8-2018

Insight into Translational Activation in Yeast Mitochondria

Julia Lynn Jones
Rowan University

INSIGHT INTO TRANSLATIONAL ACTIVATION IN YEAST MITOCHONDRIA

Julia Lynn Jones, B.A.

A Dissertation submitted to the
Graduate School of Biomedical Sciences, Rowan University
in partial fulfillment of the requirements for the Ph.D. Degree.

Stratford, New Jersey 08084

August 2018

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to first thank my mentor, Dr. Michael Anikin, for not only teaching me how to become a scientist, but for providing me with the best example possible on how to mentor a student. I thoroughly enjoyed each day of my training and I am tremendously grateful for the time, commitment, and patience given to me while in his laboratory.

I am grateful to the members of my thesis committee: Dr. Dmitry Temiakov, Dr. Michael Henry, Dr. Natalia Shcherbik, and Dr. Salvatore Caradonna for their assistance, time, guidance, criticism, and praise throughout my doctoral training.

I am thankful for all of the past and present members of our laboratory group who would stop anything to help me and who shared in my joys of successful experiments and my wallowing of failed experiments. You all made each day truly enjoyable.

To Shawna Rotoli, Jennifer Romer-Siebert, and Shin-Yi Lin, thank you for all of the encouragement and support- always.

I am thankful to Rowan University and the GSBS for providing me with this opportunity and for all of the assistance along the way.

This work was supported, in part, by a New Jersey Health Foundation grant (to MA) and from the Society of Research Scholars and the Rowan University GSBS (to JJ).

I want to thank my late Grandfather, Albert M. Forte, Jr., and my Mother for instilling in me the passion to fight and work hard for whatever I wanted in life. I am eternally grateful for my husband, Jeffrey Jones, who has loved, encouraged, and supported me in every step of the way. As for my children, you mean the absolute world to me and no matter what I may accomplish in my life; nothing will be greater than the existence of you. I am grateful to my amazing in-laws, Jeffrey and Shawn Jones, who selflessly helped me in any way possible to ensure that my degree was completed.

Lastly, but most importantly, I am grateful to God for providing such an amazing support system and for giving me the knowledge and ability to complete this degree.

ABSTRACT

Mitochondrial function depends on over a thousand proteins, of which the majority are nuclear DNA-encoded and approximately one percent are mitochondrial DNA-encoded. The mitochondrial DNA of *Saccharomyces cerevisiae* contains eight protein-encoding genes, seven of which are required for proper function of the respiratory complexes and one encodes a ribosomal protein. The bigenomic nature of the oxidative phosphorylation complexes requires coordinated expression and regulation from both the nuclear and the mitochondrial genomes. It is currently unclear how this regulatory network operates. However, it is thought that nuclear genome-encoded messengers localized to the mitochondria aid in this coordination.

A family of proteins termed mitochondrial translational activators has been shown to control the expression of all mitochondrial-encoded protein genes in a gene-specific manner. Evidently, each mitochondrial mRNA is regulated by a specific protein or a subset of proteins that permits translation of that transcript. If these factors are absent, there will be no synthesis of the polypeptide; leading to a respiratory deficient cell. Although a functional link between translational activators and the mitochondrial genes they control has long been established genetically, the activation mechanism is entirely unknown. This study focuses on the mechanism of activation of two representative members of this translational activator family, Pet111p and Cbp1p. These activators are required for the expression of *COX2* and *COB,* respectively. Translational activators have been shown to have multiple

functions including, transcript stabilization and mRNA localization to the membrane, as well as to the translation machinery. Current genetic data suggests that both Pet111p and Cbp1p interact with RNA targets in the 5′-untranslated regions of *COX2* and *COB,* respectively. However, neither the identity of these sites, nor has the ability of these proteins to interact with RNA ever been demonstrated. The objective of this study is to characterize the functional mechanism of translational activation for both Pet111p and Cbp1p and ultimately learn how they aid in the coordination of dual-genomic expression.

INTRODUCTION

A. Mitochondrial function

The origin and evolution of mitochondria in eukaryotic cells has long captivated the mind of scientists. There are many theories as to how eukaryotic cells obtained the ability to respire. The endosymbiotic theory is predominate in the field (*1*). It proposes that eukaryotes arose from an ancestral prokaryote that engulfed an aerobic heterotrophic archaea. This symbiotic relationship proved to be beneficial to both species and led to a mitochondrion in ancestral heterotrophic eukaryotes. Over time, they evolved into the cells of species today, including fungi and animals (*2*). Whatever the origin of the modern day mitochondrion in eukaryotes, this interdependent relationship is essential to life as we know it today. Throughout this evolutionary process mitochondria have maintained essential characteristics such as, a double membrane and the ability to produce adenosine triphosphate (ATP). However, many changes have also occurred, including a transfer of genetic material to the nuclear genome, along with acquiring many more essential biochemical functions within the cell (*2-4*).

All of the structural features of the mitochondrion are necessary to its various functions. The phospholipid double membrane is composed of an outer and inner membrane. The outer membrane surrounds the organelle and is exposed to the cytoplasm. Metabolites are first shuttled between the cytosol and the intermembrane space via aqueous pores. Transporters are then required for these metabolites to transverse the inner membrane into the

matrix. Larger molecules, such as proteins, utilize a multifaceted protein import system that spans to the matrix (*5*). The inner membrane is composed of complex folds called cristae which encompass the mitochondrial matrix. The mitochondrial matrix is where multiple biochemical processes occur, including the citric acid (TCA) cycle, a part of the urea cycle, and fatty acid metabolism. The cristae contain the components of oxidative phosphorylation (OXPHOS) and serve as an anchor for transcription and translational processes. Between the outer and inner membranes lies the intermembrane space. The innermembrane space harbors the protein cytochrome c, which carries electrons between complexes III and IV of the respiratory chain. This space also contains components needed for fatty acid metabolism (*2, 6, 7*).

Other biological processes, such as energy production, storage of essential calcium ions, thermogenesis, and programmed cell death are all functions attributed to mitochondria (*8*). However, the most vital and well-known function is its ability to produce energy. Most eukaryotic organisms produce energy aerobically by utilizing oxygen. Respiration occurs in mitochondria when oxygen is present, and produces more energy than fermentation, which takes place in the cytosol and utilizes fermentable sugars. Yeast monitor the amount of oxygen and fermentable sugars present, internally and externally, to determine which energetic pathway they should use (*9*). The ability to switch between the two types of energy production is greatly beneficial to this species, as it allows it to survive a variety of environmental conditions. This trait also makes yeast an ideal model

organism to use when studying various processes, especially the process of respiration.

Production of energy from aerobic respiration begins with the breakdown of various substrates into acetyl-CoA or TCA cycle intermediates. The TCA cycle oxidizes these molecules to produce GTP, nicotinamide-adenine dinucleotide (NAD), and flavin adenine dinucleotide ($FADH_2$). The reduced electron carriers then pass the electrons into a chain of electron acceptor compounds, known as the electron transport chain. As the electrons migrate through the OXPHOS complexes, protons are pumped from the matrix into the intermembrane space, creating an electrochemical gradient across the inner mitochondrial membrane. Finally, this proton gradient enables ATP synthase (complex V) to produce biochemical energy in the form of ATP.

In *S. cerevisiae* OXPHOS is composed of complexes II, III, IV, and V. Each complex contains multiple subunits and catalyzes the oxidation of a corresponding substrate. Complex II, succinate dehydrogenase, acts to donate electrons from $FADH_2$ to reduce ubiquinone to ubiquinol. The electrons are transferred to coenzyme Q; however, no protons are transported to the intermembrane space. Complex III, cytochrome *c* reductase, catalyzes the reduction of cytochrome *c* conjugated with the oxidation of coenzyme Q, and translocates four protons into the intermembrane space. Complex IV, cytochrome *c* oxidase, catalyzes the last reaction of the electron transport chain, in which the electrons from

cytochrome *c* are transferred to molecular oxygen. Complex IV contributes to the proton gradient by translocating four protons per reduced O_2 molecule. The last complex of OXPHOS, complex V or ATP synthase, is composed of two subcomplexes, F_o and F_1. The F_o sector captures energy from the proton gradient. This energy causes the F_1 sector, which projects into the mitochondrial matrix, to rotate driving the conversion of adenosine diphosphate (ADP) and phosphate (Pi) to ATP (Figure 1) (*10*).

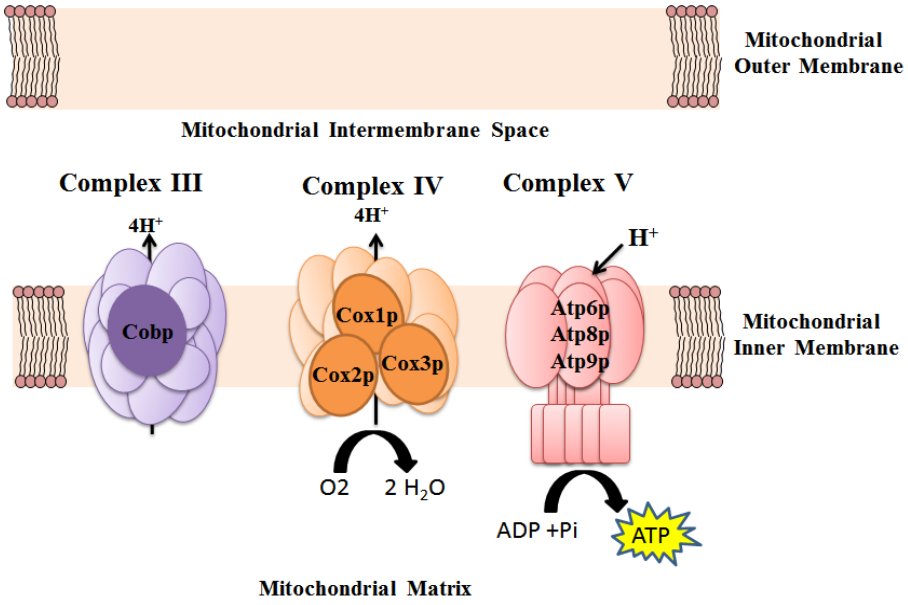

Figure 1. The inner mitochondrial membrane contains OXPHOS complexes that produce a proton gradient. Complexes III, IV, and V of OXPHOS are embedded in the inner mitochondrial membrane. Reactions from complexes III and IV produce a proton gradient (as shown) that ultimately drives ATP synthesis by complex V; providing fuel for the entire cell.

The utilization of molecular oxygen as the final electron acceptor is beneficial due to its high oxidative potential; however, it results in the

production of reactive oxygen species (ROS) as byproducts. These byproducts include superoxide, hydrogen peroxide, and hydroxyl radicals, which are very damaging to the cell. Since their destructive potential depends on their concentrations, the process of OXPHOS must be highly regulated to keep them to a minimum (*11*). This regulation occurs at multiple levels, including transcription and translation of the mitochondrial OXPHOS genes, to ensure that the stoichiometry of the complexes is maintained.

In recent decades the scientific community has taken greater interest in studies regarding mitochondria due to the many links that have been found between mitochondrial dysfunction and human disease. Initially, most mitochondrial dysfunction was correlated to defects in the mitochondrial genome and respiratory chain components. However, additional research has focused on defects in other aspects of mitochondrial function linked to cancers, metabolic disorders, and neurodegenerative diseases (*12-14*).

B. Mitochondrial gene expression

The overall function of the mitochondria depends on over 1,000 proteins, the majority of which are nuclear genome-encoded. Approximately 1% of the proteins required for mitochondrial function are encoded by the mitochondrial genome (*15*). Mitochondria carry a residual circular genome, from which most of the genes have either been translocated to the nuclear genome or no longer exist. Therefore, neither the gene products from the nuclear genome nor those from the mitochondrial genome alone are sufficient

to support mitochondrial energy production in eukaryotic cells. The mitochondrial ribosome and each complex of yeast OXPHOS, except complex II, are composed of both mitochondrial and nuclear genome-encoded subunits.

To ensure the concerted assembly of subunits from each genome, the required genes must be expressed stoichiometrically and the subunits must synchronously associate with each other at the inner mitochondrial membrane. This coordination is not restricted to the subunits of OXPHOS; the nuclear genome also provides mitochondrial import systems, chaperones, assembly factors, proteins involved in the biosynthesis and incorporation of prosthetic groups, mitochondrial inner membrane insertion factors, and proteins that aid in the disposal of unassembled subunits (*16*). Moreover, both the nuclear and mitochondrial genetic systems have separate transcription, transcript maturation, and translation pathways that must be coordinated (*15, 17*). Although there are a few examples in the literature that describe this coordination at specific steps, how it functions globally is still largely unknown.

It is known that the nucleus sends messengers in the form of proteins to the mitochondria (*18*). Once received, these proteins are believed to have multiple functions, one of which is to regulate mitochondrial gene expression (*18*). Consequently, mitochondria send signals to the nucleus reporting back on their current state (*19*). An example of such concerted genome expression can be seen in the biogenesis of cytochrome *c* oxidase. Complex IV involves three mitochondrial DNA-encoded subunits (Cox1p, Cox2p, and Cox3p) that

form the catalytic core of the enzyme and an additional nine nuclear DNA-encoded subunits. There are twenty auxiliary nuclear DNA-encoded factors that aid in the import, processing, and assembly of this complex at the inner mitochondrial membrane (*17*). All of the various subunits and accessory proteins must have a coordinated expression and localization to ensure proper complex biogenesis. Therefore, nuclear and mitochondrial gene expression, as well as the import of nuclear gene products into the mitochondria must be highly concerted and regulated. This example shows that OXPHOS complex formation is very intricate.

Another well studied case of coordinated bigenomic gene expression is the biogenesis of the *S. cerevisiae* ATP synthase complex. All of the F_1 subcomplex subunits are imported, whereas three out of the five hydrophobic core proteins of the F_o subcomplex (Atp6p, Atp8p, and Atp9p) are encoded by the mitochondrial genome (*15, 20*). Although the proteins have different origins, they all come together and assemble flawlessly due to the tight regulation of the process.

Atp6p and Atp8p are encoded by a mitochondrial message, the products of the *ATP22* and *AEP3* nuclear genes act as specific translational activators to control their synthesis *(21, 22)*. In addition, expression of both *ATP6* and *ATP8* requires the presence of the F_1 complex or an F_1 pre-assembly complex, which enables translation of the bicistronic *ATP8/ATP6* mRNA by targeting the untranslated regions (UTRs). This interaction creates a feedback loop that controls the synthesis of Atp6p and Atp8p in conjunction

with the cytosolic synthesis of the components of the F_1 complex (*18, 23*). The significance of this regulatory feedback loop is that it maintains the levels of the mitochondrial DNA-encoded Atp6p and Atp8p equal to the levels of the assembled nuclear DNA-encoded F_1 subunits. This is a clear example of how nuclear and mitochondrial gene expression is tightly coordinated to carry out cellular processes. As shown in Figure 2, complexes III and IV of OXPHOS are also composed of both nuclear genome-encoded (white subunits) and mitochondrial genome-encoded (colored subunits) proteins, therefore their biogenesis must involve coordinated bigenomic gene expression as well. These examples illustrate only a small part of the intricacies of mitochondrial and nuclear genomic cross-talk. This process of multifaceted coordinated gene expression is still largely uncharacterized.

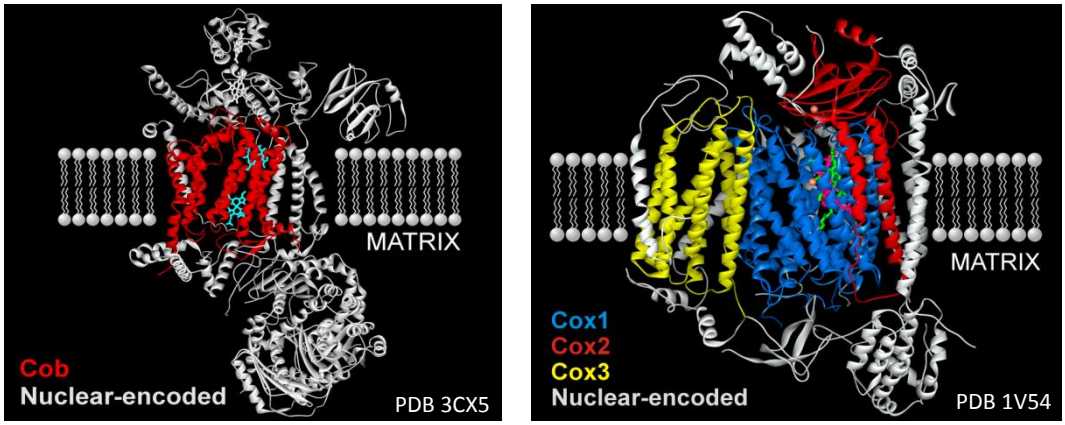

Figure 2. Mitochondrial OXPHOS complexes III and IV are composed of both nuclear and mitochondrial gene products. The left image shows a crystal structure of yeast complex III (PDB 3CX5) and the right image shows a crystal structure of bovine complex IV (PDB 1V54). These complexes, along with complex V (not shown), contain products of both nuclear (shown in

white) and mitochondrial genes (colored). Their biogenesis therefore requires a concerted expression of genes in both genomes (*24*).

B.1. Organization of the mitochondrial genome in *S. cerevisiae*

The *S. cerevisiae* mitochondrial genes are assembled in an 85.8 kilobase genome (Figure 3). The genome is maintained by nuclear-encoded factors that aid in replication, recombination, repair, and packaging. The mitochondrial DNA (mtDNA) is organized into nucleoprotein complexes called nucleoids. These nucleoids can contain up to ten genomes and can be dynamic structures depending on the cells environment (*24, 25*). In respiring conditions, nucleoids form a more open complex allowing replication, whereas under starvation conditions the opposite occurs (*26*).

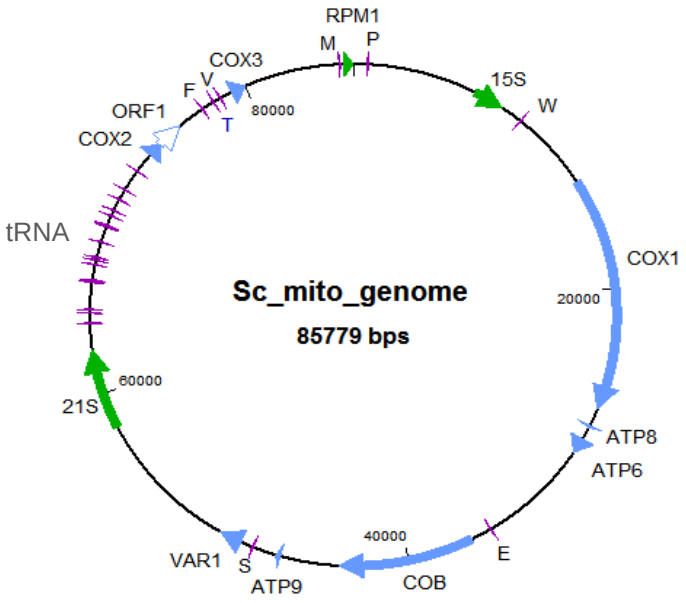

Figure 3. The mitochondrial genome of *S. cerevisiae* contains essential genes for respiration. The *S. cerevisiae* mitochondrial genome is believed to have undergone extensive reduction and transfer of genes to the nucleus.

Almost every gene remaining in the mitochondrial genome encodes a highly hydrophobic protein essential for OXPHOS. Having these proteins readily available when needed aids in the regulation of energy production by the mitochondria (*2, 27*). Protein genes are denoted in light blue, structural RNA in green, and tRNAs in purple.

Although the yeast mitochondrial genome is relatively large, it has a low gene density and contains widespread non-coding regions (*28*). There are eight proteins encoded in the mitochondrial genome and synthesized within the organelle. Seven of them are a part of the OXPHOS chain and are highly hydrophobic. These include apocytochrome b of complex III, encoded by *COB*, subunits I, II, and III of complex IV, encoded by *COX1*, *COX2,* and *COX3*, respectively, and subunits 6, 8, and 9 of complex V, encoded by *ATP6*, *ATP8,* and *ATP9*, respectively. The genome also encodes a hydrophilic polypeptide of the small subunit of the mitochondrial ribosome, encoded by *VAR1* (*15, 24*) (see Figure 3). As shown in Figure 4, in addition to these eight genes, a complete set of transfer RNAs (tRNAs), both ribosomal RNAs (rRNAs), the RNA component of RNaseP (RPM1), and non-essential open reading frames are encoded.

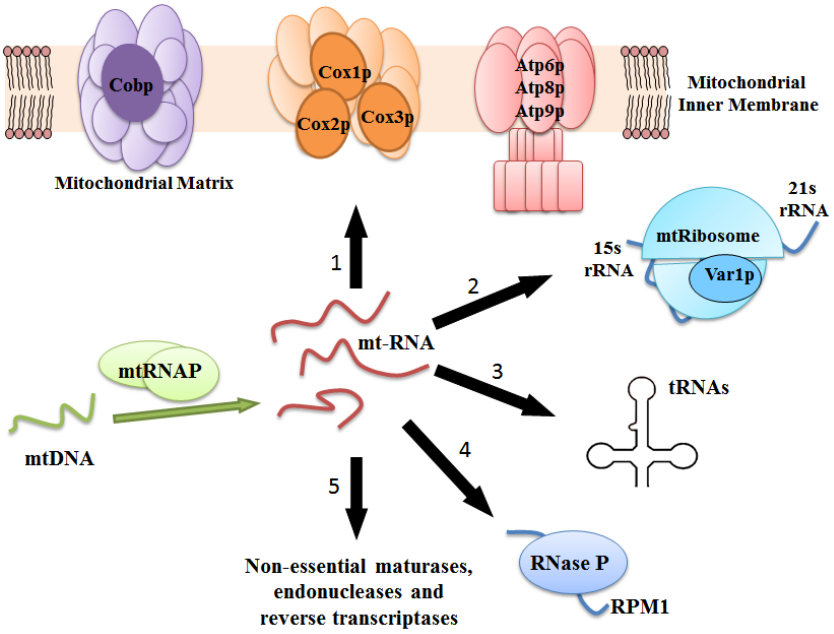

Figure 4. The mitochondrial genome encodes for essential respiratory proteins and RNAs. The mitochondrial genome is transcribed by a nuclear DNA-encoded RNA polymerase composed of the Rpo41p and Mtf1p proteins. The mt-RNA codes for: (1) OXPHOS proteins Atp6p, Atp8p, and Atp9p of complex V, Cox1p, Cox2p, and Cox3p of complex IV, and Cobp of complex III; (2) Var1p, 21S rRNA, and 15S rRNA of the mitochondrial ribosome; (3) a full set of tRNAs; (4) RPM1, the RNA component of RNaseP, and (5) a group of intron-encoded proteins including maturases, endonucleases, and reverse transcriptases, all of which are non-essential in intron-free strains (*15*).

B.2. Transcription in yeast mitochondria

The transcriptional machinery in the mitochondria is distinct from that of the nucleus. The mitochondrial genome is transcribed by a mitochondrial RNA polymerase (mtRNAP) composed of two nuclear genome-encoded subunits, the Rpo41p polymerase, and its transcription initiation factor, Mtf1p. Rpo41p has the ability to bind DNA; however, this interaction is nonspecific.

When Rpo41p is in complex with Mtf1p, the binding affinity for promoter sequences is increased (*29*). When compared to promoters of other RNA polymerases, the yeast mitochondrial promoter is relatively short, 9 base pairs (Table 1). The promoter is AT-rich with a general consensus sequence of: 5′-(-8) ATATAAGTA (+1)-3′ (underlined in Table 1), where +1 is the transcription start site (*30, 31*). However, several promoters are known to have substitutions (shown in red) relative to the consensus sequence. The strength of the promoter depends on the nucleotides that flank the consensus sequence. Variations in sequences upstream of the consensus may result in a fifteen- to twenty-fold change in promoter activity (*27, 32*). There are 14 known promoters in the yeast mitochondrial genome; 11 account for the expression of all of the mitochondrial genes, while 3 are associated with origins of replication (Table 1) (*33*).

Tight association of Rpo41p and Mtf1p with a promoter does not guarantee transcription initiation. It is believed that transcription initiation at some promoters is controlled by the concentration of ATP in the matrix (*34*). Besides this promoter-selective transcription control, there are no other known factors that control yeast mitochondrial transcription either in a gene-specific or global manner. It is currently unknown whether any other factors are necessary for elongation. The mechanism of termination is also unknown.

Table 1. Yeast mitochondrial promoters.

Primary Transcripts	Nonanucleotide Promoter
15s_rRNA, Trp	-8 +1 5'-TATTATATAAGTAATAA-3'
COX1, ATP8, ATP6	ATTGATATAAGTAATAG
Glu, *COB*	TATTATATAAGTAATAT
ATP9, Ser2, *VAR1*	ATTAATATAAGTAATAT
21s_rRNA, Thr, Cys, His	ATATATATAAGTAGTAA
Leu, Gln, Lys, Arg, Gly, Asp, Ser, Arg2	TGTTATATAAGTAATAA
Ala, Ile, Tyr, Asn, Met	TTTATTATAAGTAATAA
Phe, Val, *COX3*	TATATTATAAGTAATAA
Thr2	ATTTTTATAAGTAGTAT
fMet, RPM1, Pro	GATTTTATAAGTAATAT
COX2, Q0255	TTAATTAAAAGTAGTAT
ori2	AAATATATAAGTAATAA
ori3	AGATATATAAGTAATAG
ori5	AAATATATAAGTAATAG

Within the mitochondria, most protein and structural RNA genes are co-transcribed with tRNA(s) and some transcripts contain more than one protein coding sequence. These primary transcripts (mt-mRNA) undergo multiple steps of processing before they are presented to the mitochondrial ribosome for translation. These events include: gene separation, tRNA excision, splicing, and maturation (*27*). Current data suggest that most of the regulation of mitochondrial gene expression occurs post-transcriptionally (*24*).

B.3. Processing of mitochondrial mRNAs

The rate of synthesis and degradation of mt-mRNA is very high (*35*). In order for a transcript to be translated in a regulated manner, there must be systems in place to ensure that degradation and synthesis are finely balanced. Unlike in humans, *S. cerevisiae* mt-mRNA does not contain a

polyadenylated 3′-tail; a feature known to stabilize and aid in translation of the transcript (*36*). Instead, 5′- and 3′-UTRs flank the open reading frame (ORF) of mt-mRNAs to provide stability and translation control. The UTRs are generally very rich in secondary and tertiary structures and can range from 54 to over 1000 nucleotides in length. The structuring of the UTR is innate to RNA and may play a role in protein binding, provide protection from ribonucleases, and enable association with the ribosome (*37*).

Recent data has shown that there are hundreds of modifications that RNA molecules, including mRNA, can undergo to promote or inhibit various functions (*38*). It has been shown that the 5′-end of mRNAs can be chemically modified with different molecules. A modification of transcripts with NAD^+ attached to the 5′-end was first identified in *E. coli* and has since been found in both the cytosol and mitochondria matrix of *S. cerevisiae*. Although this modification appears to have a “capping” function, it has yet to be determined whether it plays the same role as a 5′-7-methylguanosine cap. This modification is believed to occur because Rpo41p utilizes NAD^+ as the initiating nucleotide during transcription initiation. A 5′-NAD^+ modification has been found on four yeast mitochondrial transcripts, *ATP9*, *COX2*, and the 15S and 21S rRNAs. The biological significance of such an addition still needs to be investigated (*39*).

Various studies have shown that mRNA UTRs are essential for proper expression of proteins in mitochondria. Each mitochondrial mRNA is believed to have a set of protein factors that protect it from degradation and activate its

translation by specifically interacting with the 5′-UTR. These mRNA specific nuclear genome-encoded proteins are clearly required for respiration; because without them mRNA is rapidly degraded and protein synthesis does not occur (see section B.4. for specific examples) (*18*). The 3′-UTRs of all yeast mitochondrial mRNAs carry a conserved dodecameric sequence (DS), which has been linked to 3′-processing, maturation, and stabilization of the messages (*40*). A group of proteins has been shown to interact with the DS (*41, 42*); however, the identities of the factors in this complex are still unknown. This interaction is thought to protect mRNAs from 3′-exonucleolytic degradation by the mitochondrial degradosome (mtExo) and establish the end point of 3′-end trimming (*40*).

Most primary transcripts have been shown to undergo maturation at the 3′- and 5′-UTRs prior to translation. The maturation at each end is performed by different exonucleolytic complexes. Maturation at the 5′-UTR was genetically shown to depend on the nuclear *PET127* gene (*43-45*); however, it is yet to be identified if this is the enzyme responsible for the maturation. The activity of Pet127p was evident from previous *in vivo* data, in which the processing of the *COB* primary transcript did not occur in the absence of *PET127* (*45*). Similarly, the 5′-end processing of *VAR1*, *ATP8/6,* and the *15S* rRNA is blocked in Δ*PET127* cells (*43*). These results independently implicate Pet127p in 5′-exonucleolytic mRNA maturation. Although Pet127p is targeted to the mitochondria (*43*) and is genetically

linked to 5′ to 3′ exoribonuclease activity (*44*), experimental evidence directly demonstrating this activity is still missing.

Processing at 3′-UTRs appears to have the same mechanism for most mt-mRNAs. In the early 1980's Osigna et al. showed that almost all of the mitochondrial protein coding regions, and some non-coding sequences, contain the same DS downstream of their open reading frames (5′-AAUAAUAUUCUU-3′). In the one exception, *ATP9*, this sequence is altered by one nucleotide (5′-AAUAACAUUCUU-3′). However, this substitution does not affect the function of the DS (*40*). Interestingly, each transcript is trimmed back to exactly two nucleotides downstream from this sequence (*46*), consistent with the idea that this sequence marks the point of 3′-end processing.

How this maturation occurs at the mechanistic level and how the DS is integrated in the process is still unclear. Although a protein factor has been isolated that binds to the DS signal with subnanomolar affinity, the identity of this factor remains unknown (*41, 42*). It is believed that this protein, alone or as a part of a group of proteins, may mark the maturation point and provide stability when bound to the DS. Trimming at the 3′-end during maturation is thought to be performed by mtExo, which is an ATP-dependent two-subunit 3′-exoriboexonuclease composed of the Dss1p nuclease and Suv3p RNA helicase (*47*). However, the ability of the isolated DS binding protein to block 3′-degradation by mtExo has never been demonstrated experimentally.

It is worth noting that the maturation of at least one primary transcript, the *COX1-ATP8-ATP6* mRNA, must be mechanistically distinct. This transcript is processed at the DS following the *COX1* ORF to separate it from the bicistronic *ATP6-ATP8* message. As a result, the 3′-end of mature *COX1* is generated at a point just upstream from the 5′-end of the longer form of the mature *ATP6-ATP8* transcript. This implies that the separation must occur endonucleolytically at a point specified by the DS in the *COX1* 3′-UTR. It is currently unclear if the putative DS binding protein can recruit an endonuclease to this processing site or whether this type of processing is limited to the *COX1-ATP8-ATP6* primary transcript (*40*).

Maturation at the UTRs represents only one processing event that mt-mRNAs precursors undergo. Most mitochondrial transcripts either contain additional ORFs or tRNAs present in addition to coding sequences (see Table 1). These transcripts must be cleaved in order for individual mRNAs to be translated and the tRNAs to become functional. tRNAs are first excised from their primary transcripts and then processed at both the 5′- and 3′-ends. The 5′-tRNA processing is carried out by RNaseP, which consists of nuclear genome-encoded Rpm2p and mitochondrially encoded RPM1 RNA. The 3′-tRNA processing and separation of genes is performed by RNase Z *(24, 48)*. As discussed above, cleavage by an unknown endonuclease at the DS may separate the coding regions in some polycistronic transcripts, allowing downstream processing events to occur independently on an individual transcript level (*40*).

Introns are present in three primary mitochondrial transcripts, pre-*COX1* (7 introns), pre-*COB* (5 introns), and pre-21S rRNA (1 intron), therefore, the processing of some transcripts includes splicing. They all belong to group I or II introns which undergo protein-assisted autocatalytic splicing *in vivo (27, 49)*. Introns are believed to be excised utilizing intron-encoded maturases and nuclear genome-encoded splicing factors; some of these factors act globally on all introns, while others are specific to particular transcripts (*24*). Ten of the introns encode ORFs that are in frame with the preceding exons and encode 10 additional polypeptides. Some of these products are non-essential endonucleases and reverse transcriptases (Figure 5). However, the intron-encoded RNA maturases ensure proper splicing and are only non-essential in intronless strains (*27*). The intronless yeast strains do not display significant phenotypes, showing that no aspect of mitochondrial function besides RNA splicing depends on the intron-encoded gene products (*25*).

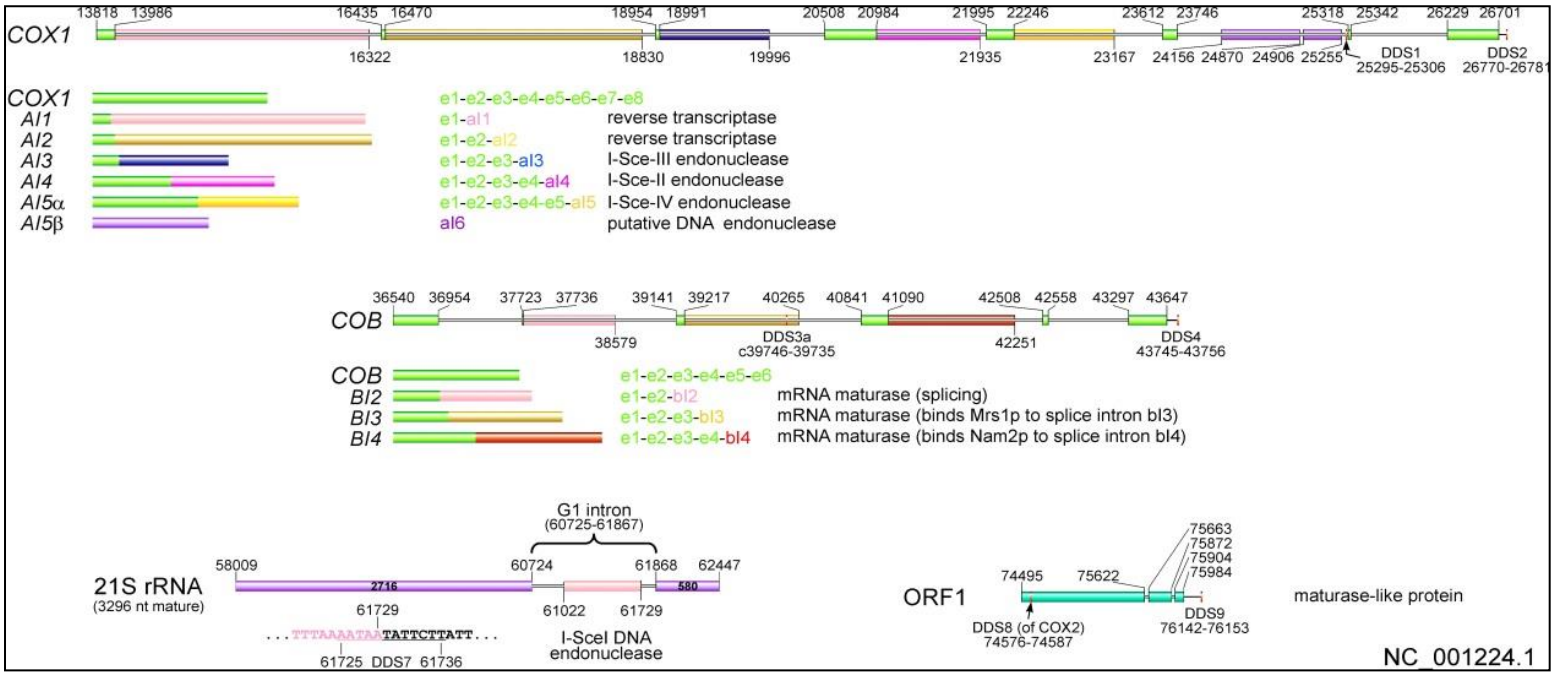
COX1
13818 13986 16435 16470 18954 18991 20508 20984 21995 22246 23612 23746 25318 25342 26229 26701
16322 18830 19996 21935 23167 24156 24870 24906 25255
DDS1 25295-25306
DDS2 26770-26781
COX1 e1-e2-e3-e4-e5-e6-e7-e8
AI1 e1-aI1 reverse transcriptase
AI2 e1-e2-aI2 reverse transcriptase
AI3 e1-e2-e3-aI3 I-Sce-III endonuclease
AI4 e1-e2-e3-e4-aI4 I-Sce-II endonuclease
AI5α e1-e2-e3-e4-e5-aI5 I-Sce-IV endonuclease
AI5β aI6 putative DNA endonuclease
COB
36540 36954 37723 37736 39141 39217 40265 40841 41090 42508 42558 43297 43647
38579 DDS3a c39746-39735 42251 DDS4 43745-43756
COB e1-e2-e3-e4-e5-e6
BI2 e1-e2-bI2 mRNA maturase (splicing)
BI3 e1-e2-e3-bI3 mRNA maturase (binds Mrs1p to splice intron bI3)
BI4 e1-e2-e3-e4-bI4 mRNA maturase (binds Nam2p to splice intron bI4)
G1 intron (60725-61867)
21S rRNA (3296 nt mature)
58009 60724 61868 62447
2716 580
61729 61022 61729
...TTTAAAATAA TATTCTTATT...
61725 DDS7 61736
I-SceI DNA endonuclease
ORF1
74495 75622 75663 75872 75904 75984
DDS8 (of COX2) 74576-74587
DDS9 76142-76153
maturase-like protein
NC_001224.1

Figure 5. The introns contained in the mitochondrial transcripts encode 10 additional polypeptides. Intronic ORFs have been found to encode maturases, endonucleases, and reverse transcriptases as shown in the scheme. The RNA maturases are required to process *COB* and *COX1* mRNA precursors and the reverse transcriptases can alter genomic DNA by transposition (*27, 50*).

B.4. Activation of translation

After mitochondrial transcripts are fully processed, translation takes place. Translation in the mitochondria is performed by a specialized mitochondrial ribosome (mitoribosome) that is distinct from the cytosolic ribosome. Almost all of the components of the mitochondrial translation system are encoded in the nuclear genome and imported into the mitochondria, with the exception of Var1p, 2 ribosomal rRNAs (21S and 15S), and a full set of mitochondrial tRNAs (Figure 4). Mitoribosomes are composed of a small and large subunit, with roughly 77 proteins and two rRNAs (*18*). In yeast, the rRNA components are largely buried within the mitoribosome, leading to the idea that the interactions that aid in translation may not be RNA-to-RNA, such as in prokaryotic ribosomes, but rather protein-to-protein (*51*). Mitoribosomes have been found tethered to the inner membrane of the mitochondria; most likely due to their role in co-translational insertion of highly hydrophobic OXPHOS proteins into the membrane (*52, 53*).

The mechanism by which mature mRNAs are presented to the mitoribosome is currently unknown. Recent studies have identified a vast network of proteins involved in replication, transcription, transcript processing,

and translation; possibly giving insight into how all of these processes work in conjunction. Many laboratories have been investigating the organization and links between these phases of gene expression and have hypothesized that there is a large higher-order complex involved (*54*). Interestingly, in 2015 Kehrein et al. classified the interactome of the yeast mitoribosome utilizing biochemical fractionation and superresolution microscopy. This study showed that a variety of proteins involved in DNA metabolism, organization of the nucleoid, mRNA maturation and processing, translational activation, translation, and OXPHOS assembly were all assembled in a higher-order complex. This complex was termed the mitochondrial organization of gene expression (MIOREX) complex. Moreover, it was shown that there are two fundamentally different assemblies of organized complexes, those found in close proximity to the nucleoid (nucleoid-MIOREX complex) and those found distributed along the mitochondrial cristae (peripheral MIOREX complexes) (*53*). This finding was significant as it provided a physical link between the components at each phase of mitochondrial gene expression. Although it is still unclear how exactly each process leads to the next, these complexes show that spatial organization is a key component to regulation.

Activation of translation requires a mitochondrial initiation factor 2 (IF2p). This factor binds to the initiator tRNA and positions it over the initiation codon on the ribosome in a GTP-dependent reaction. IF2p is evolutionarily conserved across bacteria, eukaryotic organelles, archaea, and the eukaryotic cytoplasmic translation systems (*55, 56*). In addition to this

initiation factor, members of a particular class of gene-specific nuclear genome-encoded proteins, termed translational activators, are also required in order for mitochondrial translation to occur. Translational activators were first identified in genetic analyses of nuclear genes essential for respiration. Further examination of respiratory deficient mutants pointed to a subset of genes that were needed specifically for translation of mitochondrial transcripts (*57*). Deletion of one of these genes resulted in the loss of translation of a particular mitochondrial gene product, resulting in respiratory deficiency (*58*). So far, at least one gene-specific translational activator has been classified for each of the 8 mt-mRNAs, however, this does not exclude the possibility that there may be more unidentified factors (*18*). Significantly, a *COX1*-specific translational activator TACO1 has been found to function in mammalian mitochondria (*59, 60*). Interestingly, but not surprisingly, many yeast translational activators were found to be a part of the MIOREX complex (*53*). Translational activators are believed to play different roles in translational activation as depicted in Figure 6 (*18*).

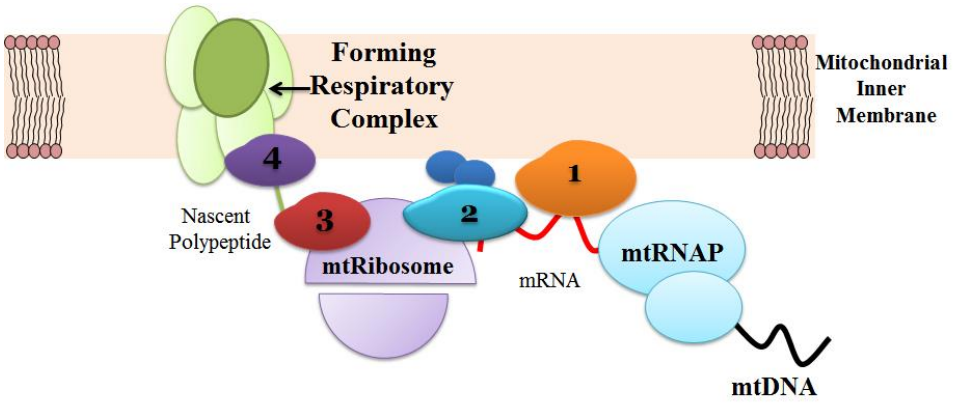

Figure 6. Translational activators act at several distinct stages of gene expression. Translational activators are thought to be tightly associated with the inner membrane. It is unclear whether they are tethered via other proteins or are directly inserted into the inner membrane themselves. A translational activator can stabilize nascent mRNA and protect it from degradation (1), provide a link between the mRNA and the ribosome (2), stabilize nascent polypeptides (3), and aid in complex assembly at the inner membrane and localize the ribosome to the proper forming respiratory complex (4). Each mRNA has a specific subset of translational activators. Some mRNAs only have one known activator that may stabilize the mRNA, whereas others have all four classes of activators that must work in concert to ensure proper protein expression (*18*).

Due to the challenges of low abundance *in vivo*, difficulty of expression and purification of these proteins for *in vitro* analysis (*61, 62*), and the lack of a mitochondrial *in vitro* translation system, little is known of how each translational activator works with its specific mRNA partner. Thus far, most if not all of the data about this family of proteins has been generated by *in vivo* genetic studies and biochemistry. To date 19 translational activators have

been identified (Table 2). A brief description of the function of the most studied representatives of this protein group is presented below.

Table 2. Yeast mitochondrial translational activators.

Mitochondrial transcript	Translational activator proteins
ATP8-ATP6	Aep3p
ATP6	Atp22p
ATP9	Aep1p, Aep2p, Atp25p
COB	Cbp1p, Cbt1p, Cbp3p, Cbp6p, Cbs1p, Cbs2p
COX1	Pet309p, Mss51p, Mam33p
COX3	Pet54p, Pet122p, Pet494p
VAR1	Sov1p
COX2	Pet111p

Translational activation of COB: The *COB* mRNA is believed to have at least six different nuclear genome-encoded proteins that aid in its translational activation. Their specific functional mechanisms are still largely unknown. Mutations in the nuclear genes *cbs1* and *cbs2* were found to cause a loss of the mature form of the *COB* transcript and inactivation of translation of the mRNA. Cbs1p and Cbs2p have also been shown to be associated with ribosomes and may provide a physical link between the mRNA and the translational machinery (*63-65*). Another nuclear genome-encoded factor, Cbp1p, has also been shown to be essential for the production of apocytochrome *b*. In the absence of Cbp1p, the *COB* mRNA is rapidly degraded. When degradation is prevented, there is still no translation of *COB*, demonstrating that Cbp1p plays a role in both stabilization and translational activation (*45*) presumably by interacting with a specific CCG element in the 5′-proximal region of the 5′-UTR (*66*). Similarly, nuclear genome-encoded

factor, Cbt1p, was found to promote pre-*COB* maturation, stabilize processed *COB,* and rescue a respiratory deficient phenotype in a strain carrying a mutation in the Cbp1p putative binding target, the CCG element (*67, 68*). This lead to the hypothesis that Cbt1p may act in conjunction with Cbp1p to promote processing and to ensure the stability of the mRNA, as well as facilitate the interaction of Cbp1p with its target in *COB* (*68, 69*). Lastly, Cbp3p and Cbp6p dimerize and act post-translationally by association with the nascent Cobp. The Cbp3p/Cbp6p dimer also interacts with the mitochondrial ribosome to promote *COB* translation *(70, 71).* Upon emergence of Cobp, the Cbp3p/Cbp6p dimer remains bound to Cobp and forms a complex with Cbp4p, an assembly factor that aids in the proper integration of the translation product into complex III (*72, 73*).

Translational activation of ATP6, ATP8, and ATP9: Although many proteins have been implicated in the translational activation of the ATPase subunits, the role that each activator plays is poorly understood. *ATP8* and *ATP6* are on the same primary transcript; however, different translational activators have been shown to be required for their expression. Nuclear DNA-encoded Atp22p has been implicated in the stability and expression of *ATP6*, presumably by recognizing a target in the 5′-UTR of *ATP8/ATP6 (74).* Similarly, a recent study has shown that nuclear encoded Aep3p acts as a translational activator specific for *ATP8* (*22*). Aep3p may also have a role in the processing of the *COX1/ATP8/ATP6* precursor mRNA (*75*). Lastly, *ATP9*

mRNA has three potential activators that were shown to be required for translation, Aep1p (*76*), Aep2p (*77*), and Atp25p (*78*). Mutations in Atp25p showed a lack of *ATP9* mRNA and of its translation product, thereby showing that Atp25p is required for *ATP9* transcript stability (*78*). Not only are these factors required for proper expression, there is also a feedback loop that regulates the levels of the ATPase subunits (see section B. Mitochondrial gene expression for more details).

Translational activation of COX1, COX2, and COX3: Although translational activation of the three mitochondrial encoded components of Complex IV is not completely characterized, there is a decent amount of literature describing the mechanistic details of this process. The expression of *COX1* relies on three known translational activators, Pet309p, Mss51p, and Mam33p. In a *pet309* null mutant there is a defect in *COX1* pre-RNA and also in the translation of *COX1* (*79*). Pet309p is believed to target the 5′-UTR of *COX1* and is needed for translation, but it is unclear if it also plays a role in stability. Interaction of Pet309p with the *COX1* mRNA was shown *in vivo* via an RNA coimmunoprecipitation; however, from this experiment it was unclear if this interaction is direct *(80).* Mss51p has a more extensive role as a translational regulator as it is a part of a feedback loop that regulates Cox1p synthesis. Cox1p contains heme and copper prosthetic groups, which contribute to the active center of Complex IV. Therefore, full complex assembly relies heavily on the status of Cox1p production. In the feedback loop, Mss51p binds to

nascent Cox1p forming a transient complex that associates with other assembly factors. This interaction limits the availability of Mss51p to act in *COX1* translation activation via interaction with the 5′-UTR (*18, 81*). Upon release of Mss51p, a new round of Cox1p translation can occur. Moreover, Mss51p has recently been implicated as a heme binding protein and has been demonstrated to contain heme regulatory motifs. The activity of Mss51p is linked to the redox environment in mitochondria, which ultimately regulates the levels of Cox1p synthesis (*82*). Recently, a third protein factor, Mam33p, has been shown to be required for Cox1p synthesis due to a role in translational activation of *COX1*. This factor is needed for cells to adapt efficiently from fermentation to respiratory conditions (*83*). *COX2* translational depends on only one known translational activator Pet111p that is believed to specifically interact with the 5′-UTR of *COX2* (*15*), whereas translational of *COX3* depends on three translational activators Pet54p, Pet122p, and Pet494p. The activators of the COX family of proteins have all been shown to associate with each other, along with the mitochondrial RNA polymerase (*84*) and the ribosome (*53*).

Translational activation of VAR1: Out of all of the mt-mRNAs the least is known about the translational activation of *VAR1*, the only non-OXPHOS complex related transcript. Sanchirico has reported that Sov1p is a translational activator of the Var1p (*85*). However, it is currently unknown how

this role is performed, if this factor provides stability along with translational activation, and if additional proteins may play a role (*18*).

The accumulating data on the known translational activators implies that they can work in many ways and have their own specific functions. These functions may be independent or rely on a subset of proteins. The process of activating translation is multifaceted and nuclear genome-encoded proteins are at the forefront of all regulatory processes. Although there are only eight proteins encoded in the mitochondrial genome, the regulation of gene expression in this system is very intricate and still not well understood. Further experimental work with this family of proteins is needed to enable a better understanding of the different mechanistic aspects of translational activation.

B.5. mRNA regulated translational repression

In 1990, Papadopoulou et al. identified a 40 kDa protein (p40) that specifically bound to the 5′-UTR of *COX2* in *S. cerevisiae*. Using a diethylpyrocarbonate (DEPC) modification assay, this study identified a 10 nucleotide segment that p40 was able to protect from modification, i.e. the location in which it interacts. The binding site was located 12 nucleotides upstream from the *COX2* start codon, and deletion of these nucleotides completely abolished the interaction with p40. Upon altering a larger region in the 5′-UTR, which ultimately altered the secondary structure but not the

binding site, the p40 binding was lost (*86*). This indicated that an element of the native secondary structure may be essential for this interaction. The protein was later shown to interact with all eight mitochondrial mRNAs in *S. cerevisiae* and a consensus p40 binding site was identified as a relatively short RNA duplex found in *COX1*, *COX2,* and *ATP9* 5′-leaders (*87*). It appears that a portion of single-stranded RNA must adjoin the consensus duplex for efficient binding of p40, although the sequence of the single-stranded element is not important. However, DEPC footprinting showed that this single-stranded element was protected by p40 and therefore is a part of the binding site (*86, 87*).

P40 was identified as the Krebs cycle enzyme, NAD^+-dependent isocitrate dehydrogenase (IDH) (*88*). IDH is composed of two non-identical proteins encoded by the nuclear genes *IDH1* and *IDH2*. Together they create an octamer, composed of four heterodimers, which is an allosterically regulated enzyme (*89*). Since IDH is essential to the mitochondrial metabolism, and a deletion of this gene led to a respiratory deficient strain, it was very challenging to study the role of the IDH/mt-mRNA interactions *in vivo*. However, using IDH disruption mutants in which both *idh1* and *idh2* were disrupted, an increase in the overall translation of mt-mRNAs was observed. Based on this finding, it was hypothesized that IDH may suppress aberrant translation of mitochondrial mRNAs in the bulk of the matrix unless interaction with the appropriate translational activator brings the message in proximity to the membrane (*90*).

C. Many yeast translational activators are putative pentatricopeptide repeat proteins

A specific subset of *S. cerevisiae* mitochondrial translational activators is presumed to perform their roles by interacting with RNA directly; however, the mechanism of this interaction is not characterized (*18*). Several of these factors, along with a few other proteins involved in RNA metabolism, have been classified as putative pentatricopeptide repeat (PPR) proteins (*91*). PPR proteins are a large family of RNA binding proteins that were first discovered upon sequencing and bioinformatic analysis of the *Arabidopsis thaliana* genome (*92*). Although PPR proteins are found in many eukaryotic species, they are the most abundant in land plants (*93*). They are believed to have diverged from prokaryotic tetratricopeptide repeat (TPR) proteins. Most PPR proteins localize to either the mitochondria or chloroplasts and their main function has been characterized as modulating gene expression at the RNA level (*62*). All current data suggests that they are single-stranded RNA binding proteins functioning in RNA metabolism, including transcription, RNA processing, splicing, stability, editing, and translation (*93*). To perform these many roles, these proteins may act non-catalytically, by mediating interactions, or they may utilize catalytic domains to facilitate specific reactions, such as RNA editing (*62*). Most of these factors are essential to the organelles in which they belong and their deletion can result in a loss in organellar function (*92*).

Each PPR protein contains PPR motifs which are present as tandem arrays of 35 amino acid sequence repeats (the “P Class”), with anywhere from 2 to 30 repeats forming a PPR domain. PPR domains have also been found to have longer or shorter than 35 amino acid motifs (the “PLS Class”). The motifs can be directly adjacent to one another or there can be gaps between them, depending on the function of the protein. Proteins may either contain only the PPR domain, or another functional module may be added, such as an endonuclease or editing domain (*62*).

The structure of a PPR motif features two alpha helices that are antiparallel to each other and interact to form a helix-turn-helix element (*94*). Each motif has certain residues that enable the protein to specifically recognize and interact with target RNA sequences. In plant PPR proteins, each motif has been shown to align with one nucleotide in the RNA target. Emerging studies and the use of crystallography data have identified individual amino acids in each motif that recognize specific ribonucleotide bases (Figure 7) (*95*).

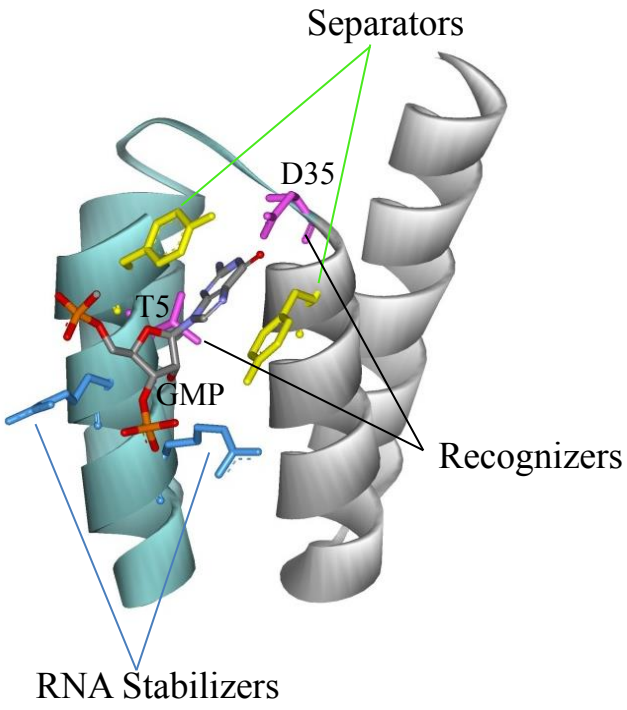

Figure 7. Modular recognition of a ribonucleotide by a PPR protein. A partial crystal structure of THA8 from *B. distachyon* (PDB code 4N2Q) showing modular recognition of GMP by one PPR motif (*96*). The structure shows positively charged amino acids that help stabilize the electronegative backbone of the RNA chain (RNA stabilizers; blue). There are two separating amino acids, usually aromatic or hydrophobic, one on each side of the base that help spatially organize the consecutive bases (Separators; yellow). Lastly, there are two amino acids, in positions 5 and 35, which specifically recognize the ribonucleotide base (Recognizers; purple). Depending on the identity of the amino acids in positions 5 and 35 of the motif, the base being recognized can vary (*97*).

The set of principles employed for this modular recognition of ribonucleotide bases has been termed the “PPR Code” (*98*). It remains unclear whether this code is universal among PPR proteins across species.

Upon analysis of the available PPR protein crystal structures, complexed with their target RNAs, some deviations from the proposed PPR code have been noticed. This result shows that this code may not be followed for the recognition of all nucleotides in the target (*99*). Many studies have shown that altering the amino acid side chains in the PPR motifs changes the specificity of ribonucleotide binding. Therefore, this specificity has the potential to become predictable and programmable (*100*). Recent studies have sought to create artificial PPR proteins with the ability to function in different biotechnological applications. For instance, by linking a designed PPR domain with another domain, such as an RNA base modification domain or an endonuclease domain, various site-specific RNA editing tools can be created (*101, 102*).

Although the bulk of the PPR protein research has been performed in land plants, computational analyses have identified analogous sequences in the genomes of other species. In *S. cerevisiae,* 3 proteins were found that matched the plant PPR motif consensus sequence, Pet309p, Aep3p, and Dmr1p. Considering that the PPR motif could have diverged differently in different kingdoms of life, previously identified PPR sequences from various yeast species were analyzed by a Hidden Markov model algorithm to define a yeast specific PPR consensus (*71*). The analysis returned a sequence pattern that was then searched for in the fungal genome and refined iteratively. This led to the identification of 12 additional PPR proteins in *S. cerevisiae* (Table 3) *(91, 103).* Despite their sequence resemblance, no yeast protein has yet

been shown to mechanistically recognize RNA bases in a modular manner as observed with plant PPR proteins *(91).*

Table 3. Putative PPR proteins in *S. cerevisiae.*

***S. Cerevisiae* Putative PPR Protein**	**RNA Metabolism Step**
Rpo41p	mtRNAP
Cbp1p	Stabilization/translation of *COB*
Pet309p	Stabilization/translation of *COX1*
Rpm2p	tRNA processing
Rmd9p	All mRNA stabilization
Dmr1p	Stabilization of 15s rRNA
Sov1p	Stabilization of *VAR1*
Aep3p	Translation of *ATP8*
Aep2p	Stabilization of *ATP9*
Pet111p	Translation of *COX2*
Atp22p	Translation of *ATP8/6*
Aep1p	Translation of *ATP9*
Rmd9Lp	Unknown
Msc6p	Unknown
Yer077c	Unknown

Understanding this family of proteins is very significant to the research set forth because the related yeast mitochondrial translational activators may theoretically exhibit common functional and mechanistic patterns. Thus far, however, all yeast PPR proteins must be considered “putative”. There is no experimental evidence to indicate that they parallel the known plant PPR proteins in terms of structural organization, ability to interact with RNA directly, or perform modular recognition of RNA targets in a single-stranded context and in accordance to a PPR code.

RATIONALE

Mitochondria are essential to cellular respiration and overall cellular homeostasis. Aberrant mitochondrial gene expression has been linked to various human diseases. Therefore, gene expression is tightly regulated to ensure that essential respiratory components are produced at the appropriate levels at the proper time. The regulatory system that controls these processes is multifaceted and its function is poorly understood at the mechanistic level. It has been shown that most regulation occurs post-transcriptionally and that there are multiple nuclear genome-encoded factors essential to the function of this regulatory network. Yeast genetic studies have identified a family of nuclear gene products, termed “mitochondrial translation activators,” that play a key role in the control of mitochondrial gene expression. A small subset of these factors have been extensively investigated, both genetically and biochemically; however, a clear picture of how they mechanistically function is still missing.

Using the model organism *S. cerevisiae* this study aims to explore two representatives of the translational activator family, Pet111p and Cbp1p, with the goal to elucidate the mechanisms involved in nuclear control of mitochondrial gene expression. By providing additional characterization of this regulatory system, this study may advance our understanding of mitochondrial translational activation across different species and potentially aid in the understanding of mitochondria-related pathologies.

CHAPTER I

INSIGHT INTO TRANSLATIONAL ACTIVATION OF *COX2* BY PET111P

I.A. Abstract

Previous genetic studies have shown that nuclear genome-encoded Pet111p acts as a translational activator for mitochondrial genome-encoded *COX2.* These studies showed that when Pet111p is not present, Cox2p is not synthesized and therefore there is no cellular respiration. However, how this activation of translation occurs is unknown. It has been inferred that Pet111p may directly contact the 5′-untranslated region of *COX2*, aiding in translational activation by a possible interaction with the ribosome. Although Pet111p has been implicated in recognizing a sequence in the 5′-untranslated region of *COX2*, the identity of this site has never been demonstrated, nor has the ability of the protein to interact with RNA. I sought to determine if Pet111p can interact with a site in *COX2* and how this interaction could lead to the activation of translation. In this current work, I confirmed that Pet111p is indeed an RNA binding protein that binds sequence-specifically to two distinct binding targets, one of which is located in the 5′-untranslated leader and the other is found in the coding sequence in the proximity of the initiation region. These interactions were also shown to be required for the stability of *COX2*. My data, combined with the available genetic data, allowed me to propose plausible mechanisms of translational activation by Pet111p. This study also explored the interactome of Pet111p in an attempt to identify other essential factors that are a part of the transcription-translation coupled mechanism of the *COX2* gene expression.

I.B. Introduction

Complex IV, *cytochrome c oxidase* (COX), is a membrane-embedded, multisubunit protein complex that catalyzes the last reaction of the electron transport chain. In this reaction, the oxidation of cytochrome *c* by molecular oxygen is coupled with the translocation of protons from the mitochondrial matrix to the intermembrane space. This ultimately contributes to the electrochemical gradient that drives the synthesis of ATP. The yeast COX complex is an assembly of 11 (*104*) or 12 (*105*) subunits, of which only three, Cox1p, Cox2p, and Cox3p, are coded for by the mitochondrial DNA, while the others are the products of nuclear genes. Therefore, a regulatory system must be in place to ensure coordinated expression from both genomes, such that all of the subunits are produced synchronously and stoichiometrically for the correct assembly of the complex (*17*).

Cox2p is essential for respiration, as it is a part of the catalytic core of COX and contains a metal binding site that anchors two essential copper ions (*17*). In the early 1970's a phenomenon was discovered that a mutation in a single nuclear gene could lead to a massive disruption of mitochondrial function and a loss of respiration (*106*). This phenomenon was observed when nuclear genome-encoded *pet111* was mutated and the expression of mitochondrial genome-encoded Cox2p was lost (*107, 108*). Poutre et al. showed that when Pet111p is not expressed, the *COX2* mRNA is decreased substantially; however, Cox2p expression is completely abolished despite the presence of the *COX2* transcript *(107)*. This study thereby characterized

Pet111p as a translational activator for *COX2 (107)* and therefore a component of the regulatory system that ensures coordinated expression of the bigenomic complex IV.

To further probe the mechanism of Pet111p dependent translational activation of *COX2*, various genetic studies were performed. At the time, the idea that the 5′-UTR was significant for translation was becoming apparent (*109, 110*). Therefore, one notion was that Pet111p may act through the 5′-UTR of *COX2.* This theory was proven when chimeric mRNAs were generated and it was shown that Pet111p could control the translation of a chimeric message, in which the *COX2* 5′-UTR preceded the *COX3* reading frame. Consistently, translation of the reciprocal RNA construct, where the *COX2* reading frame was preceded by the *COX3* 5′-leader, did not depend on Pet111p and was instead controlled by a *COX3*-specific translational activator, Pet494p (*111*). The ability of Pet111p to activate translation with only the *COX2* 5′-UTR present further supported that Pet111p directs the translation of *COX2* by acting within the 5′-UTR of the mRNA.

To further characterize the genetic interaction between Pet111p and the 5′-UTR of *COX2*, a mutational analysis of the 5′-UTR was performed. Through deletion analysis, it was discovered that the steady state level of the mRNA was generally not affected when certain regions were mutated, whereas translation was greatly diminished. Remarkably, even single point mutations appeared to affect the translation of *COX2* (*112*). When Pet111p was overexpressed it was able to suppress some of the phenotypes

associated with the mutations. These studies led to the idea that there was a close functional interaction between the 54-nt 5′-UTR of *COX2* and Pet111p. Additional genetic manipulations found that this functional interaction is confined to a 31-nucleotide segment essential for respiration and believed to harbor a direct Pet111p point of contact (*113*).

Unlike other mitochondrial gene products, Cox2p is first synthesized as a precursor protein (pre-Cox2p) carrying a 15-amino-acid leader sequence. Upon deletion of the sequence coding for the leader peptide, the mature Cox2p failed to accumulate. Significantly, altering the amino acid sequence of this leader, by shifting the reading frame, had no effect on respiration as long as the nucleotide sequence was unchanged. It was also shown that overexpression of Pet111p can partially rescue the phenotypes induced by mutations in this region (*114*); however, how Pet111p affects this is currently unknown. Therefore, this study supported the notion that Pet111p may have a direct functional relationship with this region of the *COX2* ORF, which has yet to be experimentally confirmed.

The function of Pet111p as a *COX2*-specific translational activator appears to be conserved in related yeasts, despite relatively high amino acid sequence divergence, as the production of Cox2p was also blocked in the absence of Pet111p orthologs in both *S. kluyveri* and *K. lactis* (*115*). When present at physiological levels, Pet111p is a limiting factor in the translation of *COX2.* This was evident from a two-fold decrease, relative to WT, in the efficiency of the expression of a reporter gene placed under the control of the

COX2 5′-UTR in heterozygous Δ*PET111* diploid cells. Conversely, a significant increase in the expression of Cox2p was observed when Pet111p was overexpressed (*116*). These important observations confirmed that Pet111p has the capacity to attenuate the production of Cox2p. As is the case with most yeast mitochondrial gene-specific translational activators (*18*), Pet111p is anchored to the matrix surface of the inner mitochondrial membrane, thus enabling co-translational insertion of the nascent Cox2p into the membrane (*116*). *COX2* translation at the membrane surface was shown to be of critical importance as demonstrated by an experiment in which translation of the *COX2* reading frame was placed under the control of the *VAR1* 5′-UTR. Var1p, a hydrophilic protein, is known to be synthesized in the matrix. In this experiment, although Cox2p was synthesized, it was immediately degraded due to the fact that it was not localized near essential components at the inner membrane (*117*).

Remarkably, Pet111p appears to be a part of a membrane-associated complex, which also includes several other translational activators specific to *COX1*, *COX3*, and *COB* (*54, 84*). The existence of such a network suggests an additional function for the translational activators, which is to ensure proper spatial coordination during the production and assembly of corresponding respiratory complexes (*84*). Moreover, Mtf2p, a protein associated with the mitochondrial RNA polymerase N-terminal domain (*118*), was also reported to interact with Pet111p (*84*); bridging the network of translational activators to the transcription machinery. All of these interactions

presumably take place in the framework of the recently discovered MIOREX super complex, which is responsible for the organization of mitochondrial gene expression (*53*).

This current study sets forth to determine the mechanistic role of Pet111p in *COX2* translation. Previous studies have shown a close functional relationship between Pet111p and the 5′-UTR of *COX2.* However, a direct binding of Pet111p to *COX2*, or any other RNA, has not been demonstrated. Also, it is still unclear how this functional interaction leads to the activation of *COX2* translation and if other mitochondrial proteins are involved. With the 5′-UTR being only 54 nt in length and the functional relationship between Pet111p and *COX2* extensively characterized at the genetic level, this system is ideally suited for studying the function of a translational activator *in vitro*. In my study, recombinant Pet111p and fragments of *COX2* were examined in an *in vitro* biochemical system to determine: (*i)* if Pet111p can directly bind to the *COX2* mRNA, (*ii)* if so, what is the region of the interaction(s), and *(iii)* if any other mitochondrial factors are directly involved in this interaction. This information will help to determine Pet111p's role in translational activation of *COX2*, and perhaps how this type of activation works in general.

I.C. Materials and Methods

I.C.1. *Oligonucleotides*. Synthetic DNA oligonucleotides were purchased from Integrated DNA Technology (IDT) and synthetic RNA was purchased from GE Healthcare Dharmacon and IDT. Sequences of all oligonucleotides used in the study are listed in Table 4. The synthetic RNA probes used in the EMSA and RNase footprinting experiments were 5′-labeled with T4 PNK (New England Biolabs) in the presence of γ-[^{32}P]-ATP (3000 Ci/mmol, Perkin Elmer). After labeling, the reactions were run on 7 M urea 16% polyacrylamide gels (19:1, acrylamide:bisacrylamide). The bands containing the target RNA were located on the gels by autoradiography, excised, crushed into small pieces, and extracted by shaking in 0.3 ml of 20 mM Tris·HCl pH 7.9 and 1 M NaCl for 1 h at room temperature. The RNA was recovered from the extracts using RNA Clean & Concentrator-5 columns (Zymo Research) as directed by the manufacturer's instruction manual, except that the concentration of ethanol was increased to 75% during the binding to the columns. The purified RNA was quantitated by UV/Vis spectrophotometry, incubated at 65 °C for 3 min, and then snap-cooled in ice.

I.C.2. *Primer extension analysis.* For the determination of the 5′-end of mature *COX2,* BY4743 yeast cells were grown in YP medium in the presence of 3% glycerol and 5 μg/ml tetracycline. Mitochondria were isolated by following a previously published protocol (*119*) with slight modifications. The cells were collected by centrifugation, washed with water, resuspended in a reduction

buffer (100 mM Tris·SO_4 pH 9.4, 72 mM β-mercaptoethanol) at a density of 3 g of wet cells per 10 ml buffer, and incubated for 45 min at 30 °C. The buffer was replaced with 20 mM K-phosphate pH 7.4 and 1.2 M sorbitol. Zymolyase 100T (Amsbio) was added at a ratio of 1.5 mg/g wet cells, and the suspension was incubated for 1 h at 30 °C to allow the formation of spheroplasts. Spheroplast formation was checked by resuspending 20 µl of the suspension into 1 ml H_2O and measuring the OD_{600}. When the OD_{600} reached a constant level, signifying total spheroplast formation, mitochondria were extracted by breaking the spheroplasts by osmotic shock in SHP buffer (0.6 M sorbitol, 20 mM HEPES·KOH, pH 7.4, 1 mM PMSF) followed by homogenization and differential centrifugation as described (*119*). The crude mitochondrial pellet was resuspended in mitochondrial resuspension buffer (250 mM sucrose, 10 mM HEPES·NaOH pH 7.4, 1 mM EDTA) and fractionated on a discontinuous sucrose density gradient, 60%, 32%, 23%, and 15% sucrose in 10 mM HEPES·NaOH pH 7.4, 1 mM EDTA for 1 h (134,000 × g, 4 °C) using a SW 41 Ti rotor (Beckman) as described (*120*). Purified mitochondria were collected from the 32% sucrose layer, resuspended in 5 times the volume of mitochondrial resuspension buffer, and distributed into 1.5 ml tubes. The mitochondria were pelleted by centrifugation (20 min, 12,000 × g, 4 °C), flash-frozen in liquid nitrogen, and stored at -80 °C.

Total mitochondrial RNA was isolated from approximately 100 µl of the settled mitochondria using TRI REAGENT (Molecular Research Center). Mitochondria were resuspended in 0.9 ml TRI REAGENT, homogenized with

10 strokes in a small Dounce homogenizer, and incubated at room temperature for 5 min. The homogenate was vortexed for 15 seconds with 0.2 ml of chloroform (Fisher) and the mixture was incubated for 15 min at room temperature. The phases of the suspension were then subjected by centrifugation (15 min, 12,000 × g, 4 °C) and the aqueous phase was transferred to a 1.5 ml tube. RNA was precipitated by isopropanol (Fisher), the pellet was resuspended in water, and subjected to DNase I (NEB) treatment. RNA was recovered from the solution using RNA Clean & Concentrator-5 columns (Zymo Research) and quantified by UV/Vis spectroscopy. A 5′-[^{32}P]-labeled DNA primer *COX2*-96-76 (0.1 μM) was added to 3 μg of the purified RNA in 10 μl of water and hybridized by heating to 70 °C for 5 minutes followed by snap cooling on ice for 10 min. To extend the primer, 10 μl of the hybridization solution was incubated with 10 μl of a reaction mixture containing 80 U of ProtoScript® II reverse transcriptase (New England Biolabs), 1× ProtoScript® II reverse transcriptase buffer (NEB), 8 U Murine RNase inhibitor (NEB), 10 mM DTT, 0.5 mM 2′-dNTPs, and incubated for 30 min at 42 °C. To generate DNA size markers, a DNA fragment spanning the positions from -7 to +109 relative to the *COX2* promoter start site (+1) was PCR-amplified by Phusion DNA polymerase (NEB) using plasmid pJJ10 (see Appendix section A.1. for generation of pJJ10) as a template and primers *COX2*R and *COX2*-(-61)-(-39). Due to the high A/T content of the template, the temperature during the annealing and extension steps was lowered to 50 °C and 60 °C, respectively, and the elongation time

was extended to 2 min. The fragment was then purified by GeneJET PCR purification kit (Thermo Scientific) and used as a template (0.2 μM) to extend the 5′-[^{32}P]-*COX2*-96-76 primer (0.15 μM) with 1.5 U of *Taq* DNA polymerase (NEB) in the presence of 2′,3′,-ddNTPs. The extension reactions were carried out in a thermocycler in six cycles with the temperature of melting and annealing/elongation set at 94 °C (1 min) and 58 °C (5 min), respectively, in 50 μl of Standard *Taq* reaction buffer (NEB) in the presence of dNTPs (4.8 μM each). The concentrations of ddNTPs present in the reactions were 400 μM (ddATP), 300 μM (ddCTP), and 50 μM (ddGTP). The products of extension of the primer were resolved in a 6% (19:1, acrylamide:bisacrylamide) gradient-thickness gel in the presence of 7 M urea and visualized by phosphor imaging with a Typhoon 9410 scanner (GE Healthcare).

For *COX2* primer extension analysis in wild type and Δ*PET111* W303 cells, the above experimental procedure was used with the following modifications. Wild type and Δ*PET111* W303 yeast cells were grown in 50 ml of YP medium in the presence of 2% galactose and 5 μg/ml tetracycline to an OD_{600} of 0.5. Cells were harvested by centrifugation (5,000 × g, 5 min), washed with 1 ml of water, and resuspended in 300 μl of TRI REAGENT with the addition of 300 μl of 0.5 mm zirconia silica beads (Biospec). The cells were lysed by vortexing this solution for 10 min at 4 °C. The volume of the

mixture was adjusted to 1 ml with TRI REAGENT and the lysate was clarified by centrifugation (12,000 × g, 10 min, 4 °C).

I.C.3. *Identification of the N-terminus of mature Pet111p and Mtf2p.* To identify the N-terminus of mature Pet111p, Y258 yeast cells containing a BG1805 plasmid encoding tagged Pet111p (see Figure 9 for tag composition) were purchased from GE Healthcare Open Biosystems. After an initial attempt at Pet111p isolation failed the plasmid was isolated from the cells and the DNA was sequenced (GENEWIZ). Sequencing revealed a 1634A>T substitution in the coding sequence leading to a K545M mutation in Pet111p. The mutation was corrected by site-directed mutagenesis using the QuikChange II kit (Agilent) and mutagenesis primers BG1805Pet111_fix and BG1805Pet111_fixc with the standard protocol. The corrected plasmid was transformed into BY4743 cells by the lithium method adapted from Ito et al. (*121*). Lyophilized salmon sperm DNA (Boehringer Mannheim, Germany) was resuspended to 10 mg/ml in water and shook overnight at room temperature. The mixture was boiled for 10 min to denature the dsDNA, the resulting ssDNA was flash-chilled in an ice-water mixture, and frozen at -20 °C in aliquots to prevent renaturation. 100 μl of frozen BY4347 cells were thawed, rinsed with 200 μl of TE buffer, pelleted at 2,500 × g, and resuspended in 50 μl of TE buffer. To the suspension, 2 μg of plasmid DNA, 5 μg of ssDNA, 270 μl 50% PEG 3350, and 30 μl of 1 M LiAcO pH 7.5 was added, the suspension was mixed well by gentle pipetting, and incubated at 30 °C for 30 min. After

incubation, 40 μl DMSO was added and the cells were incubated at 42 °C for 15 min. The cells were next spun down at 2,500 × g for 2 min, washed with 1 ml of TE, pelleted, resuspended in 200 μl of TE, and plated on agar plates containing SD-ura medium (yeast nitrogen base with ammonium sulfate and synthetic drop-out mixture -uracil, US Biological), 2% galactose, and 5 μg/ml tetracycline.

The cells carrying the corrected plasmid were grown from a single colony at 30 °C to an OD_{600} of 9 in 3 L of SD-ura medium, 2% galactose, and 5 μg/ml tetracycline. Mitochondria were isolated as described in section I.C.2. Approximately 0.3 ml of settled mitochondria were resuspended in 1.2 ml of lysis solution (7 M guanidine·HCl, 100 mM NaCl, 15 mM imidazole, 3 mM β-mercaptoethanol) and the suspension was sonicated on ice with five 30 sec pulses intermitted with 30 sec pauses on a F60 sonic dismembrator (Fisher Scientific). Insoluble matter was removed by centrifugation (20,000 × g, 15 min) and the supernatant was incubated with 20 μl of Ni-IDA agarose beads (Gold Biotechnology) in a tumbling tube overnight at room temperature. The beads were washed three times with the lysis solution and two times with water using 1.4 ml per wash. The bound proteins were eluted with 80 μl of 1 × LDS gel loading buffer (Novex, Life technologies), a 20 μl sample of the eluate was separated by 4-12% LDS-PAGE, and the protein bands were visualized by Coomassie staining.

The proteins in the bands were identified by peptide mass fingerprinting (PMF). The protein bands were excised from the gel and successively subjected to an in-gel reduction with 2 mM TCEP (20 min, 65 °C), an alkylation with 55 mM iodoacetamide (20 min, room temperature), and a digestion with mass-spectrometry grade trypsin (Promega) (37 °C, overnight) in 100 mM ammonium bicarbonate as previously described (*122*). The products of the digestion were extracted from the gel with 0.1% TFA and 80% acetonitrile and analyzed by MALDI-TOF mass spectrometry on a Bruker Microflex LRF instrument operated in a positive-ion field reflectron mode and using HCCA (α-Cyano-4-hydroxycinnamic acid; 10 mg/ml in 50% acetonitrile and 0.1% TFA) as the matrix. To identify the proteins, the collected spectral data were submitted to Mascot searches. A protein band with an apparent mobility of 97 kDa was identified as containing Pet111p with a Mascot score of 188. The proteins eluted from the Ni-beads were separated by PAGE once again, transferred onto an Immobilon-P^{SQ} PVDF membrane (EMD Millipore) by electroblotting, and the 97 kDa protein band was cut out from the membrane and submitted for N-terminal protein sequencing (Midwest Analytical).

Determination of the N-terminus of mature Mtf2p was performed in a similar manner with a BG1805-*MTF2* plasmid from the same ORF library (GE Healthcare Open Biosystems).

I.C.4. *Cloning, expression, and purification of Pet111p.* To purify recombinant Pet111p, a DNA fragment encoding His_6-tagged full-length Pet111p was PCR-amplified from *S. cerevisiae* BY4743 genomic DNA using primers Pet111_Nco and Pet111_Xho. The amplicon was ligated into the pT7Blue vector (Novagen) to form plasmid pMA40, from which the gene was excised and ligated into pTrcHisA (Life Technologies) at the *Nco*I and *Xho*I restriction sites. In the resulting plasmid, pMA41, the sequence coding for the 34 N-terminal amino acids of Pet111p was deleted by site-directed mutagenesis using QuikChange II kit (Agilent Technologies) and mutagenesis primers Pet111_del34 and Pet111_del34c. The deletion resulted in plasmid pGD1 expressing Δ34-Pet111p fused to an N-terminal purification tag MAHHHHHH, which in this report is referenced to as Pet111p. The integrity of the target gene in pGD1 was confirmed by DNA sequencing (GENEWIZ).

E. coli XJb(DE3) cells (Zymo Research) carrying pGD1 were grown at 37 °C in 6 L of LB medium supplemented with 3 mM L-arabinose, 1 mM $MgCl_2$, and 100 µg/ml ampicillin to an OD_{600} of 0.8-0.9. The culture was then cooled to 12 °C. Expression of Pet111p was induced with 0.2 mM IPTG and the cells were incubated at 12 °C for an additional 20 hrs. Post-induction, cells were collected by centrifugation and resuspended in a lysis buffer containing 20 mM HEPES·NaOH pH 7.0, 300 mM NaCl, 5 mM β-mercaptoethanol, and 1 mM PMSF at a density of 1 g wet cell pellet per 10 ml buffer. To induce cell lysis, the suspension was frozen overnight at -80 °C and then allowed to thaw in the presence of fresh PMSF added to a concentration

of 1 mM. The lysate was supplemented with 10 mM $MgCl_2$, bovine pancreatic DNase I (Roche) was added to a concentration of 2 µg/ml, and the mixture was stirred for 15 min at room temperature. The lysate was then cleared by centrifugation for 30 min (36,000 × g, 4 °C). The concentration of NaCl in the supernatant was adjusted to 1.5 M and 15 mM imidazole was added. The lysate was then passed through a column packed with 2 ml of Ni-IDA agarose beads (Gold Biotechnology) equilibrated with wash buffer (20 mM HEPES·NaOH pH 7.0, 1.5 M NaCl, 15 mM imidazole). The beads were washed by passing 30 ml of wash buffer through the column and the protein was released with 5 ml of elution buffer (20 mM HEPES·NaOH pH 7.0, 900 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, 225 mM imidazole, 5 mM β-mercaptoethanol). The eluate was diluted two-fold with a buffer containing 20 mM HEPES·NaOH pH 7.0, 10% glycerol, and 5 mM β-mercaptoethanol to reduce the concentration of NaCl to 450 mM and loaded on an 8 × 7.5 TSKgel Heparin-5PW (Tosoh Bioscience) column. The protein was eluted from the column with a linear NaCl concentration gradient (450 mM to 1 M) in a buffer containing 20 mM HEPES·NaOH pH 7.0, 10 mM $MgCl_2$, 10% glycerol, and 5 mM β-mercaptoethanol. Fractions containing Pet111p were combined, concentrated on an Ultracel®-50K centrifugal filter (Merck Millipore) to a volume of 2 ml, loaded on a HiLoad 16/600 Superdex 200 PG column, and the protein was eluted with gel filtration buffer (20 mM HEPES·NaOH pH 7.0, 600 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, 5 mM β-mercaptoethanol). Fractions containing pure Pet111p were pooled and concentrated on a

centrifugal filter to increase the concentration of the protein to approximately 5 mg/ml. The concentrate was diluted two-fold with glycerol, distributed into small aliquots, and stored frozen at -80 °C. A typical yield of purified Pet111p was 5 mg.

I.C.5. *Cloning, expression, and purification of Mtf2p.* A DNA fragment encoding the full-length *MTF2* gene was amplified from *S. cerevisiae* BY4743 genomic DNA using primers MA58 and MA59. This PCR product was inserted into the cloning vector pTrcHis-A at *Bam*HI and *Sac*I sites to produce plasmid pMA1. Due to the presence of an internal *Sac*I restriction site in *MTF2*, a limited digestion of the amplicon was performed. The larger product of the digestion, containing the full-length *MTF2* coding sequence, was purified in a 0.8% agarose gel before using it in the ligation reaction. The plasmid pMA1 encoded a protein form, in which Met1 of the Mtf2p precursor was fused to an N-terminal sequence, MGGSHHHHHHGMASMTGGQQMGRTLYDDDD**K**DRWGS, originating from the vector. This leader sequence contained a hexahistidine purification tag followed by an enterokinase (EK) recognition site (underlined), such that most of the leader in the purified protein could be removed by EK cleavage after the K residue (bold). The sequence in pMA1 coding for the amino acids following the hexahistidine tag of the leader and the 24 N-terminal amino acids of Mtf2p was deleted by site-directed mutagenesis using QuikChange II kit (Agilent Technologies) and mutagenesis primers EE1 and EE2. The

deletion resulted in plasmid pEE1 expressing a histidine-tagged form of Δ24Mtf2p. Although the N-terminus of mature Mtf2p was determined to be at amino acid 24, an additional amino acid, L24, was removed in the mutagenesis reaction to reduce hydrophobicity close to the purification tag. The integrity of the target gene in pEE1 was confirmed by DNA sequencing (GENEWIZ).

pEE1 was transformed into *E. coli* XJb(DE3) cells (Zymo Research) and grown as described above for expression of Pet111p (section I.C.4) with the following changes. The induction of Mtf2p was performed at 16 °C and the cells were resuspended in 40 mM Tris·HCl pH 7.9, 300 mM NaCl, 15 mM imidazole, 1mM PMSF and 5 mM β-mercaptoethanol. Mtf2p was purified from the cell lysate using nickel affinity chromatography, heparin affinity chromatography and gel filtration as in section I.C.4, with the following modifications. After binding to the protein, nickel beads were washed by passing 10 ml of high salt wash buffer (50 mM Tris·HCl pH 7.9, 1.5 M NaCl, 15 mM imidazole) through the column, followed by a second wash with low salt wash buffer (40 mM Tris·HCl pH 7.9, 300 mM NaCl, 15 mM imidazole). Bound protein was released with 3 ml of elution buffer (40 mM Tris·HCl pH 7.9, 250 mM NaCl, 5% glycerol, 250 mM imidazole, 5 mM β-mercaptoethanol). The eluate was loaded onto a 16 × 25 mm HiTrap Heparin HP column (GE Healthcare Life Sciences). The protein was eluted from the column with a linear NaCl concentration gradient (100 mM to 750 mM) in a buffer containing 50 mM Tris·HCl pH 7.9, 5% glycerol, and 5 mM β-

mercaptoethanol. Mtf2p was further purified by gel filtration on a 10/300 Superdex 200 GL column in a buffer containing 100 mM Tris·HCl pH 7.9, 200 mM NaCl, 5% glycerol, and 5 mM β-mercaptoethanol. Fractions containing pure Mtf2p were pooled, concentrated, diluted two-fold with glycerol, distributed into small aliquots, and stored frozen at -80 °C.

I.C.6. *Electrophoretic mobility shift assays*. Taken at the concentrations specified in individual figure legends, proteins were combined with 5′-[^{32}P]-labeled RNA probes in 10 μl of binding buffer (20 mM Tris·Cl pH 7.2, 50 mM NaCl, 5 mM $MgCl_2$, 5% glycerol). Yeast tRNA (Sigma) was added to reduce nonspecific binding (at a concentration indicated in the figure legends) and the mixtures were incubated for 20 min at 30 °C. The mixtures were then supplemented with 2 μl of 30% glycerol spiked with xylene cyanol and bromophenol blue and loaded on 7% or 8% (as specified in figure legends) 37.5:1 acrylamide:bisacrylamide native polyacrylamide gels cast in the presence of 1 × TBE buffer using the Mini-PROTEAN® Tetra Cell System (Bio-Rad). Electrophoresis was performed at room temperature in 0.5 × TBE running buffer for 15 min at 175 V. Radioactive bands corresponding to protein/RNA complexes and unbound RNA were visualized using storage phosphor screens and a Typhoon 9410 scanner (GE Healthcare) and quantified using ImageQuant 5.2 software (Molecular Dynamics). For Figure 23C, the apparent Kd values were calculated assuming a 1:1 RNA:protein ratio in the complex by the equation Kd^{App}= [R] × [P] / [RP], where [R], [P] and

[RP] are the molar concentrations of free RNA, free Pet111p and the RNA/Pet111p complex (all at equilibrium). The equilibrium concentration of the complex was determined as $[RP] = R_0 \times F_s$, where R_0 is the total concentration of RNA in the mixture and F_s is the fraction of shifted RNA. F_s was calculated using the data obtained by quantitation of the gel image as $F_s = I_b / (I_b + I_f)$, where I_b and I_f are the intensities of the bands corresponding to the bound and free RNA, respectively. The equilibrium concentrations of the free RNA and protein were calculated as $[R] = R_0 - [RP]$ and $[P] = P_0 - [RP]$, respectively, where P_0 is the total concentration of the protein in the mixture.

I.C.7. *RNase footprinting.* 5′-[^{32}P]-labeled RNA probes were incubated with Pet111p and/or Mtf2p (at the concentration indicated in the figure legends) in 10 μl of binding buffer (20 mM Tris·HCl pH 7.2, 50 mM NaCl, 5 mM $MgCl_2$, 5% glycerol) for 20 min at 30 °C to allow the formation of complexes. Yeast tRNA (at the concentration indicated in the figure legends) was present during the incubation as a nonspecific competitor. To induce limited RNA digestion, RNase I (New England Biolabs), RNase A (Qiagen), or RNase T1 (Boehringer Mannheim) was added to a concentration indicated in individual figure legends. All incubations were done for 15 min at 30 °C unless specified otherwise in the respective figure legend. The reactions were stopped by mixing them with 10 μl of gel loading buffer (50 mM EDTA in 95% formamide spiked with xylene cyanol and bromophenol blue) and heating them at 95 °C for 5 min. The products of digestion were separated by 7 M urea PAGE at a

percentage of 19:1 acrylamide:bisacrylamide as indicated in the figure legends. Gradient gels (thickness 0.4 mm to 1.2 mm) were used where indicated. The radioactive RNA species were visualized by phosphor imaging with a Typhoon 9410 scanner (GE Healthcare) and quantified using ImageQuant 5.2 software (Molecular Dynamics).

I.C.8. *In vitro transcription*. Utilizing SP6 RNA polymerase, an RNA probe was created containing the 5′-UTR and a portion of the *COX2* ORF. Using primers F *COX2* 5′-UTR -59 to -33 SP6 and R *COX2* +61 to +41, a region of DNA was amplified using the pJJ10 plasmid which contained a fragment of the *COX2* transcript (see Appendix section A.1) as a template. The DNA template was amplified to contain the sequence of the SP6 RNA polymerase P2 promoter (5′- AATAATTAGGTGACACTATAG -3′) (*123*), which was added at the 5′-end of the forward amplification primer. PCR amplification was performed in a volume of 250 μl using Phusion DNA polymerase (NEB) with a modified protocol. A two-step annealing cycle was used at temperatures 55 °C and 50 °C (10 sec each) and the elongation temperature was lowered to 60 °C and the time was extended to 2 min. The fragment was then purified and concentrated by the GeneJET PCR purification kit (Thermo Scientific).

In vitro transcription took place with purified SP6 RNA polymerase (see Appendix section A.3 for purification protocol). The reactions contained transcription buffer (20 mM Tris·Cl pH 7.9, 15 mM $MgCl_2$, 5mM β-mercaptoethanol), 0.2 μM PCR amplified DNA, 500 μM ATP, 100 μM GTP,

300 μM CTP, 300 μM UTP, 300 μM radiolabeled guanosine (labeled by PNK in the presence of γ-[^{32}P]-ATP; 3000 Ci/mmol, Perkin Elmer), 2 μg/μl BSA, and 1 μM SP6 polymerase. The transcription reaction took place at 37 °C for 1 h. The RNA was then gel purified as described in section I.C.1.

I.C.9. *Gel filtration of protein complexes***.** Gel filtration was performed using a Superdex 200 Increase 10/300 GL preparative gel filtration column on an ÄKTA pure chromatography system (GE Life Sciences) with highly purified Pet111p, Rpo41p, and Mtf2p. Pet111p and Mtf2p were purified as in section I.C.4 and I.C.5, respectively. The Rpo41p purification protocol is described in Appendix section A.3. In the experiment shown in Figure 24, Pet111p and Mtf2p were combined at a concentration of 16 μM each in a 100 μl volume of 20 mM HEPES·NaOH pH 7.0, 200 mM NaCl, 5% glycerol, and 5 mM β-mercaptoethanol. Where indicated a 42 nt RNA, *COX2*-1-42, was also added at a 16 μM concentration. The proteins were eluted with analytical gel filtration buffer (20 mM HEPES·NaOH pH 7.0, 200 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol) at a flow rate of 0.5 ml/min. In the experiment shown in Figure 25, Rpo41p and Pet111p were used at a concentration of 10 μM and Mtf2p was at a two-fold molar excess (20 μM). The proteins were mixed in the combinations indicated in Figure 25 in a total volume of 150 μl, incubated for 15 min at 30 °C, and loaded onto the gel filtration column. The elution was performed with 20 mM HEPES·NaOH pH 7.0, 150 mM NaCl, 5% glycerol,

and 5 mM β-mercaptoethanol at a flow rate of 0.5 ml/min. The chromatograms were analyzed using UNICORN 7 (GE Life Sciences) and graphed using Origin 6.1.

I.D. Results

I.D.1. *The 5′-end of mature COX2 is at the transcriptional start site.* The functional link between Pet111p and the translation of *COX2* has been established by genetic studies (*107, 108*). These studies have shown that the translational activation function of the protein is associated with a sequence and/or a structure in the 5′-UTR of *COX2.* This notion renders the action of Pet111p gene-specific (*116*). However, it was not clear whether Pet111p acts directly or indirectly on the 5′-UTR as previous studies were performed *in vivo* and had other mitochondrial factors present. To investigate the relationship between Pet111p and *COX2* further, an *in vitro* biochemical system needed to be established; consisting of highly purified Pet111p and various *COX2* RNA oligonucleotides. Prior to designing the synthetic RNA oligonucleotides, it was necessary to verify the position of the 5′-end in mature *COX2*.

At least three yeast mitochondrial mRNAs, *ATP8/6*, *COB*, and *VAR1*, undergo Pet127p-dependent trimming at the 5′-end, following endonucleolytic processing of the primary transcripts (*43*). However, due to a combination of the limited resolution of the method used in the study and the small length of the 5′-UTR, it was not possible to conclude whether *COX2* was also shortened in a similar way. The position of the 5′-end was later mapped to coincide with the transcription start site of the *COX2* promoter (*124*). The mapping was performed at a single-nucleotide resolution by extension of a DNA primer hybridized to *COX2* with reverse transcriptase (Figure 8A).

However, the 3′-end of the primer used in the study corresponded to position 33 inside the 5′-UTR (Figure 8B). Therefore if a trimming event occurred where this primer was annealed, the experiment would fail to detect the maturation product.

To address this, I repeated the primer extension experiment with the entire primer shifted to a location inside the coding sequence (Figure 8C). In this assay, a radiolabeled primer complementary to a sequence within the *COX2* reading frame was annealed to the *COX2* mRNA in total mitochondrial RNA isolate. This hybrid was then extended by reverse transcriptase (RT). The extension halted at the 5′-end of the mRNA (i.e. the maturation site). As shown in Figure 8D, extension of the primer did not terminate near the area of the initiation codon but produced a much larger single product. As determined by comparison with a set of DNA size markers generated by a sequencing reaction, the length of the product corresponded to the +1 position of the primary transcript (transcription start site). This result demonstrated that *COX2* does not undergo a 5′-end trimming, consistent with published data (*124*), and therefore the entire 54-nt sequence of its 5′-UTR should be included in the evaluation of the interaction with Pet111p.

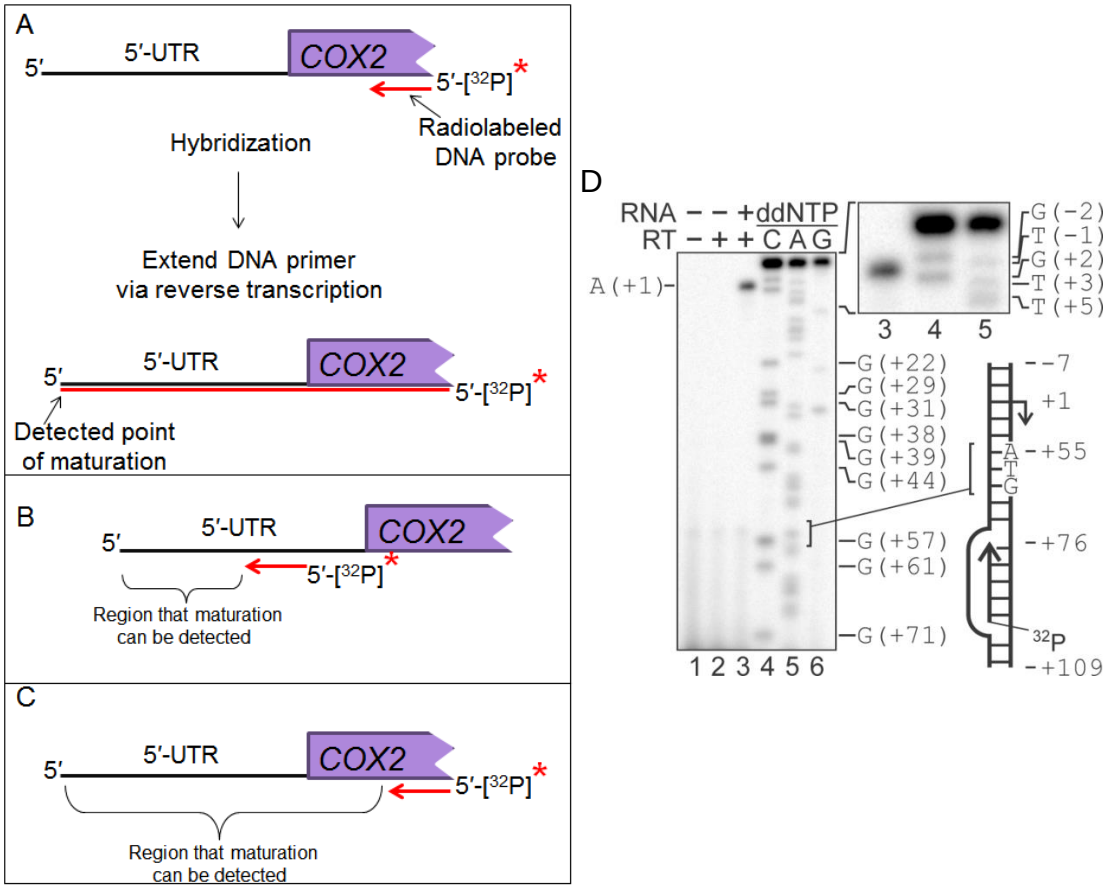

Figure 8. The 5′-end of mature *COX2* is at the transcriptional start site, +1. A. A scheme of the primer extension analysis experiment showing a radiolabeled DNA primer (red) annealed to the *COX2* mRNA. Using reverse transcription, the primer is extended to the maturation point at the 5'-end of the mRNA. B. Previous mapping performed by Coruzzi et al., utilized a primer annealed within the *COX2* 5′-UTR, therefore limiting the maturation region that can be detected to the terminal 5′-end of the *COX2* 5′-UTR (as indicated by the bracket) (*124*). C. The set up for my experiment in panel D shifted the radiolabeled primer into the ORF of *COX2*, such that the maturation region that can be detected by this method includes the entire 5′-UTR (as indicated by the bracket). D. A phosphor imaging scan of an area of a gradient-thickness 7 M urea 6% polyacrylamide gel showing extension of a 5′-[^{32}P]-labeled DNA primer *COX2*-96-76. The primer was hybridized to *COX2* mRNA in a total mitochondrial RNA isolate and extended with RT (lane 3). Control

lanes 1 and 2 contained, respectively, the unextended primer (which appears below the bottom boundary of the image and is not visible) and primer extension in the absence of mitochondrial RNA. The length of the product of extension in lane 3 was determined by comparison with DNA size markers (lanes 4-6). The markers were generated by extension of the primer by Taq DNA polymerase in the presence of ddNTPs as specified above the image. The double-stranded DNA template used in the sequencing reactions contained a region of the *COX2* promoter between positions -7 and +109 relative to the start-site (+1), as explained by the scheme to the right. The bands in lane 4 are assigned on the right of the image and the position on the gel that corresponds to the initiation codon is indicated by brackets. An area on the top of the image is enlarged to show the slow-running bands more clearly and the bands in lanes 4 and 5 are assigned on the right. Comparison with the size markers indicates that the 3′-end of the product of primer extension in lane 3 corresponds to position +1, and thus the 5′-end of mature *COX2* mRNA coincides with the promoter start-site.

I.D.2. *The N-terminus of mature Pet111p is at amino acid 35.* Prior to expressing Pet111p heterologically, the N-terminus of the mature protein needed to be identified since most proteins entering the mitochondria from the cytosol have a mitochondrial import signal that is cleaved upon entry. It is known that most proteins targeted to the mitochondrial matrix undergo a one-step cleavage by mitochondrial processing peptidase (MPP). However, there may be a secondary processing event, in which either an additional octapeptide or a single amino acid is removed from the N-terminus by Oct1p or Icp55p, respectively (*125-128*). Mitochondrial proteins expressed

heterologically in their full-length form, are often insoluble, while their mature forms are generally soluble (*129*).

I sought to create a recombinant form of the mature protein; however, the processing site of Pet111p has not been experimentally established. To determine the N-terminus of the mature Pet111p, a *PET111* expression plasmid from a yeast BG1805 open reading frame (ORF) library was acquired. This plasmid provided a C-terminally fused hexahistidine tag to facilitate metal affinity purification (see Figure 9). Tagged Pet111p was overexpressed in yeast, the mitochondrial fraction was isolated, and the processed form of the protein was partially purified using Ni-IDA beads under denaturing conditions.

Analysis of the protein preparation by SDS-PAGE (Figure 9) revealed a major protein band running at approximately 97 kDa. This protein was further identified as Pet111p by peptide mass fingerprinting (PMF) (*130*). Edman degradation analysis of the protein returned two overlapping sequences (see Figure 9). It was therefore concluded that Pet111p undergoes a two-step processing with MPP cleaving a 33 amino acid peptide at Y33 and Icp55p removing residue Y34 (see Figure 9).

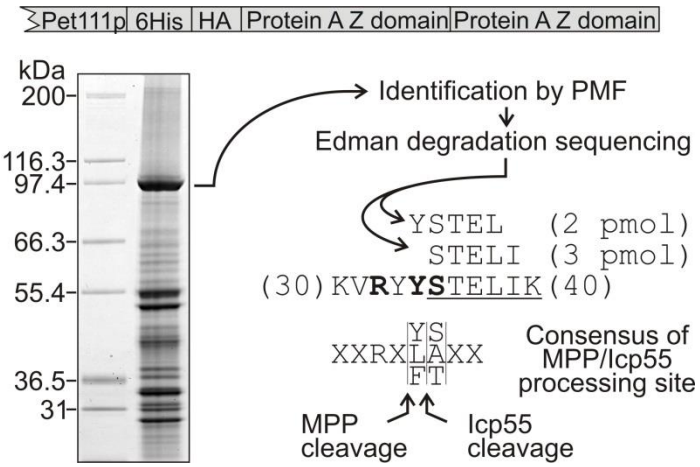

Figure 9. The mature N-terminus of Pet111p is at amino acid 35. An image of a Coomassie-stained LDS 4-12% polyacrylamide gel showing a preparation of tagged Pet111p partially purified from yeast mitochondria. The composition of the C-terminally fused purification tag is shown above the image. The positions of the hexahistidine (6His) and hemagglutinin epitope (HA) tags are indicated. A protease 3C recognition sequence exists between the HA and first Protein A Z domain as shown by the black arrow. The protein in the indicated band was identified by PMF as Pet111p. Edman degradation analysis of the protein revealed two sequence reads, pointed to by the bent arrows, which were present at a ratio of 2:3. The reads are aligned with a portion of the Pet111p sequence (numbers indicate positions of the flanking amino acid in the sequence of the Pet111p precursor), in which the determined N-terminal sequence of the mature protein is underlined. The amino acids in boldface match the consensus of the MPP/Icp55p processing site (*128*), which is shown below the alignment. The sites of cleavage by MPP and Icp55p are indicated by arrows.

I.D.3. *Expression and purification of Pet111p.* Based on the Edman degradation data, *PET111* lacking the sequence coding for the first 34 amino

acids was cloned into an *E. coli* pTrcHis-A expression vector to generate an N-terminal histidine purification tag (MAH_6) fused to amino acid S35 of the protein (plasmid pGD1). Expression of pGD1 was optimized in *E. coli* XJb(DE3) cells grown at a reduced temperature (12 °C) with a low level of induction to produce a soluble Pet111p. As judged by Coomassie staining of an SDS-PAGE (Figure 10A), the protein was purified to homogeneity by a preliminary enrichment with Ni-IDA beads followed by heparin affinity chromatography and gel filtration.

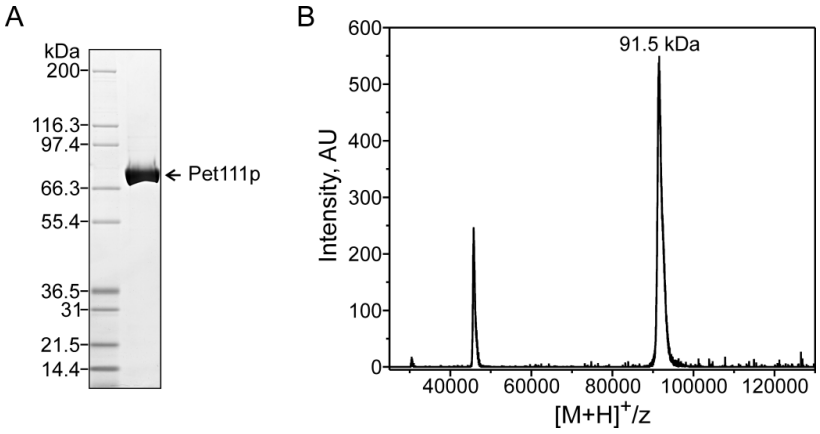

Figure 10. Purification of Pet111p. A. An image of a Coomassie-stained LDS 4-12% polyacrylamide gel showing the purity of a typical preparation of Pet111p. B. A MALDI-TOF mass spectrum of purified Pet111p showing that the observed molecular weight of the protein is in agreement with the expected. The spectrum was taken on a Bruker Microflex LRF spectrometer in the positive-ion linear mode using sinapinic acid as the MALDI matrix. The spectrum was externally calibrated using BSA as a standard.

A notable characteristic of this protein is its abnormal mobility in a denaturing polyacrylamide gel. The band of purified Pet111p corresponded to

a molecular weight of approximately 70 kDa (Figure 10A), which was substantially lower than the calculated value of 91.3 kDa. To exclude that this shift was due to a C-terminal truncation that might have occurred during the expression, the precise molecular weight of the protein in the preparation was determined by MALDI-TOF mass spectrometry. As shown in Figure 10B, the observed molecular weight of Pet111p was in agreement with the expected value.

I.D.4. *Pet111p specifically recognizes a sequence within the 5′-UTR of COX2.* Previous genetic data has implicated Pet111p in a direct functional relationship with the 5′-UTR of *COX2* (*107, 108*). However, a direct interaction has yet to be shown *in vitro*. This was tested using my highly purified recombinant Pet111p and *COX2* 5′-UTR fragments in an electrophoretic mobility shift assay (EMSA). In this experiment radioactively labeled RNA probes of a defined sequence within the *COX2* 5′-UTR were incubated with Pet111p to allow the formation of complexes, which were then separated from free RNA by native PAGE. Pet111p is a basic protein (theoretical pI 8.6) and therefore it could interact electrostatically with RNA of any sequence. To suppress nonspecific binding of the labeled probes, a molar excess of yeast tRNA was present as a competitor in all experiments. First, the entire *COX2* 5′-UTR (*COX2*-1-54) was compared with a nonspecific control (NSC) that represented a sequence downstream from the coding sequence (see Figure 11A scheme). As shown in Figure 11A, EMSA revealed prominent bands

corresponding to the *COX2*-1-54/Pet111p complex, whereas the NSC failed to complex appreciably with the protein. This led to the conclusion that Pet111p can discriminate between RNA molecules independently of other mitochondrial proteins, and that the 5′-UTR harbors a binding target of the protein.

A relatively stable stem-loop structure has been predicted to form 19 nt upstream of the reading frame of *COX2* in *S. cerevisiae* and similar structures were suggested in corresponding mRNAs of *S. kluyveri* and *K. lactis* at comparable locations (*115*). Moreover, in *S. cerevisiae*, the structure resides within a 31 nt region, which has been shown to contain residues required for respiration (*113*). These data suggest that this structure might be a target of Pet111p. To evaluate this possibility, a pair of overlapping RNA probes was used, which together cover the entire sequence of the 5′-UTR. The two probes overlapped in such a way that the stem-loop structure could not form in either of them (see scheme in Figure 11B). If the hairpin was required for Pet111p binding, one would expect to see a loss of the complex formation with either probe. However, as Figure 11B shows the 5′-end proximal probe, *COX2*-1-30, retained the ability to interact with Pet111p while the 3′-end proximal probe, *COX2*-23-54, failed to bind to the protein. These results indicate that a binding target of Pet111p resides within the sequence of *COX2*-1-30 and argue against the involvement of the stem-loop as a major binding or specificity determinant. Considering that no secondary structure is

expected to form within *COX2*-1-30, Pet111p can be classified as a sequence-specific single-stranded RNA binding protein.

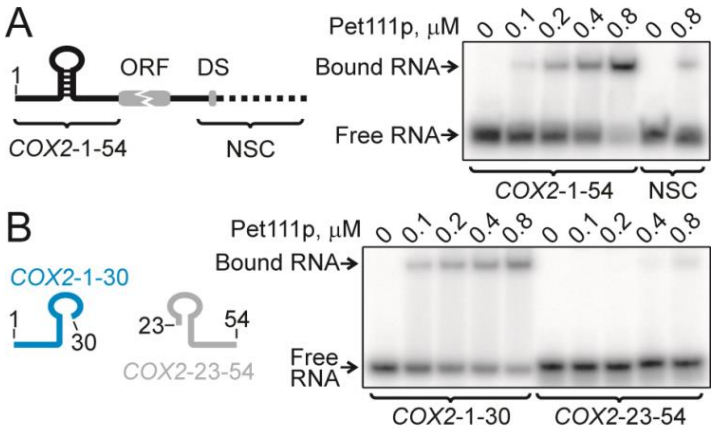

Figure 11. Pet111p binds to the 5′-proximal end of the *COX2* 5′-UTR. 5′-[^{32}P]-labeled RNA probes (0.1 μM) were incubated with varying concentrations of Pet111p and resolved in 8% polyacrylamide gels under native conditions. The bands corresponding to unbound and bound RNA (as indicated) were visualized by phosphor imaging. During binding, competitor tRNA was present in the solutions at two-fold molar excess over the protein. A. The binding efficiency of two probes was compared, one representing the entire *COX2* 5′-UTR and the other corresponding to a sequence downstream of the reading frame (non-specific control, NSC) as illustrated by the scheme to the left. DS indicates the position of a conserved dodecameric sequence present at the 3′-ends in all yeast mitochondrial mRNAs (*40*). The dashed line indicates the region not present in the mature *COX2* (*40*). B. Two overlapping probes were designed to cover the full sequence of the *COX2* 5′-UTR as illustrated in the scheme and the efficiencies of their interaction with Pet111p were compared.

The EMSA results indicated the presence of a Pet111p binding target in the RNA probe *COX2*-1-54 and more specifically a location within the

COX2-1-30 probe. With the discovery that Pet111p can directly interact with the 5′-UTR of *COX2* in the absence of any other mitochondrial factors, the next step was to identify the particular ribonucleotides involved in this interaction. To define the target within the probes with greater precision, RNase I footprinting was employed (*131*). Both probes (*COX2*-1-54 and *COX2*-1-30) were [^{32}P]-labeled at the 5′-end, gel purified, and subjected to a limited digestion with RNase I. The products of digestion were separated at a single-nucleotide resolution by denaturing PAGE (Figure 12). RNase I cleaves every single-stranded RNA base, therefore if there was duplex RNA in the probes it would not be cleaved. Each probe was incubated with Pet111p in the presence of excess yeast tRNA to allow the formation of a specific complex, the complex was then treated with RNase I in the same way as the free probe, and the formed products of digestion were separated. Figure 12A shows cleavage of the full-length 5′-UTR. The products of digestion were quantified and for each RNA fragment the effect of Pet111p was plotted as a logarithm of the intensity ratio in lanes 3 and 2 (bar plot in Figure 12A). The positive-value bars in the plot indicate the residues in the RNA at which RNase I cleavage was partially blocked by Pet111p. The cleavage pattern in lane 3 clearly indicated that the 3′-proximal end of the RNA is highly structured between nucleotides 18 and 54. Due to the presence of this structure, the Pet111p footprint could only partially be determined as having a 5′-boundary at nucleotide 7. The 3′-boundary appeared to extend at

least to position 18; however, it was too close to the uncleaved structured section to be designated more precisely.

For the *COX2*-1-30 RNA, the pattern of the cleavage products was reasonably uniform throughout the probe, except at the purine-rich 3′-end where the cleavage was somewhat less prominent. This suggested that the probe was free of secondary structures (Figure 12B, lane 2). RNase I is known to cleave poorly at GMP residues (*132*), and thus the weak RNA band corresponding to the product of cleavage at G22 was readily identifiable and provided a reference point to assign other RNA bands in each digestion (Figure 12B). The pattern of the RNase I cleavage was clearly different in the presence of Pet111p. As evident from the traces representing the distribution of radioactivity in lanes 2 and 3, the cleavage efficiency decreased drastically in a region close to the middle of the probe, while increasing in proximity towards the ends of the probe. The most prominent protection was observed in a 13 nt region between positions 7 and 19 (bar plot in Figure 12B), suggesting the presence of a Pet111p binding target within these boundaries. This region of protection generally corresponds to that observed with probe *COX2*-1-54.

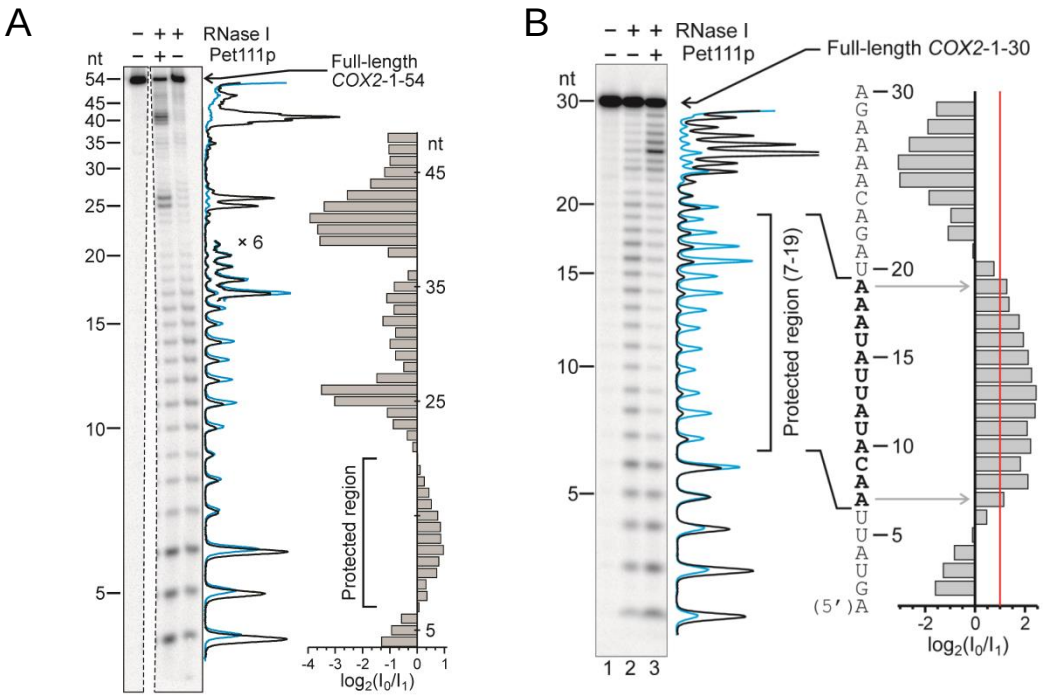

Figure 12. Pet111p protects a region between positions 7 and 19 in the *COX2* 5′-UTR. Phosphor imaging scans of 7 M urea gradient thickness 20% polyacrylamide gels showing products of limited RNase digestion of RNA probes. A. RNase I (8 u/ml) was used to digest the 5′-[^{32}P]-labeled *COX2*-1-54 RNA probe (1 µM) in the absence (right lane) or presence (center lane) of Pet111p (0.6 µM). The digests contained yeast tRNA as a nonspecific competitor (1.2 µM). The numbers to the left of the image indicate the length of the digestion products. The traces on the right show the distribution of radioactivity along the center and the right lanes (black and blue, respectively). The bar graph illustrates the changes in intensity of individual RNA bands induced by Pet111p. For each RNA fragment, intensities of the corresponding RNA bands were measured in lanes 3 (I_0) and 2 (I_1) and presented in the plot as $\log_2(I_0/I_1)$. B. 5′-[^{32}P]-labeled *COX2*-1-30 RNA (0.6 µM) was subjected to a limited digestion with RNase I (8 U/ml) in the absence or presence of Pet111p (0.8 µM). The reactions were carried out in the presence of 1 µM yeast tRNA as a nonspecific competitor. The traces next to the gel image represent the distribution of radioactivity within lanes 2 (blue)

and 3 (black). The bar graph illustrating the changes in intensity of individual RNA bands induced by Pet111p was plotted as in panel A. Positions in the RNA at which the bars in the plot exceeded 55% of the average value of all positive bars (red line) were considered protected by Pet111p and the corresponding nucleotides are indicated by boldface in the sequence of *COX2*-1-30.

I.D.5. *Pet111p can induce structural changes in the COX2 5′-UTR upon binding.* The RNase I digestion pattern observed for the full-length 5′-UTR (Figure 12A), indicates that this RNA is highly structured at the 3′-proximal end *in vitro*. Pet111p binding also induced hypersensitivity in the 3′-proximal region of the probe (Figure 12A, lane 2, upper more intense RNA bands). This hypersensitivity is indicative of double-stranded RNA bases becoming available for cleavage, i.e. conversion to single-stranded RNA upon Pet111p binding. To examine this effect further, the *COX2*-1-54 probe was exposed to three different RNases in the absence or presence of Pet111p. The 5′-UTR of the *COX2* mRNA was previously predicted (*113*) to form a stable stem-loop structure between positions 20 and 35 as illustrated in Figure 13C. Interestingly, the Pet111p 5′-UTR binding site is located just upstream from the putative stem-loop and this structure does not affect binding, as shown in the EMSA and footprinting experiments in Figure 11A and Figure 12A. However, the mode of interaction may differ when the stem-loop is present (such as it is in probe *COX2*-1-54). The stem-loop structure may provide additional stabilizing contacts with the protein or, conversely, sterically

obstruct a portion of the recognition sequence. To evaluate these possibilities, the structural organization of *COX2*-1-54 was analyzed by limited RNase digestion. The patterns of digestion of *COX2*-1-54 with RNases A, I, and T1 in the absence of the protein (Figure 13A,B) indicated that the probe contained regions of double-stranded RNA.

RNase A, which cleaves at pyrimidine residues followed by a purine, readily cleaved all sites present within the 16 5′-terminal nucleotides and also at the positions U48 and U51 near the 3′-end, suggesting that these regions are single-stranded. At the same time, the cleavage was poor at four predicted sites, U20, C24, U34, and U41, suggesting the presence of a secondary structure in the middle of the probe. The single-stranded nature of the RNA in the 5′-end proximal region was confirmed by digestion with RNase I, which also cleaved the probe efficiently between positions 3 and 18 (Figure 13A). However, the cleavage was weak beyond position 18 towards the 3′-end, suggesting the presence of a structure in this region. Probing with RNase T1 (cleaves single-stranded G residues; Figure 13B) revealed that positions G2 (not shown) and G29 were the only sites in the RNA sensitive to the RNase, while there was much less efficient cleavage at positions G22, G31, G38, G39, and G44; indicative of possible involvement of these nucleotides in a secondary structure. Taken together, the observed inefficient digestion by RNase A and RNase T1 between nucleotides 20 and 34 can be attributed to the formation of the stem of the previously proposed stem-loop structure (*113*) (Figure 13C). The cleavage patterns also indicated that the

overall folding of the RNA is more complex. The loop of the stem-loop structure was generally resistant to cleavage by RNase I and thus, even though RNase T1 could cleave it well at position G29, this region of the RNA must be at least partially engaged in base pairing. Moreover, as summarized in Figure 13C, an area just downstream from the stem-loop appeared to be structured as well, resulting in poor cleavage by RNase I, RNase A, and RNase T1. Overall, the results of this analysis of *COX2*-1-54 leads to the conclusion that the probe forms a complex secondary structure within the region from position 19 to at least position 44 and that this structure is likely to involve the predicted stem-loop (*113*) as a component.

It was puzzling why the 3′-proximal end of the probe was sensitive to RNase A while seemingly more resistant to RNase I. This inconsistency was attributed to the poor ability of RNase I to cleave RNA close to the 3′-end of a probe, which was observed with this and other RNA probes used in this study (see Figure 12B). As evident from Figure 12A, the footprint of Pet111p using the *COX2*-1-54 probe proved to be difficult to define due to the structuring of the probe and the fact that RNase I cannot cleave double-stranded RNA. Nevertheless, the region in which the protein partially protected the probe from RNase I digestion was clearly visible (Figure12A) and was generally the same as the one identified using the *COX2*-1-30 probe (Figure 12B). Thus, it was concluded that the presence of the secondary structure in the *COX2*-1-54 had little effect, if any, on the positioning of the protein on the RNA.

Unexpectedly, the binding of Pet111p induced significant changes in the digestion pattern of the probe. The cleavage efficiency at nucleotides A25 and A26 in the upstream portion of the loop in the stem-loop structure increased dramatically in the presence of Pet111p, signifying a transition of this region into a single-stranded state (Figure 12A). Similarly, the enhanced cleavage within positions 39-42 was indicative of base unpairing in that region as well. The cleavage pattern of *COX2*-1-54 by RNase T1 also changed dramatically in the presence of Pet111p (Figure 13B). Consistent with the effect observed with RNase I (Figure 12A), digestion at nucleotides G38, G39, and G44 became significantly more efficient. Conversely, the cleavage at position G29 in the loop of the stem-loop structure was suppressed. The twist of the 5 bp stem of the structure is expected to rotate the loop by approximately 180°, which should place nucleotide G29 in a position facing the binding site of the protein. Therefore, cleavage suppression at G29 may result from a steric hindrance by the bound Pet111p. Alternatively, G29 might become base paired as a result of the reorganization of the secondary structure of the probe induced by Pet111p binding.

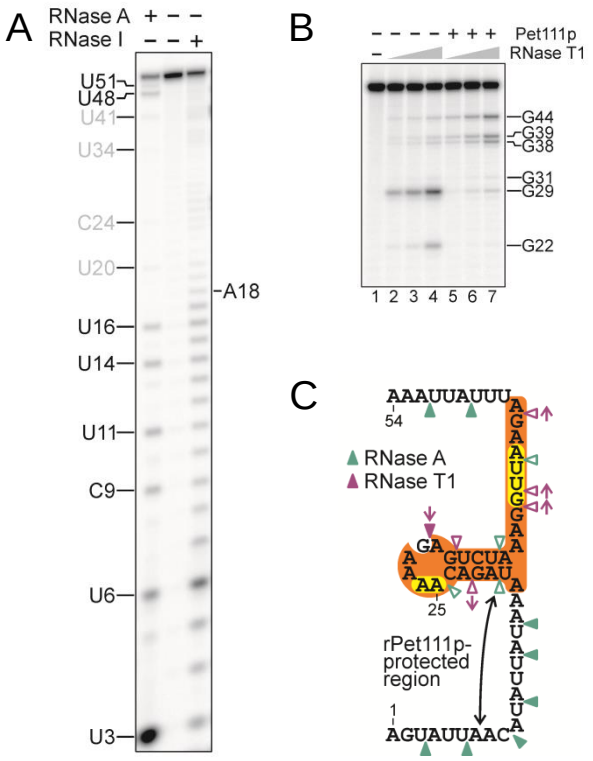

Figure 13. An analysis of the interaction between Pet111p and the entire 5′-UTR of *COX2*. Phosphor imaging scans of polyacrylamide gels showing products of limited RNase digestion of a 5′-[^{32}P]-labeled *COX2*-1-54 RNA probe. To better reveal the products, the intensity of the images were enhanced such that bands corresponding to the full-length probe (in the center lane of panel A, the outer lanes of panel B, and in panel C) and the 3-nt product in the left lane of panel A are overexposed. A. The RNA probe was digested with RNase A (10 ng/ml for 3 min, left lane) and RNase I (25 u/ml, right lane). The expected RNase A cleavage sites are indicated to the left of the image; the sites at which inefficient cleavage was observed are shown in gray. The RNase I cleavage was much less efficient in the region from approximately position 18 (indicated on right) to the 3′-end of the probe. B. The probe (200 nM) was digested with RNase T1 in the absence (lanes 2, 3, and 4) or presence (lanes 5, 6, and 7) of Pet111p. The RNase concentrations

used were 4 u/ml (lanes 2, 5), 20 u/ml (lanes 3, 6), and 100 u/ml (lanes 4, 7). The assignment of the RNA bands to corresponding cleavage sites in the RNA is shown at right of the image. C. A scheme summarizing the results of the experiments in Figure 12 and in panels A and B shown here. The teal and purple triangles point to RNase A and RNase T1 cleavage sites, respectively, and are filled at the sites where efficient cleavage occurred. Purple arrows indicate whether the efficiency of RNase T1 cleavage increased (upward arrows) or decreased (downward arrows) in the presence of Pet111p. The RNA residues that appear to be involved in a secondary structure are highlighted in orange and those that become hypersensitive to RNase I in the presence of Pet111p are highlighted in yellow. The double-headed arrow indicates the region protected from RNase I cleavage by Pet111p. The numbers signify the nucleotide positions in the probe relative to the transcription start site (+1).

I.D.6. *Pet111p interacts with a sequence in the COX2 open reading frame.* A previous genetic study provided evidence that a positively acting translational control element was present within the first 14 codons of the *COX2* ORF (*114*). Specifically, codons 2 to 6 were critical for the function of this element. Significantly, following a frame shift that altered the amino acid sequence no deleterious phenotypes were observed. However, point mutations within the element caused a loss in respiration; showing that the nucleotide sequence was critical. Overexpression of Pet111p restored respiration, thereby inferring that this protein may have a relationship with this portion of the ORF (*114*). This prompted me to examine whether this region of the mRNA had an additional binding target for Pet111p. The probe used in this experiment

(*COX2*-58-95) extended from just downstream of the initiation codon to the descending strand of a predicted stem-loop structure (*114*) located in the beginning of the coding sequence (Figure 14 scheme). EMSA revealed that this probe could readily complex with Pet111p with an affinity comparable to that of the *COX2*-1-30 probe, which was used as a positive control (Figure 14). Therefore it was concluded that at least two Pet111p binding targets are present in the 5′-end proximal region of *COX2*, one in the beginning of the 5′-UTR and one downstream from the initiation codon, and that Pet111p can recognize both targets independently of other mitochondrial factors.

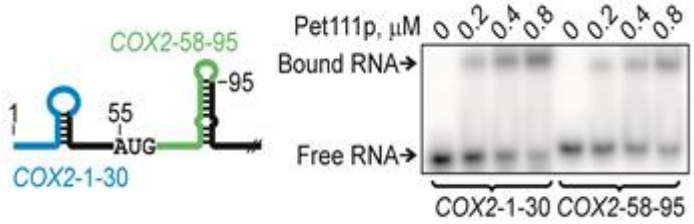

Figure 14. Pet111p binds to a target in the *COX2* ORF. 5′-[^{32}P]-labeled RNA probes (0.1 µM) were incubated with varying concentrations of Pet111p and resolved in an 8% polyacrylamide gel under native conditions. The bands corresponding to unbound and bound RNA (as indicated) were visualized by phosphor imaging. During binding, competitor tRNA was present in the solutions at two-fold molar excess over the protein. The binding efficiency of the *COX2*-1-30 probe was compared to that of a probe, *COX2*-58-95, starting downstream of the initiation codon and extending to just past the loop of a putative stem-loop structure as shown in the scheme to the left.

To further characterize the ORF interaction, RNase I footprinting was utilized with the *COX2*-58-95 RNA probe (Figure 15) in the absence or

presence of Pet111p. To assign the bands of the RNA fragments, a set of size markers was generated by digestion of the probe with RNase A. This experiment revealed a region between positions 66 and 86 as partially protected by the protein from RNase I digestion, which defined the boundaries of a second binding target of Pet111p.

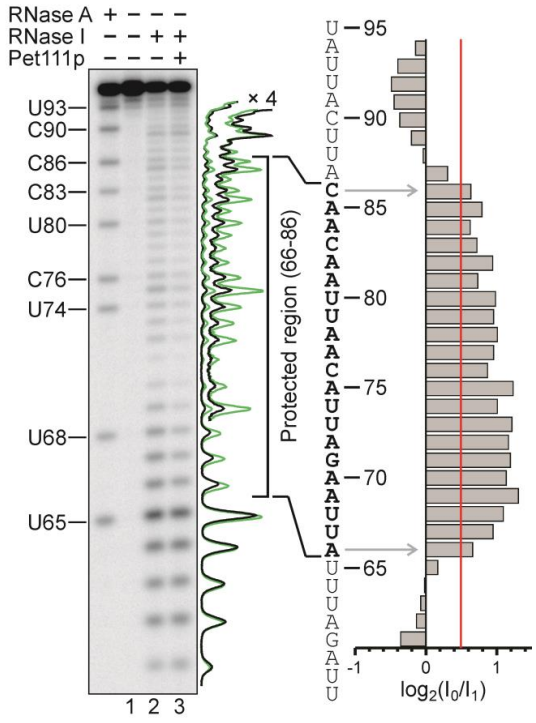

Figure 15. Pet111p interacts with a specific RNA sequence within the *COX2* ORF. 5′-[^{32}P]-labeled *COX2*-58-95 was treated with RNase I (8 U/ml) in the presence of 1.2 µM tRNA and the products of digestion were resolved and visualized as in Figure 12. Pet111p (0.6 µM) was present in the digest in lane 3. Digestion with RNase A (0.8 ng/ml) was also performed to generate RNA size markers (shown in the left lane and identified on the left). The traces and the bar plot on the right side are as in Figure 12. The green and black traces correspond to lanes 2 and 3, respectively. Intensities of the corresponding RNA bands were measured in lanes 2 (I_0) and 3 (I_1) and presented in the plot as $\log_2(I_0/I_1)$. Positions in the RNA, at which the bars in the plot exceeded

55% of the average value of all positive bars (red line) were considered protected by Pet111p and the corresponding nucleotides are indicated by boldface in the sequence of *COX2*-58-95.

The finding that the two mapped targets of Pet111p differ substantially in length (13 nt versus 21 nt in the 5′-UTR and ORF targets, respectively) and share little sequence similarity was surprising. One explanation for this observation is that Pet111p might harbor two different RNA binding sites, each recognizing one of the two targets. If the two binding sites worked independently, the protein could then associate with both targets at the same time. To test if this was the case an EMSA was used to resolve complexes formed when Pet111p was added to a mixture containing both *COX2*-1-30 and *COX2*-58-95 RNA probes (Figure 16). In the first set of reactions, radiolabeled *COX2*-1-30 RNA was present at a fixed concentration while the concentration of unlabeled *COX2*-58-95 was gradually increased. A band corresponding to a ternary complex was not observed either with this set of samples or in a reciprocal experiment where the radioactive probe *COX2*-58-95 was used at a fixed concentration and the unlabeled probe *COX2*-1-30 was present at increasing concentrations (Figure 16, right side). Moreover, each probe appeared to outcompete the other one when added in excess, with the target in the 5′-UTR exhibiting a considerably higher affinity to the protein than the one in the ORF. This data suggests that Pet111p can only bind to one of the two RNA targets at a time.

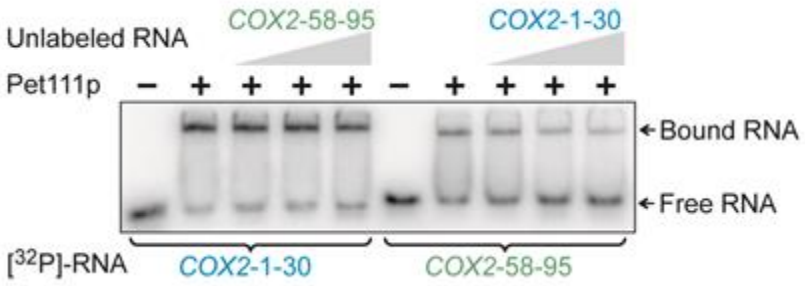

Figure 16. Pet111p utilizes one RNA binding site. 5′-[^{32}P]-labeled *COX2*-1-30 was incubated with 0.8 µM Pet111p. The formed complex was resolved from the unbound probe in a native gel and visualized by phosphor imaging. Where indicated, unlabeled specific competitor *COX2*-58-95 was present in the mixtures at increasing concentrations of 0.2, 0.4, and 0.8 µM. In a reciprocal experiment, the labeled probe was *COX2*-58-95 and the unlabeled competitor was *COX2*-1-30 (the right side of the image). The bands corresponding to free and Pet111p-bound RNA are indicated on the right. No protein was added in the control lanes, as indicated.

I.D.7. *Pet111p interacts with an RNA probe containing both binding targets with high efficiency.* My experiments show that Pet111p binds to only one of its RNA targets at a time. This led me to question how Pet111p associates with the native *COX2* transcript and whether both sites have equal affinity. To begin to answer this question I synthesized an RNA probe (*COX2*-(-59)-61) by *in vitro* SP6 RNA polymerase transcription to contain both Pet111p binding sites. The sequence of the probe corresponded to the interval in the *COX2* gene from -58 to +61 relative to the transcription start site (+1). An extra G residue was added to this sequence at the 5′-end of the probe due to the requirement of the SP6 promoter. The presence of two free *COX2*-(-59)-61

RNA bands (Figure 17, lane 1) in the EMSA analysis showed that the native RNA can assume different folds. This RNA restructuring may play a regulatory role in translational activation. The EMSA also revealed that Pet111p has the ability to shift this probe with high efficiency. The formation of a super-shifted RNA complex (lane 5), suggests that more than one Pet111p molecule can bind to this RNA simultaneously.

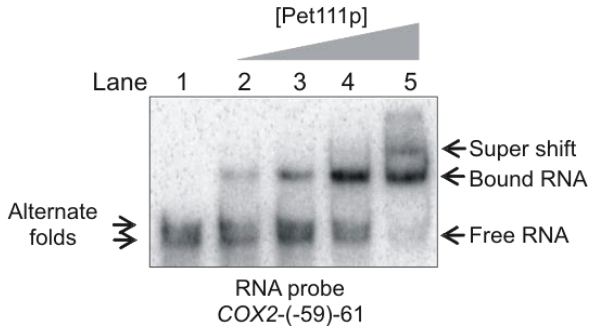

Figure 17. Pet111p strongly interacts with an RNA probe containing both the 5′-UTR and the ORF binding targets. 5′-[^{32}P]-G labeled *COX2*-(-59)-61 (low nM range) was incubated with 0.05, 0.1, 0.2, or 0.4 mM Pet111p in lanes 2, 3, 4, and 5, respectively. The formed complex was resolved from the unbound probe in a native 7% gel and visualized by phosphor imaging. 1 μM of yeast tRNA was present in all lanes as a non-specific competitor. The unbound RNA appears on the gel as two distinct bands (the arrows to the left), suggesting that there may be alternate structuring of the probe. The arrows on the right point at the different RNA species that migrated into the gel: unbound RNA (bottom), RNA presumably bound to a single molecule of Pet111p (middle), and RNA bound to more than one molecules of Pet111p (top). No protein was added in the control, lane 1, as indicated.

Previous structural analysis of the 5'-UTR of *COX2* (*COX2*-1-54 probe) revealed that it is highly structured at the 3′-end (Figure 13). However, this

probe lacked the initiation codon and sequence within the ORF. Therefore, I sought to determine whether the initiation region of *COX2* was structured as well. To do this, an RNase T1 limited digestion was performed on probe *COX2*-(-59)-61. Significantly, this digestion revealed that G57 in the initiating "AUG" along with the surrounding region was inaccessible for RNase T1 digestion. This data suggests that this part of the probe is indeed structured, at least *in vitro* (Figure 18).

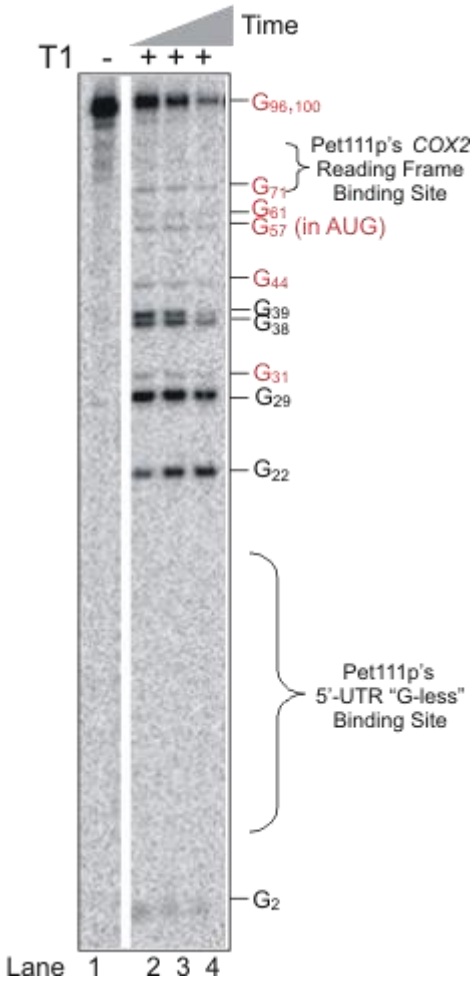

Figure 18. The AUG region appears to be structured in the *COX2*-(-59)-(61) RNA probe. A phosphor imaging scan of a 7 M urea 16%

polyacrylamide gel showing products of limited RNase digestion of the 5′-[^{32}P]-G labeled *COX2*-(-59)-(61) RNA probe. To reveal the products more clearly, the intensity of the image was enhanced such that bands corresponding to the full-length probe (the first three lanes) are overexposed. The RNA probe was digested with 500 u/ml of RNase T1 in lanes 2, 3, and 4, for 2, 4, and 8 minutes, respectively. The assignment of the RNA bands to corresponding cleavage sites in the RNA is shown to the right of the image. Red nucleotides indicate an absence of cleavage. The regions of both Pet111p binding sites are labeled by brackets.

I.D.8. *Pet111p is needed for the stability of COX2.* A majority of the mitochondrial transcripts are regulated by multiple protein factors implicated in both mRNA stability and translational activation through potential interactions with 5′-UTRs (*18*). Pet111p is the only factor known to interact at the 5′-UTR of *COX2*. However, its role has been limited to the translational activation of *COX2* mRNA and not to stability (*107*). The amount of *COX2* was previously quantitated in different Pet111p mutant cell lines as well as in the wild type by northern blot analysis (*107*). The results of this analysis showed that the level of *COX2* was decreased in Δ*PET111* cells by 3-10 fold relative to wild type. Knowing that substantial residual amounts of *COX2* in the absence of Pet111p were present led me to inquire if there are additional stability factors that interact with the 5′-UTR of *COX2*. To assess this, I performed a primer extension assay with a *PET111* deletion strain. The analysis was done with primer extension analysis as shown in Figure 8 using total RNA isolated from wild type and Δ*PET111* W303 yeast. If another

stability factor were bound to the 5′-UTR of *COX2* downstream from Pet111p, the 5′-end of the mRNA would have shifted downstream in the mutant cells and a new, shorter product of primer extension would be generated (see bottom right schematic in Figure 19). However, the primer annealed to *COX2* in the mutant cell RNA was extended to the same point as with the wild type RNA (Figure 19). This result indicated that any other proteins that may hypothetically be bound to the 5′-UTR were not able to block 5′-exonucleolytic degradation of *COX2* and thus cannot be regarded as significant mRNA stability factors in the absence of Pet111p. However, this data does not imply that there may not be other protein(s) bound to the 5′-UTR in the presence of Pet111p. Therefore, one can conclude from this experiment that there are no additional factors bound to the 5′-UTR that may independently act as significant stability factors in the absence of Pet111p. Pet111p thus remains the only known factor that is important for the stability of the *COX2* mRNA.

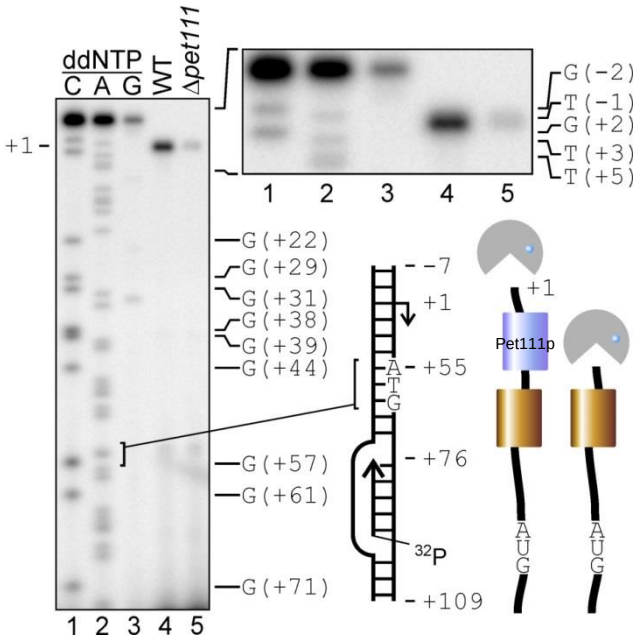

Figure 19. Pet111p provides stability for *COX2*. A phosphor imaging scan of an area of a gradient-thickness 7 M urea 6% polyacrylamide gel showing extension of a 5′-[^{32}P]-labeled DNA primer *COX2*-96-76. The primer was hybridized to *COX2* mRNA in total cell RNA isolated from both wild type and Δ*PET111* cells and extended with RT (lanes 4 and 5). The lengths of the products of extension in lane 4 and 5 were determined by comparison with DNA size markers (lanes 1-3). The markers were generated by extension of the primer by Taq DNA polymerase in the presence of ddNTPs as specified above the image and explained in Figure 8. An area on the top of the image is enlarged to show the slow-running DNA bands more clearly. The schematic at the bottom right describes the hypothesis that another stability factor (shown in gold) can interact with the 5′-UTR closer to the initiation region.

I.D.9. *The mitochondrial protein Mtf2p modulates the interaction between Pet111p and the 5′-UTR target.* Mitochondrial gene expression in yeast is believed to be carried out by a ribosome-containing supercomplex termed MIOREX (mitochondrial organization of gene expression) (*53*). In this

complex, the steps of RNA maturation, translation, and degradation are executed in a sequential and spatially confined manner. To accept RNA molecules for processing and translation, the MIOREX complex is coupled, presumably transiently, with the mitochondrial nucleoid via the transcribing RNA polymerase. It is unclear whether MIOREX is a uniform complex capable of expressing all mitochondrial mRNAs or if it represents a collection of gene-specific (or respiratory complex-specific) supercomplexes that are similar in organization and function. Among numerous proteins composing the MIOREX complex, a set of gene-specific translational activators has been identified, including Pet111p (*53*). A body of published data suggests that some of these factors interact with each other in a network, which presumably exists within the MIOREX complex. In addition to Pet111p, these proteins include translational activators (Pet494p, Pet54p, Pet122p, and Pet309p) of the other mitochondrial DNA-encoded complex IV genes (*COX1* and *COX3*). This network was shown to be linked with the small ribosomal subunit through Pet122p (*133-135*) and with the transcription machinery through the Mtf2p/Sls1p subcomplex (*136-138*).

Mtf2p (also called Nam1p) is a soluble matrix protein, required for respiration (*139, 140*). It was originally identified as a multicopy suppressor of intron splicing defects in *COB* and *COX1* transcripts induced by mutations in mtDNA (*91, 140*). Further genetic analysis showed that inactivation of the gene greatly reduced the level of mitochondrial translation and specifically blocked the synthesis of Cox1p, presumably due to a defect in *COX1* splicing

(*141, 142*). Independently, a nuclear DNA point mutation causing a temperature-sensitive petite phenotype that appeared to affect the overall levels of mitochondrial transcripts was mapped inside the *MTF2* gene (*139, 140, 143*). Due to this observation, a role of Mtf2p as a transcription factor was suggested. However, further analysis excluded this role as a general transcription factor for Mtf2p (*144*) and showed that Mtf2p is required for the stability of the *ATP8*/*ATP6* mRNA and to a lesser extent, the 21S rRNA, as well as for processing of the *COB* and *COX1* mRNAs. In addition, Mtf2p was suggested to facilitate movement of newly transcribed mRNA to the proper site of translation at the inner membrane (*145, 146*). Most recently, Mtf2p was shown to associate with translational activators Pet111p, Pet309p, and Pet494p, which control the expression of complex IV mitochondrial genes (*84*).

Taking all of this data together, it appears as though Mtf2p is a part of a link between transcription and translation (Figure 20). Considering that Mtf2p has been characterized as facilitating the movement of mRNA transcripts (*145, 146*) and also genetically and physically interacting with Pet111p (*84*), the following experiments were carried out to probe whether Mtf2p had a direct effect on the formation and/or stability of the Pet111p/*COX2* complex *in vitro*.

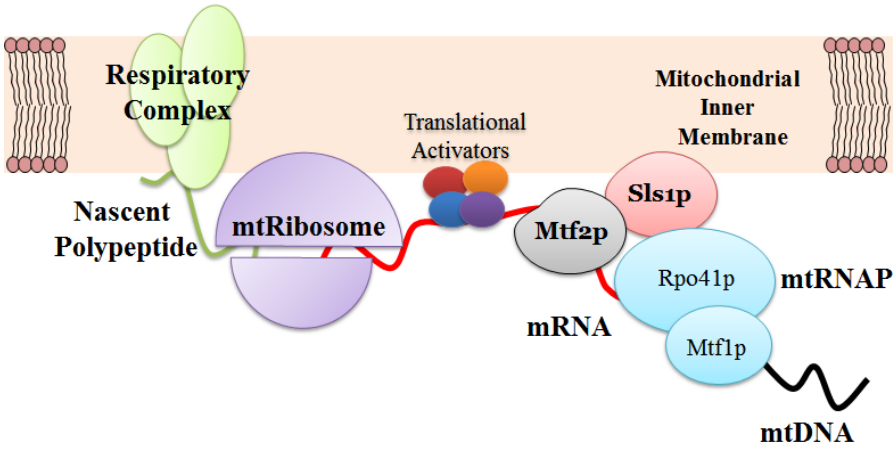

Figure 20. A transcription-coupled translation network at the inner mitochondrial membrane. Previous studies have established a network of genetic and physical interactions among a set of mitochondrial translational activators. In addition, the activators were shown to interact with the mitochondrial RNA polymerase, composed of Rpo41p and Mtf1p, via Mtf2p and Sls1p. These observations led to a model of mitochondrial gene expression, in which translational activators and Mtf2p/Sls1p act as a link between the transcription and translation machineries (*15, 84*).

To investigate if there was a direct and stable interaction between Pet111p and Mtf2p *in vitro*, recombinant Mtf2p needed to be obtained. The full-length coding sequence of the gene was cloned into a pTrcHis-A vector and expression in bacteria was attempted, however the resulting protein was not soluble. Therefore, as was the case with Pet111p, the N-terminus of the mature Mtf2p needed to be identified prior to heterologic protein expression. To do this, a yeast cell line carrying the BG1805-*MTF2* plasmid was grown and tagged Mtf2p was overexpressed and purified from isolated mitochondria under denaturing conditions by nickel affinity chromatography. The protein band corresponding to Mtf2p was subjected to Edman degradation analysis

(Figure 21), which revealed two different sequence reads, IH?IH and LNEAN in a ratio of 1:4. This pattern is indicative of two consecutive processing events. The first cleavage step, performed by MPP, appears to occur before amino acid I16 and is followed by a secondary cleavage after amino acid L23, which is carried out by Oct1p. Overall, the processing results in the mature form of Mtf2p lacking the first 23 N-terminal amino acids. The sequence of the Mtf2p precursor at the site of processing generally agrees with the previously reported consensus (*128, 129*).

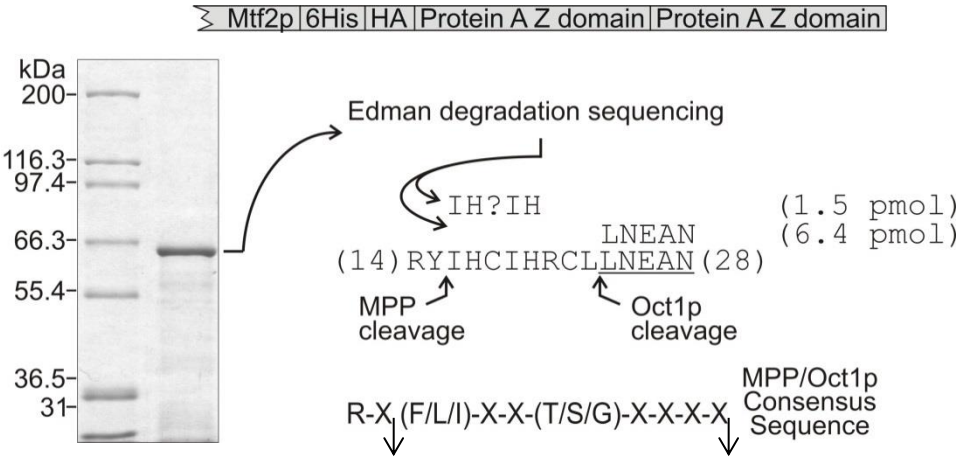

Figure 21. The N-terminus of mature Mtf2p is at amino acid 24. An image of a Coomassie-stained LDS 4-12% polyacrylamide gel showing a preparation of tagged Mtf2p purified under denaturing conditions from yeast mitochondria. The composition of the C-terminally fused purification tag is as shown above. Edman degradation analysis of the protein revealed two sequence reads. The reads are aligned with a portion of the Mtf2p sequence (numbers indicate positions of the flanking amino acid in the sequence of the Mtf2p precursor), in which the determined N-terminal sequence of the mature protein is underlined. The sites of cleavage by MPP and Oct1p are indicated by arrows. The consensus sequence for the MPP/Oct1p processing site is

shown at the bottom (*128*). In this sequence, MPP cleavage occurs at the left downward arrow and Oct1p cleavage occurs at the right downward arrow.

A pTrcHis-A expression vector derivative, carrying the coding sequence of the mature Mtf2p lacking its hydrophobic N-terminal amino acid and fused to a hexahistidine tag was constructed. This plasmid was transformed into *E. coli* XJb(DE3) cells. Low-temperature expression of the protein produced soluble Mtf2p, which was purified to homogeneity by nickel affinity chromatography followed by heparin affinity chromatography and gel filtration (Figure 22). Gel filtration showed that Mtf2p is a dimer as it eluted from the column earlier than a typical 49.4 kDa protein would (see Figure 24 which shows Mtf2p eluted at the same volume as 91 kDa Pet111p).

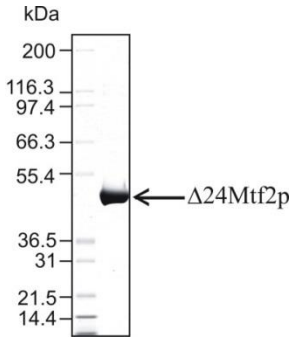

Figure 22. Purification of recombinant Mtf2p. An image of a Coomassie-stained LDS 4-12% polyacrylamide gel showing the purity of a typical preparation of Mtf2p following gel filtration chromatography.

Using purified Mtf2p, the effect of this protein on the RNA binding activity of Pet111p was examined. For this analysis an EMSA was employed in which both proteins were present and two different RNA probes were analyzed; *COX2*-1-54 and *COX2*-1-30 (Figure 23).

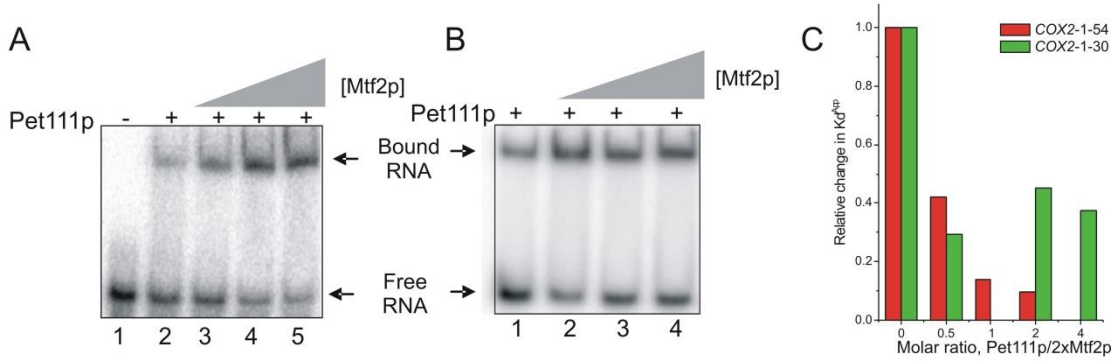

Figure 23. Mtf2p increases the apparent binding affinity of Pet111p for the *COX2*-1-54 and *COX2*-1-30 RNA probes. A. 5′-[^{32}P]-labeled *COX2*-1-54 RNA probe (100 nM) was incubated with 200 nM Pet111p in lanes 2-5 with increasing concentrations of Mtf2p (200, 400, and 800 nM in lanes 3, 4, and 5, respectively). During binding, competitor tRNA was present in the solutions at a 1.2 μM concentration. B. 5′-[^{32}P]-labeled *COX2*-1-30 RNA probe (200 nM) was incubated with 200 nM Pet111p in lanes 1-4 with increasing concentrations of Mtf2p (200, 800, and 1600 nM in lanes 2, 3, and 4, respectively). During binding, competitor tRNA was present in the solutions at a 1 μM concentration. Both sets of reaction products were resolved in 8% polyacrylamide gels under native conditions. The bands corresponding to unbound and bound RNA (as indicated) were visualized by phosphor imaging. C. The relative change in apparent Kd of the Pet111p/RNA complexes induced by Mtf2p is plotted for both RNA probes.

Significantly, this experiment shows that the apparent Kd of the Pet111p/*COX2*-1-54 complex decreased by as much as ten-fold in the presence of Mtf2p and by three-fold for the Pet111p/*COX2*-1-30 complex. No appreciable change in the binding efficiency was observed in control experiments, in which the EMSA was run in the presence of BSA or all of the components present in the Mtf2p storage buffer (data not shown) in the

absence of Mtf2p; implying that the observed decrease was strictly due to the presence of Mtf2p. Importantly, the EMSA did not reveal any signs of an interaction between Mtf2p and RNA in either the presence or absence of competitor tRNA (data not shown).

The observed effect of Mtf2p on the RNA binding activity of Pet111p is consistent with results of previously published yeast two-hybrid studies that showed that the two proteins may interact *in vivo* (*84*). However, EMSA analysis has limited utility when studying protein/protein interactions. The complications arise because the mobility of protein complexes in native gels heavily depends on the charge, size, and shape of the complex. Therefore, from this experiment it was unclear whether Mtf2p was forming a ternary complex with Pet111p and the RNA probes. To further identify if this interaction was occurring *in vitro*, analytical gel filtration was utilized.

Analytical gel filtration is an effective way to determine if two proteins interact. The retention of analytes by gel filtration columns primarily depends on the size of a protein or a complex while their charge does not play a significant role. Therefore, if two proteins interact, the higher molecular weight complex would be expected to elute earlier than either of the proteins alone. Using gel filtration analysis, I attempted to detect the formation of a complex between Pet111p and Mtf2p. As shown in Figure 24A, Pet111p and Mtf2p do not appear to interact under the experimental conditions, as a peak corresponding to the complex was not observed at a lower elution volume.

Instead, a more intense peak appears where both individual proteins co-elute; signifying that they are unlikely to be complexed.

According to my EMSA studies, Mtf2p increased the efficiency of complex formation between Pet111p and RNA probes containing the 5′-UTR binding target of the protein (Figure 23). Therefore, the next step was to examine if the formation of a complex between Pet111p and Mtf2p might be promoted by the presence of a specific RNA probe. Analysis by gel filtration was repeated in similar manner as that shown in Figure 24A; however, the RNA probe *COX2*-1-42 was added to the proteins. This 42nt probe corresponds to the 5′-end of the *COX2* 5′-UTR, which contains the binding target of Pet111p and the predicted downstream stem-loop (*115*). The latter structural feature was incorporated into the probe because Mtf2p has previously been suggested to be a stem-loop RNA binding protein (*145*). Figure 24B shows that Pet111p readily interacted with the RNA probe under the conditions of the experiment, causing the majority of the *COX2*-1-42 peak to elute earlier due to the formation of a complex. However, the RNA peak did not undergo a similar shift in the presence of Mtf2p, indicating that Mtf2p could not interact with this RNA. Once again complex formation between Pet111p and Mtf2p was not observed, even in the presence of the Pet111p binding target.

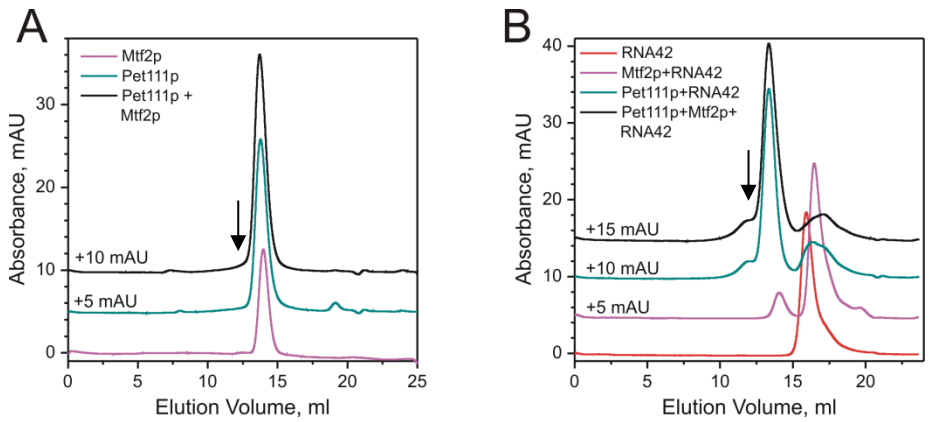

Figure 24. Analytical gel filtration shows that Pet111p and Mtf2p do not interact *in vitro*. Analytical gel filtration elution profiles were recorded for Pet111p, Mtf2p, and a mixture of the two proteins with or without the *COX2*-1-42 RNA probe. A. Mtf2p (purple line) and Pet111p (teal line) both elute at an elution volume of 14 ml, suggesting a dimeric nature of Mtf2p. A mixture of the two proteins (black line) co-eluted at 14 ml, indicating that a Pet111p/Mtf2p complex, which would be expected to elute at a smaller volume (black arrow), did not form. B. A 42 nt RNA (*COX2*-1-42) containing a Pet111p binding target sequence was used to test if the interaction between Pet111p and Mtf2p was RNA-dependent. Free *COX2*-1-42 (red line) eluted from the column at a volume of 17 ml. Taken together, Pet111p and *COX2*-1-42 (teal line) co-eluted as a complex, as indicated by the shift of the *COX2*-1-42 peak to an elution volume of 14 ml. The residual absorbance present at 17 ml was likely due to an excess of the RNA probe. A mixture of Mtf2p and *COX2*-1-42 (purple line) revealed two distinct peaks corresponding to the protein (14 ml) and RNA (17 ml); showing a lack of interaction. The addition of Pet111p to this mixture (black line) shifted the peak corresponding to the RNA to a volume of 14 ml, as expected, but did not lead to the formation of a ternary Pet111p/Mtf2p/*COX2*-1-42 complex, which would be expected to elute at a position indicated by the black arrow.

The analytical gel filtration experiments revealed that there was no interaction between Pet111p and Mtf2p in the absence or presence of *COX2*-1-42. In this experiment, Mtf2p did not interact with RNA at all, therefore characterizing it as a protein that facilitates the movement of mRNAs between the points of transcription and translation (*145, 146*) may not reflect its direct role. Overall, the gel filtration studies were not consistent with a previously reported Mtf2p/Pet111p two-hybrid interaction (*84*). This however, does not account for the clear observation that Mtf2p promoted the formation of complexes between Pet111p and the RNA molecules carrying its 5′-UTR binding target sequence. To examine this phenomenon further, the effect of Mtf2p on the Pet111p footprint was evaluated. Mtf2p was added to an RNase I footprinting assay with the *COX2*-1-30 RNA probe and changes in the Pet111p protection pattern were graphed (Figure 25).

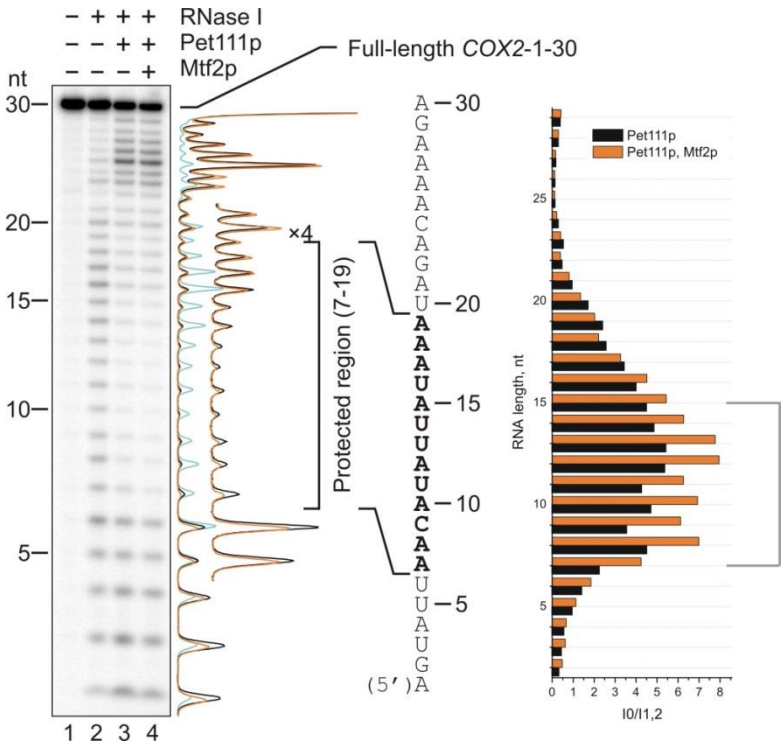

Figure 25. The 5′-UTR target of Pet111p is greater protected from RNase I digestion in the presence of Mtf2p. A phosphor imaging scan of a 7 M urea gradient thickness 20% polyacrylamide gel showing products of RNase digestion. 5′-[^{32}P]-labeled *COX2*-1-30 RNA (0.6 µM) was subjected to a digestion with RNase I (8 U/ml) in the absence or presence of Pet111p (0.8 µM) and Mtf2p (1.6 µM, lane 4). The reactions were carried out in the presence of 1 µM yeast tRNA as a nonspecific competitor. The traces next to the gel image represent the distribution of radioactivity within lanes 2 (blue), 3 (black) and 4 (orange). The bar graph illustrates the changes in the intensity of the individual RNA bands induced by the presence of Pet111p and Pet111p/Mtf2p. For each RNA fragment 2 to 29 nt in length, intensities of the corresponding RNA bands were measured in lanes 2 and 3 (black bars, Pet111p alone) and lanes 2 and 4 (orange bars, Pet111p and Mtf2p) and presented as a ratio of unprotected (without any protein) / protected (with Pet111p or Pet111p/Mtf2p). The grey bracket indicates the region in the footprint in which Mtf2p affected the RNase protection the most.

In the presence of Mtf2p, some of the nucleotides within the footprint were protected to a greater extent than when Pet111p only was present. Moreover, the protected region stayed within the same boundaries (Figure 25). From these observations, consistent with the EMSA data, it appears as though Mtf2p has the ability to strengthen the interaction between Pet111p and its 5′-UTR target.

Yeast two-hybrid studies have shown that Mtf2p interacts with the N-terminus of Rpo41p and this interaction was proposed to aid in coupling transcription to translation (*118*). Mutations in the N-terminal domain of Rpo41p caused various phenotypes; including the loss of *COX1* and *COB* mRNAs (*118, 146*) and an overall decrease in translation (*141*). This was linked to the fact that Mtf2p could not interact with Rpo41p in some of these mutants, as an *mtf2*-inactivating mutant has a similar phenotype (*146*). Interestingly, thermosensitive petite *rpo41* mutants were rescued by overexpression of Mtf2p (*118*). These data suggested that there is a physical and functional interaction between Rpo41p and Mtf2p. In order to test if this interaction occurs *in vitro* and also to determine if Rpo41p's presence is needed for the interaction between Mtf2p and translational activators, such as Pet111p, analytical gel filtration was performed. Utilizing highly purified proteins, an analytical gel filtration experiment showed that under these experimental conditions Rpo41p does not interact with Mtf2p, Pet111p, or both proteins together (Figure 26).

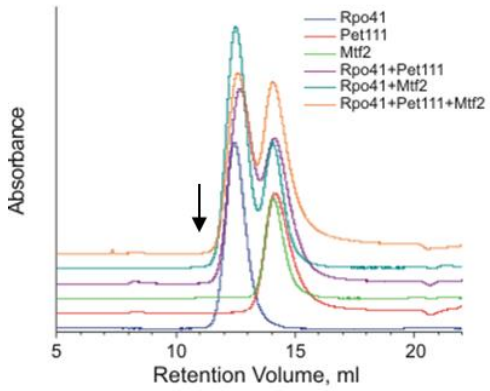

Figure 26. Pet111p, Rpo41p, and Mtf2p do not interact *in vitro*. Utilizing analytical gel filtration the interactions between Pet111p, Mtf2p and Rpo41p were analyzed. Rpo41p (blue line), Pet111p (red line), and Mtf2p (green line) were each run individually. Each of the latter proteins was tested for an interaction with Rpo41p: Rpo41p/Pet111p (purple line) and Rpo41p/Mtf2p (blue line). A mixture of all three proteins was also analyzed (orange line). If complex formation occurred in any of these experiments, a peak corresponding to a complex would have been observed at a lower elution volume indicated by the black arrow.

Despite previously characterized physical interactions between Rpo41p and Mtf2p (*118*) and Mtf2p and Pet111p (*84*) in yeast two-hybrid experiments, I was unable to detect these interactions *in vitro*. My current data suggests that the existing models (*84*) in which Mtf2p functions as a bridge between Rpo41p and translational activators, such as Pet111p, may be oversimplified and other unknown factors may be involved. Despite the lack of a stable direct interaction, Mtf2p can still enhance the apparent binding affinity of Pet111p for its 5′-UTR target.

I.E. Discussion

Seven of the eight yeast mitochondrial gene products are a part of the oxidative phosphorylation (OXPHOS) process and one, Var1p, is a soluble ribosomal protein. These seven proteins are membrane embedded and highly hydrophobic. Sanchirico et al. showed that at least two of them, Cox2p and Cox3p, need to be synthesized close to the membrane. In that study, *COX2*, which encodes a membrane protein of Complex IV, was inserted between the sequences corresponding to the 5′- and 3′-UTRs of *VAR1*. Due to the hydrophilic nature of Var1p, it has been suggested that its synthesis does not require membrane localization and that it is translated within the matrix (*117*). This type of genetic manipulation ensured that *COX2* was under the translational control mechanism of *VAR1*. Metabolic labeling experiments showed that *COX2* was translated in this foreign context. However, there was a lack of accumulation of Cox2p over time. A similar experiment was performed with *COX3* with the same outcome. This demonstrated that the translation of the membrane embedded components of OXPHOS needs to be localized to the matrix side of the inner membrane for immediate insertion (Figure 27); otherwise the translation products are rapidly degraded (*117*).

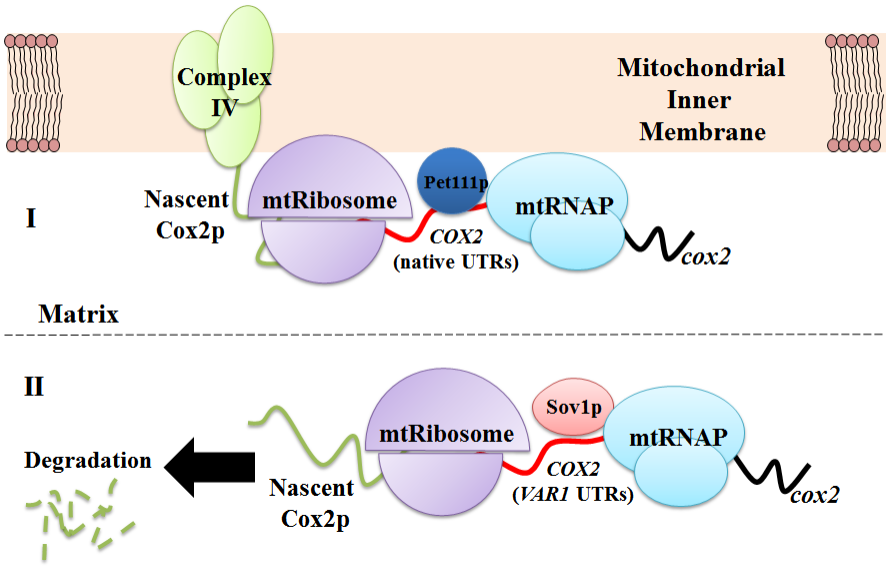

Figure 27. Mitochondrial genome-encoded membrane embedded proteins need to be synthesized at the inner mitochondrial membrane for co-translational insertion. Under normal conditions, Cox2p is synthesized near the mitochondrial membrane and co-translationally inserted (I). When the *COX2* ORF is put under the control of *VAR1*'s translational regulatory elements, the protein is made in the matrix; however, it is soon degraded without making it to the membrane for insertion (II) (*117*). This is most likely due to the lack of an interaction between the mutant mRNA and the membrane-associated translational activator Pet111p which aids in membrane localization. Thereby showing that membrane embedded proteins must be synthesized at the mitochondrial membrane for stability.

How the processes of mitochondrial transcription and translation are localized to the membrane is a current topic of study. These processes are intricate and require multiple protein factors. It has been suggested that these factors are compartmentalized in large complexes and tethered to the membrane (*18, 147*). There have been several groups that have identified different subcomplexes that are associated with transcription, post-

transcriptional processing, and translation within the yeast mitochondria. Krause et al. isolated a 900 kDa complex in which at least two gene-specific translational activators, Cbp1p and Pet309p, both essential to respiration, were present (*54*). This complex was seen as an integrator that links transcription and translation. More recently, the idea of this "integrator complex" was expanded, when Kehrein et al. identified protein components that co-purify with the mitochondrial ribosome (*53*). The study identified over one hundred proteins that are a part of this mega complex, termed the MIOREX (mitochondrial organization of gene expression) complex. Proteins found in MIOREX have functions in DNA stability and maintenance, RNA processing (splicing, maturation, and degradation), translational activation, and translation processes. Therefore, it appears that the majority of the proteins needed for each phase of gene expression are linked together in a large complex. This complex is tethered to the membrane and aids in the localization and proper membrane insertion of the mitochondrial genome-encoded subunits of OXPHOS (*53*).

Within MIOREX, each respiratory complex is thought to have a subcomplex in which all of the necessary elements for post-transcriptional processing, translation, and insertion into the membrane are localized. In addition to these elements, the individual OXPHOS complexes are thought to be assembled as a supercomplex termed the respirasome (*148*). This is a current hypothesis on how the organization of gene expression may work in general; the actual organization of these complexes and mechanistic details

of their function are still unknown. However, one essential constituent is a group of proteins, termed translational activators, which were shown to be essential for the translation of mitochondrial mRNAs.

Each mitochondrial genome-encoded transcript has a specific set of nuclear genome-encoded translational activators that are presumed to have multiple roles. Some of these activators are believed to interact with the 5′-UTR of their respective mRNAs. This interaction is thought to reduce 5′-exonucleolytic degradation, contributing to mRNA stability, and to aid in mRNA localization. Most yeast translational activators were shown to associate with the matrix surface of the inner mitochondrial membrane to enable co-translational membrane insertion of the nascent hydrophobic proteins (*18*). One of these translational activators, Pet111p, is specifically required for the translation of *COX2* mRNA (*107, 108*), which encodes subunit II of the respiratory complex IV. Pet111p was suggested to act by a direct association with the 5′-UTR of the mRNA (*111*). However, this suggestion has not yet been verified experimentally, nor has a mechanism been proposed by which the presumed interaction could lead to the activation of the mRNA's translation.

In this study, I set out to examine the relationship between Pet111p and *COX2* in greater detail. I first wanted to determine if Pet111p could directly associate with *COX2 in vitro* in the absence of any other mitochondrial factors. To answer this question, I utilized highly purified recombinant Pet111p (Figure 10) and tested its interactions with various

synthetic RNA probes *in vitro*. This is the first example of this class of proteins being expressed recombinantly in a soluble form. EMSA analysis showed that Pet111p can specifically bind to a probe comprising the entire sequence of the 5′-UTR in the absence of any other mitochondrial factors (Figure 11A). More specifically, Pet111p interacted within a region encompassing the 30 5'-end proximal nucleotides of the UTR, but not closer to the translation initiation site (Figure 11B).

A previous study alluded that Pet111p may have some type of functional relationship within the first 14 codons of the ORF of *COX2* (*114*). Prompted by this data, I tested if Pet111p could recognize an additional binding target in this region *in vitro*. An EMSA analysis (Figure 15) showed that Pet111p was able to specifically interact with an RNA probe comprising of nucleotides 4-41 of the coding sequence (nucleotides 58-95 in correlation to the transcriptional start site being +1). Therefore, it appears that Pet111p targets at least two distinct binding sites in *COX2*, one close to the 5'-end of the mRNA (5'-UTR target) and the other in the beginning of the coding sequence (ORF target). To define the boundaries of the two binding targets with a greater precision, I carried out RNase footprinting analyses of the Pet111p/RNA complexes. As shown in Figure 12B, the protein protected a 13-nt sequence in an RNA probe carrying the 5'-UTR target. The target mapped between positions 7 and 19 relative to the transcription start site. The region of protection of the probe containing the ORF target was considerably wider, 21 nt, and encompassed the nucleotides between positions 66 and 86

(Figure 15). Not all nucleotides within the mapped boundaries may constitute the corresponding Pet111p recognition signals. The sequences that are actually recognized by the protein are expected to be somewhat shorter than specified by the indicated intervals and may be discontinuous. The finding that Pet111p physically interacts with two distinct regions in *COX2* was very surprising. Significantly, across closely related yeast species both target sequences are highly conserved, as shown in the sequence alignment in Figure 28. This gives insight that the mechanism utilizing the two binding sites that are presumably involved in translational activation by Pet111p may be conserved in these species as well.

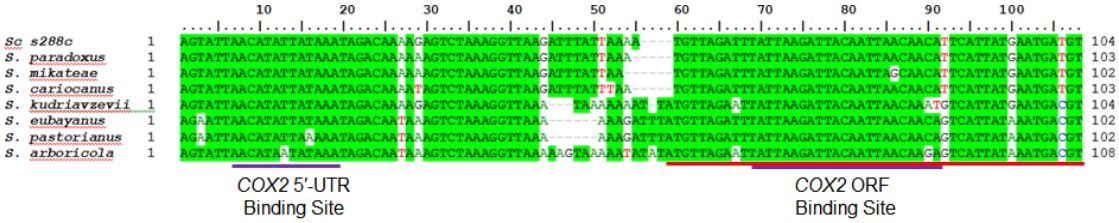

Figure 28. Both Pet111p binding targets are highly conserved across closely related species. Shown is a sequence alignment of the first 100 nucleotides of the *COX2* genes in the indicated yeast species. The two Pet111p binding sites, signified by the purple lines are highly conserved. The red line underlines the sequence corresponding to the ORF. Species and accession numbers for sequences used: *Sc. s288c* (NC_001224), *S. paradoxus* (NC_018044), *S. mikateae* (NC_031185), *S. cariocanus* (NC_031514, *S. kudriavzevii* (NC_031184), *S. eubayanus* (NW_017264706), *S. pastorianus* (NC_031515), and *S. arboricola* (KX657740).

The two identified binding targets of Pet111p vary in length and do not contain an apparent common sequence pattern that would be stringent enough to set these signals apart from the rest of the mitochondrial RNA. It is not obvious how Pet111p can recognize these two diverse targets, but this data attests to the ability of the protein to discriminate in favor of certain sequences in a single-stranded RNA context. One possibility was that Pet111p could harbor two independent RNA recognition sites, one for each target. If this were the case, the protein should be able to associate with both RNA targets simultaneously. However, utilizing an EMSA analysis in which Pet111p was incubated with both RNA probes each harboring one of the two binding targets (Figure 16), it was found that one molecule of the protein could only bind to one of the targets at a time, as no super-shifted species were observed. Therefore, the binding of RNA at the two putative recognition sites in Pet111p would be mutually exclusive, either due to a steric overlap or because some of the elements in the protein are engaged during the recognition of either target. Alternatively, Pet111p may undergo structural reorganization to adapt to certain targets, but not to others, by an induced fit mechanism. In this mode, the same structural elements in the protein may participate in binding to different RNA bases depending on the sequence context of the target and the structural configuration of the corresponding recognition modules in the protein. Whatever the mechanism is, by which Pet111p discriminates in favor of certain single-stranded RNA sequences, a comprehensive mutagenesis analysis of the RNA bases inside the two

identified binding regions is needed. This analysis would determine which bases are critical for the specific interaction to take place and thus constitute the actual recognition target of the protein. Additionally, crystallization studies can be performed with Pet111p and each of the two *COX2* binding targets.

It is currently unclear how the binding of Pet111p to *COX2* at the sites mapped in this study promotes the translation of the mRNA. The protein may be able to modify the secondary structure of the message, facilitate the recruitment of the ribosome at the site of initiation, or stabilize the association of the mRNA with other yet unknown factors needed for efficient translation. A recent study has shown that the mitochondrial ribosome cannot associate with *COX2* in Δ*PET111* cells, whereas the binding to other mRNAs is not significantly affected (*16*). However, the mechanism behind this effect is not known. It has been previously predicted that the 5'-end proximal region of *COX2* is highly structured (*86*) and that it may assume the fold shown in Figure 29A. In this structure, the RNA bases in the vicinity of the translation initiation codon are base paired with upstream RNA; this structuring would prevent the binding of the ribosome at the site of initiation. Remarkably, most of the sequence that forms the duplex with the initiation region overlaps with the 5'-UTR target of Pet111p mapped in this study. Therefore, binding of Pet111p to this target would prevent the duplex from forming, thus making the translation initiation region single-stranded and accessible to the ribosome (Figure 29A).

A comprehensive mutagenesis analysis of the *COX2* 5'-UTR has revealed a 31-nt region, located between positions 9 and 39, where substitutions led to a loss of translation of the mutant mRNA and cellular respiration (*113*). The Pet111p 5'-UTR binding target that was mapped here substantially overlaps with the upstream part of this region (Figure 29A). Curiously, the reported mutants demonstrated two clearly distinct phenotypes. Substitutions downstream from position A23 did not significantly destabilize the mutant mRNAs, but inhibited their translation. Conversely, a drastic decrease in the mRNA levels was added to the translational defects when the mutagenesis involved the nucleotides upstream from position C24. The mutations in this region alter the mapped Pet111p binding target and would be expected to interfere with the binding of the protein. Since the deletion of *PET111* caused a three- to ten-fold decrease in the *COX2* level (*107*), the observed destabilization of the mutant mRNAs is likely due to the loss of protection by Pet111p. The loss of translation resulting from the substitutions (*113*) and deletions (*112*) introduced into the 5'-leader downstream from the Pet111p binding target suggests that a second element may be present there that is required for the efficient translation of *COX2*. This element may function at the RNA level or serve as a binding site for an additional translational activator. Interestingly, I found that binding of Pet111p drastically alters the organization of the downstream RNA *in vitro*, at least in the context of the isolated 5'-UTR (Figure 13). I also showed that no additional *COX2* stability enhancing proteins were bound to the 5'-UTR in the absence of

Pet111p (Figure 19). This indicates that the function of the putative second translation control element may depend on whether or not Pet111p is bound immediately upstream. Taking all of the above considerations into account, it appears that my mapping of the Pet111p 5'-UTR binding target is in a perfect agreement with the results of the previous genetic, *in vivo* biochemistry, and structural prediction analysis of the 5'-leader (*112, 113*).

The finding that Pet111p contacts *COX2* inside the reading frame was somewhat surprising, as none of the mRNA-specific yeast mitochondrial translational activators were previously reported to act outside of the 5'-UTRs. Remarkably, the only known mammalian mitochondrial translational activator, TACO1, has recently been shown to bind selectively to the ORF of *COX1* (*60*). My data thus indicate that a similar phenomenon may also exist in yeast, and regulation of yeast mitochondrial translation may involve binding of activators inside ORFs. Noticeably, translation of a chimeric mRNA, in which the *COX2* ORF was fused to an upstream *COX3* 5'-UTR sequence, was possible in the absence of Pet111p (*112*). Therefore, if the binding of Pet111p to the ORF target mapped here is required for *COX2* translation in the wild type background, this requirement must be dictated by an element present in the *COX2* 5'-UTR and is bypassed when the entire leader is replaced. Previous genetic work has shown that a positively acting translation control element is embedded in the region of *COX2* that encodes the first 14 amino acids of the Cox2p precursor and that the sequence comprising codons 2 to 6 is critical for the function of this element (*114*). The sequence of the codons 7

to 10 was also important for translation when codon 6 was mutated. Remarkably, the ORF binding target of Pet111p that I mapped, which approximately corresponds to codons 5 to 11, is partially overlapping in sequence with this control element (Figure 29B). The reported translational and respiratory defects induced by substitutions and/or deletions within the first 10 codons of *COX2* may therefore result from weakening the interaction between Pet111p and its ORF target. Consistently, these defects were relieved in cells overexpressing Pet111p (*114*). The 5′-proximal region of the *COX2* ORF was suggested to fold into a secondary structure (*114*), in which the nucleotides at the site of translation initiation are base paired with a downstream sequence and thus unavailable for the ribosome to bind (Figure 29B). Noticeably, all of this downstream sequence is included in the ORF binding target of Pet111p. Therefore, similarly to the mechanism suggested in Figure 29A, binding of the protein at the beginning of the *COX2* ORF should be expected to prevent the base pairing of the initiation region, making it available for association with the ribosome (Figure 29B). The two putative structures shown in Figure 29 do not appear to be compatible and thus the mRNA most likely alternates between the two states. Thus, simultaneous binding of two Pet111p molecules at both targets may be necessary to prevent base pairing of the initiation codon region and to enable translation.

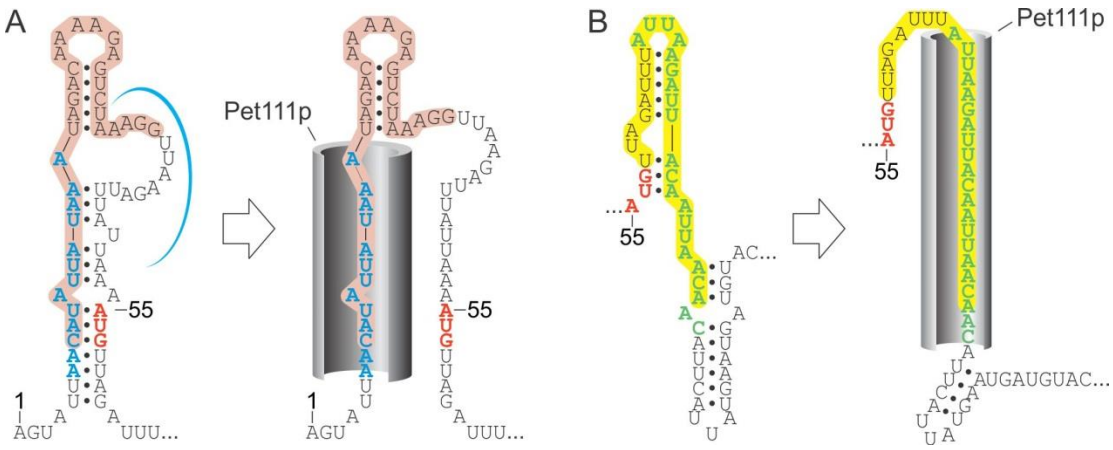

Figure 29. A possible mechanism of activation of *COX2* translation by Pet111p. A. A structure formed by the *COX2* 5′-UTR and a region in the beginning of the reading frame as previously proposed (*86*) and experimentally probed is shown on the left. The 5′-nucleotide (indicated by 1) corresponds to the transcription start site. In this structure, the translational start codon (shown in red) is base paired with a portion of upstream mRNA, most of which overlaps with the Pet111p target region (indicated by blue lettering). The right panel illustrates how binding of Pet111p (grey hollow cylinder) would prevent the RNA structure from forming making the initiation region single-stranded and available for translation. Association of Pet111p with *COX2* may also interfere with the binding of IDH to a specific site (indicated by the blue arc), which was previously suggested to inhibit translation (*90*). The portion of the 5′-UTR highlighted in pink was shown to contain residues required for respiration (*113*). B. A previously suggested alternative structure involving the start codon of *COX2* extending into the reading frame is shown on the left (*114*). In this structure, the translation initiation region is base paired to a downstream sequence, which includes a part of the Pet111p binding target mapped in this study (green lettering). The translation initiation region would remain single-stranded when Pet111p (grey cylinder) is bound to its target. In addition, binding of Pet111p at this site is expected to interfere with a downstream stem-loop structure. The stability of this structure was reported to affect the efficiency of translation of *COX2 in*

vivo (*111*).The initiation codon (shown in red) starts at position 55 relative to the transcription start site as indicated. Deletions or substitutions of specific nucleotides within the sequence highlighted in yellow, either in the WT or a 70A>C;72A>U background, were previously found to disrupt translation *in vivo* (*114*).

The mitochondrial RNA landscape is highly structured and this proposed mechanism gives insight into how gene expression regulatory proteins, such as translational activators, can overcome this innate structuring. Pet111p is known to be in limiting concentrations within the mitochondria (*116*), therefore the need for two Pet111p molecules may give insight into its regulatory role as well. If there are extremely low concentrations of Pet111p present, only one site may be occupied and thereby no translation would occur, whereas if more Pet111p is present, both RNA targets can be engaged with the protein and translation of *COX2* can start. Alternatively *COX2* expression can be regulated in a manner in which only the *COX2* mRNAs that are bound to two Pet111p molecules are translated, thereby limiting the Cox2p product. The timing in which Pet111p interacts with *COX2* is still unknown. Pet111p may interact with the mRNA co-transcriptionally, therefore preventing inhibitory structures from ever forming, or post-transcriptionally, in which case it would have to alter the structure of an already synthesized mRNA to make it suitable for ribosomal access. With this regard, it is remarkable that Pet111p could interact with the *in vitro* synthesized *COX2*-(-59)-61 RNA, which contained both binding targets

(Figure 17). This indicated that Pet111p can bind to a portion of *COX2* once it is synthesized in isolation and thus must be folded. Therefore, it is possible that this interaction may happen in either manner *in vivo* depending on the availability of Pet111p and the need for Cox2p. Future studies with a nascent RNA probe from *in vitro* mtRNAP transcription containing both Pet111p binding targets can be performed to elucidate this mechanism further.

Aside from the mechanism of translational activation suggested above, Pet111p could also offset the action of mitochondrial isocitrate dehydrogenase (IDH). IDH has previously been shown to specifically bind within the 5′-UTRs of all mitochondrial mRNAs (*86-88*). Disruption of *IDH1* leads to a substantial gain in the rate of mitochondrial translation, which is outweighed by accelerated degradation of the translation products. IDH was thus implicated as a repressor that inhibits aberrant translation away from the surface of the mitochondrial membrane (*90*). A consensus IDH binding site was determined to be a variable short stretch of single-stranded RNA adjacent to a conserved 6-bp downstream duplex (*87*). In *COX2*, remarkably, a part of the 5′-UTR target of Pet111p is expected to contribute to the duplex component of the consensus site (Figure 29A). Association of Pet111p with this target would therefore prevent formation of the IDH binding site and eliminate the translational repression of *COX2* by IDH.

Based on bioinformatic analysis, Pet111p has been placed in the pentatricopeptide repeat (PPR) protein family (*103, 149*). This protein family was first identified and is especially numerous in plant organelles (*93*). Most

PPR proteins function in post-transcriptional RNA metabolism by directly associating with predominantly single-stranded RNA targets in a sequence-specific manner (*62*). The basic structural feature of these proteins is the 35-amino acid PPR motif, which is folded into a helix-turn-helix element (*150*). Multiple PPR motifs can stack together in a protein to form the superhelical structures, termed PPR domains, which have the capacity to identify RNA sequences in a single-stranded context (*62*). Plant PPR domains recognize their RNA targets in a modular mode. They utilize a set of certain amino acids at two specific locations within their motifs to identify the bases at a ratio of one PPR motif per one RNA base (see Figure 7). Thus, there is a relationship that links a combination of the two discriminating amino acids in a PPR motif to the RNA base that the motif recognizes. The combinations that allow the identification of all four bases are collectively known as the PPR code (*98*). Only three *S. cerevisiae* proteins, Aep3p, Dmr1p, and Pet309p, were found to contain the sequence motifs that matched the plant PPR pattern (*103*). Of these three mitochondrial proteins, only Dmr1p has been reported to bind RNA directly and in a sequence-specific manner (*151*). Although the *COX1* mRNA has been found to specifically co-immunoprecipitate with Pet309p (*80*), it remains to be determined if this interaction was direct or mediated by other factors. A group of 12 additional PPR proteins has been bioinformatically identified in *S. cerevisiae* (*103*), including Pet111p and a set of other gene-specific activators of mitochondrial translation. The general functional profile of the proteins in this group appears to be consistent with

that of the plant PPR proteins. However, the sequence pattern of their PPR motifs is distinct and represents a separate “yeast” PPR motif signature (*91*). It is therefore not clear whether the yeast PPR proteins utilize the same mechanistic patterns as those in plants. My data provide the first evidence that a protein from the yeast PPR subfamily can recognize single-stranded RNA directly and sequence-specifically. However, it remains to be determined whether this recognition employs a PPR code resembling that of plants or if it is modular in nature. To better understand the mechanistic principles behind the function of yeast PPR proteins, more information on the corresponding protein-RNA interacting pairs is needed. Other gene-specific translational activators that are presumed to selectively interact with the 5'-UTRs of yeast mitochondrial mRNAs have also been identified as PPR proteins (*103*). The approach presented in this work opens a way to map the targets of these proteins with high precision.

Thus far I was able to define how Pet111p potentially activates the translation of *COX2* and although no other protein factors have been implicated to work in conjunction with Pet111p, it is clear that Pet111p cannot act solely. Therefore I next sought to identify other protein components of the *COX2* translation pathway. A review of the current literature led me to consider the Mtf2p protein. Significantly, Mtf2p is tightly associated with MIOREX (*53*). Considering all of the activities proposed for Mtf2p, this localization is fitting. Mtf2p was shown genetically to be required for proper splicing of the *COX1* and *COB* primary transcripts and for ensuring their

accumulation (*142*); as well as for *ATP8/6* mRNA and the 21S rRNA stability (*144*). Intriguingly, Mtf2p appears to physically associate with both Pet111p and the N-terminal domain of Rpo41p (*84, 118*). These interactions lead to the hypothesis that Mtf2p may facilitate the movement of newly made transcripts to the proper translational activators to ensure that translation takes place (*84*).

In this study, the role of Mtf2p in the Cox2p translational pathway was investigated using an *in vitro* approach. Protein/protein interactions between purified Mtf2p and both Rpo41p and Pet111p were analyzed by analytical gel filtration. The Mtf2p and Rpo41p binding experiments showed that these two proteins do not interact under the conditions used here (Figure 26). This finding contradicted the conclusion from a previously published yeast two-hybrid experiment, which detected a physical interaction between the N-terminal domain of Rpo41p and Mtf2p (*118*). To evaluate the possible reasons for this discrepancy, it is important to keep in mind that the yeast two-hybrid experiment is extramitochondrial. Expression of the reporter gene, which is under the control of a Gal4p-regulated promoter, is a nuclear event. Accordingly, the two interacting mitochondrial proteins (in this case, Mtf2p and Rpo41p) fused to the DNA binding and activator domains of Gal4p, must act on the reporter gene promoter in the nucleus and not in their physiological environment, the mitochondria. Consequently, the interaction of the Mtf2p and Rpo41p fusions in the nucleus was not dependent on any other mitochondrial factors in the two-hybrid experiment. Taking this into account, the likely

reasons for the observed discrepancy between my data and the previously published two-hybrid study could be a missing post-translational modification in the gel filtrations studies that was present in the two-hybrid study or the presence of a nonspecific factor that enhanced the two-hybrid interaction, such as nuclear DNA or RNA but was absent from my *in vitro* study. Nevertheless, a possibility that the yeast two-hybrid study may have given a false positive hit should be considered. Finally, Pet111p was added to this experiment to see if a translational activator could facilitate an interaction between Rpo41p and Mtf2p. However, there was still no interaction when all three proteins were present (Figure 26).

The interaction between Pet111p and Mtf2p was also probed via analytical gel filtration (Figure 24A). Despite a previous yeast two-hybrid experiment showing this interaction (*84*), these two proteins did not interact in the gel filtration experiments carried out in the absence or presence of an RNA containing the Pet111p's 5′-UTR binding target sequence (Figure 24B). This lack of an interaction may be for the same reasons as mentioned above. However very significantly, Mtf2p had a clear effect on the affinity of Pet111p to its 5′-UTR binding target. This was evident from the observation that the apparent binding affinity of Pet111p to either *COX2*-1-54 or *COX2*-1-30 increased by ten-fold and three-fold, respectively in EMSA experiments when Mtf2p was added (Figure 23). Consistently, the efficiency of protection of the *COX2*-1-30 RNA by Pet111p improved in the presence of Mtf2p in the RNase I footprinting experiment (Figure 25). Remarkably, only a portion of the

Pet111p-protected region saw an increase in protection, which was most substantial at the 5′-end of the footprint, between nucleotides 7-16. These findings are very significant, especially since Mtf2p and Pet111p do not form a stable complex, according to the analysis by gel filtration. Taken together, these data therefore suggest that Mtf2p has the ability to strengthen the interaction of Pet111p with its RNA target in a transient manner. It may act as a loading protein that helps the Pet111p/RNA complex reach its lowest energy conformation perhaps by rearranging the RNA and/or the protein's structure to ultimately stabilize the interaction. Future experiments with Mtf2p and Pet111p must be performed to determine how this effect occurs. An example could be to use a zero length cross-linker to determine if these proteins physically interact with one another.

Although the set of experiments presented here do not establish that Rpo41p, Mtf2p, and/or Pet111p can form stable complexes *in vitro*, many other observations were made that can lead to new avenues to study. Mtf2p is believed to be a facilitator in the movement of nascent mRNA to translational activators (*145*). However, a very interesting observation was that Mtf2p was purified via heparin affinity chromatography, showing a preference for nucleotides, yet it was unable to interact with RNA, even non-specifically. This may indicate that Mtf2p has a preference for DNA instead. If Mtf2p binds DNA, it may interact with Rpo41p in a DNA-dependent manner. This would open up a whole new way to think about and reevaluate the previously published data on genetic interactions of Mtf2p. Another intriguing

result was that Mtf2p only strengthened the RNA protection by Pet111p within a particular region of the footprint. The influence of Mtf2p on the interaction of Pet111p with the second binding target inside the ORF of *COX2* has yet to be determined. This target gave a wider region protected by Pet111p; therefore it would be interesting to see if Mtf2p might narrow the nucleotide interval.

Finally, the *in vitro* approach described in this work to analyze the bridge between mitochondrial transcription and translation can be expanded to include other protein factors previously linked to this mRNA pathway. This may ultimately help define the subcomplex that aids in the transfer from the transcription machinery to the activation of *COX2* mRNA translation. The most likely candidate for this subcomplex is Sls1p (Figure 20), which has been shown by multiple studies to physically interact with Mtf2p under various experimental conditions (*136-138*). Sls1p is an integral membrane protein that is required for the protein assembly of respiratory complexes III and IV, and is therefore essential for respiration (*146*). Genetic studies have also shown a link between Sls1p and Mtf2p and that the two proteins act in the same RNA handling pathway without being functionally redundant (*146*). Remarkably, there have been indications of an interaction between Sls1p and the components of the transcription machinery. Sls1p was reported to co-purify with Rpo41p (*152*) and also with the mitochondrial nucleoid in an Rpo41p-dependent fashion (*141*). Consistently, a respiratory deficient phenotype resulting from a point mutation in the N-terminal domain of Rpo41p could be suppressed by the overexpression of Sls1p (*146*). Also of importance were

the observations that overexpression of Sls1p could alleviate mitochondrial translation and respiratory defects in an *mtf2* null mutant (*141, 146*), rescue the loss of the *COB* and *COX1* mRNAs caused by a mutation in Rpo41p (*146*), and partially restore the synthesis of Cox2p in the background of multiple Rpo41p truncation and point mutations (*141*). Taken together, these data strongly suggest that Sls1p may closely interact with Mtf2p, both physically and functionally, in facilitating the transfer of mitochondrial mRNAs between the phases of transcription and translation. Therefore, it is worth determining if the two proteins can interact *in vitro* and if Sls1p can stabilize interactions between Mtf2p and Rpo41p and/or between Mtf2p and translational activators such as Pet111p.

The work presented here introduced a mechanistic background into how Pet111p is able to activate the translation of *COX2*. I also probed into the Pet111p/*COX2* interactome to begin to explore what other mitochondrial factors are a part of this gene expression pathway. This work has greatly advanced what is known about Pet111p and also highlighted future studies that will ultimately characterize this system to a fuller extent.

I.F. Supplemental Information

Table 4. Sequences of the oligonucleotides used in the study.

Oligonucleotide	Sequence (5′ to 3′)
DNA oligonucleotides	
Pet111_Nco	CCATGGCACATCATCACCATCATCATTTACAACGGAGATTTATATCCTC
Pet111_Xho	CTCGAGTTATTACTCCTCCTCCTTTTTATTCTCTTC
Pet111_del34	GCACATCATCACCATCATCATTCAACTGAGTTGATCAAAAAAAAGC
Pet111_del34c	GCTTTTTTTTGATCAACTCAGTTGAATGATGATGGTGATGATGTGC
BG1805Pet111_fix	GGTGATTTCGAATAGTTTGAAGGAGGGCATTGCGCC
BG1805Pet111_fixc	GGCGCAATGCCCTCCTTCAAACTATTCGAAATCACC
COX2-96-76 Lab name: “COX +22- +43”	CATAATGAATGTTGTTAATTG
COX2R	TTGGTACATCATTCATAATG
COX2-(-61)-(-39) Lab name: “COX2 -39- -61”	TAAAAGTAGTATTAACATATTA
Pet111Bg1805-ZZdomain	ATGTTCCTGATTATGCTGGTTAATGATTAGAAGTTCTTTTTCAAGG
RevCom Pet111Bg1805-ZZdomain	CCTTGAAAAAGAACTTCTAATCATTAACCAGCATAATCAGGAACAT
FCOX2 5′-UTR -59 to -33 +SP6	AATAATTAGGTGACACTATAGAAGTAGTATTAACATATTATAAAT
R COX2 +61 to +41	AAGGTGTTGGTACATCATTC
FCOX2 5′-UTR -59 to -33 SP6	AATAATTAGGTGACACTATAGAAGTAGTATTAACATATTATAAATAG
R COX2 +61 to +41	AAGGTGTTGGTACATCATTC
MA58	CTGATCCGGATCCATGATCAGAACATCATCTATATTAAAAAACTG
MA59	CTGATCCGAGCTCCTACCTGTCTCTAGTTAAATTTTTTTTGAGGC
EE1	GTTCTCATCATCATCATCATCATAATGAAGCCAATTTGAAAGATAG

EE2	CTATCTTTCAAATTGGCTTCATTATGATGATGATGATGAGAAC
RNA oligonucleotides	
COX2-1-54	AGUAUUAACAUAUUAUAAAUAGACAAAAGAGUCUAAAGGUUAAGAUUUAUUAAA
NSC	AAUAAUUAUAAUAAUAUUCUUAAAUAUAAUAAAGAUAUAGAUUUAUAUUCUAUU
COX2-1-30	AGUAUUAACAUAUUAUAAAUAGACAAAAGA
COX2-23-54	CAAAAGAGUCUAAAGGUUAAGAUUUAUUAAA
COX2-58-95	UUAGAUUUAUUAAGAUUACAAUUAACAACAUUCAUUAU
COX2-1-42	AGUAUUAACAUAUUAUAAAUAGACAAAAGAGUCUAAAGGUUA
COX2-(-59)-(61)	GAAGUAGUAUUAACAUAUUAUAAAUAGACAAAAGAGUCUAAAGGUUAAGAUUUAUUAAAAUGUUAGAUUUAUUAAGAUUACAAUUAACAACAUUCAUUAUGAAUGAUGUACCAACACCUU

CHAPTER II

INSIGHT INTO TRANSLATIONAL ACTIVATION OF *COB* BY CBP1P

II.A. Abstract

Genetic studies have implicated the nuclear genome-encoded Cbp1p as a translational activator for the mitochondrial genome-encoded *COB.* In the absence of Cbp1p, Cobp is not synthesized due to the loss of *COB* mRNA. Therefore Cbp1p is thought to act as an mRNA-specific stability factor. Genetic studies have suggested that Cbp1p performs this function by directly interacting with the 5′-UTR of *COB*. However, this interaction has never been demonstrated nor has a specific region in *COB* been persuasively shown to harbor a target for this interaction.

This work sets forth to define the roles that Cbp1p plays as a translational activator by studying the biochemical properties of this protein. An expression and purification protocol has been developed to provide a soluble and active Cbp1p. RNA probes spanning the 951 nt 5′-UTR of *COB* were made and tested for the ability to form complexes with Cbp1p *in vitro*. It was found that Cbp1p can indeed interact with RNA at the 5′-end proximal region of the *COB* 5′-UTR.

II.B. Introduction

The mitochondrial *COB* gene encodes for the apocytochrome *b* protein, which undergoes a post-translational modification to become a part of complex III of the respiratory chain. The primary transcript of this gene is heavily processed, including excision of $tRNA^{Glu}$, removal of introns, and 5′-end maturation (see Figure 30A) (*27, 153*). In an attempt to identify whether a protein was involved in the maturation and/or the translation of *COB*, Dieckmann et al. scanned multiple mutants deficient for Cobp expression. Utilizing blot hybridization techniques, this study identified a complex pattern of *COB* maturation intermediates differing from those in wild type cells. The gene associated with this phenotype was designated *CBP1* (*154*). DNA sequencing of the fragment revealed a single reading frame encoding Cbp1p (*155*).

Since the identification of Cbp1p, it has been linked to multiple roles in the production of Cobp. Original blot hybridization studies showed that Cbp1p may be involved in the RNA maturation process, as the mature form of *COB* did not accumulate when Cbp1p was not present and intermediates at various points of maturation were observed. These intermediates were prone to nucleolytic degradation (*156*). Subsequent studies by Staples et al., showed that when *cbp1* was mutated the levels of $tRNA^{Glu}$ were the same as the wild type, pre-*COB* was approximately 20% of the wild type level, and the mature *COB* message was not detectable (*157*). Based on these results, Cbp1p was linked to both the maturation and the stability of the *COB* precursor.

Cbp1p has also been shown to affect the translation of *COB (45)*. It has been genetically established that trimming of the 5′-end during maturation of *COB* and several other mitochondrial mRNAs relies on the *PET127* gene (*43*). In order to further characterize the roles of Cbp1p and Pet127p in *COB* maturation, Islas-Osuna et al. employed Δ*PET127* cells to show that while *COB* was not fully maturated in this background, a pre-*COB* transcript could still be translated. Interestingly, translation did not occur in a double mutant Δ*CBP1:*Δ*PET127*, although the pre-*COB* mRNA was present. This showed that in addition to its role in maintaining the stability of *COB*, Cbp1p is also required for the activation of translation of the mRNA (*45*).

Further genetic manipulations and analyses showed that Cbp1p function is dependent upon sequences within the 5′-UTR of *COB*. Utilizing hybrid and tribrid gene constructs composed of sequences from *ATP9* and *COB*, primary transcripts were created with different mutations in the 5′-UTR of *COB (158).* These manipulations and growth in different backgrounds (wild type or Δ*CBP1)* delimited Cbp1p's functionality to the 5′-end proximal third of the 5′-UTR (*159*). Examination of the mutants further defined a 63 nt sequence that contained the maturation point at nucleotide -951/-950 (numbering used in this chapter is transposed from the LL20 strain that was used in previously published studies (*160*) to the S228c reference sequence) to be sufficient for Cbp1p function (*160*). Genetic manipulations within this 63 nt sequence identified an 11 nt region that was needed for two Cbp1p-dependent functions. One was positioning the maturation cleavage site at

nucleotide -951/-950 and the second was providing stability for *COB* after the processing (*161*). Finally, site-directed mutagenesis within the 11 nt region refined an essential CCG sequence (positions -940 to -938) that was required for the stability of the *COB* transcript. This CCG element is presumed to be in the loop of a predicted stem-loop structure. In this structure, the specific nucleotide sequence in the loop is required for the putative interaction with Cbp1p while the composition of the stem is not as important (*66*). Mutations of any one of the three CCG residues led to a Δ*CBP1* phenotype. From this study it was concluded that Cbp1p recognizes and binds directly to a target containing the CCG sequence and that this interaction protects *COB* from degradation (*66*).

Although the involvement of Cbp1p in the maturation, stability, and translation of *COB* has been thoroughly characterized genetically; it is not the only nuclear gene product essential for Cobp production. Other translational activators have been defined, some of which have been proposed to interact with the 5′-UTR (Figure 30B), while others have been suggested to function post-translationally. Mutations in the nuclear genes *cbs1* and *cbs2* were found to cause cells to be deficient in the *COB* transcript (*63*). It was suggested that Cbs1p and Cbs2p interact with *COB* at a region just upstream of the initiating AUG (*162*). Association of these factors with the inner membrane (*163*) and with the ribosome (*64, 65*) is thought to provide a physical link between the mRNA and the translational machinery and to aid in membrane localization of the translation event. Overexpression of a different nuclear genome-encoded

factor, Cbt1p, was found to rescue a respiratory deficient phenotype in a strain carrying a mutation in the CCG element, leading to a suggestion that Cbt1p facilitates the interaction of Cbp1p with its CCG element target (*68*). Consistently, a significant inhibition of pre-*COB* maturation and destabilization of the processed *COB* was observed in Δ*CBT1* cells (*67, 68*). This signified that Cbt1p may act with Cbp1p to promote processing and to ensure the stability of the mRNA. A direct genetic link between the two factors was also evident from the observation that overexpression of Cbt1p suppressed a *cbp1*-associated temperature sensitive petite phenotype (*69*). Lastly, Cbp3p and Cbp6p have been found to form a heterodimer that associates with the ribosome and the nascent Cobp promoting translation (*70, 71*). Upon emergence of Cobp, the Cbp3p/Cbp6p dimer remains bound to Cobp and then forms a complex with the Cbp4p assembly factor, which aids in proper assembly into the cytochrome bc1 complex (*72*).

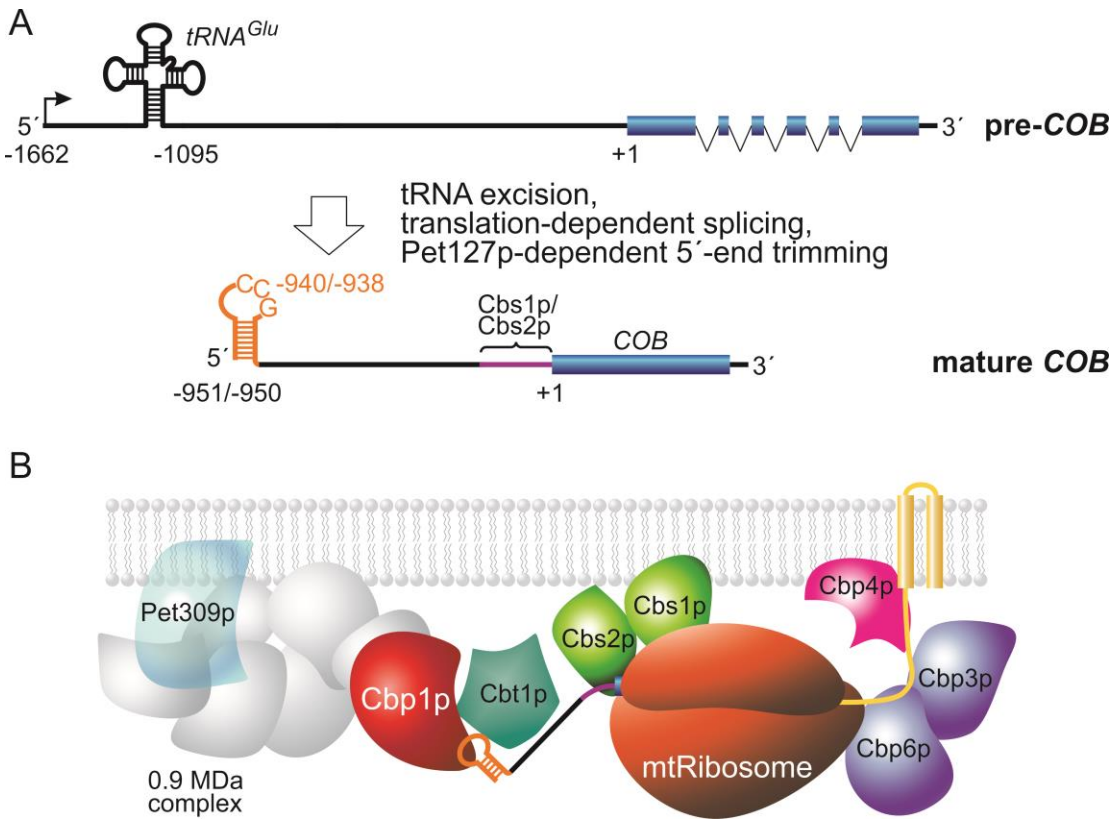

Figure 30. A schematic of post-transcriptional processing and translation of *COB*. A. The pre-*COB* mRNA is co-transcribed with $tRNA^{Glu}$, which is excised by RNase P and RNase Z to leave a 5′-end at position -1095 (*153*). The resulting *COB* intermediate is further matured by excision of five introns and a Pet127p-dependent trimming of the 5′-end region to position -951/-950 (*44*). Translation of the *COB* reading frame (blue cylinder) depends on at least two elements in the 5′-UTR. First, the 5′-end proximal CCG element (shown in orange) was suggested to interact with Cbp1p (*66*). Secondly, a sequence just upstream from the reading frame (magenta) is a presumed site of binding for the Cbs1p and Cbs2p translational activators (*162*). B. Cbp1p (red) is part of a large membrane-associated protein complex (grey) that also includes a *COX1* translational activator, Pet309p (light blue) (*54*). Cbp1p attracts *COB* to the matrix side of the inner membrane by association with the CCG element (orange stem-loop structure) (*66*). The interaction between Cbp1p and its CCG target appears to be stabilized by

Cbt1p (teal) (*68*). Two additional factors, Mtf2p and Sls1p (not pictured), have been shown to be essential for processing and/or the stability of *COB* (*118, 141, 144*). Another group of factors that ensures the membrane localization of *COB* translation includes translational activators Cbs1p and Cbs2p (green) (*63, 110*). Post-translation, the ribosome-associated Cbp3p/Cbp6p complex (purple) acts in accord with Cbp4p to aid the assembly complex III (*70-73*).

The roles that Cbp1p plays in *COB* stability and translation have been greatly characterized at the genetic level. However, to date there have not been any biochemical studies to verify that Cbp1p interacts directly with the 5′-UTR of *COB* and if this interaction relies solely on Cbp1p and no other mitochondrial factors, such as Cbt1p. To further elucidate the functions of Cbp1p I utilized an *in vitro* biochemical system, in which recombinant Cbp1p and RNA fragments representing the *COB* 5′-UTR were studied. Using this system, the aims of this study were to determine if Cbp1p can directly interact with a target in *COB* and if so, to identify the region(s) of the interactions. This study will further the knowledge of translational activation of *COB* and specifically characterize how Cbp1p aids in this process for this transcript.

II.C. Materials and Methods

II.C.1. *Oligonucleotides*. Synthetic DNA oligonucleotides and synthetic RNA were purchased from Integrated DNA Technology (IDT). Sequences of all oligonucleotides used in this portion of the study are listed in Supplemental Information Table 6. Supplemental Information Table 7 describes the positioning of each oligonucleotide in correspondence to the translational start site of *COB* being +1. The RNA probes used in the EMSA and RNase footprinting experiments were either 5′-labeled with T4 PNK kinase (New England Biolabs) in the presence of γ-[^{32}P]-ATP (3000 Ci/mmol, Perkin Elmer) or for *in vitro* transcribed RNA, labeled randomly at U resides by incorporation of α-[^{32}P]-UTP (3000Ci/mmol, Perkin Elmer) during transcription. Gel purification of the RNA, where specified, was performed as in section I.C.1.

II.C.2. *Identification of the N-terminus of mature Cbp1p.* Yeast Y258 cells carrying plasmid BG1805-*CBP1* were purchased from GE Healthcare Open Biosystems. The plasmid encoded Cbp1p fused at the C-terminus to a hexahistidine-containing purification tag as explained in Figure 31. After several unsuccessful attempts to purify the tagged Cbp1p from the mitochondrial fraction of the commercial cells, the plasmid was isolated and analyzed by DNA sequencing. It was determined that the plasmid contained *CDC25* instead of the desired *CBP1*. To construct the correct BG1805-*CBP1* plasmid, *CBP1* was PCR-amplified from yeast genomic DNA using primers

Cbp1-BG-5 and Cbp1-BG-3. The amplicon was ligated into a pT7Blue vector (Novagen) to give plasmid pAC7. The fragment containing the *CBP1* gene was then excised from pAC7 by *Bsr*GI digestion (New England Biolabs), purified from a 0.8% agarose gel, and ligated with the BG1805 vector, which was prepared from the BG1805-*CDC25* plasmid by digestion with *Bsr*GI, dephosphorylation, and agarose gel purification. The integrity of the target coding sequence in the resulting BG1805-*CBP1* plasmid was confirmed by DNA sequencing. BY4743 yeast cells containing BG1805-*CBP1* were grown from a single colony at 30 °C to an OD_{600} of 9 in 3 L of SD-ura medium (yeast nitrogen base with ammonium sulfate and synthetic drop-out mixture -uracil, US Biological), 2% galactose, and 5 µg/ml tetracycline. Mitochondria were then isolated as described in section I.C.2, lysed under denaturing conditions, and the tagged Cbp1p was partially purified using Ni-affinity beads as described in section I.C.3. The proteins in the preparation were further resolved by SDS-PAGE, transferred onto an Immobilon-P^{SQ} PVDF membrane (EMD Millipore) by electroblotting, and the band corresponding to Cbp1p (apparent molecular weight ~75 kDa) was cut from the membrane and submitted for N-terminal protein sequencing (Midwest Analytical).

II.C.3. *Cloning, expression, and purification of Cbp1p.* The Cbp1p coding sequence was amplified by PCR using primers Cbp1Trc-5 and Cbp1Trc-3 and pAC7 as a template. The amplicon was digested with *Nco*I and *Xho*I and ligated into a pTrcHis-A expression vector (Life Technologies) cut with the

same restriction enzymes. The resulting plasmid, pAC9, encoded the full-length Cbp1p fused to an N-terminal histidine purification tag (MAH_6). Two derivatives of pAC9 were then generated by site-directed mutagenesis using the QuikChange II kit (Agilent Technologies) and mutagenesis primer pairs: Cbp1-del25/Cbp1-del25c and Cbp1-del33/Cbp1-del33c. The mutagenesis resulted in the deletion of sequences encoding 25 and 33 N-terminal amino acids of Cbp1p producing plasmids pAC20 and pAC21, respectively. The integrity of the target gene in both pAC20 and pAC21 plasmids was confirmed by DNA sequencing (GENEWIZ).

E. coli XJb(DE3) cells (Zymo Research) carrying pAC21 were grown at 37 °C in 2 L of LB medium supplemented with 3 mM L-arabinose, 1 mM $MgCl_2$, and 100 µg/ml ampicillin to OD_{600} 0.5. The culture was then cooled to 16 °C, expression of Cbp1p was induced with 0.2 mM IPTG, and the cells were incubated at 16 °C for an additional 20 hrs. Cells were collected by centrifugation and resuspended in a lysis buffer containing 100 mM Tris·HCl pH 7.5, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, and 1 mM PMSF at a density of 1 g wet cell pellet per 10 ml buffer. To induce cell lysis, the suspension was frozen overnight at -80 °C and then allowed to thaw in the presence of fresh 1 mM PMSF, 2.5 mM pepstatin, and 3 mM benzamidine. The lysate was treated with DNase I and prepared for affinity purification as in section I.C.4. The cleared lysate was passed through a column packed with 2.5 ml of Ni-IDA agarose beads (Gold Biotechnology) equilibrated with wash buffer (100 mM Tris·HCl pH 7.5, 300 mM NaCl, 10 mM

imidazole). The beads were washed by passing 30 ml of high salt wash buffer through the column (100 mM Tris·HCl pH 7.5, 900 mM NaCl, 10 mM imidazole) and the protein was released with four sequential elutions of 2 ml each in elution buffer (100 mM Tris·HCl pH 7.5, 900 mM NaCl, 10% glycerol, 225 mM imidazole, 5 mM β-mercaptoethanol). The eluate in fractions 2 and 3 was combined, diluted two-fold with a buffer containing 100 mM Tris·HCl pH 7.5, 10% glycerol, and 5 mM β-mercaptoethanol to reduce the concentration of NaCl to 450 mM, and loaded on a 16 × 25 mm HiTrap Heparin HP column (GE Healthcare Life Sciences). The protein was eluted from the column with a linear NaCl concentration gradient (250 mM to 1 M) in a buffer containing 100 mM Tris·HCl pH 7.5, 10% glycerol, and 5 mM β-mercaptoethanol. Fractions containing Cbp1p were pooled, concentrated on an Ultracel®-50K centrifugal filter (Merck Millipore) to a volume of 500 μl. The concentrate was diluted two-fold with glycerol, distributed into small aliquots, and stored frozen at -80 °C. A typical yield of purified Cbp1p was 2 mg.

II.C.4. *Generation of RNA probes.* The *COB* CCG 27 RNA probe was chemically synthesized, while the others were synthesized via *in vitro* transcription by SP6 RNA polymerase. A region of yeast mitochondrial DNA including the 3′-end of the $tRNA^{Glu}$, the whole 5′-UTR of pre-*COB*, and a portion of the *COB* ORF was cloned into a pT7Blue vector (Novagen) to generate the pJJ05 plasmid (see Appendix section A.1). Using pJJ05 as a template, several specific DNA regions were PCR amplified to produce DNA

templates for *in vitro* transcription. The amplification reactions were performed in a total volume of 250 μl (50 μl per tube) using Phusion DNA polymerase (NEB) and generally following the manufacturer's protocol. To optimize the amplification, the temperature during the annealing and extension steps was lowered to 53 °C and 55 °C, respectively, and the elongation time was extended to 2 min due to the high A/T content of the amplified sequences. The DNA templates were amplified to contain the sequence of the SP6 RNA polymerase P2 promoter (5′- AATAATTAGGTGACACTATAG -3′) (*123*), which was added at the 5′-ends of the forward amplification primers. The primers used to amplify the DNA templates and corresponding RNA probes are listed below in Table 5. Due to the requirements of the promoter, a single GMP residue was added at the 5′-end of each RNA probe, except for RNA1 in which a GMP followed by an AMP residue was added, all of which are not present in the *COB* sequence. The amplified DNA fragments were then purified by the GeneJET PCR purification kit (Thermo Scientific).

Table 5. Primers used to amplify *COB* DNA templates for *in vitro* transcription of corresponding RNA probes.

DNA template	Amplification primers (forward / reverse)	Transcription product	Location of RNA in pre-*COB* relative to the translation start site (+1)
A	F *COB* SP6 1-29 / R *COB* 200-171	RNA1	-1096 to -896
B	F *COB* SP6 131-158 / R *COB* 414-392	RNA2	-965 to -682
C	F *COB* SP6 340-369 / R *COB* 594-568	RNA3	-756 to -502
D	F *COB* SP6 502-535 / R *COB* 733-707	RNA4	-594 to -363
E	F *COB* SP6 652-683 / R *COB* 890-864	RNA5	-444 to -206
F	F *COB* SP6 816-841 / R *COB* 1021-992	RNA6	-280 to -75
G	F *COB* SP6 907-934 / R *COB* 1190-1166	RNA7	-189 to 95
H	F *COB* SP6 131-158 / R *COB* 200-171	RNA8	-965 to -896
I	F *COB* SP6 340-369 / R *COB* 414-392	RNA9	-756 to -682

In vitro transcription was performed using purified recombinant SP6 RNA polymerase (see Appendix A. 3). The reactions were carried out for 45 min at 41 °C in a volume of 20 μl of transcription buffer (20 mM Tris·HCl pH 7.9, 15 mM $MgCl_2$, 2 μg/μl BSA, 5 mM β-mercaptoethanol). Each reaction contained 75 nM DNA template, 500 nM SP6 RNA polymerase, 500 μM ATP, 500 μM GTP, 300 μM CTP, 300 μM UTP, and 0.25 μM [α-^{32}P]-UTP (Perkin Elmer). The products of transcription were separated by 7 M urea PAGE in an 8% gel (19:1 acrylamide:bisacrylamide). The radioactive RNA species were

visualized by phosphor imaging with a Typhoon 9410 scanner (GE Healthcare).

II.C.5. *Electrophoretic mobility shift assays*. Taken at the concentrations specified in individual figure legends, Cbp1p was combined with 5′-[^{32}P]-labeled RNA probes or probes labeled throughout with [α-^{32}P]-UTP in 10 μl of binding buffer (40 mM Tris·HCl pH 7.2, 50 mM NaCl, 10 mM $MgCl_2$, 5% glycerol). Yeast tRNA (Sigma) was added at concentrations indicated in the figure legends to reduce nonspecific binding and the mixtures were incubated for 15 min at 30 °C. The mixtures were then supplemented with 2 μl of 30% glycerol spiked with xylene cyanol and bromophenol blue, and loaded on 6 or 7% (as specified in figure legends; 37.5:1 acrylamide:bisacrylamide) native polyacrylamide gels cast in the presence of 1 × TBE buffer using the Mini-PROTEAN® Tetra Cell System (Bio-Rad). Electrophoresis was performed at room temperature in 0.5 × TBE running buffer for 20 min at 175 V. Radioactive bands corresponding to protein/RNA complexes and unbound RNA were visualized using storage phosphor screens and a Typhoon 9410 scanner (GE Healthcare).

II.D. Results

II.D.1. *The N-terminus of mature Cbp1p is at amino acid 26.* To examine if Cbp1p is engaged in direct contacts with *COB in vitro*, the Cbp1p protein was heterologically expressed and purified. As discussed in section I.D.2, nuclear genome-encoded mitochondrial proteins generally carry a mitochondrial import signal that is cleaved from the N-terminus of the protein upon entry into the mitochondria. There is an initial cleavage by mitochondrial processing peptidase (MPP), which may be the only processing event. However, a subset of proteins undergo a secondary cleavage by either octapeptidyl aminopeptidase (Oct1p) or intermediate cleavage peptidase 55 (Icp55p) (*125, 128*). After cleavage, a protein is considered mature and able to assume a correct folding and perform its proper role(s) within the mitochondria (*129*). In order to study Cbp1p in an *in vitro* biochemical system, it is preferable to use the mature protein, as this will ensure that this system is as close to native as reasonably achievable. However, the processing site of Cbp1p has yet to be experimentally established.

Using the BG1805-*CBP1* plasmid encoding the full-length Cbp1p, the mature N-terminus was identified. In this plasmid, the reading frame of *CBP1* was followed in-frame by a sequence that encoded a trifunctional purification tag. This tag included a hexahistidine segment that enabled metal affinity purification of the protein. Tagged Cbp1p was overexpressed in yeast, the mitochondrial fraction was isolated, and the processed form of the protein

was partially purified using Ni-IDA beads under denaturing conditions (Figure 31).

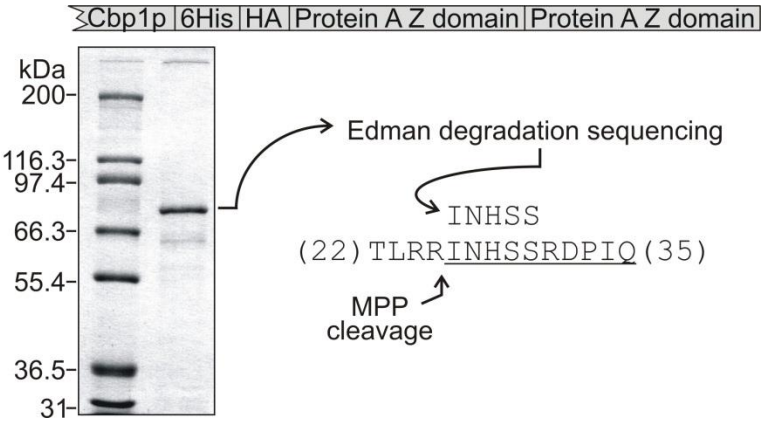

Figure 31. The N-terminus of mature Cbp1p is at amino acid 26. An image of a Coomassie-stained LDS 4-12% polyacrylamide gel showing a preparation of tagged Cbp1p purified from yeast mitochondria. The composition of the C-terminally fused purification tag is shown above the image. Edman degradation analysis of the protein in the indicated band revealed a single sequence read, pointed to by the bent arrow. The read is aligned with a portion of the Cbp1p sequence (numbers indicate positions of the flanking amino acids in the sequence of the Cbp1p precursor), in which the determined N-terminal sequence of the mature protein is underlined. The MPP processing site is indicated by the arrow (*128*).

Analysis of the preparation by a polyacrylamide gel (Figure 31) revealed a major band at approximately 75 kDa. The protein in the band was submitted for Edman degradation analysis which resulted in one sequence read, INHSS, corresponding to amino acids 26-30 in the Cbp1p precursor. This observation is consistent with a one-step processing mechanism, in which MPP cleaves the Cbp1p precursor after amino acid R25 (*128*).

II.D.2. *Expression and purification of Cbp1p.* Although the mature N-terminus of Cbp1p was determined to be at amino acid 26, an alternate deletion mutant, Δ33*CBP1,* was used in the following experiments due to its increased stability (data not shown). A fragment of *CBP1* was cloned into an *E. coli* pTrcHis-A expression vector to generate Δ33Cbp1p with an N-terminal histidine purification tag (MAH_6) fused to amino acid I34 of the protein (plasmid pAC21). From this point on, this form of the protein will be referred to as Cbp1p. Expression of the protein was optimized in *E. coli* XJb(DE3) cells at a reduced temperature (16°C) and a low induction level. This allowed for Cbp1p to be obtained in a soluble form. As shown in Figure 32, the protein could be purified to homogeneity by a preliminary enrichment with Ni-IDA beads followed by heparin affinity chromatography.

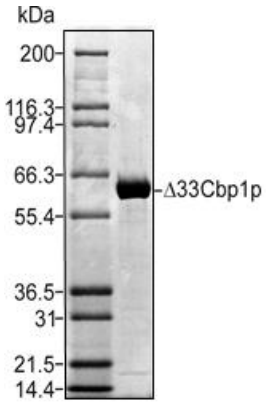

Figure 32. Purification of Cbp1p. An image of a Coomassie-stained LDS 4-12% polyacrylamide gel showing the purity of a preparation of Cbp1p.

Remarkably, the apparent mobility of Cbp1p in a denaturing gel corresponded to approximately 62 kDa, while the expected molecular weight of this form of the protein is 73 kDa (Figure 32). A similar abnormality in electrophoretic mobility was mentioned in a previous study involving Cbp1p as well (*61*). Thus as was the case with a *COX2* translational activator, Pet111p (see Figure 10), it appears as though Cbp1p has an unusually high mobility during electrophoresis. This common characteristic between Pet111p and Cbp1p is intriguing and suggests that they may share chemical and/or structural features. Further characterization of other yeast mitochondrial translational activators is needed to see if they may represent a new group of high electrophoretic mobility proteins.

II.D.3. *Cbp1p interacts directly with the 5′-end proximal region of COB.* Utilizing the highly purified recombinant Cbp1p, the next step was to determine if this protein interacts with RNA and more specifically *COB*. Previous genetic studies proposed that Cbp1p directly interacts with *COB* close to the 5′-end (*66, 159*). However, to date this protein has not been shown to associate with RNA directly *in vitro*. In this study, to determine if and where Cbp1p interacted with *COB*, I used an EMSA analysis to scan the entire 5′-UTR for a binding target(s) of Cbp1p. The 5′-UTR of *COB* is 951 nucleotides, therefore an RNA probe containing the entire 5′-UTR would be too large to utilize in an EMSA. To overcome this, seven smaller RNA probes were generated. They represented partially overlapping fragments of the

entire sequence of this region (Figure 33A). The first RNA probe (RNA1) contained some sequence from upstream of the maturation site present only in pre-*COB.* This region was included in the analysis because as a factor implicated in maturation of *COB (157)*, Cbp1p may theoretically bind to targets on both sides of the cleavage site during the maturation process. Similarly, RNA7 included sequence from the beginning of the ORF of *COB*. This region was included because the Pet111p translational activator associates within the *COX2* ORF close to the translation initiation site (see Chapter I section I.D.6). Therefore by analogy, the region in the beginning of the *COB* ORF could also harbor a potential Cbp1p target.

To generate the RNA probes, seven DNA templates were prepared by PCR amplification, during which an SP6 RNA polymerase promoter sequence was added into the upstream region of each template by inclusion at the 5′-ends of the forward primers. Using these DNA templates, a set of RNA probes was synthetized by *in vitro* transcription with SP6 RNA polymerase (Figure 33B).

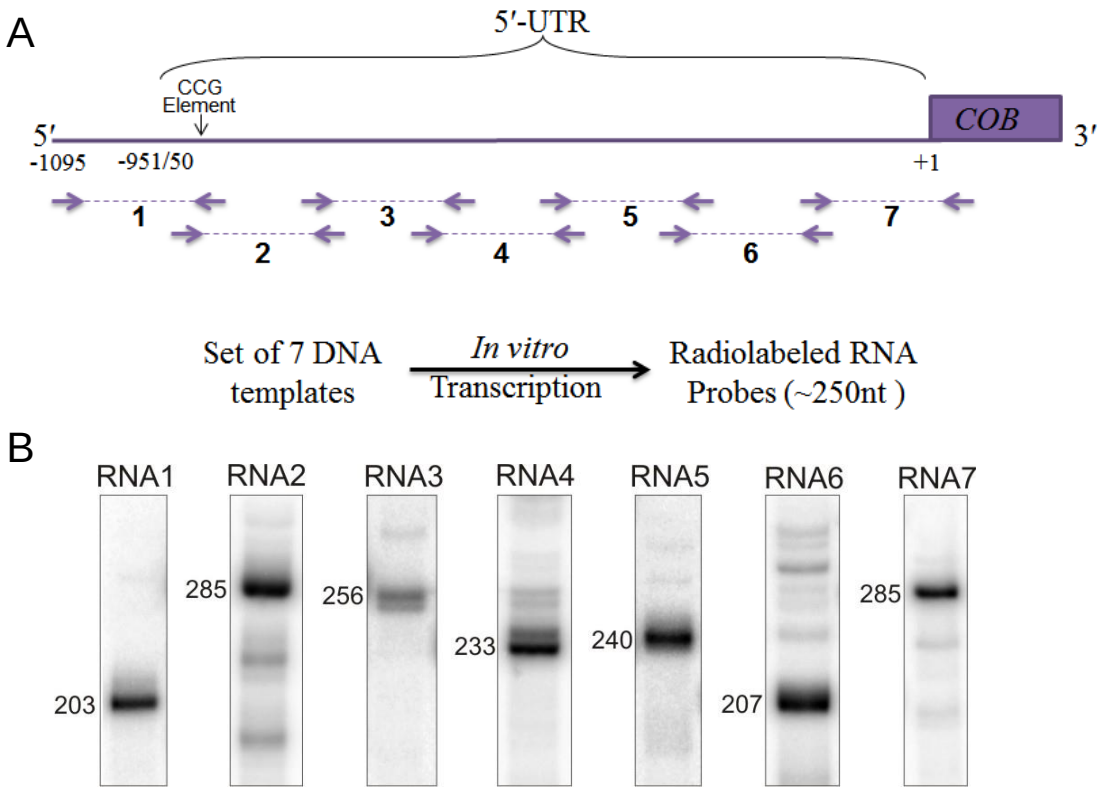

Figure 33. Generation of RNA probes that span the *COB* 5′-UTR. A. A schematic showing the location of seven RNA probes along the 5′-UTR of *COB*. Each probe was approximately 200-300 nucleotides in length and overlapped in sequence with the neighboring probes by at least 75 nucleotides. The probes spanned the region from -1096 to +95, relative to the translational start site (+1), with -951/-950 being the 5′-end of mature *COB*. To synthetize the probes by *in vitro* transcription, corresponding DNA templates were PCR-amplified using pairs of primers positioned as indicated by the solid arrows. An SP6 RNA polymerase promoter sequence was added to each of the forward primers to introduce the promoters into the DNA templates. B. Portions of a phosphor imaging scan of a 7 M urea 8% polyacrylamide gel showing products of *in vitro* transcription reactions that generated the seven RNA probes. The RNA was radiolabeled by random incorporation of [α-^{32}P]-UTP. Expected sizes of run-off transcription products

are indicated to the left. The intensity in each image was adjusted independently to show the transcription products more clearly.

As shown in Figure 33B, the seven RNA probes were successfully generated by *in vitro* SP6 RNA polymerase transcription. Utilizing the RNA products, an EMSA was performed to see if Cbp1p could interact with the *COB* fragments and if so, to determine the proximal location of this interaction. In the presence of Cbp1p, all of probe RNA1 was found in a complex with the protein (Figure 34A). Probe RNA2 also associated with Cbp1p, although with less efficiency. This result shows that Cbp1p can establish a direct physical interaction with RNA in the absence of any other mitochondrial factors. The protein also clearly exhibits specificity for the sequence contained within probes RNA1 and RNA2. These two RNAs represent a region 146 nt upstream and 268 nt downstream of the maturation site (nt -951). Significantly, this observation is consistent with previously published genetic data in which the site of functional activity of Cbp1p was localized to the 5′-end proximal third of *COB* (*159*).

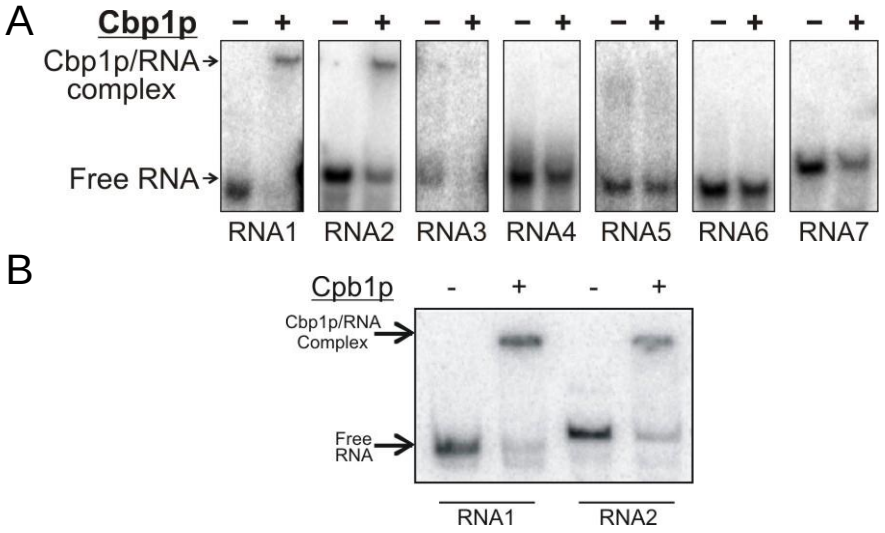

Figure 34. Cbp1p interacts with the 5′-end proximal region of the *COB* 5′-UTR. A. [^{32}P]-labeled RNA probes were incubated with or without 500 nM Cbp1p and resolved in a 7% polyacrylamide gel under native conditions. The bands corresponding to free and complexed RNA (as indicated) were visualized by phosphor imaging. During binding, 3 μM competitor tRNA was present in the solutions. The images were derived from the same native gel and the intensity was adjusted independently in each image to show the bands more clearly. B. [^{32}P]-labeled RNA1 and RNA2 were gel purified and incubated with 500 nM Cbp1p, where indicated, in the presence of 1 μM competitor tRNA. Cbp1p/RNA complexes were separated from unbound RNA in a 6% polyacrylamide gel under native conditions and the corresponding bands were visualized by phosphor imaging.

Unlike RNA1 and RNA2, other RNA probes did not appear to form specific complexes with Cpb1p. Interestingly, there was a loss of intensity for the free RNA in the lanes with RNA3 and RNA7 when Cbp1p was present; however, shifted RNA bands were not observed (Figure 34A). Therefore, it does not appear that Cbp1p complexed with either of these two RNAs at a

1:1 ratio. However, it cannot be completely excluded that both of these probes form large irregular complexes, in which the highly basic Cbp1p is present at a molar excess over the RNA. Under the conditions of the experiment, such complexes would not enter the gel.

To ensure that the observed RNA/Cbp1p complexes were not due the presence of shorter, aberrant RNA species that may have been generated in the *in vitro* transcription reactions, probes RNA1 and RNA2 were gel purified and tested again for an interaction with Cbp1p. As shown in Figure 34B, Cbp1p could form complexes with both purified RNA1 and RNA2 probes with a similarly high affinity. This signified that there is a region within each of these two RNA probes that Cbp1p can bind to specifically. Since the two probes shared the same 70 nt overlapping region, I presumed that this region is where a Cbp1p binding target may be harbored. Interestingly, this is the region that also contains the CCG element (nucleotides -940 to -938 relative to the translational start, +1) that was genetically implicated to be the point of interaction for Cbp1p (*66*). To test if Cbp1p can interact with this region, an additional RNA probe, RNA8, was generated by *in vitro* transcription to contain the sequence of the RNA1 and RNA2 overlap (nucleotides -965 to -896). An EMSA assay was performed to compare the specificity of the association of Cbp1p with the probes RNA8 and RNA9 (Figure 35). RNA9 encompassed a 76 nt region (nucleotides -756 to -682) downstream from RNA8. RNA9 was used in the assay as a nonspecific control because it

represented a part of the probe RNA3, which failed to form a specific complex with Cbp1p (see Figure 34A).

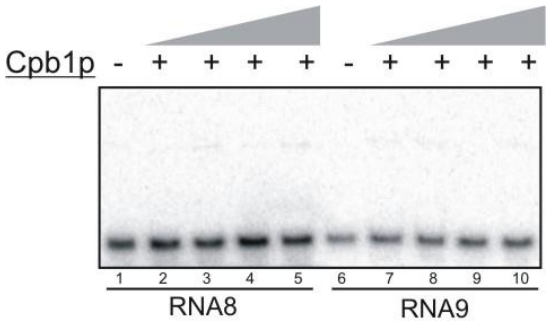

Figure 35. Cbp1p does not interact with an RNA probe comprising the sequence between nucleotides -965 and -896 in the 5′-UTR of pre-*COB*. [^{32}P]-labeled RNA probes were incubated with increasing amounts of Cbp1p (200, 400, 600, and 800 nM in lanes 2 and 7, 3 and 8, 4 and 9, and 5 and 10, respectively) and resolved in a 7% polyacrylamide gel under native conditions. During binding, competitor tRNA was present in the solutions at a two-fold molar excess over the protein. The bands were visualized by phosphor imaging.

To my surprise, Cbp1p did not show specificity for the probe RNA8. This was an interesting finding as it implied that either *(i)* Cbp1p recognized two distinct sequences at the 5′-end of RNA1 and the 3′-end of RNA2, neither of which was a part of RNA8, or *(ii)* a structure that can only form in the context of RNA8 obscured the binding target of Cbp1p.

Previous genetic analysis has suggested that Cbp1p may directly interact with a target containing a “CCG element” (*66*). In the context of the RNA8 probe, this sequence is expected to be a part of a stable stem-loop

structure (Figure 36A), which Cbp1p could not recognize as a specific target (see Figure 35). However, the structure of the region may not necessarily be the same in a larger context of the entire *COB* mRNA. To test whether Cbp1p can specifically bind to this region when it is not structured, I utilized a 27 nt probe (*COB* CCG 27) that was designed to contain the CCG element along with the flanking sequences in such a way that the stem of the predicted structure was destabilized. The interaction of Cbp1p with the *COB* CCG 27 probe was evaluated in an EMSA (Figure 36B).

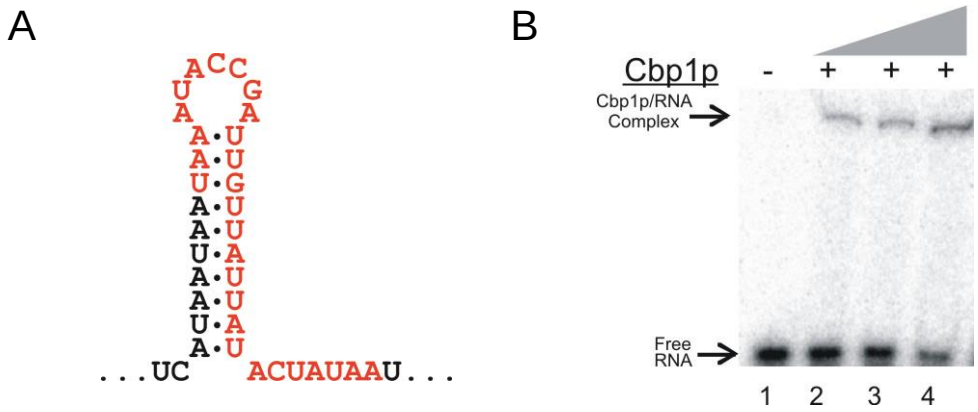

Figure 36. Cbp1p exhibits a low-affinity specific interaction with an unstructured RNA probe containing the CCG element. A. A 27-nt RNA probe, *COB* CCG 27, was synthetized to incorporate the sequence in the vicinity of the CCG element while disrupting a stable stem-loop structure predicted to form in this region of *COB*. The nucleotides included in the probe are shown in red. B. 5′-[^{32}P]-labeled *COB* CCG 27 (0.1 µM) was incubated with varying concentrations of Cbp1p (200, 400, and 800 nM in lanes 2, 3, and 4, respectively) and the mixtures were resolved in a 7% polyacrylamide gel under native conditions. The bands corresponding to unbound and Cbp1p-bound RNA, as indicated, were visualized by phosphor imaging.

During binding, competitor tRNA was present in the solutions at a two-fold molar excess over the protein.

Significantly, this experiment demonstrated that Cbp1p can indeed form a complex with the *COB* CCG 27 in the presence of a two-fold molar excess of competitor tRNA. This data shows that Cbp1p can specifically and directly target the CCG-containing region in a single-stranded context. Nevertheless, the efficiency of the formation of the *COB* CCG 27 complex appeared to be lower compared to those observed with either RNA1 or RNA2 (Figure 34B). Since the exact sequence and the boundaries of the Cbp1p binding target are not known, it is possible that the *COB* CCG 27 probe did not contain all of the necessary residues to establish a stronger binding with the protein. This could explain the considerable drop in the affinity that was observed in this experiment. Further analysis of the interaction of Cbp1p with this and related RNA probes is necessary in order to determine the specific binding target of the protein.

II.E. Discussion

As discussed in Chapter I (section I.E.), mitochondrial gene products that are a part of the respiratory chain are synthesized close to the inner membrane (*117*). This localization ensures that each hydrophobic subunit is co-translationally inserted into the proper respiratory complex (*18, 147*). An effort to study this coordination has resulted in the identification of a large protein complex, termed the MIOREX complex. MIOREX components perform at multiple steps involved in gene expression, including: transcription, post-transcriptional processing, translational activation, translation, and post-translational mRNA degradation (*53*). These MIOREX complexes are thought to be segmented, in that, all required components for gene expression of each individual respiratory complex is localized into one unit termed a "respirasome" (*148*).

This chapter of the study aimed to define the roles of Cbp1p in Cobp expression; the only mitochondrial gene product in complex III of the respiratory chain. The other subunits of complex III, along with the various regulatory and assembly proteins, are nuclear genome-encoded. Cobp must be available for membrane insertion in a manner that is coordinated with the nucleus-encoded components; therefore, gene expression from both genomes is coordinated (*24*). Although little is known of this coordination, it is well-established that there is a group of nuclear genome-encoded proteins that aid in Cobp expression and ensure that respiration occurs (*18*). These proteins have been implicated at various steps of gene expression including,

mRNA maturation, mRNA stabilization, translational activation, nascent polypeptide stabilization, and cotranslational insertion of the newly synthesized Cobp (*18*) (see Figure 30).

As a member of the PPR protein family (*103*), Cbp1p is predicted to act as a sequence-specific single-stranded RNA binding protein. Therefore, to characterize Cbp1p function biochemically, I sought to study its relationship with *COB*. Cbp1p has long been genetically implicated in the maturation, stability, and translation of the *COB* transcript (*154, 156, 157*). Genetic manipulations of the 5′-UTR of *COB* defined a region in which Cbp1p was believed to act directly. This region was first delineated to the 5′-end proximal third of the 5′-UTR (*159*), and then narrowed down to an 11 nt sequence (*160*), containing at least three essential nucleotides, CCG (nt -940 to -938) (*66*). All of these previous studies were performed *in vivo* with other mitochondrial factors present. Therefore, whether it was Cbp1p or another protein responsible for the physical interaction with *COB* has been unknown. Previous attempts to isolate this protein heterologically were proven to be unsuccessful (*61*) and therefore it has never been studied in an *in vitro* system. However, I was able to clone and express a soluble recombinant Cbp1p (Figure 32) suitable for biochemical characterization.

The 5′-UTR of *COB* has been experimentally determined to be 951 nt (*161*), which made it a challenge to study the full-length RNA *in vitro*. To overcome this, I created seven smaller RNA probes that spanned the 5′-UTR and contained overlapping sequences. The overlapping regions ensured that

potential binding targets were not destroyed in the process of separating the sequence. In an initial EMSA with the seven RNA probes and highly purified Cbp1p I determined that: *(i)* Cbp1p was able to form protein/RNA complexes, signifying that this protein could interact with RNA in the absence of any other mitochondrial factors and *(ii)* it directly contacts the *COB* mRNA as previously proposed (*66*). This experiment showed that Cbp1p was able to interact with both the RNA1 and RNA2 probes, with the RNA1 probe interaction being more efficient (Figure 34). These two probes comprised of a region 146 nt upstream and 268 nt downstream of the maturation site (nt -951) (Figure 33). This finding was significant as it was consistent with previously published genetic data that suggested a functional involvement of Cbp1p with the 5′-end proximal third of the 5′-UTR of *COB* (*159*).

The next step was to refine this binding target further. Since the two probes shared an overlapping sequence it was logical to assume that this region may harbor the Cbp1p interaction site. To test this, probe RNA8 was created to contain the complete sequence of the overlapping region. Interestingly, this probe also contained the CCG element that was previously proposed to be the point of interaction for Cbp1p (*66*). An EMSA with probe RNA8 showed that Cbp1p was not able to form any appreciable complexes with this RNA (Figure 35). This result was puzzling, as it suggested that: *(i)* Cbp1p may interact with two alternate sequences either at the 5′- end of RNA1 or the 3′-end of RNA2, or *(ii)* probe RNA8 folded into an aberrant structure in which the binding target of Cbp1p was engaged in base pairing.

Further *in silico* analysis showed that most of the nucleotides in RNA8 may indeed be involved in secondary structuring (Figure 37); therefore, any potential Cbp1p target(s) would be inaccessible for an interaction.

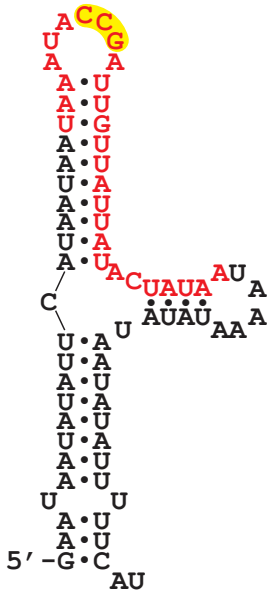

Figure 37. Probe RNA8 is predicted to be highly structured. *In silico* modeling of the RNA8 probe revealed that it may be folded in a tight structure with no regions available for an interaction with a single-stranded RNA binding protein, such as Cbp1p. The CCG element is highlighted in yellow and the position of the *COB* CCG 27 probe (see below) is shown in red. This analysis was performed using the Mfold server (The RNA Institute, College of Arts and Sciences, State University of New York at Albany).

To eliminate the possibility of base pairing, I designed a shortened RNA probe, *COB* CCG 27, which contained the CCG element and flanking sequence in a single-stranded context (Figure 36). EMSA analysis with *COB* CCG 27 showed that Cbp1p could form a complex with this unstructured probe; however, the efficiency of complex formation was lower when

compared to that observed with probe RNA1 (Figure 34B). This led to the notion that the *COB* CCG 27 probe did not contain all of the necessary residues to establish a stronger binding with the protein.

Interestingly, there is a sequence repeat (5′-UAUUCAUAAUAAUAAAUA-3′) that appears twice in the 5′-UTR of the *COB* precursor. The positioning of these sequences is quite significant in the context of the *COB* maturation process. The first repeat is directly at the 5′-end of pre-*COB* (nt -1097 to -1080) formed by excision of $tRNA^{Glu}$ (Figure 30A). The second repeat (nt -958 to -941) is found at the site of the final maturation step directly upstream of the CCG element. The 5′-end of mature *COB* (nt -951) is therefore positioned within this repeat. Previous genetic studies implicated Cbp1p in *COB* maturation as multiple maturation intermediates at various points accumulated in Δ*CBP1* cells (*156*). These intermediates were thought to form due to a loss in stability provided to the primary transcript by Cbp1p (*157*). Taking all of this into consideration, I hypothesize that Cbp1p may function by interacting with this repeat sequence. While my initial design of *COB* CCG 27 only included a portion of the repeat sequence, the probe could still interact with Cbp1p, albeit weakly. More studies with this sequence must be performed to determine if it represents a binding target of Cbp1p and is significant to the function of the protein. Specifically, EMSA analysis and RNase footprinting with fragments of probe RNA1 can help define the region which Cbp1p recognized specifically.

Overall, this study has shown for the first time that Cbp1p can be studied *in vitro*. I was able to identify a relatively small region of the 5′-UTR that harbors the Cbp1p interaction target(s), which is consistent with previously published genetic data (*66, 159*). The fact that other translational activators of *COB* have been implicated in interacting with the 5′-UTR (Figure 30) suggests that Cbp1p may not be able to efficiently interact with its target in the absence of some of these factors. Nevertheless, Cbp1p was able to complex with probe RNA1 very efficiently, showing that this system contained everything needed for this interaction to occur *in vitro*. This however does not exclude that there may be other factors that enhance this interaction *in vivo*. For instance there may be a protein that aids in loading Cbp1p to *COB* such as, Mtf2p, which was shown to substantially increase the affinity of Pet111p for the 5′-UTR of *COX2* (Figure 23)**.** Alternatively, other translational activators of *COB*, such as Cbt1p, may associate with the Cbp1p/RNA complex providing additional stability or even additional RNA contacts. Once the binding region of Cbp1p is defined, experiments with other *COB*-specific mitochondrial factors can be performed to characterize the translational activation process to a greater extent.

II.F. Supplemental Information

Table 6. Sequences of the oligonucleotides used in the study.

Oligonucleotide	Sequence (5′ to 3′)
DNA oligonucleotides	
Cbp1-BG-5	GTACAAAAAAGCAGGCTACAAAATGTTTTTACCTCGTCTCGTTCGGTACAGG
Cbp1-BG-3	CACATCAACCACTTTGTACAAGAAAGCTGGGTTTCTTAAGTAACGTTTGACAGC
Cbp1Trc-5	CATAATCCATGGCTCATCATCATCATCATCATATGTTTTTACCTCGTCTCGTTCG
Cbp1Trc-3	GCTAATCTCGAGTTATCATCTTAAGTAACGTTTGACAGCC
Cbp1-del25	GCTCATCATCATCATCATCATATCAACCACAGCAGCAGG
Cbp1-del25c	CCTGCTGCTGTGGTTGATATGATGATGATGATGATGAGC
Cbp1-del33	GCTCATCATCATCATCATCATATTCAAAAACAGGTCTTGG
Cbp1-del33c	CCAAGACCTGTTTTTGAATATGATGATGATGATGATGAGC
F *COB* SP6 1-29	AATAATTAGGTGACACTATAGAATTCATAATAATAAATATTTGTAAAAAAAG
R *COB* 200-171	ATGAAAAATATATTATATATTTTATTATAG
F *COB* SP6 131-158	AATAATTAGGTGACACTATAGAATAATATATTCATAATAATAAATACCG
R *COB* 414-392	TAATTTATATTCTCCTTTCGGGG
F *COB* SP6 340-369	AATAATTAGGTGACACTATAGAATTTATTTTATTTTTTTTTTATAGTTCCGG
R *COB* 594-568	GACGATTATATTAATATTAATGATATC
F *COB* SP6 502-535	AATAATTAGGTGACACTATAGAATATTATATTTATATAATATTAATATAAAAATC
R *COB* 733-707	CGTGGATAATATTAATTATATTATATC
F *COB* SP6 652-683	AATAATTAGGTGACACTATAGAATATTTTTAAATATTATATTATATTATTAAC
R *COB* 890-864	AATTATTAAAAAAAAAAGATATTTCTC
F *COB* SP6 816-841	AATAATTAGGTGACACTATAGAATCTATTAACTTTTTTTTTTAATGG
R *COB* 1021-992	AATATATAAATATATATTATTAATTAATAC
F *COB* SP6 907-934	AATAATTAGGTGACACTATAGAAAATATGTATATATAATAAAAAAATAG
R *COB* 1190-1166	ATATTTCATCAATAATTAATTGATG

RNA oligonucleotide	
COB CCG 27	UAAAUACCGAUUGUUAUUAUACUAUAA

Table 7. Correspondence of the oligonucleotides used in the study to regions in the *COB* gene.

Primer Name	Positions of the 5′- and 3′-ends of the primer, respectively, relative to the A nucleotide of the start codon (+1) in *COB*
F *COB* SP6 1-29	-1096 to -1067
R *COB* 200-171	-925 to -896
F *COB* SP6 131-158	-965 to-938
R *COB* 414-392	-704 to -682
F *COB* SP6 340-369	-756 to -727
R *COB* 594-568	-528 to -502
F *COB* SP6 502-535	-594 to -561
R *COB* 733-707	-389 to -363
F *COB* SP6 652-683	-444 to -413
R *COB* 890-864	-232 to -206
F *COB* SP6 816-841	-280 to -255
R *COB* 1021-992	-104 to -75
F *COB* SP6 907-934	-189 to -162
R *COB* 1190-1166	71 to 95

SUMMARY AND CONCLUSIONS

All of the components needed for the various phases of mitochondrial gene expression have been found to be compartmentalized in large complexes tethered to the inner membrane of the mitochondria. These complexes are very intricate and composed of both nuclear- and mitochondrial-genome encoded components. This study set forth to characterize one such component, nucleus-encoded gene-specific translational activators, which have been identified and proven essential in many species including, *S. cerevisiae* and *H. sapiens.*

Part of the data presented here has revealed that *S. cerevisiae's* Pet111p, a translational activator for *COX2*, interacts sequence-specifically at two distinct locations in the mRNA; one within the 5′-UTR and a second within the coding region. Due to the innate inhibitory structuring of the region of initiation, translation of *COX2* is not possible. Based on my findings, I hypothesize that Pet111p must be associated with both binding targets to preclude the structural inhibition of *COX2* translation. Since RNA is highly structured in general, the mechanism proposed, in which a translational activator can alter and/or prevent inhibitory RNA structures from forming, can be transposed to other gene expression systems.

The finding that Pet111p may perform at least part of its function from within the ORF is quite significant. Although this is the first example of this type of translational control in *S. cerevisiae*, it has been previously found in *M. musculus* that a gene-specific translational activator of *COX1* (TACO1) acts

by interacting specifically within the ORF. Mutations that disrupt this function led to human pathologies, including Leigh syndrome (*59*), showing the critical importance of this type of translational control. It has been presumed that in *S. cerevisiae* translational activation only relied on UTRs, and since *H. sapiens* mitochondrial transcripts do not contain extensive UTRs, the mechanisms of translational activation were thought to be distinct. However, considering these new findings, further characterization of this process in *S. cerevisiae* can lead to a deeper understanding of mitochondrial translational activation in general and aid in the development of therapeutics for various human mitochondrial diseases.

REFERENCES

1. Martin, W. F., Garg, S., and Zimorski, V. (2015) Endosymbiotic theories for eukaryote origin, *Philos Trans R Soc Lond B Biol Sci 370*, 20140330.

2. Friedman, J. R., and Nunnari, J. (2014) Mitochondrial form and function, *Nature 505*, 335-343.

3. Zimorski, V., Ku, C., Martin, W. F., and Gould, S. B. (2014) Endosymbiotic theory for organelle origins, *Curr Opin Microbiol 22*, 38-48.

4. Margulis, L. (1971) Symbiosis and evolution, *Sci Am 225*, 48-57.

5. Dudek, J., Rehling, P., and van der Laan, M. (2013) Mitochondrial protein import: common principles and physiological networks, *Biochim Biophys Acta 1833*, 274-285.

6. Kuhlbrandt, W. (2015) Structure and function of mitochondrial membrane protein complexes, *BMC Biol 13*, 89.

7. Daum, G., Bohni, P. C., and Schatz, G. (1982) Import of proteins into mitochondria. Cytochrome b2 and cytochrome c peroxidase are located in the intermembrane space of yeast mitochondria, *J Biol Chem 257*, 13028-13033.

8. McBride, H. M., Neuspiel, M., and Wasiak, S. (2006) Mitochondria: more than just a powerhouse, *Curr Biol 16*, R551-560.

9. Otterstedt, K., Larsson, C., Bill, R. M., Stahlberg, A., Boles, E., Hohmann, S., and Gustafsson, L. (2004) Switching the mode of metabolism in the yeast Saccharomyces cerevisiae, *EMBO Rep 5*, 532-537.

10. Rak, M., Zeng, X., Briere, J. J., and Tzagoloff, A. (2009) Assembly of F0 in Saccharomyces cerevisiae, *Biochim Biophys Acta 1793*, 108-116.

11. Hancock, J. T., Desikan, R., and Neill, S. J. (2001) Role of reactive oxygen species in cell signalling pathways, *Biochem Soc Trans 29*, 345-350.

12. Halliwell, B. (1994) Free radicals, antioxidants, and human disease: curiosity, cause, or consequence?, *Lancet 344*, 721-724.

13. Nunnari, J., and Suomalainen, A. (2012) Mitochondria: in sickness and in health, *Cell 148*, 1145-1159.

14. Duchen, M. R., and Szabadkai, G. (2010) Roles of mitochondria in human disease, *Essays Biochem 47*, 115-137.

15. Fox, T. D. (2012) Mitochondrial protein synthesis, import, and assembly, *Genetics 192*, 1203-1234.

16. Couvillion, M. T., Soto, I. C., Shipkovenska, G., and Churchman, L. S. (2016) Synchronized mitochondrial and cytosolic translation programs, *Nature 533*, 499-503.

17. Fontanesi, F., Soto, I. C., Horn, D., and Barrientos, A. (2006) Assembly of mitochondrial cytochrome c-oxidase, a complicated and highly regulated cellular process, *Am J Physiol Cell Physiol 291*, C1129-1147.

18. Herrmann, J. M., Woellhaf, M. W., and Bonnefoy, N. (2013) Control of protein synthesis in yeast mitochondria: the concept of translational activators, *Biochim Biophys Acta 1833*, 286-294.

19. Liu, Z., and Butow, R. A. (2006) Mitochondrial retrograde signaling, *Annu Rev Genet 40*, 159-185.

20. Arnold, I., Pfeiffer, K., Neupert, W., Stuart, R. A., and Schagger, H. (1998) Yeast mitochondrial F1F0-ATP synthase exists as a dimer: identification of three dimer-specific subunits, *Embo J 17*, 7170-7178.

21. Helfenbein, K. G., Ellis, T. P., Dieckmann, C. L., and Tzagoloff, A. (2003) ATP22, a nuclear gene required for expression of the F-0 sector of mitochondrial ATPase in Saccharomyces cerevisiae, *Journal of Biological Chemistry 278*, 19751-19756.

22. Barros, M. H., and Tzagoloff, A. (2017) Aep3p-dependent translation of yeast mitochondrial ATP8, *Mol Biol Cell 28*, 1426-1434.

23. Rak, M., and Tzagoloff, A. (2009) F1-dependent translation of mitochondrially encoded Atp6p and Atp8p subunits of yeast ATP synthase, *Proc Natl Acad Sci U S A 106*, 18509-18514.

24. Lipinski, K. A., Kaniak-Golik, A., and Golik, P. (2010) Maintenance and expression of the S. cerevisiae mitochondrial genome--from genetics to evolution and systems biology, *Biochim Biophys Acta 1797*, 1086-1098.

25. Williamson, D. (2002) The curious history of yeast mitochondrial DNA, *Nat Rev Genet 3*, 475-481.

26. Kucej, M., Kucejova, B., Subramanian, R., Chen, X. J., and Butow, R. A. (2008) Mitochondrial nucleoids undergo remodeling in response to metabolic cues, *J Cell Sci 121*, 1861-1868.

27. Turk, E. M., Das, V., Seibert, R. D., and Andrulis, E. D. (2013) The mitochondrial RNA landscape of Saccharomyces cerevisiae, *PLoS One 8*, e78105.

28. Petersen, R. F., Langkjaer, R. B., Hvidtfeldt, J., Gartner, J., Palmen, W., Ussery, D. W., and Piskur, J. (2002) Inheritance and organisation of the mitochondrial genome differ between two Saccharomyces yeasts, *J Mol Biol 318*, 627-636.

29. Tang, G. Q., Deshpande, A. P., and Patel, S. S. (2011) Transcription factor-dependent DNA bending governs promoter recognition by the mitochondrial RNA polymerase, *J Biol Chem 286*, 38805-38813.

30. Schinkel, A. H., Groot Koerkamp, M. J., Van der Horst, G. T., Touw, E. P., Osinga, K. A., Van der Bliek, A. M., Veeneman, G. H., Van Boom, J. H., and Tabak, H. F. (1986) Characterization of the promoter of the large ribosomal RNA gene in yeast mitochondria and separation of mitochondrial RNA polymerase into two different functional components, *Embo J 5*, 1041-1047.

31. Biswas, T. K., Edwards, J. C., Rabinowitz, M., and Getz, G. S. (1985) Characterization of a yeast mitochondrial promoter by deletion mutagenesis, *Proc Natl Acad Sci U S A 82*, 1954-1958.

32. Wettstein-Edwards, J., Ticho, B. S., Martin, N. C., Najarian, D., and Getz, G. S. (1986) In vitro transcription and promoter strength analysis of five mitochondrial tRNA promoters in yeast, *J Biol Chem 261*, 2905-2911.

33. Deshpande, A. P., and Patel, S. S. (2012) Mechanism of transcription initiation by the yeast mitochondrial RNA polymerase, *Biochim Biophys Acta 1819*, 930-938.

34. Amiott, E. A., and Jaehning, J. A. (2006) Mitochondrial transcription is regulated via an ATP "sensing" mechanism that couples RNA abundance to respiration, *Mol Cell 22*, 329-338.

35. Parker, R. (2012) RNA degradation in Saccharomyces cerevisae, *Genetics 191*, 671-702.

36. Chang, J. H., and Tong, L. (2012) Mitochondrial poly(A) polymerase and polyadenylation, *Biochim Biophys Acta 1819*, 992-997.

37. Kertesz, M., Wan, Y., Mazor, E., Rinn, J. L., Nutter, R. C., Chang, H. Y., and Segal, E. (2010) Genome-wide measurement of RNA secondary structure in yeast, *Nature 467*, 103-107.

38. Gilbert, W. V., Bell, T. A., and Schaening, C. (2016) Messenger RNA modifications: Form, distribution, and function, *Science 352*, 1408-1412.

39. Walters, R. W., Matheny, T., Mizoue, L. S., Rao, B. S., Muhlrad, D., and Parker, R. (2017) Identification of NAD+ capped mRNAs in Saccharomyces cerevisiae, *Proc Natl Acad Sci U S A 114*, 480-485.

40. Osinga, K. A., De Vries, E., Van der Horst, G., and Tabak, H. F. (1984) Processing of yeast mitochondrial messenger RNAs at a conserved dodecamer sequence, *Embo J 3*, 829-834.

41. Min, J., and Zassenhaus, H. P. (1993) Identification of a protein complex that binds to a dodecamer sequence found at the 3' ends of yeast mitochondrial mRNAs, *Mol Cell Biol 13*, 4167-4173.

42. Li, H., and Zassenhaus, H. P. (1999) Purification and characterization of an RNA dodecamer sequence binding protein from mitochondria of Saccharomyces cerevisiae, *Biochem Biophys Res Commun 261*, 740-745.

43. Wiesenberger, G., and Fox, T. D. (1997) Pet127p, a membrane-associated protein involved in stability and processing of Saccharomyces cerevisiae mitochondrial RNAs, *Mol Cell Biol 17*, 2816-2824.

44. Fekete, Z., Ellis, T. P., Schonauer, M. S., and Dieckmann, C. L. (2008) Pet127 governs a 5' -> 3'-exonuclease important in maturation of apocytochrome b mRNA in Saccharomyces cerevisiae, *J Biol Chem 283*, 3767-3772.

45. Islas-Osuna, M. A., Ellis, T. P., Marnell, L. L., Mittelmeier, T. M., and Dieckmann, C. L. (2002) Cbp1 is required for translation of the mitochondrial cytochrome b mRNA of Saccharomyces cerevisiae, *J Biol Chem 277*, 37987-37990.

46. Hofmann, T. J., Min, J., and Zassenhaus, H. P. (1993) Formation of the 3' end of yeast mitochondrial mRNAs occurs by site-specific cleavage two bases downstream of a conserved dodecamer sequence, *Yeast 9*, 1319-1330.

47. Dziembowski, A., Piwowarski, J., Hoser, R., Minczuk, M., Dmochowska, A., Siep, M., van der Spek, H., Grivell, L., and Stepien, P. P. (2003) The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism, *J Biol Chem 278*, 1603-1611.

48. Hollingsworth, M. J., and Martin, N. C. (1986) RNase P activity in the mitochondria of Saccharomyces cerevisiae depends on both mitochondrion and nucleus-encoded components, *Mol Cell Biol 6*, 1058-1064.

49. Lambowitz, A. M., and Zimmerly, S. (2011) Group II introns: mobile ribozymes that invade DNA, *Cold Spring Harb Perspect Biol 3*, a003616.

50. Eickbush, T. H. (1999) Mobile introns: retrohoming by complete reverse splicing, *Curr Biol 9*, R11-14.

51. Amunts, A., Brown, A., Bai, X. C., Llacer, J. L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S. H. W., and Ramakrishnan, V. (2014) Structure of the yeast mitochondrial large ribosomal subunit, *Science 343*, 1485-1489.

52. Watson, K. (1972) The organization of ribosomal granules within mitochondrial structures of aerobic and anaerobic cells of Saccharomyces cerevisae, *J Cell Biol 55*, 721-726.

53. Kehrein, K., Schilling, R., Moller-Hergt, B. V., Wurm, C. A., Jakobs, S., Lamkemeyer, T., Langer, T., and Ott, M. (2015) Organization of Mitochondrial Gene Expression in Two Distinct Ribosome-Containing Assemblies, *Cell Rep*.

54. Krause, K., Lopes de Souza, R., Roberts, D. G., and Dieckmann, C. L. (2004) The mitochondrial message-specific mRNA protectors Cbp1 and Pet309 are associated in a high-molecular weight complex, *Mol Biol Cell 15*, 2674-2683.

55. Spencer, A. C., and Spremulli, L. L. (2004) Interaction of mitochondrial initiation factor 2 with mitochondrial fMet-tRNA, *Nucleic Acids Res 32*, 5464-5470.

56. Towpik, J. (2005) Regulation of mitochondrial translation in yeast, *Cell Mol Biol Lett 10*, 571-594.

57. Tzagoloff, A., and Dieckmann, C. L. (1990) PET genes of Saccharomyces cerevisiae, *Microbiol Rev 54*, 211-225.

58. Costanzo, M. C., and Fox, T. D. (1990) Control of mitochondrial gene expression in Saccharomyces cerevisiae, *Annu Rev Genet 24*, 91-113.

59. Weraarpachai, W., Antonicka, H., Sasarman, F., Seeger, J., Schrank, B., Kolesar, J. E., Lochmuller, H., Chevrette, M., Kaufman, B. A., Horvath, R., and Shoubridge, E. A. (2009) Mutation in TACO1, encoding a translational activator of COX I, results in cytochrome c oxidase deficiency and late-onset Leigh syndrome, *Nat Genet 41*, 833-837.

60. Richman, T. R., Spahr, H., Ermer, J. A., Davies, S. M., Viola, H. M., Bates, K. A., Papadimitriou, J., Hool, L. C., Rodger, J., Larsson, N. G., Rackham, O., and Filipovska, A. (2016) Loss of the RNA-binding protein TACO1 causes late-onset mitochondrial dysfunction in mice, *Nat Commun 7*, 11884.

61. Weber, E. R., and Dieckmann, C. L. (1990) Identification of the CBP1 polypeptide in mitochondrial extracts from Saccharomyces cerevisiae, *J Biol Chem 265*, 1594-1600.

62. Manna, S. (2015) An overview of pentatricopeptide repeat proteins and their applications, *Biochimie 113*, 93-99.

63. Rodel, G. (1986) Two yeast nuclear genes, CBS1 and CBS2, are required for translation of mitochondrial transcripts bearing the 5'-untranslated COB leader, *Curr Genet 11*, 41-45.

64. Krause-Buchholz, U., Barth, K., Dombrowski, C., and Rodel, G. (2004) Saccharomyces cerevisiae translational activator Cbs2p is associated with mitochondrial ribosomes, *Curr Genet 46*, 20-28.

65. Krause-Buchholz, U., Schobel, K., Lauffer, S., and Rodel, G. (2005) Saccharomyces cerevisiae translational activator Cbs1p is associated with translationally active mitochondrial ribosomes, *Biol Chem 386*, 407-415.

66. Chen, W., and Dieckmann, C. L. (1997) Genetic evidence for interaction between Cbp1 and specific nucleotides in the 5' untranslated region of mitochondrial cytochrome b mRNA in Saccharomyces cerevisiae, *Mol Cell Biol 17*, 6203-6211.

67. Rieger, K. J., Aljinovic, G., Lazowska, J., Pohl, T. M., and Slonimski, P. P. (1997) A novel nuclear gene, CBT1, essential for mitochondrial cytochrome b formation: terminal processing of mRNA and intron dependence, *Curr Genet 32*, 163-174.

68. Ellis, T. P., Schonauer, M. S., and Dieckmann, C. L. (2005) CBT1 interacts genetically with CBP1 and the mitochondrially encoded cytochrome b gene and is required to stabilize the mature cytochrome b mRNA of Saccharomyces cerevisiae, *Genetics 171*, 949-957.

69. Staples, R. R., and Dieckmann, C. L. (1994) Suppressor analyses of temperature-sensitive cbp1 strains of Saccharomyces cerevisiae: the product of the nuclear gene SOC1 affects mitochondrial cytochrome b mRNA post-transcriptionally, *Genetics 138*, 565-575.

70. Dieckmann, C. L., and Tzagoloff, A. (1985) Assembly of the mitochondrial membrane system. CBP6, a yeast nuclear gene necessary for synthesis of cytochrome b, *J Biol Chem 260*, 1513-1520.

71. Kronekova, Z., and Rodel, G. (2005) Organization of assembly factors Cbp3p and Cbp4p and their effect on bc(1) complex assembly in Saccharomyces cerevisiae, *Curr Genet 47*, 203-212.

72. Gruschke, S., Kehrein, K., Rompler, K., Grone, K., Israel, L., Imhof, A., Herrmann, J. M., and Ott, M. (2011) Cbp3-Cbp6 interacts with the yeast mitochondrial ribosomal tunnel exit and promotes cytochrome b synthesis and assembly, *J Cell Biol 193*, 1101-1114.

73. Gruschke, S., Rompler, K., Hildenbeutel, M., Kehrein, K., Kuhl, I., Bonnefoy, N., and Ott, M. (2012) The Cbp3-Cbp6 complex coordinates cytochrome b synthesis with bc(1) complex assembly in yeast mitochondria, *J Cell Biol 199*, 137-150.

74. Zeng, X., Hourset, A., and Tzagoloff, A. (2007) The Saccharomyces cerevisiae ATP22 gene codes for the mitochondrial ATPase subunit 6-specific translation factor, *Genetics 175*, 55-63.

75. Ellis, T. P., Helfenbein, K. G., Tzagoloff, A., and Dieckmann, C. L. (2004) Aep3p stabilizes the mitochondrial bicistronic mRNA encoding subunits 6 and 8 of the H+-translocating ATP synthase of Saccharomyces cerevisiae, *J Biol Chem 279*, 15728-15733.

76. Ziaja, K., Michaelis, G., and Lisowsky, T. (1993) Nuclear control of the messenger RNA expression for mitochondrial ATPase subunit 9 in a new yeast mutant, *J Mol Biol 229*, 909-916.

77. Ackerman, S. H., Gatti, D. L., Gellefors, P., Douglas, M. G., and Tzagoloff, A. (1991) ATP13, a nuclear gene of Saccharomyces cerevisiae essential for the expression of subunit 9 of the mitochondrial ATPase, *FEBS Lett 278*, 234-238.

78. Zeng, X., Barros, M. H., Shulman, T., and Tzagoloff, A. (2008) ATP25, a new nuclear gene of Saccharomyces cerevisiae required for expression and assembly of the Atp9p subunit of mitochondrial ATPase, *Mol Biol Cell 19*, 1366-1377.

79. Manthey, G. M., Przybyla-Zawislak, B. D., and McEwen, J. E. (1998) The Saccharomyces cerevisiae Pet309 protein is embedded in the mitochondrial inner membrane, *Eur J Biochem 255*, 156-161.

80. Zamudio-Ochoa, A., Camacho-Villasana, Y., Garcia-Guerrero, A. E., and Perez-Martinez, X. (2014) The Pet309 pentatricopeptide repeat motifs mediate efficient binding to the mitochondrial COX1 transcript in yeast, *RNA Biol 11*, 953-967.

81. Siep, M., van Oosterum, K., Neufeglise, H., van der Spek, H., and Grivell, L. A. (2000) Mss51p, a putative translational activator of cytochrome c oxidase subunit-1 (COX1) mRNA, is required for synthesis of Cox1p in Saccharomyces cerevisiae, *Curr Genet 37*, 213-220.

82. Soto, I. C., and Barrientos, A. (2016) Mitochondrial Cytochrome c Oxidase Biogenesis Is Regulated by the Redox State of a Heme-Binding Translational Activator, *Antioxid Redox Signal 24*, 281-298.

83. Roloff, G. A., and Henry, M. F. (2015) Mam33 promotes cytochrome c oxidase subunit I translation in Saccharomyces cerevisiae mitochondria, *Mol Biol Cell 26*, 2885-2894.

84. Naithani, S., Saracco, S. A., Butler, C. A., and Fox, T. D. (2003) Interactions among COX1, COX2, and COX3 mRNA-specific translational activator proteins on the inner surface of the mitochondrial inner membrane of Saccharomyces cerevisiae, *Mol Biol Cell 14*, 324-333.

85. M.E. Sanchirico, Understanding Mitochondrial Biogenesis through Gene Relocation, Doctoral Thesis, University of Massachusetts, Amherst, 1998.

86. Papadopoulou, B., Dekker, P., Blom, J., and Grivell, L. A. (1990) A 40 kd protein binds specifically to the 5'-untranslated regions of yeast mitochondrial mRNAs, *Embo J 9*, 4135-4143.

87. Dekker, P. J., Stuurman, J., van Oosterum, K., and Grivell, L. A. (1992) Determinants for binding of a 40 kDa protein to the leaders of yeast mitochondrial mRNAs, *Nucleic Acids Res 20*, 2647-2655.

88. Elzinga, S. D., Bednarz, A. L., van Oosterum, K., Dekker, P. J., and Grivell, L. A. (1993) Yeast mitochondrial NAD(+)-dependent isocitrate dehydrogenase is an RNA-binding protein, *Nucleic Acids Res 21*, 5328-5331.

89. Lin, A. P., Demeler, B., Minard, K. I., Anderson, S. L., Schirf, V., Galaleldeen, A., and McAlister-Henn, L. (2011) Construction and analyses of tetrameric forms of yeast NAD+-specific isocitrate dehydrogenase, *Biochemistry 50*, 230-239.

90. de Jong, L., Elzinga, S. D., McCammon, M. T., Grivell, L. A., and van der Spek, H. (2000) Increased synthesis and decreased stability of mitochondrial translation products in yeast as a result of loss of mitochondrial (NAD(+))-dependent isocitrate dehydrogenase, *FEBS Lett 483*, 62-66.

91. Herbert, C. J., Golik, P., and Bonnefoy, N. (2013) Yeast PPR proteins, watchdogs of mitochondrial gene expression, *RNA Biol 10*, 1477-1494.

92. Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., Caboche, M., Debast, C., Gualberto, J., Hoffmann, B., Lecharny, A., Le Ret, M., Martin-Magniette, M. L., Mireau, H., Peeters, N., Renou, J. P., Szurek, B., Taconnat, L., and Small, I. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis, *Plant Cell 16*, 2089-2103.

93. Schmitz-Linneweber, C., and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression, *Trends Plant Sci 13*, 663-670.

94. Delannoy, E., Stanley, W. A., Bond, C. S., and Small, I. D. (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles, *Biochem Soc Trans 35*, 1643-1647.

95. Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., Wang, Z., Long, J., He, J., Wang, H. W., Wang, J., Zhu, J. K., Shi, Y., and Yan, N. (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins, *Nature 504*, 168-171.

96. Ke, J., Chen, R. Z., Ban, T., Zhou, X. E., Gu, X., Tan, M. H., Chen, C., Kang, Y., Brunzelle, J. S., Zhu, J. K., Melcher, K., and Xu, H. E. (2013) Structural basis for RNA recognition by a dimeric PPR-protein complex, *Nat Struct Mol Biol 20*, 1377-1382.

97. Filipovska, A., and Rackham, O. (2013) Pentatricopeptide repeats: modular blocks for building RNA-binding proteins, *RNA Biol 10*, 1426-1432.

98. Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y. S., Bond, C. S., and Small, I. (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins, *PLoS Genet 8*, e1002910.

99. Yagi, Y., Hayashi, S., Kobayashi, K., Hirayama, T., and Nakamura, T. (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants, *PLoS One 8*, e57286.

100. Yagi, Y., Nakamura, T., and Small, I. (2014) The potential for manipulating RNA with pentatricopeptide repeat proteins, *Plant J 78*, 772-782.

101. Kobayashi, T., Yagi, Y., and Nakamura, T. (2016) Development of Genome Engineering Tools from Plant-Specific PPR Proteins Using Animal Cultured Cells, *Methods Mol Biol 1469*, 147-155.

102. Coquille, S., Filipovska, A., Chia, T., Rajappa, L., Lingford, J. P., Razif, M. F., Thore, S., and Rackham, O. (2014) An artificial PPR scaffold for programmable RNA recognition, *Nat Commun 5*, 5729.

103. Lipinski, K. A., Puchta, O., Surendranath, V., Kudla, M., and Golik, P. (2011) Revisiting the yeast PPR proteins--application of an Iterative Hidden Markov Model algorithm reveals new members of the rapidly evolving family, *Mol Biol Evol 28*, 2935-2948.

104. Geier, B. M., Schagger, H., Ortwein, C., Link, T. A., Hagen, W. R., Brandt, U., and Von Jagow, G. (1995) Kinetic properties and ligand binding of the eleven-subunit cytochrome-c oxidase from Saccharomyces cerevisiae isolated with a novel large-scale purification method, *Eur J Biochem 227*, 296-302.

105. Strecker, V., Kadeer, Z., Heidler, J., Cruciat, C. M., Angerer, H., Giese, H., Pfeiffer, K., Stuart, R. A., and Wittig, I. (2016) Supercomplex-associated Cox26 protein binds to cytochrome c oxidase, *Biochim Biophys Acta 1863*, 1643-1652.

106. Ebner, E., Mennucci, L., and Schatz, G. (1973) Mitochondrial assembly in respiration-deficient mutants of Saccharomyces cerevisiae. I. Effect of nuclear mutations on mitochondrial protein synthesis, *J Biol Chem 248*, 5360-5368.

107. Poutre, C. G., and Fox, T. D. (1987) PET111, a Saccharomyces cerevisiae nuclear gene required for translation of the mitochondrial mRNA encoding cytochrome c oxidase subunit II, *Genetics 115*, 637-647.

108. Strick, C. A., and Fox, T. D. (1987) Saccharomyces cerevisiae positive regulatory gene PET111 encodes a mitochondrial protein that is translated from an mRNA with a long 5' leader, *Mol Cell Biol 7*, 2728-2734.

109. Costanzo, M. C., and Fox, T. D. (1988) Specific translational activation by nuclear gene products occurs in the 5' untranslated leader of a yeast mitochondrial mRNA, *Proc Natl Acad Sci U S A 85*, 2677-2681.

110. Rodel, G., and Fox, T. D. (1987) The yeast nuclear gene CBS1 is required for translation of mitochondrial mRNAs bearing the cob 5' untranslated leader, *Mol Gen Genet 206*, 45-50.

111. Mulero, J. J., and Fox, T. D. (1993) PET111 acts in the 5'-leader of the Saccharomyces cerevisiae mitochondrial COX2 mRNA to promote its translation, *Genetics 133*, 509-516.

112. Mulero, J. J., and Fox, T. D. (1993) Alteration of the Saccharomyces cerevisiae COX2 mRNA 5'-untranslated leader by mitochondrial gene replacement and functional interaction with the translational activator protein PET111, *Mol Biol Cell 4*, 1327-1335.

113. Dunstan, H. M., Green-Willms, N. S., and Fox, T. D. (1997) In vivo analysis of Saccharomyces cerevisiae COX2 mRNA 5'-untranslated leader functions in mitochondrial translation initiation and translational activation, *Genetics 147*, 87-100.

114. Bonnefoy, N., Bsat, N., and Fox, T. D. (2001) Mitochondrial translation of Saccharomyces cerevisiae COX2 mRNA is controlled by the nucleotide sequence specifying the pre-Cox2p leader peptide, *Mol Cell Biol 21*, 2359-2372.

115. Costanzo, M. C., Bonnefoy, N., Williams, E. H., Clark-Walker, G. D., and Fox, T. D. (2000) Highly diverged homologs of Saccharomyces cerevisiae mitochondrial mRNA-specific translational activators have orthologous functions in other budding yeasts, *Genetics 154*, 999-1012.

116. Green-Willms, N. S., Butler, C. A., Dunstan, H. M., and Fox, T. D. (2001) Pet111p, an inner membrane-bound translational activator that limits expression of the Saccharomyces cerevisiae mitochondrial gene COX2, *J Biol Chem 276*, 6392-6397.

117. Sanchirico, M. E., Fox, T. D., and Mason, T. L. (1998) Accumulation of mitochondrially synthesized Saccharomyces cerevisiae Cox2p and Cox3p depends on targeting information in untranslated portions of their mRNAs, *Embo J 17*, 5796-5804.

118. Rodeheffer, M. S., Boone, B. E., Bryan, A. C., and Shadel, G. S. (2001) Nam1p, a protein involved in RNA processing and translation, is coupled to transcription through an interaction with yeast mitochondrial RNA polymerase, *J Biol Chem 276*, 8616-8622.

119. Diekert K, I.P.M. de Kroon A, Kispal G, & Lill R (2001) Isolation and subfractionation of mitochondria from the yeast Saccharomyces cerevisiae. *Methods in Cell Biology*, eds Pon LA, Schon EA (Academic Press, San Diego), Vol 65, pp 37-51.

120. Meisinger, C., Sommer, T., and Pfanner, N. (2000) Purification of Saccharomcyes cerevisiae mitochondria devoid of microsomal and cytosolic contaminations, *Anal Biochem 287*, 339-342.

121. Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations, *J Bacteriol 153*, 163-168.

122. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels, *Anal Chem 68*, 850-858.

123. Dobbins, A. T., George, M., Jr., Basham, D. A., Ford, M. E., Houtz, J. M., Pedulla, M. L., Lawrence, J. G., Hatfull, G. F., and Hendrix, R. W. (2004) Complete genomic sequence of the virulent Salmonella bacteriophage SP6, *J Bacteriol 186*, 1933-1944.

124. Coruzzi, G., Bonitz, S. G., Thalenfeld, B. E., and Tzagoloff, A. (1981) Assembly of the mitochondrial membrane system. Analysis of the nucleotide sequence and transcripts in the oxi1 region of yeast mitochondrial DNA, *J Biol Chem 256*, 12780-12787.

125. Mossmann, D., Meisinger, C., and Vogtle, F. N. (2012) Processing of mitochondrial presequences, *Biochim Biophys Acta 1819*, 1098-1106.

126. Vogtle, F. N., Wortelkamp, S., Zahedi, R. P., Becker, D., Leidhold, C., Gevaert, K., Kellermann, J., Voos, W., Sickmann, A., Pfanner, N., and Meisinger, C. (2009) Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability, *Cell 139*, 428-439.

127. Vogtle, F. N., Prinz, C., Kellermann, J., Lottspeich, F., Pfanner, N., and Meisinger, C. (2011) Mitochondrial protein turnover: role of the precursor intermediate peptidase Oct1 in protein stabilization, *Mol Biol Cell 22*, 2135-2143.

128. Fukasawa, Y., Tsuji, J., Fu, S. C., Tomii, K., Horton, P., and Imai, K. (2015) MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites, *Mol Cell Proteomics 14*, 1113-1126.

129. Gakh, O., Cavadini, P., and Isaya, G. (2002) Mitochondrial processing peptidases, *Biochim Biophys Acta 1592*, 63-77.

130. Pappin, D. J., Hojrup, P., and Bleasby, A. J. (1993) Rapid identification of proteins by peptide-mass fingerprinting, *Curr Biol 3*, 327-332.

131. Nilsen TW (2014) RNase Footprinting to Map Sites of RNA–Protein Interactions. *Cold Spring Harbor Protocols* 2014(6):pdb.prot080788.

132. Meador, J., 3rd, Cannon, B., Cannistraro, V. J., and Kennell, D. (1990) Purification and characterization of Escherichia coli RNase I. Comparisons with RNase M, *Eur J Biochem 187*, 549-553.

133. Haffter, P., McMullin, T. W., and Fox, T. D. (1990) A genetic link between an mRNA-specific translational activator and the translation system in yeast mitochondria, *Genetics 125*, 495-503.

134. Haffter, P., McMullin, T. W., and Fox, T. D. (1991) Functional interactions among two yeast mitochondrial ribosomal proteins and an mRNA-specific translational activator, *Genetics 127*, 319-326.

135. Haffter, P., and Fox, T. D. (1992) Suppression of carboxy-terminal truncations of the yeast mitochondrial mRNA-specific translational activator PET122 by mutations in two new genes, MRP17 and PET127, *Mol Gen Genet 235*, 64-73.

136. Gavin, A. C., Aloy, P., Grandi, P., Krause, R., Boesche, M., Marzioch, M., Rau, C., Jensen, L. J., Bastuck, S., Dumpelfeld, B., Edelmann, A., Heurtier, M. A., Hoffman, V., Hoefert, C., Klein, K., Hudak, M., Michon, A. M., Schelder, M., Schirle, M., Remor, M., Rudi, T., Hooper, S., Bauer, A., Bouwmeester, T., Casari, G., Drewes, G., Neubauer, G., Rick, J. M., Kuster, B., Bork, P., Russell, R. B., and Superti-Furga, G. (2006) Proteome survey reveals modularity of the yeast cell machinery, *Nature 440*, 631-636.

137. Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A. P., Punna, T., Peregrin-Alvarez, J. M., Shales, M., Zhang, X., Davey, M., Robinson, M. D., Paccanaro, A., Bray, J. E., Sheung, A., Beattie, B., Richards, D. P., Canadien, V., Lalev, A., Mena, F., Wong, P., Starostine, A., Canete, M. M., Vlasblom, J., Wu, S., Orsi, C., Collins, S. R., Chandran, S., Haw, R., Rilstone, J. J., Gandi, K., Thompson, N. J., Musso, G., St Onge, P., Ghanny, S., Lam, M. H., Butland, G., Altaf-Ul, A. M., Kanaya, S., Shilatifard, A., O'Shea, E., Weissman, J. S., Ingles, C. J., Hughes, T. R., Parkinson, J., Gerstein, M., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae, *Nature 440*, 637-643.

138. Babu, M., Vlasblom, J., Pu, S., Guo, X., Graham, C., Bean, B. D., Burston, H. E., Vizeacoumar, F. J., Snider, J., Phanse, S., Fong, V., Tam, Y. Y., Davey, M., Hnatshak, O., Bajaj, N., Chandran, S., Punna, T., Christopolous, C., Wong, V., Yu, A., Zhong, G., Li, J., Stagljar, I., Conibear, E., Wodak, S. J., Emili, A., and Greenblatt, J. F. (2012) Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae, *Nature 489*, 585-589.

139. Lisowsky, T. (1990) Molecular analysis of the mitochondrial transcription factor mtf2 of Saccharomyces cerevisiae, *Mol Gen Genet 220*, 186-190.

140. Lisowsky, T., and Michaelis, G. (1989) Mutations in the genes for mitochondrial RNA polymerase and a second mitochondrial transcription factor of Saccharomyces cerevisiae, *Mol Gen Genet 219*, 125-128.

141. Rodeheffer, M. S., and Shadel, G. S. (2003) Multiple interactions involving the amino-terminal domain of yeast mtRNA polymerase determine the efficiency of mitochondrial protein synthesis, *J Biol Chem 278*, 18695-18701.

142. Asher, E. B., Groudinsky, O., Dujardin, G., Altamura, N., Kermorgant, M., and Slonimski, P. P. (1989) Novel class of nuclear genes involved in both mRNA splicing and protein synthesis in Saccharomyces cerevisiae mitochondria, *Mol Gen Genet 215*, 517-528.

143. Lisowsky, T., Riemen, G., and Michaelis, G. (1990) Change of serine309 into proline causes temperature sensitivity of the nuclear NAM1/MTF2 gene product for yeast mitochondria, *Nucleic Acids Res 18*, 7163.

144. Groudinsky, O., Bousquet, I., Wallis, M. G., Slonimski, P. P., and Dujardin, G. (1993) The NAM1/MTF2 nuclear gene product is selectively required for the stability and/or processing of mitochondrial transcripts of the atp6 and of the mosaic, cox1 and cytb genes in Saccharomyces cerevisiae, *Mol Gen Genet 240*, 419-427.

145. Wallis, M. G., Groudinsky, O., Slonimski, P. P., and Dujardin, G. (1994) The NAM1 protein (NAM1p), which is selectively required for cox1, cytb and atp6 transcript processing/stabilisation, is located in the yeast mitochondrial matrix, *Eur J Biochem 222*, 27-32.

146. Bryan, A. C., Rodeheffer, M. S., Wearn, C. M., and Shadel, G. S. (2002) Sls1p is a membrane-bound regulator of transcription-coupled processes involved in Saccharomyces cerevisiae mitochondrial gene expression, *Genetics 160*, 75-82.

147. Shadel, G. S. (2004) Coupling the mitochondrial transcription machinery to human disease, *Trends Genet 20*, 513-519.

148. Vartak, R., Porras, C. A., and Bai, Y. (2013) Respiratory supercomplexes: structure, function and assembly, *Protein Cell 4*, 582-590.

149. Small, I. D., and Peeters, N. (2000) The PPR motif - a TPR-related motif prevalent in plant organellar proteins, *Trends Biochem Sci 25*, 46-47.

150. Ringel, R., Sologub, M., Morozov, Y. I., Litonin, D., Cramer, P., and Temiakov, D. (2011) Structure of human mitochondrial RNA polymerase, *Nature 478*, 269-273.

151. Puchta, O., Lubas, M., Lipinski, K. A., Piatkowski, J., Malecki, M., and Golik, P. (2010) DMR1 (CCM1/YGR150C) of Saccharomyces cerevisiae encodes an RNA-binding protein from the pentatricopeptide repeat family required for the maintenance of the mitochondrial 15S ribosomal RNA, *Genetics 184*, 959-973.

152. Markov, D. A., Savkina, M., Anikin, M., Del Campo, M., Ecker, K., Lambowitz, A. M., De Gnore, J. P., and McAllister, W. T. (2009) Identification of proteins associated with the yeast mitochondrial RNA polymerase by tandem affinity purification, *Yeast 26*, 423-440.

153. Bonitz, S. G., Homison, G., Thalenfeld, B. E., Tzagoloff, A., and Nobrega, F. G. (1982) Assembly of the mitochondrial membrane system. Processing of the apocytochrome b precursor RNAs in Saccharomyces cerevisiae D273-10B, *J Biol Chem 257*, 6268-6274.

154. Dieckmann, C. L., Pape, L. K., and Tzagoloff, A. (1982) Identification and cloning of a yeast nuclear gene (CBP1) involved in expression of mitochondrial cytochrome b, *Proc Natl Acad Sci U S A 79*, 1805-1809.

155. Dieckmann, C. L., Homison, G., and Tzagoloff, A. (1984) Assembly of the mitochondrial membrane system. Nucleotide sequence of a yeast nuclear gene (CBP1) involved in 5' end processing of cytochrome b pre-mRNA, *J Biol Chem 259*, 4732-4738.

156. Dieckmann, C. L., Koerner, T. J., and Tzagoloff, A. (1984) Assembly of the mitochondrial membrane system. CBP1, a yeast nuclear gene involved in 5' end processing of cytochrome b pre-mRNA, *J Biol Chem 259*, 4722-4731.

157. Staples, R. R., and Dieckmann, C. L. (1993) Generation of temperature-sensitive cbp1 strains of Saccharomyces cerevisiae by PCR mutagenesis and in vivo recombination: characteristics of the mutant strains imply that CBP1 is involved in stabilization and processing of cytochrome b pre-mRNA, *Genetics 135*, 981-991.

158. Mittelmeier, T. M., and Dieckmann, C. L. (1990) CBP1 function is required for stability of a hybrid cob-oli1 transcript in yeast mitochondria, *Curr Genet 18*, 421-428.

159. Dieckmann, C. L., and Mittelmeier, T. M. (1987) Nuclearly-encoded CBP1 interacts with the 5' end of mitochondrial cytochrome b pre-mRNA, *Curr Genet 12*, 391-397.

160. Mittelmeier, T. M., and Dieckmann, C. L. (1993) In vivo analysis of sequences necessary for CBP1-dependent accumulation of cytochrome b transcripts in yeast mitochondria, *Mol Cell Biol 13*, 4203-4213.

161. Chen, W., and Dieckmann, C. L. (1994) Cbp1p is required for message stability following 5'-processing of COB mRNA, *J Biol Chem 269*, 16574-16578.

162. Mittelmeier, T. M., and Dieckmann, C. L. (1995) In vivo analysis of sequences required for translation of cytochrome b transcripts in yeast mitochondria, *Mol Cell Biol 15*, 780-789.

163. Michaelis, U., Korte, A., and Rodel, G. (1991) Association of cytochrome b translational activator proteins with the mitochondrial membrane: implications for cytochrome b expression in yeast, *Mol Gen Genet 230*, 177-185.

164. Defontaine, A., Lecocq, F. M., and Hallet, J. N. (1991) A rapid miniprep method for the preparation of yeast mitochondrial DNA, *Nucleic Acids Res 19*, 185.

165. He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., McAllister, W. T., and Durbin, R. K. (1997) Rapid mutagenesis and purification of phage RNA polymerases, *Protein Expr Purif 9*, 142-151.

APPENDIX

A. Materials, methods, and results of supplemental experiments

A.1. *Construction of plasmids containing the 5'-UTR regions of COX2, COB, and ATP8/6.* Yeast BY4743 cells were grown in YP medium supplemented with 2% galactose and mitochondria were purified as described in section I.C.3. Mitochondrial DNA was isolated using a protocol adapted from Defontaine et al. (*164*). In brief, 100 μl of settled purified mitochondria were lysed by gentle tumbling in 400 μl of buffer containing 50 mM Tris·HCl (pH 7.8), 100 mM NaCl, 10 mM EDTA, and 1% SDS for 30 min at room temperature. 500 μl phenol:choloroform:isoamyl alcohol was added, mixed well, and incubated for 5 min with rotation at room temperature. The emulsion was centrifuged (16,000 × g, 5 min), the aqueous layer was transferred to a new tube, and the extraction was repeated with an equal volume of chloroform. The final aqueous layer was collected and diluted to a final concentration of 10% ammonium acetate and 2.5 volumes of ethanol. DNA was precipitated at -20 °C for 1 h, pelleted by centrifugation for 15 min (20,000 × g, 4 °C), washed once with ethanol, aspirated, dissolved in 10 mM Tris·HCl, pH 8.5, and quantitated by UV/Vis spectrophotometry.

To construct a plasmid carrying a *COX2* related sequence, a fragment corresponding to the region from -538 to +37 (relative to the A1 of the translational start site), was PCR-amplified by Phusion DNA polymerase (Thermo Fisher) using the isolated mtDNA as a template and primers F-

COX2 5′-UTR (-538 to -516) and R *COX2* 5′-UTR +37 to +15. A two-step annealing phase (55 °C and 50 °C, 10 sec each) and an extended elongation phase (2 min) at a temperature of 60 °C was used. The amplicon was purified by the GeneJET PCR purification kit (Thermo Scientific) and ligated into the pT7Blue vector (Novagen) in accordance with the manufacturer's protocol, resulting in plasmid pJJ10.

For the plasmid containing a region of mtDNA related to the *COB* gene*,* the same protocol was followed, except that the primers were Scmt35402-41 and Primer *COB* Exon1 reverse-2. A 1307 nt fragment was amplified including the nucleotides from -1138 to +169 relative to the A1 of the translational start codon. The fragment was ligated into the pT7Blue vector to create plasmid pJJ05.

For the plasmid containing a fragment corresponding to the *ATP8/6* 5′-UTR region, the DNA amplification was performed as above utilizing primers F-ATP8 5′-UTR (-983 to -961) and R-ATP8 5′-UTR (-5 to +17). The amplicon contained a region of mitochondrial DNA from -983 to +17 relative to the A1 of the translational start site of *ATP8*. The amplicon was ligated into the pT7Blue vector to yield plasmid pJJ08. To verify the accuracy of the PCR amplifications, the sequence of the inserts in each plasmid was determined (GENEWIZ) and found to match the corresponding sequences in the S288c yeast mtDNA (accession NC_001224).

A.2. *Expression and purification of Δ57Pet111p and Δ79Pet111p and activity assessment of Δ79Pet111p.* In an attempt to improve the stability of Pet111p in solution, two alternative forms, Δ57Pet111p and Δ79Pet111p, were created in addition to the mature form of the protein, Δ34Pet111p. For this, plasmid pGD1 was used as a template and site-directed mutagenesis was performed with the QuikChange II kit (Agilent) and mutagenic primers Pet111del57 and Pet111del57c or Pet111del79 and Pet111del79c. This resulted in plasmids pJJ11 (Δ57Pet111p) and pJJ12 (Δ79Pet111p), respectively. The sequence of the inserts in each plasmid was verified by DNA sequencing (GENEWIZ). pJJ11 and pJJ12 were transformed independently into *E. coli* XJb(DE3) cells (Zymo Research) and both Δ57Pet111p and Δ79Pet111p were expressed and purified as described in section I.C.4 with the following modifications: $MgCl_2$ was not added to the buffers during heparin affinity purification and gel filtration was not performed. The purification resulted in a homogeneous Δ79Pet111p, whereas a high purity of Δ57Pet111p was not achieved (Figure 38). A trial EMSA was performed to test the activity of Δ79Pet111p with the RNA probe, *COX2*-1-54 (see Table 4 for the RNA sequence), following the protocol in section I.C.6.

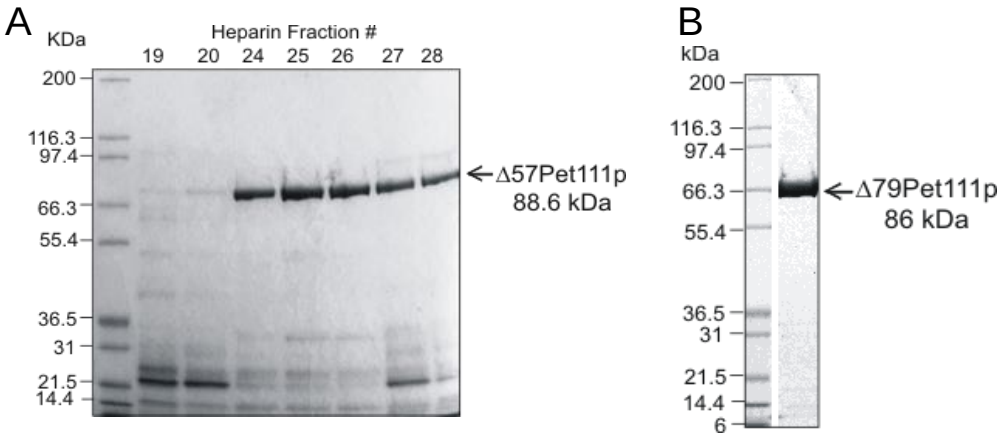

Figure 38. Results of heparin affinity purification of Δ57Pet111p and Δ79Pet111p. Images of Coomassie-stained LDS 4-12% polyacrylamide gels showing the purity of Δ57Pet111p in heparin column fractions (panel A) and the purity of a preparation of Δ79Pet111p (Panel B).

Although both proteins could be purified, Δ57Pet111p appeared to be less stable than the other two forms of Pet111p. The impurities and/or Δ57Pet111p degradation products that co-eluted with Δ57Pet111p in heparin affinity purification are indicative of an improperly folded protein that can aggregate with other proteins in the lysate (Figure 38A). However, Δ79Pet111p was purified to homogeneity and appeared to be stable at high concentrations (Figure 38B). This preparation of Δ79Pet111p was assayed for RNA binding activity by EMSA analysis (Figure 39), which was performed similarly to the experiment shown in Figure 11A.

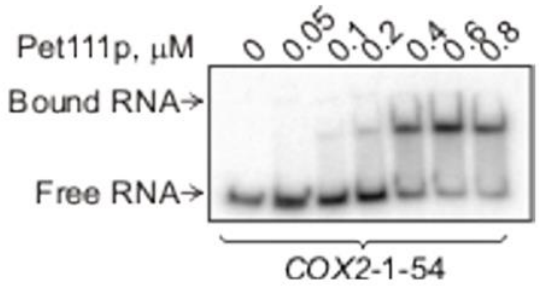

Figure 39. Δ79Pet111p interacts with *COX2*-1-54. 5′-[^{32}P]-labeled *COX2*-1-54 (0.1 μM) was incubated with varying concentrations of Δ79Pet111p and resolved in an 8% polyacrylamide gel under native conditions. The bands corresponding to unbound and bound RNA (as indicated) were visualized by phosphor imaging. During binding, competitor tRNA was present in the solutions at two-fold molar excess over the protein.

As shown in Figure 39, Δ79Pet111p could bind to the *COX2* 5′-UTR with efficiency similar to that of Δ34Pet111p (compare with Figure 11A). This signifies that the integrity of the protein was preserved upon the deletion of an additional 45 amino acids from the N-terminus.

A.3. *Expression and purification of SP6 and Rpo41p.*

SP6 RNA polymerase: Hexahistidine tagged SP6 RNA polymerase was expressed from the plasmid pBH176 (*165*) in BL21 *E. coli* cells. The cells were grown in 0.5 L of LB medium supplemented with 100 μg/ml ampicillin at 37 °C to an OD_{600} of 1.0. Cells were induced with 1 mM IPTG at 37 °C for 3 h. After induction, the cells were harvested by centrifugation (3,000 × g, 10 min,

4 °C) to yield a 3.6 g wet cell pellet. The cells were resuspended in 40 ml lysis buffer (40 mM HEPES·NaOH, 300 mM NaCl, 1 mM PMSF, 5 mM β-mercaptoethanol, 15 mM imidazole) and sonicated with a F500 sonic dismembrator (Fisher Scientific) on ice using 6 s pulses intermitted with 4 s pauses for a total of 5 min of the “on” time. Insoluble matter was removed by centrifugation (20,000 × g, 25 min, 4 °C). The concentration of NaCl in the supernatant was adjusted to 800 mM and the solution was passed through a column packed with 2 ml Ni-IDA agarose beads (Gold Biotechnology) equilibrated with wash buffer (40 mM HEPES·NaOH pH 7.0, 800 mM NaCl, 15 mM imidazole). The beads were washed by passing 45 ml of wash buffer through the column and the protein was released with 2.5 ml of elution buffer (20 mM HEPES·NaOH pH 7.0, 800 mM NaCl, 5% glycerol, 250 mM imidazole, 5 mM β-mercaptoethanol). The eluate was diluted with 20 mM HEPES·NaOH pH 7.0, 5% glycerol, 5 mM β-mercaptoethanol to reduce NaCl concentration to 300 mM) and loaded on a 16 × 25 mm HiTrap Heparin HP column (GE Healthcare Life Sciences). The protein was eluted from the column with a linear NaCl concentration gradient (150 mM to 1.2 M) in a buffer containing 40 mM HEPES·NaOH pH 7.0, 10 mM $MgCl_2$, 10% glycerol, and 5 mM β-mercaptoethanol. Fractions containing SP6 were combined, analyzed by SDS PAGE for purity (Figure 40A), concentrated on an Ultracel®-50K centrifugal filter, diluted 50% with glycerol, and frozen at -80 °C.

Rpo41p: The plasmid pEK8 encoding an N-terminal hexahistidine fusion of the mature form of Rpo41p was provided by Michael Anikin (unpublished). *E. coli* BL21(DE3) codon plus RIPL cells (Agilent) carrying pEK8 were grown at 37 °C in 4 L of LB medium supplemented with 100 µg/ml ampicillin to an OD_{600} of 0.55. The cells were induced with 0.2 mM IPTG at 16 °C for 20 h. The culture was centrifuged to yield a 16 g pellet and the cells were resuspended in 150 ml of lysis buffer containing 40 mM Tris·HCl pH 7.9, 300 mM NaCl, 10 mM imidazole, 1 mM PMSF, 5% glycerol, and 5 mM β-mercaptoethanol. The suspension was treated by ultrasound as in the previous section and the lysate was cleared by centrifugation (20,000 × g, 25 min, 4 °C). The supernatant was passed through a column packed with 2 ml of Ni-IDA agarose beads (Gold Biotechnology) equilibrated with wash buffer (40 mM Tris·HCl pH 7.9, 300 mM NaCl, 15 mM imidazole). The beads were first washed with a high salt buffer (20 ml of wash buffer supplemented with 1.5 M NaCl) and then with 20 ml of wash buffer. The protein was released with five sequential 2 ml elutions with elution buffer (40 mM Tris·HCl pH 7.9, 300 mM NaCl, 5% glycerol, 200 mM imidazole, 5 mM β-mercaptoethanol). The fractions that contained the bulk of Rpo41p were loaded on a 8 × 7.5 TSKgel Heparin-5PW (Tosoh Bioscience) column. The protein was eluted from the column with a linear NaCl concentration gradient (100 mM to 800 mM) in a buffer containing 30 mM Tris·HCl pH 7.9, 5% glycerol, and 5 mM β-mercaptoethanol. The peak fraction with a high concentration of Rpo41p (in a volume of 2 ml) was loaded on a HiLoad 16/600 Superdex 200 PG column

and the protein was eluted with gel filtration buffer (60 mM Tris·HCl pH 7.9, 300 mM NaCl 5% glycerol, 5 mM β-mercaptoethanol). The fractions containing pure Rpo41p (Figure 40B) were pooled, concentrated, aliquoted into small volumes, and frozen at -80 °C.

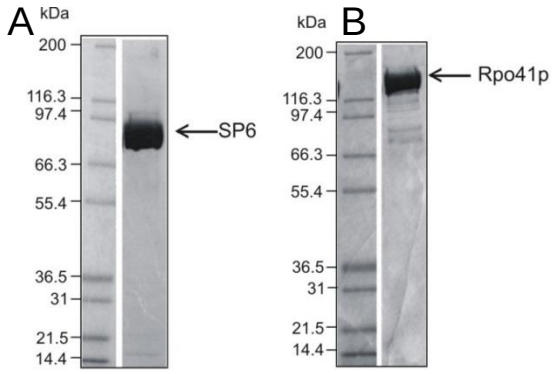

Figure 40. Purification of RNA polymerases. Images of Coomassie-stained LDS 4-12% polyacrylamide gels showing the typical products of purification of SP6 RNA polymerase (panel A) and Rpo41p (panel B).

A.4. *Preliminary crystallization trials with the Pet111p/COX2-5-21 complex.* Pet111p was purified as described in section I.C.4 except that the gel filtration step was omitted. In a total volume of 1.08 ml, the protein (65.5 μM) was combined with a 30% molar excess of *COX2*-5-21 synthetic RNA (IDT DNA; see Table 8 for the sequence) containing the 5′-UTR Pet111p binding site. To reduce the concentration of NaCl to 270 mM, the solution was diluted with 1.5 ml of a buffer containing 20 mM HEPES·NaOH pH 7.0, 10 mM $MgCl_2$, 5% glycerol, and 5 mM β-mercaptoethanol. The resulting mixture was incubated for 10 min at 22 °C and then loaded on a HiLoad Superdex 200 16/600 preparative column (GE Life Sciences). The complex was eluted with 20 mM

HEPES·NaOH pH 7.0, 250 mM NaCl, 10 mM $MgCl_2$, 5% glycerol, 5 mM β-mercaptoethanol (Figure 41A). Fractions containing the target Pet111p/*COX2*-5-21 complex were pooled and concentrated to 580 µl on a centrifugal filter (10 kDa cutoff). The complex in the concentrate was quantitated to be 6.7 mg/ml by UV/Vis spectroscopy (Figure 41B). The complex was then distributed into 25 µl aliquots, frozen in liquid nitrogen, and stored at -80 °C until further use.

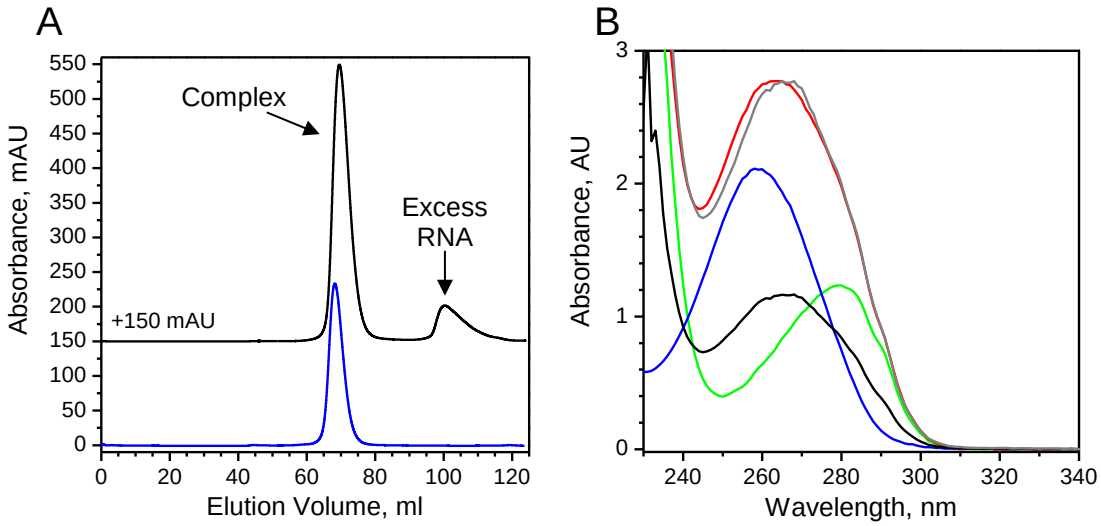

Figure 41. Preparation of the Pet111p/*COX2*-5-21 complex. A. A chromatogram showing results of gel filtration of the Pet111p/*COX2*-5-21 complex. The blue trace corresponds to free Pet111p and the black trace shows the separation of the complex from the excess of *COX2*-5-21. In the black trace, note the increase in intensity of the Pet111p peak due to the co-elution of the complexed RNA. B. The concentration and the composition of the purified complex were analyzed by UV/Vis spectrophotometry. The spectra of the free protein (at 1 mg/ml; green) and an equimolar solution of *COX2*-5-21 (blue) were added to generate a theoretical spectrum of a stoichiometric complex (red) at a concentration of 1 mg/ml (by the protein).

The purified complex was quantitated by comparing its spectrum (black) to the theoretical trace. The grey trace shows absorbance of the purified complex normalized to the theoretical curve.

The UV/Vis spectrum of the purified complex followed closely the theoretical trace corresponding to an equimolar Pet111p/*COX2*-5-21 complex, suggesting that no significant loss of the RNA occurred during the preparation (Figure 41B). The complex was used in screening for crystallization conditions with the Crystal Screen, Crystal Screen 2, and Natrix kits (Hampton Research). The screening was performed at 20 °C using the sitting drop vapor diffusion method. To form the crystallization drops, 1 μl of the complex solution was mixed with 1 μl of precipitants. The drops were sealed and incubated over 0.5 ml of corresponding precipitant solutions. Although thin needle-like crystals were observed under a number of crystallization conditions, 3-dimentional crystals suitable for structural analysis were not obtained in the screening so far. Further optimization of the stability of the protein in solution and the length of the RNA is needed in order to improve the quality of the crystals.

B. Supplemental information

Table 8. Oligonucleotides used in this section of the study.

Oligonucleotide	Sequence (5′ to 3′)
DNA Oligonucleotides	
F-*COX2* 5′-UTR (-538 to -516)	GACCCCGAAGGAGTATAAATAA
R *COX2*5′UTR +37 to +15	GGTGTTGGTACATCATTCATAA
Scmt35402-41	CCTCTTCATGTTGATAATATCGGT TCGATTCCGATTAAGG
Primer *COB* Exon1 reverse-2	AGATGAATAATGCATAGCCA
F-*ATP8* 5′-UTR (-983 to -961)	CCAGCTGTACAATCTTAAGTTA
R-*ATP8* 5′-UTR (-5 to +17)	GGAACTAATTGTGGCATTTTTA
Pet111del57	CATCATCACCATCATCATTCTTCAA CAATAAGCGGTG
Pet111del57c	CACCGCTTATTGTTGAAGAATGAT GATGGTGATGATG
Pet111del79	CATCATCACCATCATCATAGCCCC AGCAGCTTAGATATTC
Pet111del79c	GAATATCTAAGCTGCTGGGGCTAT GATGATGGTGATGATG
RNA Oligonucleotides	
COX2-5-21	UUAACAUAUUAUAAAUA

ABBREVIATIONS

TCA: the citric acid cycle

OXPHOS: oxidative phosphorylation

ROS: reactive oxygen species

UTR: untranslated region

MtDNA: mitochondrial DNA

mtRNAP: mitochondrial RNA polymerase

NAD: nicotinamide-adenine dinucleotide

mt-mRNA: mitochondrial messenger RNA or mitochondrial primary transcript

tRNA: transfer RNA

rRNA: ribosomal RNA

ORF: open reading frame

DS: dodecameric sequence

mtExo: mitochondrial degradosome

pre-*COB: COB* precursor mRNA

Mitoribosome: mitochondrial ribosome

PPR: pentatricopeptide repeat

MIOREX: mitochondrial organization of gene expression complex

COX: complex IV, cytochrome c oxidase

nt: nucleotide(s)

PMF: peptide mass fingerprinting

PAGE: polyacrylamide gel electrophoresis

EMSA: electrophoretic mobility shift assay

RT: reverse transcriptase

ATTRIBUTES

All of the figures and experimental work in this study were performed by Julia Jones and interpreted by Julia Jones and Dr. Michael Anikin with the exception of the following:

Single-stranded lyophilized salmon sperm DNA was prepared by Dr. Dmitriy Markov (section I.C.3)

The determination of the N-terminus of mature Mtf2p experiment was performed and the gel image created by Dr. Michael Anikin and Daniel Emeka Ezidiegwu (section I.C.3 and Figure 21).

Construction of the expression plasmid pGD1 was performed by Dr. Michael Anikin, Andrew T. Cowan, and Gary Devine (section I.C.4).

Cloning of *MTF2* (plasmid pEE1) was performed by Dr. Michael Anikin and Daniel Emeka Ezidiegwu (section I.C.5).

The W303 cell lines (wild type and Δ*PET111*) were graciously given to me from the laboratory of Dr. Michael Henry (Section I.D.8).

Dr. Michael Anikin and Andrew Cowan genetically modified the BG1805-*CBP1* plasmid to contain the right gene (section II.C.2).

Dr. Michael Anikin and Andrew Cowan created plasmids, pAC7, pAC9, pAC20, and pAC21 (section II.C.3).

The determination of the N-terminus of mature Cbp1p experiment was performed and the gel image was obtained by, Dr. Michael Anikin and Andrew Cowan (section II.C.2 and Figure 31).

The plasmid pEK8 was created by Ekaterina Kashkina and plasmid pBH176 was created by Biao He (section Appendix A.3).