from references 8 and 31. The stem-loop shown is to emphasize that the UTL is likely structured but direct evidence for Pet54p binding to a stem-loop in the UTL is speculative (see results section). (B) For the aI5β pre-RNA, the P1 and P10 interactions that position the exons (thick lines) into that catalytic core (thin lines with arrows) are emphasized. Interactions with the 3’ exon are shown to emphasize one possible model for Pet54p splicing function.
GCGGAATTCCGAGTAATACGACTCAGCCTATAGGGTCTCGAAACCTATATTTACTA
AA) and COX3UTLBOT (5’-
GGAAGCTTCTTGATGTCTACTTCTTTCTAAATGTGTCAT) to make pCOX3UTL. Each PCR product contained a T7 RNA polymerase promoter for in vitro transcription. The PCR products were cloned in pGFP UV vector (Clontech, Mountain View, CA.). The α5β and UTL constructs were sequenced completely.

PCR templates were used to generate transcriptsβ containing RNA subdomains. For α5β, the P1+INS template was amplified from pa5Prox using A5Top and Delta P4-5Bot (5’-GGAAGCTTATAATTATAAGAGTTTCCCCGTTTA), while the P3-3E template was amplified from a version of pa5FL deleted for the large insertion sequence downstream of P1 using Delta P1Top (5’-
GCGGAATTCTAATACGACTCAGCCTATAGGGTCTCGGCTTAAAATTCACTA) and A5 Prox. For COX3UTL, the COX3UPST template was amplified from pCOX3UTL using primers T7COX3UTL and 8A1 (5’-CTCTTTTCGACCGGATTATTTATTTTC), while the COX3DNST template was amplified from pCOX3UTL using primers 8A2 (5’-
GGTAATACGACTCAGCCTATAGGGAGAGAAATAATAATCCGGTCTCGAAAGAG) and COX3UTLBOT.

The PET54 gene was amplified from genomic DNA using the primers PET54Top (5’-CCGGATCCGATGAAGGCTTCTAGTAAAGCTATTA) and PET54Bot (5’-
GGAAGCTTCTCGACTAAGATGTTCAATTTATA). The PCR product was cloned into the pet28b vector (EMD Biosciences, Inc. San Diego, CA) downstream of a His6 tag that was used for protein purification. The Pet54p expression clone was sequenced.
Protein expression and purification.

Initial experiments showed that recombinant Pet54p was largely insoluble when expressed in *E. coli* and so protein purification was performed under denaturing conditions. Recombinant Pet54p was expressed and purified essentially as described previously for the I-AniI protein (Emerick et al., 1996) with the following exceptions. The lysis, wash and elution buffers contained 6 M guanidine-HCl. The eluted protein was renatured by dialysis against a series of buffers containing decreasing concentrations of guanidine-HCl. The final dialysis was against 20 mM Tris-HCl, pH 7.9, 500 (or 100) mM NaCl and 50% glycerol. The protein was stored at -20°C. Protein concentration was determined by absorption at 279 nm in 6 M Guanidine-HCl, 20 mM Na-phosphate buffer, pH 6.5 using the calculated extinction coefficient of 43320 M⁻¹cm⁻¹. A small contaminating peptide (~7 kDa) was present in each preparation and the A279 value was corrected for its presence prior to calculating the Pet54p concentration.

For filter binding and RNA footprinting experiments, Pet54p lacking the first 20 amino acids (Pet54Δ20p) of its N-terminus was used. This truncation results in more soluble protein allowing for purification under native conditions. Both Pet54Δ20p and recombinant Mrs1p were purified as described previously for the I-AniI protein with the exception that Pet54Δ20p was expressed at 16°C for 20 hours (Solem et al., 2002).

RNA transcription and purification

All plasmids were digested with *Hind*III. Transcription reactions were in 50 or
100 µl of reaction medium containing 2-5 µg of plasmid DNA, 5 units per µl phage T7 RNA polymerase, 25 mM NaCl, 8 mM MgCl2, 40 mM Tris-HCl, pH 8.0, 2 mM spermidine, 10 mM dithiothreitol (DTT), 0.5-1 mM NTPs and 1 unit per µl RNase OUT (Promega, Madison, WI) for 60-120 min at 37°C. 32P-labeled transcripts used in cross-linking experiments were synthesized by adding 1 µCi/µl [α-32P]UTP (3,000 Ci/mmole; ICN Biomedicals, Irvine, CA) to a transcription mix containing 0.5 mM 4-thioUTP (Ambion, Austin, TX), 0.5 mM UTP, 1 mM of each remaining NTP. High specific activity transcripts used in filter binding experiments were synthesized by adding 4 µCi/µl [α-32P]UTP (3,000 Ci/mmole; ICN Biomedicals) to a transcription mix containing 0.04 mM UTP, 0.4 mM of each remaining NTP.

To generate end-labeled adenosine- or cytosine-phosphorothioate containing RNAs, the transcription mix included 1 mM guanosine, 0.025 mM ATPαS or 0.063 mM CTPαS (Glenn Research, Sterling, VA) along with the remaining nucleotide triphosphates. For all synthesis reactions, after transcription, the DNA template was digested with DNase I (0.5 units per µl; FPLC-purified; Pharmacia Biotech Inc, Piscataway, NJ) for 20 min at 37°C. Transcripts were extracted with phenol-chloroform-isoamyl alcohol (phenol-CIA; 25:24:1 v/v/v), centrifuged through a Sephadex G-50 (Sigma Chemical Co., St. Louis, M0) spun column. Unsubstituted transcripts were prepared similarly but without phosphorothioate triphosphates.

Transcripts were 5’ end-labeled as described (Caprara et al., 2007). The end-labeled and high specific activity transcripts were further purified by denaturing gel electrophoresis. For all experiments, prior to use, RNA was heated to 90°C in H2O for 20 sec and placed on ice.
uv cross-linking

For the cross-linking experiments, internally 32P-labeled 4-thiouridine containing transcript (20 nM) was incubated in the absence or presence of Pet54p (~250 nM) in 50 mM Tris-HCl, pH 7.5, 100 mM NaCl and 7 mM MgCl2 (TNM buffer) for 10 min. For competition experiments, unlabeled competitor RNAs were included at a final concentrations of 50 to 750 nM (see Figures 2-2 and 2-3 legends). The reaction mixture was then placed in a 96-well plate on ice and irradiated at 312 nm in a uv Stratalinker 1800 (Stratagene, La Jolla, CA) for 20 min. A ribonuclease cocktail of RNaseA and T1 was added (Worthington Biochemical Corp., Lakewood, NJ) and the RNA was digested for 60 min at 42°C. The products were separated on a 10% SDS-PAGE gel. The dried gels were visualized with a phosphorimager (Amersham Biosciences, Piscataway, NJ) or exposure to X-ray film. All experiments were repeated with identical results.

RNA splicing assays

5’-32P-end-labeled αI5β pre-RNA was incubated in TMN buffer supplemented with 2 mM spermidine, 10 mM DTT, 0.7 units/µl of RNase OUT, 1 mM GTP and 0.7 µg/µl tRNA for 5 min at 37°C. Pet54p (final concentration, ~250 nM) and/or Mrs1p (final concentration 1 µM) were added to initiate the splicing reaction. Aliquots were removed at increasing times and the reaction quenched with 1µl of 0.5 M EDTA (pH 8.0) and extraction with phenol-CIA. The reaction products were separated in a denaturing 10% polyacrylamide/7M urea /5% glycerol gel and the dried gels were visualized with a phosphorimager. Titration experiments confirmed that the concentrations of Mrs1p and Pet54p were saturating for these experiments.
Delineation of Pet54p binding sites

For boundary experiments, 5' - $^{32}$P-labeled phosphorothioate substituted RNA was cleaved with iodine (80 nM final concentration) for 10 min at 42°C. Iodine was removed by ethanol precipitation and the RNA re-suspended in water. For T1 digestion, non-substituted 5' - $^{32}$P-labeled RNA was cleaved with T1 (5 units; Ambion) for 10 min at 50°C, extracted with phenol-CIA, ethanol precipitation and the RNA re-suspended in water. Pet54p (~250 nM) and RNA (10 nM) were incubated in TNM buffer at 37°C for 5 min and then filtered through a nitrocellulose filter that was subsequently washed with 3 ml of TNM buffer. Pet54p-bound RNA was recovered by extraction of the filter with phenol-CIA and ethanol precipitation. The products were separated in a denaturing 6% polyacrylamide/7M urea gel and the dried gels were visualized with a phosphorimager. The experiments were repeated with identical results.

RNA binding

For $K_d$ measurements, ~ 2 pM 32P-labeled RNA was incubated with 0 to 500 nM Pet54p in 50 µl TMN buffer containing 0.25 µg/µl heparin and 0.1 µg/µl bovine serum albumin at 37°C. The complexes were filtered after 30 min through nitrocellulose (BA85; Schleicher and Schuell, Keene, NH), washed with 3 ml of buffer, dried and counted using a Beckman-Coulter LS6500 scintillation counter (Beckman-Coulter Inc., Fullerton, CA). The data were fit to the following equation: Fraction RNA bound = A x E/(E + $K_d^{app}$)
where A is the total amplitude bound and E is the concentration of Pet54p and $K_d^{app}$ is the apparent equilibrium binding constant. Each titration was repeated two to four times and errors are expressed as standard deviations.
RNA footprinting

5’end-labeled transcripts (~3 nM) were incubated in the absence or presence of Pet54p (1.25 or 2.5 µM) in TNM buffer plus 10 ng/µl *E. coli* tRNA at 37°C for 10 min. Aliquots (15 µl) were treated with either buffer, RNase 1 (2.1 units, Ambion) or RNase V1 (0.012 units, Ambion) for 20 sec and the reactions stopped by addition of 2 µl of aurin tricarboxylic acid, 1 µl of 0.5 M EDTA and extraction with phenol-CIA. The products were separated in a series of denaturing 6% polyacrylamide/7M urea gel and the dried gels were visualized and quantified with a phosphorimager. The bands were identified based on co-migration with RNase T1 and random hydrolysis markers. Because single nt resolution is more difficult with fragments larger than 150 nts, the assignment of the ends of larger fragments may be two or three nts displaced. The data were analyzed by subtracting background (based on mock treated control samples) and band intensities were normalized for each lane with respect to the precursor RNA to account for any loading differences in each lane. Residues were scored protected by Pet54p if the band intensity was less than 50% of those in the absence of protein (Caprara et al., 2007).
Chapter III:

Pet54p interacts with the ai5β intron and COX3 5’ UTL through an unusual RRM motif.

*Benjamin J. Kaspar and Mark G. Caprara*
Abstract:

Pet54p is a dual function protein involved in translation and group I intron splicing within the mitochondria of *Saccharomyces cerevisiae*. Previously we have shown that this protein forms specific RNPs with either of its two natural substrates, the aI5β intron and COX3 5’UTL. Formation of these complexes was shown to be dependent on a shared RNA binding surface employed by Pet54p in the recognition of both substrates. In this work, we demonstrate the importance of the RNA recognition motif (RRM) of Pet54p to RNA binding. These experiments suggest that Pet54p interacts with both aI5β and the COX3 5’ UTL with unusually high specificity for a single RRM protein and that a specific RNA/protein interaction maps to the RRM region of the protein. Further this work demonstrates similarities in both amino acid sequence and RNA recognition properties between Pet54p and the RRM of the human Fox-1 protein. These similarities may explain the remarkably high specificity of Pet54p and allow us to put forth a hypothesis about a possible role for the Pet54p RRM in inducing RNA conformational change.

Introduction:

Three essential genes encoded within the yeast mitochondria are parasitized by mobile genetic elements called group I introns (Skelly and Maleszka, 1991). These selfish alleles disrupt the cytochrome oxidase (COX) 1, apocytochrome b (COB) and large ribosomal RNA genes at the DNA level, but mitigate host harm through post-transcriptional splicing from mRNA transcripts (Pel and Grivell, 1993). This type of splicing differs profoundly from nuclear RNA splicing in that no universal machinery has
evolved to carry out the process. In this regard, Group I intron splicing is not dependent
upon a shared set of conserved co-factors acting on all RNAs in this class. Instead,
a variety of adaptations have arisen in both the host and individual mobile genetic
elements to facilitate post transcriptional processing of these potentially lethal insertions.
These adaptations vary from introns that undergo catalysis without the assistance of
proteins, to others that require complex multiprotein systems (Lambowitz et al., 1999;
Bassi et al., 2002).

Many splicing co-factors are encoded within the introns themselves, existing
together as mosaic parasitic elements (Delahodde et al., 1989; Bolduc et al., 2003).
Proteins that facilitate splicing but are not encoded by group I introns are contributed by
the host nucleus. In many cases, these proteins bind with high specificity and facilitate
splicing of a single intron (Shaw et al., 1996; Shaw and Levin 1997). The persistence
of introns through dependence upon specific host protein co-factors is an evolutionary
mystery. The introns appear to be strictly selfish, providing no direct benefit to the host.
Yet the host organism, instead of eliminating these elements, has evolved mechanisms of
coping with them. In most cases, it is difficult to trace the lineage of these unique co-
factors. Most of the nuclear factors involved in group I intron splicing bear little
structural or functional homology to known proteins. Yet, there are rare exceptions when
RNA binding co-factors responsible for processes such as aminoacylation or translational
activation appear to have evolved the ability to splice group I introns as a
secondary function (Lambowitz et al., 1999). Unlike those factors that function only in
group I intron splicing, these proteins frequently possess well-characterized domains or
structures. For this reason, they may provide a means of understanding how RNA
binding proteins evolve additional functions. This subject is of paramount importance in understanding the complex and extensively coupled RNA processing events of the cellular nucleus, which, often involve single proteins in multiple roles throughout the lifetime of an mRNA.

Pet54p is a specific RNA binding protein that functions in translation and splicing within the yeast mitochondria (Costanzo and Fox, 1986; Valencik et al., 1989). Remarkably, while these functions require distinct sets of protein co-factors, both absolutely require Pet54p. Previously, we had demonstrated that Pet54p recognizes both the ai5β intron and COX3 5'UTL through the same, or an overlapping binding site (Kaspar et al., 2008). This property makes Pet54p unique among dual function group I intron splicing co-factors, the majority of which have evolved separate binding sites to facilitate splicing. In this way, Pet54p appears to have more in common with nuclear intron spliceosomal factors such as arginine/serine rich proteins. These factors, commonly known as SR proteins, contain an N-terminal RRM domain for RNA binding and a C-terminal serine/arginine rich domain. Despite their relatively small size, they have multiple roles in RNA processing and will recruit different sets of protein co-factors depending upon where they are bound to the intron RNA (Valcarcel and Gebauer, 1997; Sanford et al, 2005). Determination of the roles of these two domains in RNA binding and protein recruitment led directly to advances in understanding of alternative splice site regulation, nonsense mediated decay, and nuclear export.

Like the SR proteins, Pet54p contains an RRM domain, recognizes specific regions of RNA, and facilitates multiple functions in RNA metabolism. What is fascinating about Pet54p however is that its role in splicing of a single optional intron is
clearly a secondary function. Thus, the protein offers a unique opportunity to understand how small RNA processing factors have evolved to accomplish multiple functions.

Previously we postulated a model in which Pet54p contains a single RNA binding site and multiple effectors domains (Kaspar et al., 2008). This model brings together our own observation that Pet54p recognizes both substrates through the same binding site; with those of Valencik and McEwen (1991) demonstrating that PET54 mutations singly effecting splicing or translation could be generated. The next important step in understanding the role of this protein in RNA metabolism is to define the regions of the protein responsible for its different functions. We hypothesized that the RRM domain is important for specific RNA binding and sought to demonstrate this through mutational analysis, cross-linking and peptide cleavage experiments. Here, we show for the first time, a function for the RRM domain in specific interactions with yeast mitochondrial RNA.

Results:

Binding stoichiometry of the Pet54 protein

In order to determine the stoichiometry of Pet54p-RNA complexes, crosslinking competition experiments were performed employing internally radiolabeled 4-thio uridine-substituted aI5β or COX3 UTL RNAs as the labeled substrate. We had previously determined the $K_d$ of Pet54p to be 18.3 (+6) nM for the aI5β intron and 20.8(+/−8) nM for the COX3 5’UTL. Thus, in determining stoichiometry, Pet54p was used at 400 nM or, roughly 20 fold above the $K_d$. The amount of labeled RNA and protein in each reaction
was kept constant while increasing amounts of cold competitor RNA were added. After reaching equilibrium, reactions were crosslinked and digested with RNases before being run on a 10% SDS-Page gel and exposed on a phosphor screen. The resulting gel images are shown on the inset of the graphs in Figure 3-1A and 3-1B. The gel images show that increased concentrations of cold competitor RNA resulted in decreased crosslink formation when either the COX3 UTL or aI5β intron were used as labeled, crosslinker containing substrates.

To determine the relationship between competitor RNA concentration and crosslink formation, we compared the extent of crosslinking with increasing amounts of competitor RNA to the amount of crosslinking observed when no competitor was present. The inverse of the fraction of COX3 UTL or aI5β RNA crosslinked was plotted as a function of the concentration of competitor RNA. Figure 3-1A and B show the graphical representation of these data for the COX3 UTL and aI5β respectively. The plots indicate that the relationship between the concentration of cold competitor RNA added and the fraction of labeled RNA crosslinked was linear in both cases. The slope of the line for the experiments was 0.96 for the COX3 UTL and 0.73 for the aI5β intron, suggesting that the stoichiometry of complex formation for Pet54p and either the COX3 UTL or aI5β intron RNA was roughly 1:1.

**Pet54p contains a putative RRM domain with homology to the human Fox-1 RRM**

A protein PSI-BLAST search revealed that Pet54p contains an RNA Recognition Motif (RRM) fold at its C-terminus. This search also revealed numerous Pet54p homologs in yeast species closely related to *Saccharomyces cerevisiae*. Outside of these
Figure 3-1. Binding stoichiometry of Pet54p. Internally $^32$P-radiolabeled 4-thiouridine containing α5β (A) or COX3 UTL RNA (B) was incubated with 400nM Pet54p in the absence or presence of increasing amounts of specific competitor RNA (62 nM-1µM). Autoradiograms from each experiment are shown in the inset panels. Graphs indicate the inverse of the fraction of RNA cross-linked in the presence of various concentrations of competitor RNAs relative to the amount cross-linked in the absence of competitor RNA. Data were fit to a linear regression curve.
**Figure 3-2.** Residues conserved between the Pet54p RRM and the human Fox-1 RRM. Regions important for RNA binding include the RNP1, RNP2 and β₁α₁ domains. All three regions display high levels of conservation between the two proteins as indicated by the asterisks. The phenylalanine residues outlined in light blue represent Pet54p residues mutated in this study and their homologous positions in Fox-1.
homologs, only the human splicing factor Fox-1 displayed a level of similarity to Pet54p significant enough to be detected by the search (28% identical, 50% positives, E value = 0.068). The similarities between Pet54p and Fox-1 were confined to the RRM domains of each protein. Figure 3-2 shows a sequence alignment of the Pet54p RRM region with the Fox-1. The alignment of these sequences suggests that the Pet54p RRM and Fox-1 RRM share several features that may be important for RNA binding. As shown by the color bars and asterisks respectively, Pet54p possesses pronounced secondary structure and amino acid conservation in the RNP1 and RNP2 domains that comprise the major RNA interacting sites within the RRM. Typically, these are regions of beta sheets containing conserved hydrophobic and basic amino acids that form specific contacts with RNA. The proteins also share significant homology in the $\beta_{1}\alpha_{1}$ loop, a region that has been shown in the Fox-1 NMR structure to recognize four bases with high specificity and induce a bend in the RNA substrate.

Figure 3-3A shows the location of these conserved secondary structures and amino acids on the NMR structure of the human Fox-1 protein. Blue shading of the secondary structure indicates a conserved residue while yellow shaded regions represent conserved substitutions. Labeled residues indicate conserved residues or substitutions that have been shown to directly contact RNA in the Fox-1 crystal structure. In Figure 3-3B, residues that contact RNA in the Fox-1 NMR structure, but are not conserved between Pet54p and Fox-1 are shown in red along with conserved RNA contacting amino acids.
Figure 3-3. Residues conserved between the Pet54p and Fox-1 RRMs mapped on the NMR structure of Fox-1. (A). Conserved residues (blue) and conserved substitutions (Yellow) between Pet54p and Fox-1. Labeled residues indicate conserved residues involved directly in RNA binding. (B). All residues shown to contact Fox-1 in the NMR structure. Color coding same as in (A) except red amino acids indicate non-conserved amino acids.
A semi-conserved residue in the Pet54p RRM does not affect the RNA binding properties of Pet54p

In order to test the importance of the RRM domain in RNA recognition by Pet54p, mutagenesis experiments were performed at select amino acids hypothesized to be important for RNA binding. The first of these mutations, F241A corresponds to PHE160 on the Fox-1 structure in Figure 3-3A. The analogous mutation in the Fox-1 protein had previously been shown to result in a 30,000 fold decrease in the affinity of Fox-1 for its RNA ligand (Auweter et al., 2006). A second mutation, F244A, also targeted a conserved phenylalanine within the RNP1 domain. As shown in Figure 3-4A, the F241A mutant rescued the respiratory phenotype of a ΔPET54 strain, indicating a minimal affect on respiratory complex function. A similar result was observed for the F244A mutant (data not shown).

To determine if either mutant resulted in changes in α15β splicing, RT-PCR was performed using a primer specific for the exon downstream of α15β to generate cDNA, followed by PCR with primers specific for the un-spliced intron or ligated exons. As shown in Figure 3-4B, the PET54 deletion strain, when rescued with either the F241A or F244A mutants, resulted in complete restoration of α15β splicing.

We subsequently attempted to express and purify recombinant Pet54p F244A and F241A from E.coli. Unfortunately, we were unable to express the F244A mutant. However, the F241A mutant yielded recombinant protein.

Following expression and purification of the Pet54p F241A from E. coli, the protein was assayed for the ability to bind RNA via filterbinding. Figure 3-4C shows the K_d value of the mutant protein F244A for the α15β intron compared to the wild type
protein. Calculation of $K_d$ values revealed that the mutant protein bound with affinity similar to that of the wild type protein.

We then tested the specificity of the mutant through cross-linking. Previously, we had used this assay to demonstrate that Pet54p interacted specifically with its native substrates, the COX3 UTL and aI5β intron. In those experiments, the aI5β and COX3 UTL RNAs were shown to compete with the labeled, substituted aI5β intron when the wild type protein was used while equimolar amounts of the An Cob intron had no effect on cross-linking (Kaspar et al., 2008). In the experiment shown in Figure 3-4D, the F244A mutant displayed that same pattern of competition as the wild type protein. The band is present when no unlabeled RNA is added, but disappears in the presence of 50 nM aI5β or COX3 UTL RNA. The presence of unlabeled An COB RNA has no effect on crosslink formation. Thus this mutation did not appear to affect the ability of the protein to recognize its native substrates.

**Iodosobenzoic acid cleavage of Pet54p**

Iodosobenzoic acid is a reagent that cleaves proteins specifically at tryptophan residues (Mahoney and Hermodson, 1979). Because Pet54p contains only three tryptophans, all within the RRM domain, we reasoned that cleavage of the crosslinked protein with 0-iodosobenzoic acid, might give information about the location of the specific crosslink. The experimental design is outlined in Figure 3-5A. The protein is first crosslinked to either the aI5β intron or the COX3 UTL. The RNA is then digested away with RNases, leaving behind only the region of RNA protected by the protein. The crosslinked RNA is then treated with o-iodosobenzoic acid in order to cleave the protein.
Subsequently, the cleaved peptide fragments are run on an SDS-PAGE gel in order to determine the size of the cross-linked fragment.

Figure 3-5B shows the sizes and amino acid sequences of the expected peptide fragments. The Pet54p RRM regions, with homologies to the Fox-1 protein depicted, is shown as in Figure 3-2.

Figure 3-5C shows the result of Iodosobenzoic acid cleavage of crosslinked Pet54p. When the UTL RNA was employed as the substrate, no cleavage was observed in either the untreated or mock treated lanes, suggesting that all cleavage resulted specifically from the presence of o-iodosobenzoic acid. One cleavage band appears between 25 and 16 kDa while another occurs at less than 6 kDa. As shown in the left three lanes of the Figure, the result is similar for the experiment where the aI5β 3’TR RNA was used though only 40% of the protein was cleaved.

Given this pattern of cleavage, there are two possibilities for the identity of the smallest cross-linked fragment. It may represent the 2 kDa fragment within the RRM or the 1.6 kDa fragment. Figure 3-6 shows what each of these fragments look like on the NMR structure of Fox-1. The regions homologous with the 2 kDa fragment, shown in 3-6A and 3-6B, contain portions of the RNP2 domain as well as the conserved β1α1 loop region. This region is critical for specificity in the Fox-1 protein and contains three RNA contacting amino acids (3-6B). Figure 3-6B shows that, in Pet54p, many of these amino acids are conserved (blue) or display conserved substitutions (yellow) including two out of three RNA contacting amino acids (colored and labeled). The 1.6 kDa fragment, corresponding to β-sheet 1 in the Fox-1 structure displays less conservation (Figure 3-6D) while conserved amino acids (blue) and substitutions (yellow) are evident, there are
Figure 3-4. Analysis of Pet54p F241A and F244A mutants. (A) Pet54 F241A fully restores respiration in a ∆PET54 strain. Compare PET54 (indicating the wild type gene) with PET54 F241A. Both fully restored growth of the ∆PET54 strain on non-fermentable media (top right panel). Over-expression of F241A in the wild type background also had no effect on growth (top left panel) indicating an absence of dominant negative effects. (B) Both PET54 F241A and PET54 F244A fully restored αi5β splicing in the ∆PET54 strain. RT-PCR was performed a 3’ primer specific for the exon downstream of αi5β and two upstream primers corresponding to either the αi5β intron or the exon upstream of the intron. “Precursor” indicates COXI transcripts containing the αi5β intron while “LE” refers to the ligated exon. (C) K_d of recombinant wild type Pet54p compared with that of Pet54p F241A as determined by nitrocellulose filter binding (see Chapter 2). (D) Pet54p F244A forms a specific cross-link. Internally 32P-radiolabeled 4-thiouridine containing αi5β was incubated with Pet54p in the presence or absence of 250nM competitor RNA followed by cross-linking and digestion with RNases (see Chapter II).
**Figure 3-5.** o-iodosobenzoic acid cleavage. (A). Schematic of the experiment. UV cross-linking of Pet54p with $^{32}$P-radiolabeled 4-thiouridine containing RNA is followed by RNase digestion and cleavage with o-iodosobenzoic acid. (B). Expected sizes of Pet54p o-iodosobenzoic acid cleavage fragments. Residues within the Pet54p RRM are outlined as in figure 3-2. (C). Result of o-Iodosobenzoic acid cleavage of Pet54p in the presence of α5β intron or COX3 5'UTL RNA. No cleavage is observed in untreated or mock lanes. Addition of o-iodosobenzoic acid results in a series of cleavage bands, one between 16 and 25 kDa and another smaller band below 6 kDa.
no conserved RNA contacting amino acids in this region. Figure 3-6D shows the
location of each of these fragments within the structure of the entire Fox-1 RRM domain.
The 2 kDa fragment is shown in green while the 1.6 kDa fragment is purple. The
location of the β₁α₁ loop region within the 2 kDa fragment is boxed.

DISCUSSION:

Having previously characterized the RNA binding properties of Pet54p (Kaspar et al.,
2008) we sought to determine which regions of the protein were important for specific
RNA binding. We hypothesized that the RRM domain was critical for specific binding
based upon its homology to the structurally characterized RRM of the Fox-1 protein. Fox-
1 is a single RRM containing splicing co-factor that displays unusually high specificity for
its RNA ligand (Auweter et al., 2006). Fox-1 recognizes the sequence UGCAUGU both
through canonical RRM interactions between its beta sheet regions and through a unique
RNA binding surface within the β₁α₁ loop of the RRM. The loop region extends the RNA
binding surface of the protein allowing for increased specificity and affinity. Remarkably,
Fox-1 also induces a strong bend in its seven base RNA ligand, a unique property of this
RRM. Like Fox-1, Pet54p is able to selectively bind its natural substrates with high
specificity. The stochiometry of the Pet54p/RNA complex suggests that, even though
both the aI5β intron and COX3 5' UTL are long, remarkably AU rich sequences,
each Pet54p molecule participates in a specific interaction with a single RNA domain
within both natural ligands. In contrast, single RRM containing proteins of the hnRNP and
SR protein families generally act non-specifically, relying instead on interactions between
multiple factors and RNA binding sites of varying strength for specific interaction (Singh
Figure 3-6. Two possible identities for cross-linked Pet54p cleavage fragment mapped onto the Fox-1 RRM structure. (A). Fox-1 structure of the region homologous with the Pet54p 2 kDa fragment showing conserved residues (blue) and conserved substitutions (yellow) along with conserved RNA contacting amino acids (labeled). (B). All RNA contacting amino acids within the 2 kDa fragment region. Residues in red are amino acid contacting residues that are not conserved between Pet54p and Fox-1. (C) Fox-1 structure of the region homologous with the Pet54p 1.6 kDa fragment with conserved residues (blue) and conserved substitutions (yellow). (D). Location of 2 kDa (green) and 1.6 kDa (purple) fragments on the NMR structure of Fox-1. Residue color code is as described above. Boxed region highlights the $\beta_{1}\alpha_{1}$ loop within the 2 kDa fragment.
and Valcarcel, 2005).

To begin to explore the contributions of the Pet54p RRM to RNA/Protein interactions, we used mutagenesis and peptide mapping approaches. The mutagenesis approach targeted two phenylalanine residues within the RNP1 region of the Pet54p RRM. We chose to target this region for two reasons. First, it is the most likely region to be involved in specific RNA binding based upon the known structures of RRM containing proteins (Maris et al., 2005). In this regard, the phenylalanine corresponding to position F244 in Pet54p is particularly highly conserved and participates in base stacking interactions with RNA in many RRMs (ibid). Secondly, the F241 mutation in Pet54p correlates with position F160 in the human Fox-1 protein. When this position was mutated to alanine, the binding affinity of Fox-1 was reduced 30,000 fold. Surprisingly, the RT-PCR results suggested that these mutations did not affect splicing of the aI5β intron in vivo. Furthermore, based upon the lack of a respiratory phenotype, it seems unlikely that either mutation affected translation for the COX III 5’ UTL. Only one of these mutations, F241A, generated recombinant protein. Comparison of this protein with wild type Pet54p revealed that the mutation had no effect on the binding affinity or specificity of the protein.

These results of the mutagenesis experiments were somewhat surprising. They suggested that the targeted amino acids did not play a large role in the affinity or specificity of Pet54p. For this reason, we subsequently employed cross-linking and o-iodosobenzoic acid cleavage as a less targeted approach to determining regions important for RNA binding. This assay demonstrated that the specific RNA/Protein cross-link, formed by Pet54p only in the presence of its natural ligands mapped to a region within
the RRM. We have narrowed this region down to two fragments. The first of these fragments contains the unique $\beta_1\alpha_1$ loop that has been demonstrated in Fox-1 to be critical for specific recognition of at least four bases as well as bending of the RNA. The second fragment consists predominantly of $\beta$-sheet 2, which, generally does not participate in specific interactions with RNA. These results suggest that the RRM of Pet54p plays a role in RNA binding. However, the total contribution of the RRM to specificity and affinity remains unclear.

These data point to two possible conclusions regarding the RNA binding specificity of Pet54p and the involvement of the RRM. The first is that, like Fox-1, the protein possesses the ability to bind a unique sequence with a level of specificity unusual for single RRM containing proteins. This ability may result from the presence of non-canonical RNA binding regions within the RRM, such as the $\beta_1\alpha_1$ loop, or from regions outside of the RRM that also participate in binding. A second possibility is that the protein does not bind RNA in a highly specific manner. Rather, the secondary and tertiary structure in Pet54p’s RNA ligands mask most potential binding sites for the protein, leaving a single region or domain in each RNA available for binding.

**A hypothesis regarding the potential role of the RRM in group I intron splicing**

The ability of the Fox-1 RRM to induce a bend in its RNA substrate (Figure 3-7A) is truly remarkable and has implications for our continued study of Pet54p. It suggests that an RRM alone may be capable of inducing conformational change within its RNA substrate. In chapter II, we put forth the hypothesis that Pet54p binds to both the COX3 UTL and aI5$\beta$ intron through a common RNA binding domain yet facilitates
splicing and translation through separate effector domains (Figure 3-7B). However, if the Pet54p RRM is capable of inducing a conformational change similar to that observed in Fox-1, a separate effector domain for splicing may not be needed. Rather, the RRM alone could be responsible for both RNA recognition and stimulation of αI5β splicing (Figure 3-7C). Such a result would be truly novel as involvement of RRM binding in RNA catalysis has never been reported.

Materials and Methods:

Molecular cloning and mutagenesis

The F2441A mutant was generated through PCR based mutagenesis employing the PET54 gene cloned into vector pet28b (EMD Biosciences, Inc. San Diego, CA) as a template. PCR was used to generate two half molecules using the Primers Pet54Top (5’-CCGGATCCGATGAAGGCTTCTAGTAAAGCTATT) and F241ABot (5’-GAAAGATATCGCTCGTAGATTTCTTTTCATTTCATCATCCCA) to generate the first half molecule and Pet54Bot (5’GGAAGCTTCTTCACAAAGATGTCTTAAAT), to generate the second half. An additional PCR was performed with these two fragments as well as the Primers Pet54Top and Pet54Bot to generate the full length Pet54 F241A gene. The gene product was then digested with BamHI and HindIII and cloned into the pet28b protein expression vector.

The F2444A mutant was made in the same manner only, to generate the half molecules, Pet54Top was used with F244ABot (5’CGAGTTTTCCCGCAGATAAATCGTAGATTTCTTT) while Pet54Bot and
**Figure 3-7.** Two hypotheses regarding the role of Pet54p in splicing of the aI5β intron. (A). NMR structure of the Fox-1 protein with transparent blue arrow showing the bend induced in the RNA substrate upon binding. (B) and (C) two possible hypotheses for the function of Pet54p in splicing. (B) Pet54p recognizes the intron through the RRM domain but facilitates splicing through a separate effector domain. (C). The Pet54p RRM domain recognizes the intron and induces a conformational change leading to splicing.
F244Atop (5'-TTTATATCTGCGGAAAAACTCGCATGACGCCTACCGC) were used to generate the second fragment. Mutations were confirmed thorough sequencing.

Both mutants, as well as the PET54 wild type gene were also cloned into the 2 2 µM origin plasmid containing a leucine marker under control of the endogenous PET54 promoter for transformation into Yeast.

**Generation of a recombinant PET54 mutant**

Recombinant Pet54p F241A was expressed in *E. coli* as described in Kaspar et al, 2008.

**Yeast Transformation**

All experiments employed yeast strain BY4741 [Genotype, (genotype, *Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0)*]. Yeast transformation of PET54 wild type, F244A and F41A was performed according to the protocol of (Gietz and Schiestl, 2007). All three plasmids were transformed into wild type BY4741 as well as a PET54 deletion strain in the BY4741 background.

**uv crosslinking**

³²P-labled 4-thiouridine containing RNA transcripts were generated as described in Kaspar et al, 2008. For stoichiometry experiments, 20nM of this RNA was incubated in the absence of presence of 400 nM Pet54p in a buffer containing 50mM Tris-HCl, Ph 7.5, 100mM NaCl, and 7mM MgCl₂. Competitor RNAs were added to these reactions in concentrations ranging from 62 nM to 1 µM. Reactions were allowed to proceed to
equilibrium before being placed on a 96 well plate on ice and irradiated at 312 nM in a uv stratalinker 1900 (Stratgene, La Jolla Ca) for 20 min. Subsequently, a ribonuclease cocktail of RNaseA, and T1 was added (Worthington Biochemical Corp., Lakewood, NJ. RNA was digested for 2 hours at 37°C. Products were separated on an SDS-PAGE gel and visualized with a phosphorimager (Amersham Biosciences, Piscataway, NJ). The density of cross-linked bands was quantified and band intensities in the presence of competitor were plotted as the inverse fraction of total RNA bound in the absence of competitor on the y-axis, relative to the concentration of competitor added on the x-axis. To determine the stoichiometry of the complex, this plot was fit to a linear regression line where the slope is equivalent to the number of molecules of RNA bound to each Pet54p protein.

**Binding experiments**

Filter binding experiments employing recombinant Pet54p wild type protein and the Pet54p F241A were performed according to the protocol of Kaspar et al, 2008.

**Assay of respiratory function.**

To determine the respiratory phenotype of each transformant described in the ‘yeast transformation’ section of the materials and methods, cells were plated on either synthetic drop out media lacking leucine or YPG media (1% yeast extract, 2% peptone and 5% glycerol) and scored for growth at 30°C.

**o-Iodosobenzoic acid cleavage**
Cross-linking was performed under the conditions described in Kaspar et al 2008. After RNase treatment, cross-linked reactions were divided into three fractions. The first fraction was put aside as an untreated control, the second was incubated in 4M guanidine dissolved in 80% acetic acid and 180 mM p-cresol. The third fraction was incubated in 4M guanidine dissolved in 80% acetic acid and 180 mM p-cresol along with o-iodosobenzoic acid at a final concentration of 10mg/ml. Samples were then incubated for 12 hours at 37°C. After incubation, samples were TCA precipitated and run on a 15% SDS-Page gel. Cross-linked bands were resolved using a phosphorimager.
Chapter IV

Summary and Future direction
Summary:

Prior to our work, the nuclear encoded protein Pet54p of *Saccharomyces cerevisiae* was characterized as a protein that localized to the mitochondria where it was required for splicing of the aI5β group I intron as well as translation of COX1II mRNA. At this time, it was not known if the protein interacted directly on the RNAs involved in these processes to accomplish its function. We hypothesized that the protein formed RNP complexes with the aI5β intron and COX3 5’ UTL. Because little is known about the formation of such RNPs, we felt that pursuing this hypothesis would generate significant contributions to the fields of RNP formation, group I intron splicing, and yeast mitochondrial biology. Our work confirms this hypothesis and provides the first *in vitro* biochemical characterization of a yeast mitochondrial translation factor. Further, this work represents the first demonstration of an RRM domain participating in a specific interaction with yeast mitochondrial RNA. Lastly, our work represents the first characterization of a dual function group I intron splicing factor that employs the same surface to recognize both of its substrates.

A shared RNA binding site in the Pet54p protein is required for translational activation and group I intron splicing in yeast mitochondria

Our first goal with regard to Pet54p was to demonstrate that the protein bound RNA directly. To this end, we expressed and purified recombinant Pet54p from *E. coli* and assayed its interactions with the aI5β intron a cross-linking assay. This assay revealed that the cross-link formed upon interaction between Pet54p and the aI5β was not decreased by the addition of excess RNA from the unrelated COB group I intron of *A.*
This indicated that Pet54p formed a specific interaction with the ai5β intron. Furthermore, because the COB intron shares many conserved secondary structures with ai5β, these results suggested that Pet54p recognizes sequences or structures outside of the conserved group I intron core. Crosslinking experiments also revealed that Pet54p employs shared or overlapping binding sites to recognize both the ai5β intron and COX1II 5’UTL, suggesting that Pet54p employed the same surface in recognition of both substrates. Splicing assays revealed that, in the presence of MRS1, Pet54p stimulated a 1.5 fold increase in the amplitude of ligated exon formation from 0.04 ($\pm$0.005) min$^{-1}$ to 0.06 ($\pm$0.002) min$^{-1}$. These results suggested that the protein functions specifically in the second step of splicing.

Equilibrium binding assays revealed that Pet54p binds the ai5β intron with a $K_d^{\text{app}}$ of 18.3 ($\pm$ 6) nM. The protein bound to the COX3 5’ UTL with similar affinity ($K_d^{\text{app}}$ = 20.8 ($\pm$ 8) nM). Consistent with the results of the competition experiments, Pet54p did not detectably bind the A. nidulans COB intron. Binding studies using fragments of the ai5β intron and COX3 5’ UTL affirmed that Pet54p recognizes sequences outside of the catalytic core of the ai5β intron and also showed that the protein binds specifically within the first 302 nts of the COX3 5’UTL. The regions of RNA recognized by Pet54p were further refined through RNase footprinting and binding assays, which narrowed down the regions of interaction to 66nt and 78 nt in the ai5β and the COX3 5’ UTL respectively. Comparison of these regions revealed significant sequence similarities between regions recognized by Pet54p in the ai5β intron and the COX3 5’UTL.
**Pet54p interacts with the ai5β intron and COX3 5’ UTL through and unusual RRM motif.**

In a second study, cross-linking assays revealed that Pet54p bound to both the ai5β intron and the COX3 UTL with 1:1 stoichiometry. This result suggested that Pet54p possessed unusually high specificity for a single RRM containing protein. To explore the contribution of the RRM to RNA binding, mutational analysis and petptide cleavage experiments were employed. These experiments suggested that conserved amino acids within Pet54p did not contribute to binding. However, the specific cross-link reported by Pet54p in the presence of Pet54p was mapped to the RRM domain. This implies a function for the RRM in specific RNA recognition. Furthermore, the RRM of Pet54p was shown to display striking similarities to the unusual RRM contained within the Human Fox-1 protein.

**FUTURE DIRECTIONS**

**Distinguishing between two models of RNA recognition by Pet54p**

Thus far, we have explored the specificity of Pet54p largely in the context of its native substrates. The picture that emerges from this work is that of a protein able to bind to specific domains in these RNAs despite the presence of long stretches of highly repetitive sequence. What we do not know is how the binding specificity of Pet54p fits into the broader story of yeast mitochondrial RNA metabolism. For example, the respiratory subunit encoding mitochondrial mRNAs contain leader sequences ranging in size from 54 nucleotides for the COX1I UTL to over 1000 bases for the COB leader. These leaders
are essential for translation yet they possess no consensus translation sequences and information content is severely dampened by the AU rich nature of the RNAs. Further confusing matters, these leaders typically contain multiple AUG codons before the beginning of open reading frame, raising questions as to how the translation machinery identifies the appropriate start site. In other words, those features that are normally described as being “specific” to translation, the Shine-Dalgarno sequence and appropriately positioned start codon in prokaryotes, or the 5’ cap in eukaryotes, are simply not present in yeast mitochondrial mRNAs. Instead, each mRNA requires a series of 5’ leader specific and mutually exclusive co-factors for translation. Thus the kind of specificity observed in mitochondrial translation does not appear to be one of shared factors binding common consensus sequences. Rather, specificity is most likely derived from synergistic interactions between unique RNA sequences and structures along with multiple protein components specific for each transcript.

Like mitochondrial translation initiation leader sequences, group I intron splicing within *Saccharomyces cerevisiae* mitochondria appears to rely on specific proteins rather than a general pathway. While all group I introns contain a conserved catalytic core, they display significant variation in their peripheral RNA structure. Furthermore, many group I introns are punctuated by long open reading frames and insertions. Again, there are no consensus sequences for protein facilitated group I intron splicing. Instead, proteins may stimulate the formation or melting of secondary and tertiary structures in a manner that is unique to each intron. In these cases, the structure of the RNA often plays a significant role in the specificity of the interaction.

We can say, from genetic analysis, that Pet54p specifically facilitates translation
of COX3 or splicing of aI5β. From our own binding and competition studies, we can put
fourth the claim that the protein interacts specifically with a single region of the aI5β and
the COX3 UTL. However, we cannot say with any certainty how much of this observed
specificity is a function of Pet54p, and how much derives from the states of the RNA
ligands themselves. Put another way, we have not achieved the level of resolution in our
assays to determine the “consensus binding site” for Pet54p. Without fully understanding
the nature of this site, we are faced with two possible explanations for the observed
binding properties of Pet54p. The first is that Pet54p is able to discriminate between
many similar sequences because of the inherent specificity of the protein. The second
possibility is that the protein can bind many sequences, but, due to the high AU richness
of both RNA ligands, most potential binding sequences for Pet54p are hidden within
secondary structure and, therefore, unavailable for binding. In the first hypotheses, the
protein alone is largely responsible for specificity while in the second; the state of the
RNA also plays a large role.

Determining the nature of the consensus site has important evolutionary
implications as well. In chapter II of this thesis, we proposed that Pet54p most likely
became involved in splicing of aI5β through fortuitous interaction with sequences or
structures in the intron resembling those in the COX3 UTL. This hypothesis was based
upon footprinting and boundary assay data which indicated that the protein was
recognizing similar sequences in both RNAs. Determining the range of RNAs that
Pet54p can potentially bind will give an indication of how fortuitous the association
between aI5β and Pet54p actually is which, in turn, will help to answer the question of
why Pet54p became involved in splicing of a single optional group I intron in an intron
Here I propose to address these questions through a series of SELEX experiments. SELEX is an unbiased way to determine the RNA binding specificity of a protein through the binding and selection of high affinity RNAs from a random pool of sequences (Tuerk and Gold, 1990). The selection procedure employed will be similar to that of Tacke and Manley (1995) in which SELEX experiments were used to explore the RNA binding specificity of RRM containing proteins. In our experiments, his tag containing Pet54p will be immobilized on Ni⁺ Agarose beads and exposed to saturating amounts of RNAs containing a 20 base stretch of random sequence and a conserved primer-binding region. RNAs that bind to Pet54p are then isolated, reverse transcribed and amplified by PCR. After amplification, these DNAs are transcribed by T7 polymerase to generate RNAs that will be used in further rounds of selection (Tacke and Manley, 1995).

After multiple rounds of selection, cDNAs are sequenced, and these sequences are aligned and compared. When single RRM containing proteins from the SR family were profiled in this way, the result was generally selection for a series of similar sequences. These sequences can be aligned to produce a consensus sequence representing the base selected for with the highest frequency, on average, in each position of the randomized sequence. To ensure that the sequences derived from SELEX experiments reflect a specific interaction with Pet54p, cross-linking and affinity binding experiments will be performed with RNAs isolated from the assay.

The SELEX experiment should favor either the hypothesis that specificity is governed predominantly by the Pet54p protein, or the model whereby the RNA structure also makes significant contributions to specificity. For example, if a great deal of
sequence variation is seen in the RNAs selected for through Pet54p binding, and these variant sequences occur frequently within the α15β intron and COX3 UTL, then at least some of the observed specificity of the protein most likely comes from the high degree of secondary structure in Pet54p’s natural ligands. On the other hand, if the SELEX library heavily favors a single consensus sequence and, this sequence appears to occur only once in the intron and UTL, then the hypothesis that Pet54p possesses inherent high specificity is favored. In this way, we can begin to understand the nature of RNA binding specificity in yeast mitochondrial mRNA metabolism.

**Determining the contribution of the RRM domain to Pet54p affinity, specificity and catalysis of the α15β intron.**

With a molecular weight of only 35 kDa, Pet54p is a relatively small protein. However, this versatile factor accomplishes a great deal for its size. It binds specifically to two distinct RNA substrates though an overlapping binding site, yet it may possess distinct effector domains for its differing roles in splicing and translation. We have made progress in determining the binding properties of Pet54p, the regions and sequences of RNA recognized by the protein, and its function in splicing. Furthermore, we have mapped a specific RNA/Protein crosslink to the RRM domain. Despite these accomplishments, we know very little about how the different regions of the protein work together to facilitate splicing or translation.

In determining the roles of individual effector domains in the function of Pet54p, the most reasonable place to start is with the RRM. This is the only domain within Pet54p that has been characterized in other proteins. Furthermore, RRM domains can
often be expressed independently. For example, the closely related CstF-64 RRM domain has been expressed as a soluble peptide capable of binding with high affinity and producing a specific crosslink (Monarez et al., 2007).

For this reason, I propose to express the Pet54p RRM as an independent domain. This will allow us to answer questions about the role of this domain in splicing as well as the overall contribution of the RRM to the binding affinity and specificity of Pet54p.

The first question we will address with the recombinant RRM protein is whether RRM binding alone is sufficient to facilitate the increase in ligated exon observed in the α5β splicing assays reported in chapter II. As described in the conclusion of chapter III, the Pet54p RRM is homologous to the human Fox-1 RRM, which induces a bend in its RNA substrate. If the Pet54p RRM induces a similar conformational change, it may be sufficient to facilitate splicing. Comparison of Pet54p wild type protein with the RRM alone in splicing assays will answer this question.

The binding properties of the domain will be assayed through filter binding and crosslinking experiments similar to those described in chapters II and III. Filter binding experiments will allow for a determination of the $K_d$ value of the RRM alone. Comparison of this $K_d$ value to that of the wild type protein should allow us to calculate the contribution of the RRM domain to the overall affinity of Pet54p for its native substrates. Previously, we have used crosslinking experiments to demonstrate a specific interaction between Pet54p and its cognate RNA sequences. Comparison of the crosslinking profile of the Pet54p RRM domain alone with that of wild type protein should reveal whether RRM alone is sufficient for specific binding.

Binding and crosslinking experiments comparing the single RRM domain to the
full-length wild type protein should yield information about the overall contribution of the RRM to RNA binding and specificity. However, exploring the specific contributions to RNA recognition made by each region of Pet54p in more detail will require a more nuanced approach. Thus, another series of SELEX experiments will be conducted to more accurately gauge the total range of sequences that full length Pet54p or the RRM alone can bind with high affinity. This study will be performed in the same manner as the previously described SELEX experiments only, this time the Pet54p RRM domain alone will be used to bind RNAs.

If regions outside of the RRM play a role in determining the specificity of Pet54p, then comparison of the SELEX data for the wild type protein with that of the RRM alone will reveal this. For example, if the overall profile of RNAs recovered from the wild type protein and the RRM alone are the same, then it is unlikely that peripheral regions play a significant role in RNA binding. However, if regions outside the RRM do make a contribution to specificity, then the profile of the wild-type protein will be narrower than that of the RRM alone. It is not uncommon in RNA binding proteins for regions outside of the RRM to make contributions to binding specificity, and this would be consistent with the hypothesis that Pet54p possesses unusually high specificity for a single RRM containing protein, perhaps due to the contributions of ancillary domains.

The SELEX experiment should also allow us to more effectively interpret the RNase footprinting data described in chapter II. In those experiments, discontinuous protections were observed over a relatively long stretch of RNA in both the COX3 5’ UTL and the α15β intron. For this reason, it was impossible to separate protections resulting from specific RNA recognition from those that may have been brought on by
changes in RNA conformation upon Pet54p binding. If the SELEX result indicates a consensus sequence, and that sequence overlaps with protected regions in α5β and the COX3 UTL, we can say with reasonable certainty that those regions are protected through specific RNA binding. We can then affirm this by mutating the RNAs in these regions and assaying for loss of binding in the presence of Pet54p.
Bibliography


An allosteric-feedback mechanism for protein assisted group I intron splicing. *RNA*, 13, 211-222.


Monarez, R. R., MacDonald, C. C., and Dass, B. (2007). Polyadenylation proteins CstF-


