In presenting this dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this dissertation may be granted by the professor under whose direction it was written when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. In the absence of the professor, the dean of the Graduate School may grant permission. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

_____

Justin L. Cotney

Divergent evolution of the paralogous human mitochondrial transcription factors, h-mtTFB1 and h-mtTFB2, to fulfill unique functions in mitochondrial gene expression, biogenesis, and retrograde signaling

By

Justin L. Cotney

Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences

Program in Genetics and Molecular Biology

_____
Gerald S. Shadel, Ph.D.
Advisor

_____
Xiaodong Cheng, Ph.D.
Committee Member

_____
William G. Kelly, Ph.D.
Committee Member

_____
Ichiro Matsumura, Ph.D.
Committee Member

_____
Daniel Reines, Ph.D.
Committee Member

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

Divergent evolution of the paralogous human mitochondrial transcription factors, h-mtTFB1 and h-mtTFB2, to fulfill unique functions in mitochondrial gene expression, biogenesis, and retrograde signaling

By

Justin L. Cotney

B.S., Birmingham-Southern College, 2002

Advisor: Gerald S. Shadel, Ph.D.

An Abstract of
A dissertation submitted to the Faculty of the
Graduate School of Arts and Sciences of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences

Program in Genetics and Molecular Biology

2007

Mitochondria are essential organelles, resulting from an ancient endosymbiosis, that are found in virtually all eukaryotes. To maintain proper function, human mitochondria have maintained distinct genomes and a dedicated gene expression system. The human mitochondrial transcription system consists of three types of proteins: POLRMT, a bacteriophage-like RNA polymerase; h-mtTFA, a mitochondrial DNA binding protein; and the mtTFBs, two factors that provide a physical link between h-mtTFA and POLRMT required for transcription initiation. The human mtTFBs, h-mtTFB1 and h-mtTFB2, are unique transcription factors exhibiting homology with N6-adenine RNA dimethyltransferases. Phylogenetic analysis suggests a very early duplication of the endosymbiont KsgA gene as the source of these two paralogs. Consistent with this ancestry, I demonstrate that both h-mtTFB1 and h-mtTFB2 have maintained RNA methyltransferase activity. Overexpression of h-mtTFB1 in human cells increases the level of methylated 12S rRNA and induces mitochondrial mass. h-mtTFB2 overexpression elevates mitochondrial DNA levels, transcripts, mass, membrane potential, and surprisingly induces a coordinate increase in h-mtTFB1 expression. This indicates that h-mtTFB1 is the major 12S methyltransferase and h-mtTFB2 is involved primarily in mitochondrial transcription. These results also suggest a major role for these factors in coordinating mitochondrial biogenesis. By combining their activities, a robust remodeling of the mitochondrial compartment occurs, increasing organelle mass and greater respiratory capacity. This response is dependent specifically upon methylation activity of h-mtTFB1 and indicates that the methylation status of 12S rRNA is a metric for mitochondrial function. In support of this idea cells harboring the mtDNA mutation A1555G, a mutation in 12S rRNA linked to non-syndromic deafness, have elevated

methylated 12S rRNA and phenotypes similar to those associated with h-mtTFB1 overexpression. I propose that 12S rRNA methylation status is regularly monitored by the cell and either induces expression of h-mtTFB1 to improve overall methylation or induces mitochondrial mass in preparation for more OXPHOS complexes to be produced by fully methylated mitochondrial ribosomes. Altogether this research indicates that these factors are intimately involved in regulating mitochondrial biogenesis. Inappropriate modulation of their levels or activities might contribute to human disease by producing mitochondria deficient in transcription, translation, or respiration.

Acknowledgements

I would like to take this opportunity to thank all of those who have contributed to this work. First and foremost I must thank my advisor Dr. Gerald Shadel, who has been a constant source of support and guidance. His high standards for research have expanded my thinking and trained me to be a much better scientist. I am indebted to all those of the Shadel lab whose friendships, technical assistance, and intellectual input will be sorely missed, but I must mention a few members who have made extensive contributions to this work and my development as a scientist. Thanks to Dr. Nick Bonawitz, Dr. Jana Eaton, Maria Lebedeva and Dr. Tim Shutt for their friendship and advice both in science and in my personal life. Many thanks to Dr. Sharen McKay for her contributions directly to the work presented here both in cultivation of the A1555G cybrid line and several assays of the cybrids and cell lines I generated, her technical and intellectual input during my training, and her unenviable task of keeping the lab well stocked and properly functioning. Thanks to Dr. Zhibo Wang for providing purified POLRMT and to Dr. Min-Xin Guan for providing the A1555G and wild type cybrids used in this work. I must also thank my family for their constant love, encouragement, and freedom to pursue my educational goals. Lastly, but certainly not least, to my wife Jennifer whose undying love and support have allowed me to survive my graduate career and become a better person along the way.

Divergent evolution of the paralogous human mitochondrial transcription factors, h-mtTFB1 and h-mtTFB2, to fulfill unique functions in mitochondrial gene expression, biogenesis, and retrograde signaling

By

Justin L. Cotney

B.S., Birmingham-Southern College, 2002

Advisor: Gerald S. Shadel, Ph.D.

A dissertation submitted to the Faculty of the
Graduate School of Arts and Sciences of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences

Program in Genetics and Molecular Biology

2007

Table of Contents

List of Tables

List of Figures

Chapter I: Origins and evolution of mitochondria

Mitochondria are dual-membrane bound organelles essential for eukaryotic cells. These organelles are theorized to be the result of a mutualistic event that allowed two different cell types to live together, referred to here and throughout as endosymbiosis, leading to the formation of a eukaryotic cell. The evidence that mitochondria were once free-living bacteria is manifold. First the organelle, in most eukaryotes, contains its own genome. The mitochondrial DNA, or mtDNA, is generally thought to be much more compact than nuclear DNA and present in many cases as a circular form reminiscent of bacterial genome organization. They can vary widely in their size and content, but encode at least a few components of the eukaryotic respiratory machinery. The respiratory components are very similar both structurally and functionally to those employed by current free-living bacteria. The ribosomes present in mitochondria are more similar to bacterial ribosomes than their cytoplasmic counterparts, and are sensitive to many different types of antibiotics that target prokaryotic translation. Lastly the mitochondria house many biological reactions, including the citric acid cycle, which also take place within free-living bacteria. While mitochondria show many similarities to bacteria, particularly the α-proteobacteria, they are now totally dependent on the cell in which they reside.

In this chapter I will discuss the possible ancestors of the mitochondrion, the evidence for endosymbiosis, and the variety of genomes that are found within mitochondria of various eukaryotic species. By understanding the nature of the formation of mitochondria we can begin to understand the biology underlying the unique systems that function within mitochondria.

1.1 Origins of mitochondria

As early as the mid 1800's those that were observing cells under the light microscope had noticed different compartments within the cytoplasm. They related these compartments to organs found within multicellular organisms and referred to them as organelles. Many noted the presence of organelles with a bacterial-like appearance within plant and animal cells, later referred to as chloroplasts and mitochondria, but not until 1905 did Mereschkowsky make the suggestion that chloroplasts were either derived from or were in fact some species of bacteria (Mereschkowsky 1905). Mereschkowsky claimed that chloroplasts were "symbionts", existed independently of the nucleus, and could only be derived from preexisting chloroplasts. Wallin presented a similar idea of symbiosis in the 1920's for the presence of mitochondria, but failed to make the connection that the mitochondria were not viable outside the cell and the cell was entirely dependent on the presence of mitochondria (Wallin 1922, 1924, 1926). These ideas fell out of favor rather quickly during the period surrounding the Second World War and were largely unnoticed until the 1970's when the endosymbiosis theory was brought back under the scientific spotlight (Margulis 1970). Since that time it has become widely accepted that mitochondria and chloroplasts are the result of ancient endosymbiotic events, but exactly how and why these occurred is hotly debated. There are many theories on the origins of such organelles, ranging from the hydrogen hypothesis (Martin and Müller 1998), to syntrophy (Moreira and Lopez-Garcia 1998), and to oxygen detoxification (Vellai and Vida 1999). I will summarize one of these theories, oxygen detoxification which in my opinion has the most supporting physical evidence, in an attempt to identify the ancestor of mitochondria and the possible circumstances

surrounding the development of such an organelle.

The environment of earth nearly 4.5 billion years ago was altogether different from what we know today. The sun had not yet reached its maximum intensity, the earth had cooled from its original formation, and the atmosphere contained very little gases and virtually no free oxygen (Cavalier-Smith 2006; Kasting 1993). This early habitat was most likely totally incompatible with any present form of life. There was abundant water generated during the formation of the earth, but nearly all of it was frozen. As carbon dioxide, water vapor and hydrogen sulfide from volcanic activity built up in the atmosphere, a greenhouse effect was established. This atmospheric change allowed the earth's surface to reach temperatures that would support liquid water for an extended period of time (Sleep et al. 2001). As ice melted around the globe, the oceans that subsequently formed were filled with a wide variety of salts, minerals, and other molecules. The chemical laden oceans allowed for virtually endless interactions of various substances that eventually gave rise to groups of molecules that were able to generate more of themselves. At some point many of these assembled to form a living organism.

These primitive organisms were faced with many challenges, but none as daunting as extracting energy from their environment. It is very likely that early organisms harnessed the energy of inorganic molecules from openings in the earth's crust. The inability to store energy from these molecules would have limited life to zones near such openings (Holm and Andersson 1995; Baross and Hoffman 1985). Electrons released from molecules that organisms absorbed were used to pump protons out of the cell and generate a concentration gradient. When protons were allowed to flow back

down their concentration gradient into the cell, the flow was harnessed by mechanical movement of a molecular motor that drove the formation of an energy currency that could be used in a wide variety of reactions, ATP. This process occurred in the absence of oxygen and is referred to as anaerobic respiration. The electron donors employed by these organisms, such as hydrogen gas or hydrogen sulfide, contain relatively little energy (Lloyd 2006). Higher energy containing molecules might have been available but harnessing energy from these molecules was dangerous. Early organisms most likely utilized lower energy molecules so that chemical reactions could be controlled and to protect other components of the organism (Kooijman and Troost 2007). This allowed these organisms to only harvest small amounts of energy safely at any given time, making their metabolisms rather inefficient. Thus the first organism to obtain a relatively safe energy source that could be converted to a form that could be stored would have a selective advantage in this early environment.

Sunlight and carbon dioxide were highly abundant, so the advent of photosynthesis solved the energy source problem. Photosynthesis utilizes the energy from sunlight to generate a chemical gradient similar to that of the anaerobic organisms that can be harnessed to produce the high-energy molecule ATP. During this process carbon from CO_2 in the atmosphere is fixed to form carbohydrates and gaseous oxygen. With a steady supply of energy intact, these photosynthetic organisms, likely the ancestors of present day phototrophic purple bacteria or cyanobacteria, were able to increase in number (Blankenship 1992). The rapid production of oxygen by the proliferating photosynthetic bacteria and depletion of CO_2 produced an atmosphere very similar to what we presently observe (Kasting 1993). This dramatic change in the earth's

atmosphere had profound effects on the chemical makeup of the earth's oceans. Oxygen at high concentrations is extremely reactive with many biological molecules. Many of the anaerobic organisms present during this time were exterminated by increasing amounts of oxygen in the atmosphere due to their inability to control oxygen-based reactions within their interior. The oxygen driven extinction was never completed, however, as many species of anaerobic bacteria exist to this day. Many of the organisms that survived the onslaught of toxic oxygen developed a system of harnessing chemical potentials in a fashion similar to the anaerobes and phototrophs described above. They used the abundant oxygen as an electron acceptor to generate water and prevent oxygen from reacting with other cellular components (Andersson and Kurland 1999). The biological machinery utilized by these new organisms pumped protons out of the cell by using energy released from the oxidation of NADH to NAD^+. Proton pumps, very similar to those employed by the anaerobes, were again used to generate a pH gradient and the flow of protons was used to drive the formation of ATP by turning a similar molecular motor, the F_0F_1 ATPase. The coupling of proton pumping to generation of ATP (oxidative phosphorylation) will be discussed in greater detail below. Additionally the metabolic system that was developed by the aerobic organisms was much more efficient than anaerobic respiration and the proton-motive force generated by respiratory proton pumping was also harnessed to drive the import of sugars and generate movement by rotation of structures such as flagella (Margulis and McMenamin 1990).

The advent of aerobic respiration allowed organisms to flourish in the presence of oxygen, and production of CO_2 allowed a carbon cycle between photosynthetic and aerobically respiring organisms. As the number of all organisms increased on earth, they

competed with one another for resources in their environments. In order to obtain more nutrients, some organisms began to increase in size. Major alterations in their overall cellular structure gave these organisms the ability to engulf smaller organisms, such as cyanobacteria, and harvest all of the energy that they contained either in the form of ATP or sugars (Margulis 1971). While this strategy was advantageous when prey organisms were readily available, the enlarged size of these organisms increased their overall energy demand, decreased their ability to respire efficiently, and left them vulnerable to high levels of toxic oxygen (Bilinski 1991). Large predatory organisms would require substantially more respiratory machinery in their outer membranes to maintain the same level of potential gradient generation. Thus, they needed a mechanism to concentrate the respiratory machinery in order to produce ATP efficiently, but probably more importantly to also detoxify oxygen that had entered the cell. This would have required a high degree of specialization of membranes to achieve compartmentalization of the respiratory machinery within the cell instead of the outer membrane (Vellai and Vida 1999).

Smaller bacteria that were being harvested as a food source by larger cells utilized the same oxidative phosphorylation machinery. If larger cells could engulf aerobic bacteria, not break them down but instead allow them to coexist, the first step toward the compartmentalization of oxidative phosphorylation would be acheived. Thus the stage was set for the origins of mitochondria. A large cell would be the ancestor of the cytoplasm and plasma membrane of the modern eukaryotic cell (and possibly the nucleus as well) and the small bacteria would give rise to mitochondria (Margulis et al. 2006). The engulfment of bacteria and attempts to maintain or cultivate them by larger cells undoubtedly occurred many times (Gray 1989), but eventually a balance was struck such

that the larger cell provided nutrients and protection for smaller cells and in return the smaller cells provided oxygen detoxification. Oxygen detoxification was the most plausible first function of mitochondria because bacteria initially lacked an ATP export mechanism (Vellai and Vida 1999). This initial arrangement provided benefit for both organisms and is described by the term symbiosis, but later one organism lived within another, which requires an additional term: endosymbiosis.

1.2 Endosymbiosis

Initially the living arrangement between the host and engulfed cell would have been tenuous at best. The membranes covering the engulfed organisms were maintained, although there is some debate about its nature and how it prevented destruction of the bacteria and also generated a microenvironment for respiration to occur most efficiently. While this membrane maximized the ability of the engulfed organisms to create a pH gradient and properly reduce oxygen, it also allowed these organisms to simply re-fuse with the plasma membrane and be released from the cell (Margulis and Chapman 1998). The host organism therefore needed to develop a mechanism to prevent release of the bacteria and provide an appealing environment. Whether this mechanism involved alteration of the engulfment vesicle after entry into the cell such that the vesicle could no longer allow exit or the bacteria altered the vesicle environment such that they prevented their export is up for debate (Vesteg et al. 2006). Either way, once the bacteria had been retained within the host cell they would need to be maintained and cultivated. The two competing drives of both the host and the engulfed thus drove the control of bacterial reproduction. The host would need to limit resources available to the bacteria, while the bacteria would need to slow their growth to prevent destruction of their protective

environment. Once this initial condition had been reached and endosymbiosis had occurred the bacteria that eventually became mitochondria underwent drastic changes to form the organelles that we observe today (Gray et al. 1999).

The mitochondria in their earliest forms would have basically been fully functional bacteria that were co-opted for sweeping oxygen from the cell. However, an interesting consequence of oxygen reduction by these organisms would be production of rather high levels of ATP. Since these developing organelles no longer had the high energy demands of free-living cells they would have likely accumulated high concentrations of ATP. To maintain the ability to reduce oxygen, either the mitochondria or the host cell developed a mechanism to release ATP from the organelle. This required the active transport of ATP across both membranes via a complex identified as the adenine nucleotide transporter (ANT) (Vellai and Vida 1999). The path was set for the end of endosymbiosis and a shift to a situation more fittingly described as enslavement (Cavalier-Smith 2006). The host cell had developed an effective means of generating large amounts of ATP and needed to maintain and exploit this advancement. The host cell drove its mitochondria toward production of ATP and away from what it saw as non-essential processes such as movement or full amino acid and nucleic acid biosynthesis. The host cell could provide many of these resources and allow the mitochondria to specialize in ATP production. Eventually even this specialization was not sufficient for the host cell's growing energy demands so it required and a means of providing complete control over mitochondria by the nucleus. Over many millions of years the developing eukaryotic cell proceeded to remove nearly all nonessential process from mitochondria, transfer genes from the mitochondrial DNA to the nuclear genome, and develop

mechanisms for surveilling mitochondrial function. Removal of these processes resulted in a major streamlining of overall mitochondrial function. The cell gained control of mitochondrial transcription, translation, and even genome replication. All of the proteins involved in these processes became products of nuclear genes and were eventually lost from the mitochondrial genome. However the machinery that is now used for these processes, while similar in some regards to their bacterial counterparts (mitochondrial ribosomes) (O'Brien 2003), is quite different and appears to have been derived by piecing together machinery from other systems (mitochondrial replication and transcription) (Bonawitz et al. 2006; Shutt and Gray 2006). Many of the genes encoding proteins that comprise the oxidative phosphorylation machinery were also transferred to the nucleus, but that was not true for all of these types of genes. Ideally the mitochondrial genome would have been eliminated by the host cell, but due to hydrophobicity or possibly the instability of some proteins several genes were unable to be removed from the mitochondrial genome and remained within the organelle (Gray 1999). The maintenance of this genome in its various forms has been one of the major pieces of evidence that mitochondria were derived from ancient bacteria.

1.3 Mitochondrial genomes

DNA was not conclusively observed within the mitochondrial organelle until the mid-1960s (Rabinowitz et al. 1965). Following this discovery in chicken embryos, researchers discovered mitochondrial DNA (and chloroplast DNA) in many different eukaryotes. Mitochondrial genomes exist in both circular and linear forms, and even tandem arrays. They can vary widely in size from ~6 kbp in several apicomplexans to more than 600 kbp in subspecies of *Zea mays* (*Organelle genomes database*). However,

size and gene content do not necessarily correlate since some genomes are 100 times larger than others but contain a similar number of genes (Gray et al. 1998).

While there may be diversity among mitochondrial genomes, in humans and virtually all vertebrates the mitochondrial genome is a circular DNA molecule of ~16,000 bp. Unlike nuclear chromosomes that are present in two copies each, there are numerous mitochondrial genomes per cell ranging from hundreds to tens of thousands in humans depending on the tissue type (Lightowlers et al. 1997). Vertebrate mtDNA encodes 13 proteins, all of which are essential components of OXPHOS complexes (Figure 1). To translate the mRNAs for these proteins the mitochondrial genome also contains 22 tRNAs and 2 rRNAs, a reduction from 64 tRNAs and 3 rRNAs in other systems due to alterations in the mitochondrial genetic code, reduction in genome size, and modified ribosomes (Boore 1999). This genome is almost exclusively inherited from the maternal lineage in humans (Giles et al. 1980) and has been used to track the evolution and migration of humans from Africa and across the globe (Chen et al. 1995; Coskun et al. 2003). While this genome been a useful tool for molecular anthropologists and human geneticists, the conservation of the vertebrate mitochondrial genome alone is not sufficient to conclude that mitochondria owe their heritage to bacteria. Maintenace of biological processes such as OXPHOS was a clue, but not until the discovery of a mitochondrial genome that contained many bacterial proteins was their clear genetic evidence that such ancestry was possible. Sequencing of the mitochondrial DNA of the protozoan *Reclinomonas americana* revealed "a eubacterial genome in miniature" (Lang et al. 1997). This genome was considerably larger than vertebrate mtDNAs at ~70,000 bp, and interestingly contained 18 protein-coding genes that had never been observed in a

mitochondrial genome. These new genes included several ribosomal proteins and even four genes of a bacterial like multi-subunit RNA polymerase (Lang et al. 1997). This observation gave the mitochondrial community great confidence in the argument that the rickettsia α-proteobacteria are the closest free-living relatives of mitochondria. The gene content of this protist's mitochondria illustrates that clearly during evolution the mitochondrial genome has been pared down resulting in more advanced eukaryotes containing highly compact genomes. This genome sequence also boosted speculation that mitochondria originally utilized bacterial methods of gene expression that were later replaced by bacteriophage components (Shutt and Gray 2006).

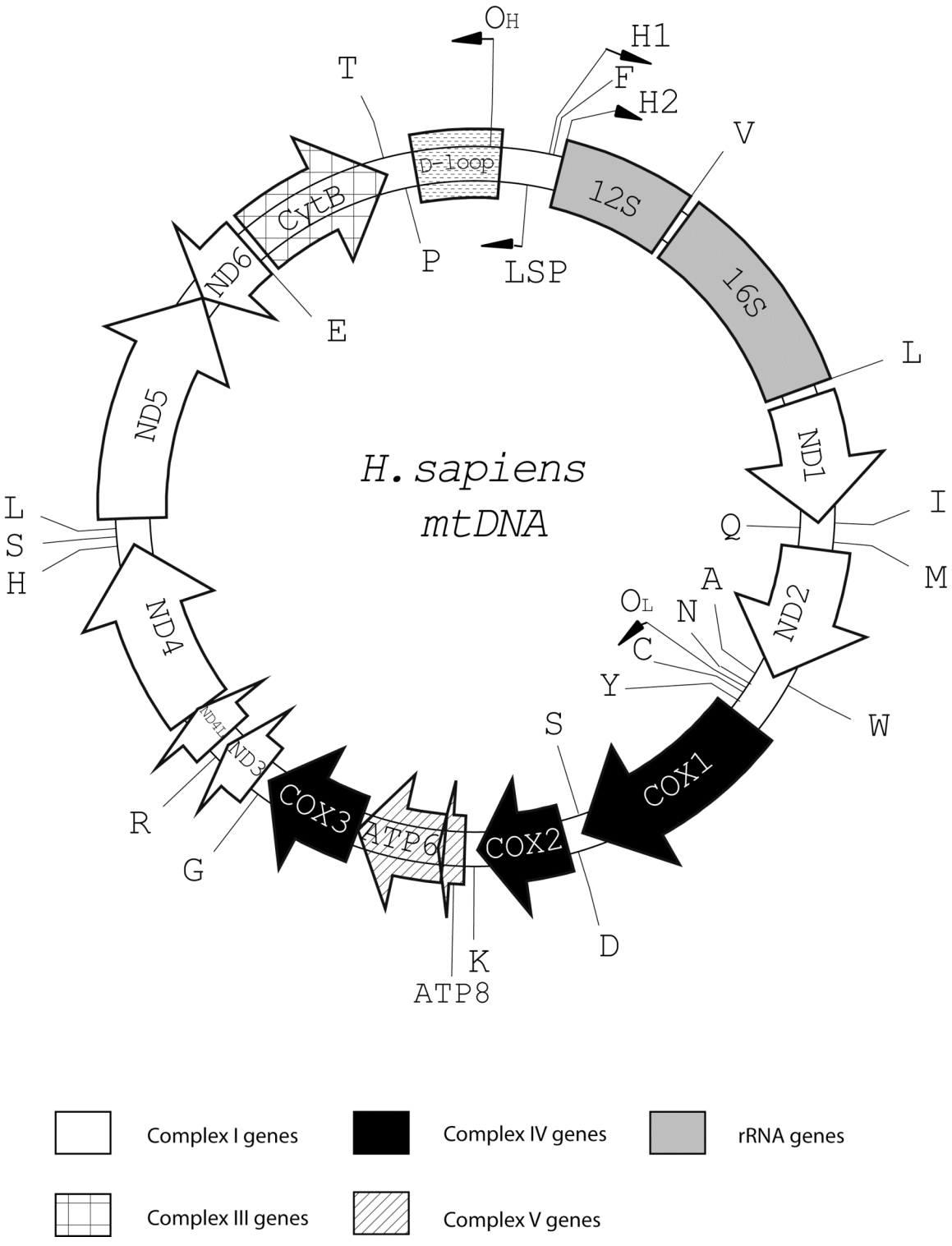

H.sapiens
mtDNA
OH
H1
F
H2
V
T
D-loop
12S
CytB
P
LSP
16S
ND6
E
L
ND5
ND1
I
L
S
H
Q
M
A
OL
N
ND2
C
Y
W
S
ND4
ND4L
ND3
COX1
R
COX3
ATP6
COX2
G
D
K
ATP8
Complex I genes
Complex IV genes
rRNA genes
Complex III genes
Complex V genes

Figure 1 Human mitochondrial DNA

Depicted is the sequence of the 16.5 kbp circular DNA molecule found in human mitochondria (Anderson et al. 1981). Previous isolation of the genome by CsCl and the sequence of the genome reveals a G/T strand bias such that one strand has a higher molecular weight than the other and is referred to as the heavy strand. The genome encodes 37 total genes, of which 13 are protein coding genes (block arrows) each being an individual component of four of the five complexes that comprise the mitochondrial oxidative phosphorylation system. The complexes that each gene is a part of are indicated in the legend. This genome also encodes all the RNA components necessary for translating the mitochondrial mRNAs: two rRNAs (gray blocks) and 22 tRNAs (single letters). There are three promoters used for transcription found in the only major non-coding region of the genome: one for the light strand (LSP) and two for the heavy strand (HSP1 and HSP2). The directionality of the depicted open reading frames shows a major strand bias in coding for proteins (clockwise = heavy strand, counterclockwise = light strand), 12 of 13 total being transcribed from the heavy strand. Other features of the human mitochondrial genome include two origins of replication (O_H and O_L) and a large non-coding region downstream of the light strand promoter (D-loop) that is bound by 7S DNA forming a triplex region.

Chapter II: Mitochondrial biology

As detailed in the previous section, mitochondria and modern day bacteria are derived from a common ancestor. Many of the metabolic reactions that take place within mitochondria are the same as those in free-living bacteria, certain process, such as DNA replication and gene expression, are unique to mitochondria. Here I will discuss several aspects of mitochondrial biology, highlighting those functions that are conserved between prokaryotes and these organelles.

2.1 Mitochondria as metabolic crossroads

2.1.1 Essential processes

Mitochondria house numerous biochemical pathways, many of which are essential to the modern eukaryotic cell. Arguably the most important set of reactions are those of the tricarboxcylic acid cycle (TCA cycle). Utilized by both bacteria and mitochondria, the TCA cycle is essential for aerobic respiration by supplying the majority of NADH and $FADH_2$ required for oxidative phosphorylation. This cycle provides intermediates for numerous other biosynthetic processes in the cell. Acetyl-CoA is the entry point for this cycle, being derived from either import of pyruvate into mitochondria and release of CO_2, or breakdown of fatty acids and amino acids within the mitochondrion. The first step of this process, once acetyl-CoA is obtained, is the production of citrate. Acetyl-CoA is unable to cross the mitochondrial membrane and yet is required for production of fatty acids and cholesterol in the cytoplasm. Citrate can be exported from mitochondria and converted back to acetyl-CoA by ATP citrate lyase. Other intermediates in this pathway, α-ketoglutarate and oxaloacetate, are required for the production of the amino acids glutamate and aspartate. Malate from the cycle is used

in the biosynthesis of glucose. Lastly for this cycle, succinyl-CoA is utilized for the production of porphyrin, which is in turn necessary for heme biosynthesis. The TCA cycle is truly a central part of cellular metabolism, is derived from the prokaryotic ancestor, and is one of the major functions of mitochondria (Williamson and Cooper 1980).

The β-oxidation of fatty acids necessary for breakdown of fatty acids to reusable components is carried out by bacteria and also housed within mitochondria. Long, medium, or short fatty acyl-CoAs transferred from the cytoplasm to mitochondria are sequentially broken down to shorter and shorter units. This process serves to provide a pool of acetyl-CoA for both the mitochondrion and the cytoplasm that can be subsequently used to synthesize new fatty acids (Schulz 1991).

The branched-chain amino acids (leucine, isoleucine, and valine) cannot be synthesized by mammals and must be acquired from food. Bacteria and plants are capable of synthesizing and catabolizing these amino acids and serve as the major sources for these molecules. When we consume foods containing these amino acids our cells take them up and use them as building blocks for proteins. If we do not consume enough of these essential amino acids or obtain more than our cells can use, cells must manage the amounts in the cell and either maintain levels present or break down excesses into usable or discardable molecules. While vertebrate mitochondria are not capable of synthesizing these amino acids, they have maintained the ability to catabolize these molecules and prevent over-accumulation. Excesses of these amino acids are not particularly toxic to most cells, but the central nervous system is highly sensitive to elevated levels of amino acids (Hutson et al. 2005).

Another essential process is the production and assembly of iron/sulfur clusters into proteins. This biosynthetic pathway is also inherited from bacteria. Proteins that harbor these clusters are involved typically in processes requiring electron transport such as various isomerization reactions. Many proteins containing these groups are localized in mitochondria, however several such proteins exist in the cytoplasm of eukaryotic cells. Mitochondria function as sites of controlled formation of these clusters and facilitate their export through ATP binding cassette (ABC) transporters to be complexed with their cognate proteins in the cytoplasm (Lill et al. 1999).

These process and many others illustrate how important mitochondria have become for the modern eukaryotic cell. Without mitochondria our cells would be unable to perform many basic metabolic tasks. Many of these pathways are thus dependent on healthy, active mitochondria. Many components of the processes listed above require ATP to drive reactions and/or mitochondrial membrane potential in order to function properly. Thus this reinforces the need for mitochondria's most famous role: housing the oxidative phosphorylation machinery.

2.1.2 OXPHOS

ATP is the most widely used cellular energy source and organisms must be able to generate a constant supply of the molecule. Bacteria are capable of producing high levels of ATP themselves via oxidative phosphorylation and, as described in the preceding chapter, were employed by eukaryotic cells to perform this task. The machinery for OXPHOS is highly conserved both structurally and mechanistically between bacteria and mitochondria. This process primarily utilizes NADH and $FADH_2$ derived from the breakdown of glucose to CO_2 and H_2O, but these electron donors can be derived from

other sources in times of scarce glucose. These molecules serve as donors of electrons that flow to oxygen through a series of molecular machines located in the inner membrane of the mitochondrion. Mitochondria during their evolution have developed many folds of their inner membranes, known as cristae, in order to maximize the total surface area of the membrane in a small space and provide the most area possible to insert the OXPHOS machinery. The machines are, in essence, electrically powered pumps that move protons from the interior of the mitochondria (the matrix) to the area between the two membranes that enclose the organelle (the intermembrane space). The pumping of protons from the matrix causes a high concentration to form between the mitochondrial membranes, creating a situation where the protons want to return to equilibrium across these membranes. The potential for movement of protons back to the matrix is then harnessed by the mitochondria in a controlled fashion by allowing the protons to flow back through a molecular motor (F_0F_1 ATPase) that physically rotates and drives the formation of ATP. The process is itself conceptually fairly simple but the actual mechanisms involved in this process are highly complex.

The OXPHOS system consists of five molecular machines, Complexes I-V. Complexes I through IV are involved in pumping of protons via oxidation of electron donors mentioned above and complex V harnesses the proton flow to drive phosphorylation of ADP, thus the term oxidative phosphorylation. The process begins by transferring of two electrons from NADH to the flavin mononucleotide group within complex I, the NADH dehydrogenase, oxidizing NADH to NAD^+. Once these electrons have been liberated they are channeled through the complex by several iron-sulfur clusters, much like an electrical wire, to the electron carrier coenzyme Q (also known as

Q or ubiquinone). The transfer of two electrons results in the pumping of four protons from the matrix and a coenzyme Q that is primed to carry electrons to the rest of the machinery. Complex I is a complicated, large biological machine and its exact mechanism of pumping protons is unknown. It is made up of proteins encoded by both nuclear and mitochondrial genes. The mtDNA-encoded components make up only a small portion of the total structure, seven proteins compared to 39 nuclear encoded proteins (Figure 1), but are all integral membrane proteins that serve as a base for which the rest of the complex is assembled around (Carroll et al. 2003). The hydrophobic portion of the complex I anchored in the membrane houses the iron-sulfur clusters mentioned above and is the site of transfer of electrons to Q. Without these proteins (ND1-4, 4L, 5 and 6) complex I is unable to function and OXPHOS is defective. Complex II, which is made up entirely of nuclear encoded proteins, is responsible for oxidation of $FADH_2$. The enzyme from the TCA cycle responsible for production of this molecule, succinate dehydrogenase, is physically part of this complex and allows for local production of this electron donor. Similar to complex I, complex II facilitates the transfer of electrons from a donor $FADH_2$ to Q, but it does not span the inner membrane and does not pump any protons during this process (Stryer 1995).

Reduced coenzyme Q, QH_2, is a highly hydrophobic molecule and moves rapidly between the lipid bilayer of the inner mitochondrial membrane where it encounters complex III, cytochrome reductase. This complex is faced with the task of moving electrons from coenzyme Q to cytochrome c. Cytochrome b, the only component of this complex encoded by mtDNA (Figure 1), has two heme groups that are necessary for removing electrons from coenzyme Q in a complex process known as the Q cycle

(Osyczka et al. 2005). These electrons are then subsequently passed to a Fe-S protein, cytochrome c_1, and finally to cytochrome c. The energy released during the electron movement results in pumping of protons, in this case one proton for each electron that reaches cytochrome c.

With an electron in its grasp, cytochrome c is free to move from complex III on the intermembrane space side of the inner membrane to complex IV, the cytochrome oxidase. This complex contains ten subunits, three of which are mtDNA encoded (Figure 1). Here, single electrons from four cytochrome c carriers are funneled to the terminal electron acceptor, oxygen, to form two molecules of water. The electrons again move along metal clusters found in several of the subunits and the energy released is harnessed to pump two protons per electron that reaches oxygen (Michel et al. 1998). As mentioned in the previous chapter the flow of electrons to molecular oxygen to form water renders the oxygen fairly inert and prevents reaction with other biological molecules.

The protons that have been pumped out of the matrix into the intermembrane space now have a desire to return to the matrix to equalize the concentration between the matrix and the intermembrane space. Protons cannot pass directly across the inner membrane and must have a physical channel through this membrane in order to reach the matrix. Complex V, the mitochondrial F_0F_1 ATPase, provides the means for protons to return to the matrix. The F_0 subunit, which contains two mtDNA encoded components (Figure 1), spans the inner membrane and provides a portal for protons to cross the membrane. Much like a water wheel used to power machines in early industry, the movement of protons back to the matrix causes rotation of the F_1 subunit of the ATPase.

The physical movement of the F_1 subunit is harnessed to form ATP thereby coupling membrane potential generated by the proton pumps to phosphorylation of ADP. However, this membrane potential is not strictly used for ATP production. Flow of protons back to the matrix is also used to power active transport of various molecules into mitochondria, transfer electrons from NADH to NADPH, and even generate heat in warm-blooded organisms (Schwerzmann and Pedersen 1986).

Normally functioning OXPHOS machinery is capable of effectively transferring electrons and fully reducing oxygen to water. However if the chain becomes clogged either by malfunctioning subunits or imbalances in the respiratory chain, electrons may dwell in a particular location for an extended period of time. This is generally thought to happen when coenzyme Q is unable to relinquish its electrons in a rapid fashion to cytochrome b of complex III or complex I maintains electrons at the FMN moiety (Wiesner et al. 2006). If this occurs, single electrons can leap from coenzyme Q and react with abundant molecular oxygen in the mitochondrial matrix. Transferring of a single electron to oxygen results in incomplete oxidation and the production of superoxide anion ($O_2^{\bullet -}$), a highly reactive molecule that can lead to the formation of various other reactive oxygen species (ROS) and damage virtually any type of biological molecule: lipid, nucleic acid, or protein. Because of the physical location of mtDNA, the mitochondrial genome is thought to be particularly susceptible to ROS driven damage (Doudican et al. 2005). If the mtDNA acquires mutations it can produce OXPHOS subunits that function improperly, leading to generation of even more ROS, and pushing the cell into a "vicious cycle" of ROS and cellular damage (Bonawitz, Rodeheffer, and Shadel 2006). To prevent this scenario it is imperative that the cell maintains

appropriate levels of mtDNA gene expression matched with nuclearly encoded OXPHOS genes to generate the safest operation of OXPHOS. However, this is not to say that all ROS are absolutely detrimental, as ROS are used to read out mitochondrial respiratory function and have documented functions in numerous signaling pathways.

2.2 Replication and inheritance of mitochondria

As cells grow and prepare to replicate, they must synthesize not only a fresh copy of their genome but also synthesize more of all of their various structures including their organelles. Mitochondria, unlike some organelles such as lysosomes, cannot be generated de novo and must be produced from preexisting organelles. This process requires the production of both more mitochondrial mass (membranes and proteins) and genetic material (mtDNA). The mechanisms involved are in broad strokes similar to bacterial growth and division, but with some aspects peculiar to mitochondria.

2.2.1 Organelle

Unlike their bacterial ancestors, mitochondria do not always exist as discrete bodies but instead form a dynamic network akin to the endoplasmic reticulum. In dividing tissues of vertebrates, the mass of this mitochondrial network steadily increases throughout the cell cycle and is then divided more or less equally between daughter cells (Lee et al. 2007). This requires the synthesis of the vast majority of components of the mitochondrion, be it lipid or protein, in the cytoplasm and import into preexisting organelle thus enlarging the structure. Once sufficient levels of organelle have been generated and the rest of the cell is ready to divide, the mitochondria must be separated for the cells of the next generation. In order to split apart the organelle mass,

mitochondria have maintained a fission system very similar to that utilized by bacteria during division (Chan 2006). Mitochondrial fission entails the recruitment of a dynamin-related GTPase to sites along the organelle by several proteins causing a nucleation event to occur. As molecules of the GTPase are assembled together they form a coil around the outer membrane of the organelle. Hydrolysis of GTP by the members of the coil causes the structure to constrict incrementally eventually allowing for both membranes to be effectively pinched through. Once the mitochondria have been divided, the dynamin-related GTPase then dissociates in order to catalyze further fission events (Osteryoung and Nunnari 2003). However, mitochondria are not constantly undergoing fission during biogenesis. Individual fragments of the organelle network are capable of meeting and ultimately fusing together. Two different mechanisms are required, one dedicated to fusion of the outer mitochondrial membrane and the other given the task of fusing the inner mitochondrial membrane; the exact workings of each are unknown (Hoppins et al. 2007). The fusion process may be critical for exchange of genetic and structural material during the biogenesis process effectively allowing mitochondrial "complementation" and the production of a more evenly functional, active network.

The replication and expansion of the mitochondrial network is also known to occur without cell division. In skeletal muscle cells, chronic exercise results in a major increase in overall mitochondrial mass to meet the ATP requirements of muscle cells at work (Wu et al. 1999). Increased mitochondrial mass is also noted in brown adipocytes of mammals when faced with cold temperatures in a process known as adaptive thermogenesis. So the ability to create mitochondria is not only important for providing

the next generation with sufficient organelle but is also necessary for specific functions of specialized tissue types in higher eukaryotes (Wu et al. 1999).

In contrast to the active process of creating mitochondrial mass, the process by which mitochondria are segregated to daughter cells appears to be quite passive. Mitochondria themselves are moved along microtubules, but whether there is a concerted effort by the cell to move equal amounts of mitochondrial mass to locations in the cell to ensure inheritance by daughter cells is not clear. In budding yeast, mitochondria are specifically transported into the budding daughter cell in hopes that part of the organellar network is captured during division (Catlett and Weisman 2000). In addition to inheritance of mass itself, the cell may have functions that ensure the organelles that are passed on are competent for respiration, but no such mechanism has been discovered.

2.2.2 mtDNA

After organelle mass has been generated, the second critical parameter for overall mitochondrial replication is production of more mitochondrial genomes. Here mitochondria diverge substantially from their bacterial ancestry. The mitochondrial genomes are localized in compact groups of several genomes (nucleoids), which are thought to be the basic unit of inheritance of the genome for daughter cells (Legros et al. 2004). In vertebrates, individual mtDNA molecules have been demonstrated to be replicated by a rather unique mechanism, that may be applicable to many other organisms that contain circular genomes.

Mitochondrial transcription from the light strand promoter (Figure 1) and subsequent processing of transcripts provides primers for replication (Xu and Clayton 1996). A single DNA polymerase used only in mitochondria (DNA Pol γ) is responsible

for replicating the genome. This DNA polymerase is homologous to the A family of DNA polymerases and requires an accessory subunit that shows similarity to tRNA synthetases (Kaguni 2004). Extension of DNA from an RNA primer at the heavy strand origin of replication results in a unidirectional replication fork that proceeds nearly two-thirds around the entire molecule where the light strand origin of replication is revealed. A second replication complex is formed and begins to polymerize the opposite strand producing what could be considered a single Okazaki fragment. A primase activity has been detected for priming of this second site of replication initiation (Wong and Clayton 1985, 1985), but the exact identity of the protein responsible for this activity remains unknown. Replication of both strands continues with one strand being completed before the other, deemed an "asynchronous" mode of replication (Brown et al. 2005). While this mechanism was considered the default model for vertebrate mtDNA replication for many years, the model has been challenged recently by a bidirectional mode of replication (Pohjoismäki et al. 2006). This model requires multiple priming events and would not be predicted to be dependent on transcription for initiation, in direct opposition to the known requirement of mitochondrial transcription for mtDNA replication. In either mode, once replication has occurred the DNA molecule must be ligated to close the circle and is then associated with other mtDNA molecules. Whether all mtDNA molecules are duplicated every cell cycle is unknown, but much like the organelle itself the mtDNA is free to replicate throughout the cell cycle (Pica-Mattoccia and Attardi 1972). Even in nondividing cells mtDNA is constantly being replicated to replace damaged molecules or those lost during degradation of whole parts of the organelle via autophagy. This requires that the mitochondrial replication machinery be able to divide

without the high concentrations of nucleotides available during nuclear genomic replication, resulting in a system that can expand or contract mtDNA numbers by changes in dNTP levels (Eaton et al. 2007; Taylor et al. 2005; Franco et al 2007).

It is well known that mtDNA is associated with the inner mitochondrial membrane (Albring et al. 1977). Proteins that participate in the formation of nucleoids may anchor the mtDNA here so that the molecule is physically close to the sites of translation allowing for OXPHOS subunit assembly line to form (Wang and Bogenhagen 2006). Some have speculated that the nucleoid is anchored to specific spots on the inner membrane, which form links to the outer membrane and connect with the rest of the cell's cytoskeleton (Meeusen and Nunnari 2003). This could provide a method by which the cell can know which portions of the organelle network contain genetic material and facilitate segregation (Boldogh et al. 2003).

2.3 Signaling and apoptosis

Owing to their central locations within various metabolic pathways and the production of the vast majority of usable energy in the cell, mitochondria are constantly being measured for activity and are also themselves integral components of various signaling pathways. There are two forms of signaling: anterograde and retrograde. Most identified pathways to date are anterograde, or in a direction to mitochondria. These signals come from various sources such as environmental nutrients, hormones, or signaling cascades generated in the cytoplasm. These pathways affect activities of proteins in the cytoplasm that can alter nuclear gene expression, causing remodeling of the mitochondrial compartment or could even interact with the organelle directly. A major signaling pathway for mitochondria is the PGC-1α pathway responsible for

generating more mitochondrial mass. In response to exercise the PGC-1α gene is transcribed producing elevated levels of the protein (Norrbom et al. 2004). PGC-1α, a transcriptional co-activator, then associates with a number of different proteins including the nuclear respiratory factors, NRF-1 and NRF-2, and induces transcription of target genes (Scarpulla 2006). These factors are known to control the expression of many mitochondrial genes and binding sites have been detected in the promoters of the mitochondrial transcription factors h-mtTFA, h-mtTFB1, and h-mtTFB2 (Gleyzer et al. 2005). Alterations of these genes' mRNA levels eventually result in elevated mitochondrial gene expression, higher levels of respiratory complexes, and increased overall mitochondrial mass, but the full nature of this response is not totally understood. Another example of this type of signaling is illustrated by the TOR kinase pathway. The TOR pathway is an integrator of nutrient signals, either through hormones such as insulin or some nutrient molecules directly. Once activated by upstream events, TOR stimulates cytoplasmic ribosome biogenesis and translation (Schmelzle and Hall 2000), but is also implicated in control of mitochondrial output. In yeast, which are able to ferment some sugars and will shift their metabolic processes to utilize such carbon sources exclusively, this involves down regulation of mitochondrial translation ultimately reducing mitochondrial respiration (Bonawitz et al. 2007). In this case yeast sense a particular type of nutrient, glucose for instance, and want to use it as quickly as possible in order to gain an advantage in their environment. By shifting to glycolysis and fermentation and away from respiration and production of mitochondria, yeast can devote all of their resources to a specific metabolic pathway. TOR is not only a generator of an anterograde signal but is also integral component of retrograde signaling pathways (Liu and Butow 2006).

Retrograde signaling describes signals that emanate from the mitochondrion itself and cause changes either directly in the cytoplasm or alter gene expression in the nucleus (Butow and Avadhani 2004). Metabolic intermediates of the TCA cycle have been implicated in signaling by affecting synthesis of biological molecules, such as amino acids that directly impinge on the functions of the TOR pathway above (Schmelzle and Hall 2000). Some TCA intermediates have also been shown to bind to G-protein coupled receptors allowing the cell to measure overall flux through this essential pathway within mitochondria. As the major ATP suppliers and the danger associated with oxygen-based respiration, the respiratory chain must be continuously monitored as well. Reactive oxygen species are thought to be a sensitive, rapid, and convenient measure of respiratory chain health. One idea is that ROS can quickly oxidize biological molecules that in turn either directly escape from the mitochondrion or initiate other events that lead out of mitochondria to the nucleus (Bonawitz and Shadel 2007). Once the signal has reached the nucleus, expression of mitochondrial genes is altered to change respiration levels or activities to produce the desired cellular outcome. Consistent with this idea, ROS levels are implicated in controlling respiration rates, generating stress responses, and are linked to overall longevity of organisms (Schulz et al. 2007).

In addition to metabolic based signaling pathways, mitochondria are now famous for their major contribution to the intrinsic pathway of apoptosis. Apoptosis initiated by mitochondria is the result of release of the mobile electron carrier, cytochrome c, from the inter-membrane space to the cytoplasm of the cell. Once this occurs cytochrome c interacts with Apaf-1 creating a complex that is able to bind and activate various caspases. These caspases then initiate the classic cascade that leads to cellular

destruction (Alberts 2002). The cell can initiate this event by activating proteins that open pores in the outer membrane that allow cytochrome c to escape, or highly dysfunctional mitochondria can damage themselves badly enough that this electron carrier is released (Duan et al. 2003). In either case fission and fusion of mitochondria appear to be critical for prevention and propagation of the apoptotic cascade (Hoppins et al. 2007). Sites of fission appear to be points of nucleation for pore formation in the outer membrane, and fusion may serve to reduce the amount of highly dysfunctional mitochondria by spreading the damage equally throughout the organelle network. Additionally, the major activator of apoptosis, p53, has been shown to localize to mitochondria and interact with several mitochondrial proteins (Achanta et al. 2005; Chen et al. 2006; de Souza-Pinto, Harris, and Bohr 2004; Yoshida et al. 2003). The exact significance of these interactions is unknown, but may point to a mechanism where p53 directly affects mitochondrial health in order to influence the decision to undergo apoptosis or not.

2.4 Cooperation of two genomes

As has been mentioned several times thus far, mitochondria contain their own genome, but owe their overall structure to proteins of nuclear origin. For many aspects of mitochondrial biology, such as fission/fusion or outer membrane composition, the mtDNA does not have a direct role. However, in two other aspects, the mitochondrial genome must coordinate intimately with the nucleus. The first such interaction is the production of the OXPHOS chain. The mtDNA-encoded subunits are generally integral membrane components and are thought to nucleate the rest of the complexes that they are a part of. This requires direct physical interaction of mitochondrial and nuclear encoded

proteins in order to produce a functional complex. The amounts of each subunit from mtDNA genes must be balanced to match subunits being imported. Without properly matching each genomes' output, imbalances in the respiratory chain can occur resulting in diminished ATP production or excessive ROS and associated consequences discussed earlier.

To properly match mitochondrial encoded subunit production with nuclear encoded subunits, translation by both cytoplasmic and mitochondrial systems must be regulated. This leads us to a second location of direct interaction of products from both the nuclear and mitochondrial genomes: mitochondrial ribosomes. As described in previous sections, the mitochondrial ribosomes are dedicated to translating mitochondrial mRNAs. The nucleus must supply all of the protein components for this machine and the mtDNA providing the RNA components of the structure (Sharma et al. 2003). Here, much like the case of respiratory complex assembly, the production of mitochondrial rRNAs must be matched to nuclear proteins to produce accurate, processive ribosomes. Lack of ribosomal proteins could cause backups in translation or production of mutated OXPHOS subunits. Improper levels of rRNAs might prevent ribosome formation altogether or result in a substantial level of wasted resources in excess transcripts. Anterograde and retrograde signaling pathways surely exist for allowing these two genomes to effectively communicate their output, but to date none have been definitely identified. The research to follow in this thesis, rather unexpectedly, begins to shed light on a possible mode of retrograde signaling used to monitor mitochondrial ribosomes and provides insight into how mitochondria as a whole can be regulated by coupling of several processes.

2.5 Roles of mitochondria in human disease

As should be clear by the description of the wide variety of metabolic process that take place in mitochondria, the probablity for a mitochondrial related disorder are great. When combining mitochondrial DNA mutations with mutations in nuclear genes that cause mitochondrial defects, nearly 1 in 8000 births are affected by some sort of mitochondrial related dysfunction (Thorburn 2004). Mitochondrial disorders vary widely in their physical manifestation, severity, and penetrance. Some disorders affect every cell in the human body, while others appear to be tissue specific. In other cases, two people may have the same pathogenic mutation but one is unaffected. Understanding such complicated pathologies is dependent on developing a more complete understanding of normal mitochondrial biology.

Many mitochondrial disorders are the result of inborn errors of metabolism: single defective genes inherited by offspring causing malfunctioning metabolic pathways. Many of the pathways in mitochondria have related diseases. Defects in branched-chain amino acid breakdown result in a toxic build up of α-keto acids and cause the potentially fatal maple syrup urine disease (MSUD) (Hutson et al. 2005). Defects in fatty acid oxidation lead to hypoglycemia and hyperammonemia (Schulz 1991). An inability to properly produce porphyrin can cause porphyria and anemia due to lack of heme synthesis (Elder 1993).

A large number of disorders are linked to ATP production of mitochondria and show the widest range of phenotypes. Since function of the respiratory chain is dependent on both genomes, nuclear and mtDNA mutations can generate disease. Several specific mtDNA point mutations have been identified over the last twenty years

that cause pathology (Ruiz-Pesini et al. 2007). However due to the operation of mitochondria in the context of a nuclear background, not all individuals that harbor a specific mutation are affected. In other cases the absolute level of mtDNA is altered forming a class of diseases known as mtDNA depletion. These are the result of defects in maintenance or replication of mtDNA due to defective replication machinery, lack of nucleotides for replication, or altered stability of the mitochondrial genome (Copeland 2007). All of these diseases are subject to threshold effects and require high levels of depletion or mutated mtDNA before pathology will manifest. Tissues with high energy demands, such as neurons or muscle, are generally the most sensitive to defects surrounding mtDNA. Many of the phenotypes associated with mtDNA disorders are displayed as neurological or muscular defects and are characterized by degeneration of each. Neurodegeneration in both Parkinson's and Alzheimer's diseases have been linked to reduced ATP levels and other mitochondrial defects (Weissman et al. 2007). Much of the tissue degeneration observed in other diseases is thought to be the result of enhanced apoptosis due to defective mitochondria, however the obvious disease to link to apoptosis is cancer. Many cancers are thought initially to use mitochondria to drive nuclear instability, but over time appear to reduce overall mitochondrial activity. This could be due to a general trend for tumors to exist in hypoxic environments and are less dependent on respiration, but may also suggest that they are reducing mitochondrial activity to prevent the organelles from initiating apoptosis (Singh 2006).

As one can see, mitochondria are significant contributors to human disease. These organelles carry a double-edged sword. On one side they provide the cell with numerous essential processes and on the other defects in a single pathway can have

dramatic effects on human health. From a human perspective, understanding how mitochondria cause disease and affect human health is the most important goal of mitochondrial research. However, in our constant desire to escape the ever-present disease of ageing, a major branch of mitochondrial research will most likely focus around the documented ability of mitochondrial function to affect lifespan of some organisms.

Chapter III: Mitochondrial gene expression

Virtually all the new research to be described later in the text is focused on understanding the biological contributions of two human mitochondrial transcription factors. In order to understand how these factors function and what impacts they could possibly have on mitochondrial activities, I must first discuss what is already known about the human mitochondrial transcription machinery. Shortly after the discovery of mtDNA, mitochondria from human cells were shown to harbor many different RNA species (Attardi and Attardi 1967; Attardi and Attardi 1968). The hybridization of RNA obtained from mitochondria to isolated mitochondrial DNA from the same organisms suggested that the mtDNA was being transcribed. During the early 1970's many aspects of the kinetics of mitochondrial RNA production and stability were uncovered (Aloni and Attardi 1971, 1972; Pica-Mattoccia and Attardi 1971), but the machinery responsible for production of these RNAs was not discovered until the mid 1980s (Fisher and Clayton 1988; Masters et al. 1987) and the exact events involved in initiation and elongation remained elusive until the 1990's.

Much of the gross biochemical characterization of mitochondrial transcription originally occurred in bovine tissue and HeLa and KB cells, but the yeast *S. cerevisiae* was the first organism to allow the identification of the genes required for this process. The genetic and biochemical tools of yeast enable the identification and cloning of the mitochondrial RNA polymerase. Consistent with a bacterial origin of mitochondria, the mitochondrial transcription system was thought to be similar to that of prokaryotes. A single yeast gene, *RPO41*, was identified as being essential for mitochondrial transcription and mtDNA maintenance (Greenleaf et al. 1986) and thought to be the primary subunit of a bacterial-like multi-subunit RNA polymerase. Surprisingly,

sequencing and analysis of the *RPO41* gene revealed a protein bearing no major homology with any subunit of the *E. coli* RNA polymerase (Masters et al. 1987). Instead, the yeast mitochondrial RNA polymerase was found to contain sequences that are highly conserved among bacteriophage RNA polymerases, namely T3 and T7. This discovery has led to much speculation about the function and origin of the mitochondrial transcription system (Shutt and Gray 2006). These types of RNA polymerases have been demonstrated to function as a single subunit, without the need for any other factors, suggesting that mitochondria might indeed utilize a very simple form of transcription.

A non-specific active single subunit was not the case however, as promoter specific mitochondrial RNA polymerase activity in both yeast (Winkley et al. 1985; Kelly and Lehman 1986) and humans (Fisher and Clayton 1985) was found to require at least two components. The second factor in yeast was identified by both biochemical and genetic means and originally named *MTF1* (Lisowsky and Michaelis 1989; Schinkel et al. 1987). This protein was shown to be required for specific yeast mitochondrial transcription *in vitro* and postulated to act in a fashion similar to bacterial sigma factors (Jang and Jaehning 1991); however the similarity both at sequence and structural levels to sigma factors was eventually discounted (Shadel and Clayton 1995; Schubot et al. 2001). The mitochondrial transcription system of yeast initially served as the basic model for human mitochondrial transcription. The same basic components were thought to operate in human mitochondria, but identification of a human protein required for specific mitochondrial transcription exhibited no homology to the fungal *MTF1*. The human mtTF1 factor appears to be a unique protein, containing two high mobility group (HMG) domains that conferred DNA binding (Parisi and Clayton 1991). A yeast protein,

Abf2p, was identified as having substantial homology to the human mtTF1 factor, but was not required for yeast mitochondrial transcription (Xu and Clayton 1992). This led to a change in nomenclature with the human mtTF1 factor being named h-mtTFA and the yeast *MTF1* factor being labeled sc-mtTFB in order to emphasize the differences between these two mitochondrial transcription systems. Attempts to study mitochondrial transcription in Xenopus led to the discovery of two factors necessary for specific transcription initiation by the mitochondrial RNA polymerase (Antoshechkin and Bogenhagen 1995). One factor was the homolog of h-mtTFA, and the other showed physical characteristics reminiscent of sc-mtTFB. Studies of the Xenopus transcription system revealed binding of xl-mtTFA to mtDNA promoters, and suggested a mitochondrial RNA polymerase: xl-mtTFB holoenzyme that was then recruited to these promoters by mtTFA (Bogenhagen 1996). This model initiated the search for other transcription factors in human mitochondria and laid the basic foundation for our understanding of mammalian mitochondrial transcription machinery.

The basic human mitochondrial transcription machinery has since been identified and indeed consists of three types of factors: the mitochondrial RNA polymerase (POLRMT), mtTFA, and mtTFB. My research has focused on the third group of this system, mtTFB, in attempts to define this class of transcription factors' contributions to mitochondrial gene expression and organelle biology. In the following sections I will discuss the known characteristics of the human mitochondrial transcription machinery to demonstrate how this system might function.

3.1 Mammalian mitochondrial transcription machinery

3.1.1 POLRMT

The human mitochondrial RNA polymerase, much like the yeast homolog, shows a high degree of similarity with single subunit bacteriophage RNA polymerases (Tiranti et al. 1997). Identification of mitochondrial RNA polymerases from other organisms has since demonstrated that use of a phage derived transcription system is conserved (Shutt and Gray 2006). The carboxy-terminal two thirds of these mitochondrial proteins contain several conserved bacteriophage RNA polymerase domains, and most have a large amino-terminal domain (ATD) extension that bears no resemblance to the phage proteins (Masters et al. 1987; Tiranti et al. 1997). In most metazoans, ten of a total of twelve phage RNA polymerase domains are conserved however most higher eukaryotic mitochondrial RNA polymerases lack domains 1A and 1B (Miller et al. 2006). The catalytic activity of T7 bacteriophage RNA polymerase resides in the carboxy-terminal three-quarters of the protein with the first quarter containing sequences necessary for stabilizing an open promoter complex (Ikeda and Richardson 1987). The extension of the amino-terminal domain, and an apparent lack of sequences utilized for initiation in some organisms, suggested that the ATD may function as an interaction point for other proteins involved in mitochondrial transcription and could impart new activities to the RNA polymerase (Masters et al. 1987). Indeed portions of the ATD are dispensable for mitochondrial transcription in yeast but are necessary for mtDNA maintenance (Wang and Shadel 1999). The ATD has also been shown to be a site of interaction with mitochondrial RNA processing proteins and translation in yeast. The arrangement couples the mitochondrial RNA polymerase to translational activators and ultimately to

translation itself (Rodeheffer and Shadel 2003; Bryan et al. 2002; Rodeheffer et al. 2001). Whether such interactions and requirements for the ATD in human POLRMT exist are unknown, but the mitochondrial RNA polymerase has recently been shown to interact with a mitochondrial ribosomal protein, MRPL12. This could indicate that mammalian mitochondrial transcription and translation are also coupled (Appendix A; Wang et al. 2007). It also stands to reason that the regions of mitochondrial RNA polymerase not directly involved in catalytic activity may be sites for interaction with other proteins demonstrated to be present in mitochondria such as p53 (Yoshida et al. 2003) or hormone receptors (Psarra, et al. 2006) and provide sophisticated mechanisms for regulation of mitochondrial transcription.

All mitochondrial RNA polymerases assayed have the ability to initiate RNA synthesis from a polydA-dT template, much like the non-specific activity of T7 RNA polymerase. When presented with double stranded DNA templates that contain a cognate mitochondrial promoter, mitochondrial RNA polymerases are unable to effectively initiate transcription alone. However, the yeast mitochondrial RNA polymerase if presented with a specific promoter template that is supercoiled or contains mismatches near the site of initiation (a premelted template) does not require any specificity factors to initiate transcription (Matsunaga et al. 2004). This indicates, at least for yeast, that the mitochondrial RNA polymerase is fully capable of specifically recognizing mitochondrial promoters, but lacks the ability to melt DNA or stabilize an open promoter complex. Such activity appears to be maintained in the human mitochondrial RNA polymerase (Gaspari et al. 2004) and gives insight into the molecular events that take place during mitochondrial promoter recognition and initiation.

Aside from mitochondrial transcription, the mitochondrial RNA polymerase in both yeast and humans has been demonstrated to have activities outside of mitochondria or not involved directly with mitochondrial gene expression. In yeast, the *RPO41* gene is synthetically lethal with a number of genes involved in cell wall synthesis (Tong et al. 2004; Lesage et al. 2005). How *RPO41* contributes to this pathway is unknown but no other genes that participate in mitochondrial transcription or mtDNA maintenance show synthetic lethality with cell wall synthesis proteins. This suggests a function function for the mitochondrial RNA polymerase independent of mitochondrial transcription. In mammalian cells an alternative splice form of the POLRMT gene has been described that causes nuclear localization of the isoform (Kravchenko and Chumakov 2005). This splice variant occurs in primates and rodents and the resulting mRNA is translated such that the mitochondrial localization sequence is absent. Reducing levels of this splice variant causes decreases in transcripts from several different nuclear genes indicating that a subset of genes are expressed by this RNA polymerase (Kravchenko et al. 2005).

3.1.2 mtTFA

mtTFA was initially identified by its promoter specific transcription stimulatory activity in a mitochondrial extract capable only of nonspecific activity (Fisher and Clayton 1985). Sequencing of the gene revealed a DNA binding protein with two HMG boxes separated by a small linker region and containing a small carboxy-terminal tail (Parisi and Clayton 1991). This protein, like many other HMG box proteins, binds double stranded DNA in both sequence specific and nonspecific fashions (Fisher and Clayton 1988). mtTFA has two high affinity sites, both of which reside near the promoters of each mtDNA strand. It also shows spaced binding throughout the D-loop

region (Fisher et al. 1992) and non-specific binding throughout the mtDNA (Ohgaki et al. 2007). mtTFA has a higher affinity for DNA near the light strand promoter (LSP) than the heavy strand promter (HSP) (Figure 1), and *in vitro* stimulates transcription from the LSP to a much greater extent (Fisher and Clayton 1985; Falkenberg et al. 2002). The stimulation of transcription by mtTFA stems from its ability to bend and wrap double stranded DNA. This places the mtDNA in a conformation that is already unwound or is more easily unwound exposing a single strand site for transcription initiation (Fisher et al. 1992). *ABF2*, the homologous yeast gene, was shown to have very similar DNA binding characteristics and was linked to maintenance of the yeast mitochondrial genome (Diffley and Stillman 1991). However, the Abf2p protein was found not to be essential for mitochondrial transcription (Xu and Clayton 1992). Sequence analysis revealed that both mtTFA and Abf2p contain two HMG boxes, but Abf2p lacked a carboxy-terminal tail. This tail is essential for mtTFA's transcriptional activity and if added to the Abf2p protein is sufficient to convert the protein to a transcriptional activator (Dairaghi et al. 1995).

The Abf2p protein does not have a role in transcription but packages mtDNA. This led many to speculate that mtTFA may have DNA packaging activities as well. Levels of mtTFA were found to be altered in a patient suffering from mtDNA depletion and greatly decreased in tissue culture cell lines lacking mtDNA (Larsson et al. 1994). This effect could not be directly linked to a DNA packaging role as human mitochondrial DNA replication and transcription are coupled. mtTFA has since been documented to be absolutely necessary for mtDNA maintenance during development (Larsson et al. 1998), some even suggesting that mtDNA is tightly packaged by mtTFA (Alam et al. 2003; Takamatsu et al. 2002), similar to histones and nuclear DNA. Recently many have begun

to discount the transcription factor activity of mtTFA and suggest that its major function is to protect and package mtDNA (Kang et al. 2007; Kanki et al. 2004). However high levels of mtTFA have been documented to negatively impact mitochondrial transcription and mtTFA is absolutely required for *in vitro* mitochondrial transcription (Falkenberg et al. 2002; Parisi et al. 1993).

Levels of mtTFA, as well as POLRMT, have been closely associated with levels of mtDNA and vice versa suggesting a mutual dependence (Seidel-Rogol and Shadel 2002; Siciliano et al. 2000). Additionally, mtDNA has for many years been used as a marker for mitochondrial mass. Pathways that are utilized for mitochondrial biogenesis (PGC-1α) are known to enhance expression of mtTFA and elevate levels of mtDNA (Wu et al. 1999). Because of the multiple links of mtTFA to mitochondrial biology, defects in mtTFA expression or activity have been linked to many different diseases including Parkinson's, diabetes, and mtDNA depletion disorders (Belin et al. 2007; Choi et al. 2001; Larsson et al. 1994). mtTFA could be an attractive target for patients suffering from these kinds of diseases once mtTFA roles in these pathologies are understood. In relation to disease pathology, mtTFA may serve as a protective agent for mtDNA and facilitate its repair (Weissman et al. 2007). mtTFA has been shown to bind with a high affinity to cisplatin damaged DNA, suggesting that the protein is acting as a recruitment factor for repair machinery in mitochondria (Yoshida et al. 2002). The binding to damaged DNA is enhanced in the presence of p53 suggesting that mtTFA may even have a direct role in apoptosis (Yoshida et al. 2003).

Much like POLRMT, other splice forms of mtTFA have been detected in both humans and mice. One form of the protein loses a significant portion of the second HMG

box (Tominaga et al. 1993) and thus has altered DNA binding capabilities (Dairaghi et al. 1995). The other splice variant has been observed in the testes of mice and humans. This isoform lacks a mitochondrial targeting sequence and is instead targeted to the nucleus (Larsson et al. 1996). The expression of mtTFA is known to change during spermatogenesis in humans (Larsson et al. 1997), suggesting a different role for mtTFA in the testes or a need to remove the protein from mitochondria and reduce the mtDNA copy number of sperm. However, it is very tempting to speculate that in the nucleus POLRMT and mtTFA participate in a mitochondrial-like transcription system under specific conditions.

3.1.3 mtTFB1 and mtTFB2

3.1.3.1 Transcription factor activity

The third and final class of proteins that are essential for human mitochondrial transcription are the mtTFB proteins. After identification of sc-mtTFB-like activity in Xenopus, searches of human genome databases turned up two homologs to the yeast mtTFB: h-mtTFB1 and h-mtTFB2 (Falkenberg et al. 2002; McCulloch et al. 2002). Cloning, expression, and purification of each recombinant protein soon followed. Addition of purified h-mtTFB1 to a mitochondrial extract that could be stimulated by h-mtTFA resulted in elevated levels of transcripts (McCulloch et al. 2002). h-mtTFB1 co-immunprecipitates with POLRMT from mitochondrial extracts (McCulloch and Shadel 2003). These interactions were also seen when mtTFB1 and mtTFB2 were found to copurify with POLRMT during production of recombinant protein (Falkenberg et al. 2002). The sc-mtTFB protein is also known to bind directly to the mitochondrial RNA

polymerase, but the residues and surfaces necessary for this interaction are spread across the Rpo41p protein and do not cluster in the ATD as might have been expected (Cliften et al. 2000; Cliften et al. 1997). Both h-mtTFB1 and h-mtTFB2 demonstrated direct binding to h-mtTFA as well; this interaction is mediated by the c-terminal tail of h-mtTFA (McCulloch and Shadel 2003). This observation would suggest the requirement for the tail of h-mtTFA for transcriptional activity is base upon its interaction with the h-mtTFB proteins. Subsequent analysis of all four recombinant proteins showed that either h-mtTFB1 or h-mtTFB2 are absolutely required for POLRMT and h-mtTFA mediated specific transcription *in vitro*. In these experiments the absolute transcriptional activity of h-mtTFB2 was reported to be at least an order of magnitude greater than h-mtTFB1 (Falkenberg et al. 2002). Additionally, similar to sc-mtTFB and xl-mtTFB, h-mtTFB1 was shown to have weak binding to promoter DNA that could easily be competed by random DNA sequence (McCulloch and Shadel 2003).

3.1.3.2 Homology with rRNA methyltransferases

These characteristics confirmed interactions predicted from the Xenopus transcription system (Bogenhagen 1996) and gave insight to the function of the mtTFB proteins in transcription. However, possibly most interesting, a crystal structure of sc-mtTFB (Schubot et al. 2001) and analysis of the primary sequence of the mammalian and fungal mtTFB cDNAs revealed significant homology with a family of N6-adenine ribosomal RNA methyltransferases (McCulloch et al. 2002). This immediately roused speculation, that like mtTFA and POLRMT, these proteins could have additional functions. Ribosomal RNA methyltransferases comprise a class of proteins that are widely dispersed in nature and catalyze modifications that are necessary for proper folding and stability of ribosomal RNA. While modification in general has been highly conserved, specific modifications have not (Decatur and Fournier 2002). One clear exception arises in the case of methylation of the 3' terminal stem-loop of small subunit rRNA. This stem-loop is one of the most highly conserved structures in all ribosomal rRNAs (Mears et al. 2002; Van Buul et al. 1984). The N6-adenine rRNA methyltransferases catalyze the transfer of two methyl groups from donor molecules, S-adenosylmethionine (SAM), to each of two adenines located in the loop portion of the conserved structure. The resulting N6-dimethyladenines are the only modified bases that are conserved in prokaryotic, eukaryotic and mitochondrial small subunit rRNAs (O'Farrell et al. 2006). This high degree of conservation might suggest that these modifications are absolutely essential to the function of the ribosome, but this is not the case.

The sole member of the rRNA methyltransferase family of proteins in *E. coli*, KsgA, can be deleted without major deleterious effects in laboratory conditions and deletion actually confers resistance to the aminoglycoside kasugamycin (Helser et al. 1971, 1972). This is not to say, however, that there are no side effects to loss of methylation of these two adenines. Elimination of KsgA activity in *E. coli* causes an extension in doubling time and increases the translational error rate (Igarashi et al. 1981; van Buul et al. 1984). Isolated ribosomes from *ksga*$^-$ strains display subunits that dissociate from one another more readily (Poldermans et al. 1980) and require higher levels of initiation factors for proper translation initiation (Poldermans et al. 1979). Kasugamycin has recently been shown to bind to the small subunit of the ribosome and effectively mimic mRNA. Binding of kasugamycin prevents mRNA from entering the P and E sites of the ribosome, thus inhibiting translation initiation (Schuwirth et al. 2006; Schluenzen et al. 2006). Interestingly, 30S small subunits derived from *ksga*$^-$ mutants are fully capable of binding kasugamycin but lack of a methylated stem-loop allows the 50S subunit to bind to the small subunit more readily in the presence of the drug (Poldermans et al. 1979; Schuwirth et al. 2006). Still the exact function of these methylation events is unknown. Researchers have shown that methylation of the two adenines serves to prevent duplex interaction of the 3-terminal sequence with other RNA and enhance formation of the stem-loop (Micura et al. 2001). This might suggest that lack of methylation of the terminal stem-loop causes the region surrounding the A and P sites to become more flexible and allow mRNA and tRNA to bind to the small subunit even in the presence of kasugamycin. This would also explain the elevated error rate of ribosomes lacking these modifications.

The properties required for catalyzing such reactions, several methylations on an evolving substrate, by a single protein may make these RNA methyltransferases amenable to multiple functions in the cell. The KsgA protein was originally only thought to function in the modification of rRNA, but recently KsgA has been identified to have a function independent of it methyltransferase activity. The exact nature of this function is unknown but may be involved in the acid shock response (Inoue et al. 2007). Also, KsgA could have a role in selecting which mRNAs are translated. The protein has been shown to autoregulate itself by binding to its own mRNA and effectively block its translation (van Gemen et al. 1989). What sequences are necessary for this interaction are unknown, but formally leaves open the possibility that other messages could be regulated in this fashion. The *DIM1* gene, a homolog of KsgA in yeast responsible for methylating the conserved stem-loop of cytoplasmic 18S rRNA, was shown to be essential in yeast (Lafontaine et al. 1994). The methylation activity was shown not to be the critical function however. The Dim1p protein was instead absolutely essential for its role in processing of 18S rRNA from a 20S molecule independent of methylation (Lafontaine et al. 1995). Additionally, Dim1p has also been implicated in the proper splicing and export of the *LID1* mRNA and perhaps others (Carnahan et al. 2005). Together these activities illustrate the possibility of these types of proteins for developing multiple functions and why they may have been prime candidates for components of the mitochondrial transcription system.

3.1.3.3 Are the mtTFBs dual functional proteins in vivo?

Having set the precedent for multiple functionality of N6-adenine rRNA methyltransferases and having identified a transcriptional role for the homologs in human

mitochondria, the question that arises is whether or not the mtTFB factors have maintained rRNA methyltransferase activity. This modification is known to exist on the small subunit rRNA of hamster mitochondria (Baer and Dubin 1980), and has also been identified in human mitochondria (Seidel-Rogol et al. 2003). mtTFB1 shows binding to the standard methyl donor for the rRNA methylation, SAM, (McCulloch et al. 2002) and elimination of binding had no effect on its ability to stimulate mitochondrial transcription *in vitro* (McCulloch and Shadel 2003). This suggests a situation, much like that of Dim1p, where the putative methyltransferase activity is not required for and is separate from any transcriptional roles. Initial attempts to measure methylation activity on *in vitro* transcribed RNA were unsuccessful, most likely due to a partially formed ribosome as the actual substrate for methylation (Desai and Rife 2006). Instead a rather simple *in vivo* approach was employed to discern whether h-mtTFB1 as an active rRNA methyltransferase or not. Owing to the highly conserved nature of the 3′-terminal stem-loop of small subunit rRNA is was postulated that if h-mtTFB1 could be expressed in bacteria lacking the KsgA gene, methylation could be restored and measured by sensitivity to kasugamycin. Indeed this was the case as h-mtTFB1 and not a mutant deficient in SAM binding was able to restore sensitivity to kasugamycin of a *ksga*$^{-}$ E. coli strain. Subsequent primer extension analysis confirmed the presence of N6-methyladenines on bacterial 16S rRNA when wild type h-mtTFB1 was expressed (Seidel-Rogol et al. 2003).

This activity identified h-mtTFB1 as a bona fide dual function protein, but whether both activities are maintained in mitochondria has yet to be determined. Shortly after cloning of the mtTFB factors, h-mtTFB1 was identified by linkage analysis to be a

nuclear modifier locus of an mtDNA point mutation associated with non-syndromic deafness (Bykhovskaya et al. 2004). Humans who harbor the mtDNA mutation A1555G are particularly susceptible to deafness, especially when treated with aminoglycoside antibiotics, and cell lines isolated containing this mutation have defects in mitochondrial translation (Guan et al. 2000). Aminoglycosides, such as kasugamycin, are known to negatively affect prokaryotic and mitochondrial translation. Most humans are resistant to high doses of these types of antibiotics, but some are rendered irreversibly deaf (Fischel-Ghodsian 1999). The hair cells of the ear appear to be the only tissue affected by these antibiotics in humans, possibly due to the concentration and persistence of many different drugs in the perilymph of the ear (Harada et al. 1986) or an inherent propensity to undergo apoptosis. The A1555G mutation is just upstream of the 3'-terminal stem-loop of the mitochondrial 12S small subunit rRNA. This site is very near the predicted target of h-mtTFB1 and h-mtTFB2 for methylation. It is hypothesized that this mutation's ability to cause deafness is linked to the methylation status of the terminal stem-loop and affecting activity of h-mtTFB1 could effectively protect patients with this mutation. No such correlation could be made for h-mtTFB2 (Bykhovskaya et al. 2004). The link to deafness and rRNA structure increases the possibility of at least h-mtTFB1 being a methyltransferase in mitochondria, but whether the h-mtTFB factors directly methylate the human 12S rRNA must be empirically determined.

In cultured cells from Drosophila, knockdown of mtTFB1 was shown to have no major effect on mitochondrial transcription, but resulted in dramatically decreased mitochondrial translation (Matsushima et al. 2005). This indicates a role for mtTFB1 in translation, most likely due to methylation of rRNA, but the exact nature of this function

is unclear. Studies of mtTFB2 in this same system demonstrated a major function in mitochondrial transcription, but a role in regulating translation was not addressed (Matsushima et al. 2004). These results initially were interpreted to mean that in Drosophila, mtTFB1 and mtTFB2 are functionally distinct and do not have overlapping functions. While it is highly suggestive that in the conditions given in cultured Drosophila cells mtTFB1 cannot complement transcription during loss of mtTFB2 and mtTFB2 cannot complement translation during loss of mtTFB1, one cannot rule out the possibility that each protein has a dual function that contributes to normal biology. The research to be described in the next chapter has focused on examining whether h-mtTFB2 is also a putative dual function protein, the evolutionary steps that occurred to generate two of these factors, and the biological roles of these factors that have necessitated the maintenance of two similar proteins in human mitochondria.

3.2 Mitochondrial transcription

In the past few years major advances have been made in our understanding of how mammalian mitochondrial transcription functions and is regulated. Totally recombinant *in vitro* transcription has been achieved identifying the components described above as the minimal set of factors required for specific transcription (Falkenberg et al. 2002). Combining equal molar amounts of either h-mtTFB1 or h-mtTFB2 with POLRMT, providing a template containing the control region of human mtDNA (sequences spanning the LSP to HSP), and addition of up to a 200 fold molar excess of h-mtTFA results in robust specific transcription from the LSP. The HSP is stimulated only minimally as has been observed from mitochondrial extracts. Based on

these simple observations and the characteristics of each factor that have been described previously one can begin to develop models of human mitochondrial transcription.

3.2.1 Models of initiation.

There are three basic models for describing how mitochondrial transcription initiation might take place. These models I will define as mtTFA directed, POLRMT directed, and sequential. These models attempt to integrate the known activities for each protein, and are for the most part mutually exclusive and can be tested in an *in vitro* setting quite easily. In all of these descriptions mtTFB can represent either mtTFB1 or mtTFB2. The first model, as the name states, places many of the initial steps of initiation in the hands of mtTFA. In this scenario mtTFA binds to its high affinity site near the LSP. Its ability to bind and wrap DNA, in the context of a circular molecule, could facilitate unwinding of the DNA near the binding site exposing the single stranded promoter. The tail of mtTFA then recruits a h-mtTFB: POLRMT holoenzyme and, based on the length of the tail and the orientation of mtTFA DNA binding, places the RNA polymerase in the proper location for initiation. Once bound to mtTFA the holoenzyme is held in close proximity and proper spacing to single stranded DNA and can initiate only at specific nucleotides it can reach. The spacing of binding sites for mtTFA near sites of initiation has been documented to be important for transcription to occur (Dairaghi et al. 1995) and one could hypothesize that the location of the mtTFA binding site is the critical factor for promoter selectivity, not the sequence of the promoter itself.

The second model places promoter selection activity firmly on the RNA polymerase. In this case, mtTFA merely acts as a tether to bring and maintain POLRMT near the mtDNA. Again an mtTFB: POLRMT holoenzyme is recruited to the mtDNA

via an interaction with the carboxy-terminal tail of mtTFA, but now mtTFA is not responsible for melting the DNA at the promoter or establishing an open complex. The RNA polymerase holoenzyme once held near the DNA can melt the DNA exposing a single strand that is then bound in a sequence specific fashion and transcription is initiated. Here both spacing of mtTFA binding sites and the sequence of the promoter itself are essential to achieve faithful transcription.

The third model specifies sequential binding of the transcription machinery before transcription can occur. As with the other two models, the first step involves binding of mtDNA by mtTFA. Once mtTFA is bound to a location near a promoter, mtTFB can then be recruited. Together mtTFA and mtTFB are able to form an open complex that POLRMT is subsequently able to recognize mitochondrial promoters and initiate transcription. Specific promoter sequence recognition could reside in the mtTFB proteins, POLRMT, or a combination of the two with the order of these binding events being the crucial parameter.

The first two models are the more likely scenarios based upon experimental data. POLRMT and h-mtTFB1 or h-mtTFB2 are known to bind to each other in the absence of any template or h-mtTFA (Falkenberg et al. 2002) making it very likely that they form a holoenzyme off template. Since h-mtTFA has many binding sites throughout the mitochondrial genome, it becomes very difficult to explain how a protein that can bind in many locations and possibly produce the same structural changes in the mtDNA at each binding site can act as the major determining factor for promoter selection. Instead it seems that h-mtTFA may induce structural changes throughout the mtDNA, possibly unwinding the mtDNA in many positions locally, and the sequence specificity for

promoter utilization lies in the h-mtTFB: POLRMT holoenzyme. In fact POLRMT has been shown to contribute critically to sequence recognition of mammalian mitochondrial promoters, but neither POLRMT or h-mtTFB1 or h-mtTFB2 are able to actually bind to the promoters alone (Gaspari et al. 2004). The mitochondrial RNA polymerase has been shown to be able to specifically activate transcription alone in yeast, but only on templates that are premelted or supercoiled exposing the single stranded nucleotides near the site of initiation. Addition of sc-mtTFB actually can inhibit this process, reducing the overall rate of transcription but producing more fully extended products (Matsunaga and Jaehning 2004). It seems very likely that POLRMT has also maintained the ability to recognize and initiate from single stranded promoters, but the function of h-mtTFB is to stabilize single-stranded DNA in an open complex, possibly even the melting of the DNA as well, thus allowing POLRMT to perform its duty. Neither of these factors is capable of binding normal, linear double stranded DNA, this function has been extended to h-mtTFA. H-mtTFA binds to its high affinity sites and induces a conformational change or even partially melts the mtDNA near human mitochondrial promoters and recruits the holoenzyme to the promoter location. Once in the proper location the holoenzyme is able to fully melt the DNA, with h-mtTFB maintaining the open complex and POLRMT making the sequence specific contacts to initiate transcription.

3.2.2 Elongation and Termination

Regardless of how promoter selection and initiation occur, the basic steps of elongation and eventually termination must follow. Like initiation, the events that occur shortly thereafter are poorly understood and are subject mostly to speculation. In other systems, once initiation has occurred many components are left at the promoter and the

RNA polymerase proceeds on its way. In yeast, sc-mtTFB maintains its interaction with Rpo41p for at least 12 nucleotides after initiation (Mangus et al. 1994). After this period of time sc-mtTFB is released and Rpo41p continues with elongation. In this regard sc-mtTFB does function similarly to a sigma factor and can be recycled upon release. This recycling could allow a small amount of sc-mtTFB to catalyze many different initiation events. It is unclear at this point whether mtTFA and/or the mtTFB proteins remain bound to POLRMT for any period of time after initiation, but it seems likely that an h-mtTFB: POLRMT holoenzyme exists for at least a short period of time during elongation.

Since human mitochondrial transcription utilizes a circular template, elongation must be terminated to both create specific types of transcripts and prevent collisions of highly processive elongating complexes with initiating complexes at promoters. How long transcripts are terminated is poorly understood, but termination of transcripts containing the two mitochondrial rRNAs has been well described. The major documented site of mitochondrial transcription termination is located just downstream of the 3′ terminus of the 16S rRNA transcript within the flanking leucine tRNA (Christianson and Clayton 1986). This termination event effectively creates an rRNA transcriptional unit allowing the effective accumulation of rRNAs versus mRNAs. The sequence associated with termination of this transcript serves as a high-affinity binding site for the mitochondrial termination factor (mTERF) (Kruse et al. 1989). This factor functions as a trimer and preferentially binds to the light strand DNA (Nam and Kang 2005), giving mTERF the ability to terminate transcription in only one direction (Shang and Clayton 1994). Transcription must also proceed through the rRNAs and through the

rest of the mtDNA in order to produce the mRNAs encoded by the light strand. The decision to terminate after the 16S RNA or to proceed appears to be determined in part by the mTERF, but not simply by binding of the factor to the termination sequence. Instead mTERF can also bind simultaneously to a region between the LSP and HSP forming a loop of DNA containing the rRNA genes (Prieto-Martín et al. 2004; Martin et al. 2005). Transcription from the HSP then initiates in one of two locations (Figure 1). If transcription initiates at H1, originally defined as the single HSP, transcription will be terminated just past the 16S rRNA. However, if transcription begins at a second site (H2), just two nucleotides upstream of the 12S rRNA, elongation will not terminate and transcription proceeds to make long, polycistronsic transcripts (Martin et al. 2005). It is not clear if mTERF is the only factor mediating this phenomenon or if other factors may also be involved. Thus, it is tempting to speculate that one of the reasons for two h-mtTFB factors could be to influence the decision of initiating transcription from either H1 or H2.

3.3 Mitochondrial ribosomes and the h-mtTFB factors

As described in earlier sections, mitochondrial contain their own dedicated ribosomes that are responsible for the translation of mRNAs encoded by the mitochondrial genome. Much of our understanding of mitochondrial ribosomes comes from studies in bacteria, but recently a structure of the bovine mitochondrial ribosome has been obtained. Mammalian mitochondrial ribosomes share many common functional characteristics, but their structures are somewhat different. While there may be differences in the overall protein and nucleic acid makeup of these two types of

ribosomes, they both require similar steps in assembly to produce the most accurate, efficient ribosomes possible.

Production of bacterial ribosomes involves a complicated multi-step process where many ribosomal proteins are assembled around an rRNA core to make a large biological machine. The large subunit (50S) and small subunit (30S) are synthesized independently then associate during translation initiation. The first step in this process is to properly transcribe and process the rRNAs. Typically nuclear rRNA genes reside close together and are initially transcribed as one large unit (Nierhaus 1991). This large poly-cistronic transcript must be then processed to form the necessary individual rRNAs. This processing does not occur simply as one cleavage followed by another. Instead the ribosomal RNAs are simultaneously being cleaved, modified, and complexed with proteins in a controlled step-wise fashion. Certain processing events must happen in a particular order and many modifications do not occur until partial assembly has been performed (Fatica and Tollervey 2002). The KsgA protein mentioned previously is a prime example of a modifying enzyme that requires a partially assembled bacterial ribosome before it can recognize its target (Desai and Rife 2006). Thus defects in any single process, be it ribosomal protein production, cleavage, or modification can lead to improperly assembled ribosomes and severe defects in cellular function.

In mitochondria these steps are not known, but most likely retain a similar order of events owing to their bacterial heritage. The ribosomal proteins are all synthesized from nuclear genes, translated in the cytoplasm, and then imported into mitochondria (O'Brien 2003). Once in mitochondria, they must be assembled with 16S and 12S rRNAs transcribed from the mtDNA. The mitochondrial ribosomes typically contain many more

proteins and less RNA so some assembly steps may have been lost and new steps gained (Sharma et al. 2003). It is not known whether any of these steps are coupled, but based on the occurrence of a specific rRNA transcriptional unit in human mitochondria, few cleavages must occur to produce the proper RNA sequences (Martin et al. 2005). The mitochondrial rRNAs of vertebrates would only theoretically require four cleavage events, all of which would be taken care of by the tRNA processing machinery so its seems unlikely that there are RNA endonucleases devoted specifically to this process. However, the processing of yeast cytoplasmic rRNA requires *DIM1* (Lafontaine et al. 1994) and two homologs of this protein, h-mtTFB1 and h-mtTFB2, are present in human mitochondria leaving open the possibility that these enzymes could play a role in the processing of the rRNAs. Additionally there appear to be very few modified bases in mammalian rRNAs suggesting many fewer steps in the assembly process (Dubin 1974). Two of the few conserved modifications in human mitochondrial rRNAs are thought to be catalyzed by h-mTFB1 and h-mtTFB2 (Seidel-Rogol et al. 2003), and could be evidence that transcription, modification, and assembly of ribosomes are coupled in human mitochondria. The exact involvement of the h-mtTFB factors in these events has yet to be established, but decreasing levels of mtTFB1 in Drosophila have negative impacts on mitochondrial translation (Matsushima et al. 2005). These observations are highly suggestive of major contributions of these factors to the overall process of synthesizing mitochondrial ribosomes. Coupling of these processes by a class of proteins involved in several steps of ribosome production might well be a target for modulating ribosome synthesis in mitochondria and the mtTFB factors could have major impacts on gene expression at both transcriptional and translation levels in human mitochondria.

Chapter IV: Results

Portions of this chapter reproduced from Cotney and Shadel. J.Mol. Evol. (2006) and Cotney, Wang, and Shadel. Nucleic Acids Res. (2007).

Identification of two homologs of the sc-mtTFB protein in human mitochondria was a great advancement to our understanding of human mitochondrial gene expression (Falkenberg et al. 2002; McCulloch et al. 2002). These proteins were identified to be essential for specific transcription along with mtTFA and POLRMT (Falkenberg et al. 2002). These factors also have homology to rRNA methyltransferases and h-mtTFB1 was demonstrated to have maintained this activity on a heterologous substrate (Seidel-Rogol et al. 2003). Having two mtTFB factors seems to be a common theme amongst vertebrates and higher eukaryotes, but the presence of a single homolog in lower eukaryotes, yeast and worms led many to speculate that a duplication event had occurred relatively late during eukaryotic evolution (Rantanen et al. 2003). Additionally, despite the possibility of being dual function proteins, studies in Drosophila suggested that the mtTFB factors had non-overlapping and totally separate functions (Matsushima et al. 2005; Matsushima et al. 2004). These observations seem to paint two different pictures of the h-mtTFB factors. One being an evolutionary event that occurred early and allowed maintenance of two activities in both proteins and the other being a later duplication event that created two proteins with totally different functions *in vivo*. The outcomes of such models would impact how mitochondrial transcription would be predicted to function in various organisms and affect our ability to extend what we learn in model systems like yeast, worms, and flies to mammalian mitochondrial gene expression. This research aims to understand how these genes have evolved and predict how transcription might function in mitochondria of humans and other organisms. Homologs of the h-mtTFB factors were identified in many different species including bacteria, archaea, and both mitochondrial and cytoplasmic homologs in eukaryotes. Comparisons of primary

sequence coupled with crystal structure data and known functions of several members of each group of proteins analyzed allowed us to make predictions about activities of many proteins, including h-mtTFB2 and sc-mtTFB. To test predictions made for the functions of h-mtTFB2 and sc-mtTFB and what roles they might have in mitochondrial gene expression, both h-mtTFB2 and sc-mtTFB were assayed for rRNA methyltransferase activity. Finally, having information about the evolution and activities of the mtTFB factors in hand, research was conducted to understand the individual contributions of each h-mtTFB protein to rRNA methylation, transcription, and various components of mitochondrial biology. Together these studies have given us a more complete view of the roles and activities of h-mtTFB1 and h-mtTFB2 *in vivo* and point to functions in a previously undescribed mitochondrial signaling pathway.

4.1 Divergent evolution of the mitochondrial transcription factor B family

4.1.1 Duplication of the ancestral rRNA methyltransferase

Using a refined sequence alignment (Supplemental Figure 1) of representative sequences for prokaryotic, eukaryotic, and archaeal dimethyltransferases, as well as known and predicted mtTFB orthologs (Table 1), I performed parsimony analysis using the PHYLIP package (Felsenstein 2005) to determine a consensus phylogenetic tree for this family of proteins (Figure 2). Consistent with the findings of Shutt and Gray (2006), I found that the bacterial and eukaryotic cytoplasmic KsgA orthologs form distinct clades that branch from a common ancestral point and that archaeal orthologs branch off early from the eukaryotic lineage.

4.1.2 Lineages of mtTFBs

Likewise, I found that the mtTFB1 and mtTFB2 groups also form distinct lineages. Unexpectedly, the mtTFB1 and mtTFB2 proteins do not form a single clade that is distinct from the fungal mtTFBs. Instead, the branch point between mtTFB1 and mtTFB2 occurs such that the fungal mtTFBs and mtTFB2 form a clade (Bootstrapping value of 100) that is separated from the mtTFB1 clade, consistent with a much earlier gene duplication event than proposed previously (Rantanen et al. 2003). Additionally, branching of the fungal mtTFBs from the mtTFB2 group to form a separate clade is highly supported (Bootstrapping value of 91).

To investigate these observations in more detail and eliminate the possibility of other functionally distinct sequences artificially producing this branching order, a second round of phylogenetic analysis was performed using only the mitochondrial transcription B sequences and E. coli KsgA as an outgroup. The protist mtTFB homologs were excluded from this alignment in order to reduce the possibility of lateral gene transfer events complicating tree construction. Sequences were again aligned using Clustal and manually inspected. The final refined alignment (Supplemental Figure 2) was then subjected to phylogenetic analysis using Tree-Puzzle and the Mega 3.1 software package.

Trees generated by four methods showed very similar branching patterns (Figure 3). In this phylogenetic analysis three distinct clades are observed: mtTFB1, mtTFB2, and fungal mtTFB. However, where previously the branching order of these clades had not been resolved (Shutt and Gray 2006), I again observed mtTFB2 and fungal mtTFB sequences forming a clade distinct from mtTFB1. This branching order had very high support values by all methods (Figure 3).

Table 1 Accession numbers of amino acid sequences used in this study

Species	Dimethyltransferase	mtTFB1	mtTFB2	mtTFB
Thermus thermophilus	YP_005890			
Clostridium difficile	ZP_01030144			
Bacillus cereus	EAL11808			
Thermosynechococcus elongates	NP_681448			
Cyanobacteria bacterium Yellowstone	YP_474382			
Nitrococcus mobilis	EAR23300			
Bordetella pertussis	NP_881867			
Neisseria meningitides	NP_273739			
Burkholdena mallei	YP_102046			
Nitrosomonas eutropha	EAO18317			
Thiobacillus denitrificans	YP_316098			
Legionella pneumophila	YP_128180			
Xylella fastidiosa	NP_299427			
Pseudomonas putida	EAP49827			
Shewanella putrefaciens	EAO93098			
Yersinia pestis	NP_404135			
Chlorobium chlorochromatii	YP_379498			
Escherichia coli	NP_285748			
Salmonella typhimurium	NP_459095			
Methanosarcina mazei	Q8PU18			
Methanopyrus kandleri	NP_614217			
Pyrococcus abyssi	NP_560585			
Methanocaldococcus jannashii	NP_248023			
Ferroplasma acidarmanus	EAM94385			
Pyrobaculum aerophilum	NP_126070			
Methanospirillum hungatei	YP_504435			
Archaeoglobus fulgidus	NP_070611			
Haloarcula marismortui	YP_135020			
Caenorhabditis elegans	E02H1.1	T03F1.7		
Caenorhabditis briggsae	CAE57911	CAE66633		
Apis melifermis	XP_624425	XP_395897		
Anopheles gambiae	ENSANGG00000014455	ENSANGG00000006874	ENSANGG00000009115	
Drosophila melanogaster	NP_651660	NP_996062	NP_649971	
Ciona intestinalis	ENSCING00000009730	ENSCING00000003901		
Tetraodon nigroviridis	GSTENG00028675001	GSTENG00031327001	GSTENT00032827001	
Takifugu rubripes	SINFRUG00000142444	SINFRUP00000139638	SINFRUG00000150836	
Danio rerio	ENSDARG00000005057	ENSDARG00000040727	XM_692861	
Xenopus laevis	AAI06333	AAH56010		
Xenopus tropicalis	ENSXETG00000014830	NP_001016494	ENSXETG00000013456	
Gallus gallus	XM_424746	XP_426165	XP_419535	
Monodelphis domestica	ENSMODG00000019610	ENSMODG00000008169		
Rattus norvegicus	XP_215477	NP_852139	NP_001008294	
Mus musculus	NM_025447	NM_146074	NP_032275	
Bos taurus	XP_586750	ENSBTAP00000013261	ENSBTAG00000000072	
Canis familiaris	XP_535250	ENSCAFP00000000817	XM_537224	
Pan paniscus		CAH93207		
Pan troglodytes	XM_517743		ENSPTRG00000002188	
Homo sapiens	AAC97955	NP_057104	NP_071761	
Schizossaccharomyces pombe	NP_596122			NP_593495
Debaryomyces hansenii				XP_461682
Candida albicans				EAL03685
Eremothecium gossypii				NP_986111
Kluyveromyces lactis	P78697			XP_451899
Candida glabrata				XP_449831
Saccharomyces kluyveri	YM479-Contig2343			AAC49738
Saccharomyces douglasii				Q9Y829
Saccharomyces cerevisiae	NP_015057			CAA90199
Dictyostelium discoideum	EAL65509			EAL64275
Trypanosoma brucei	XP_825820			AAM97535
Leishmania major				XP_843186

* denotes sequences from Ensemble.org, all other sequences retrieved from NCBI.

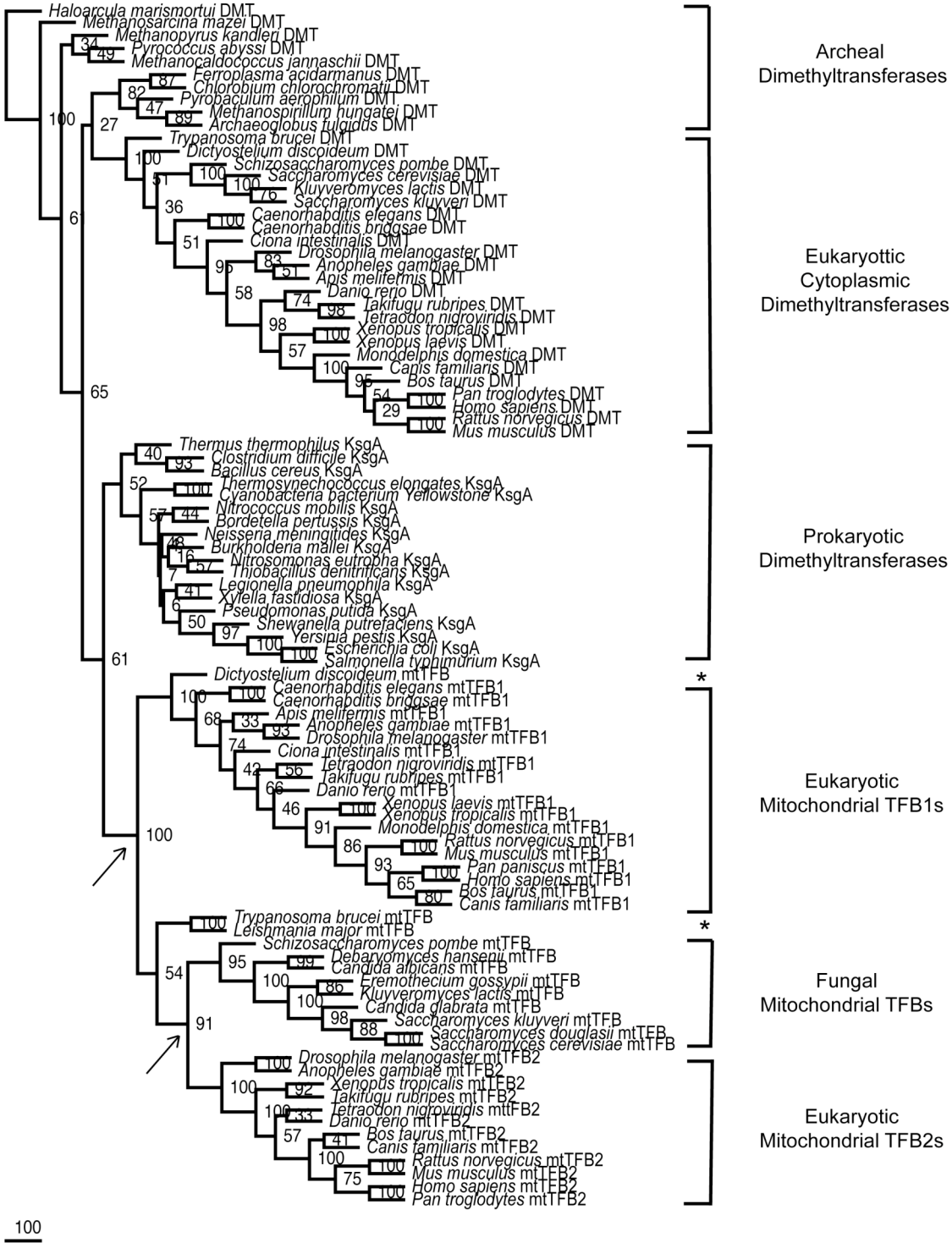

Haloarcula marismortui DMT
Methanosarcina mazei DMT
Methanopyrus kandleri DMT
Pyrococcus abyssi DMT
Methanocaldococcus jannaschii DMT
Ferroplasma acidarmanus DMT
Chlorobium chlorochromatii DMT
Pyrobaculum aerophilum DMT
Methanospirillum hungatei DMT
Archaeoglobus fulgidus DMT
Trypanosoma brucei DMT
Dictyostelium discoideum DMT
Schizosaccharomyces pombe DMT
Saccharomyces cerevisiae DMT
Kluyveromyces lactis DMT
Saccharomyces kluyveri DMT
Caenorhabditis elegans DMT
Caenorhabditis briggsae DMT
Ciona intestinalis DMT
Drosophila melanogaster DMT
Anopheles gambiae DMT
Apis mellifermis DMT
Danio rerio DMT
Takifugu rubripes DMT
Tetraodon nigroviridis DMT
Xenopus tropicalis DMT
Xenopus laevis DMT
Monodelphis domestica DMT
Canis familiaris DMT
Bos taurus DMT
Pan troglodytes DMT
Homo sapiens DMT
Rattus norvegicus DMT
Mus musculus DMT
Thermus thermophilus KsgA
Clostridium difficile KsgA
Bacillus cereus KsgA
Thermosynechococcus elongates KsgA
Cyanobacteria bacterium Yellowstone KsgA
Nitrococcus mobilis KsgA
Bordetella pertussis KsgA
Neisseria meningitides KsgA
Burkholderia mallei KsgA
Nitrosomonas eutropha KsgA
Thiobacillus denitrificans KsgA
Legionella pneumophila KsgA
Xylella fastidiosa KsgA
Pseudomonas putida KsgA
Shewanella putrefaciens KsgA
Yersinia pestis KsgA
Escherichia coli KsgA
Salmonella typhimurium KsgA
Dictyostelium discoideum mtTFB
Caenorhabditis elegans mtTFB1
Caenorhabditis briggsae mtTFB1
Apis mellifermis mtTFB1
Anopheles gambiae mtTFB1
Drosophila melanogaster mtTFB1
Ciona intestinalis mtTFB1
Tetraodon nigroviridis mtTFB1
Takifugu rubripes mtTFB1
Danio rerio mtTFB1
Xenopus laevis mtTFB1
Xenopus tropicalis mtTFB1
Monodelphis domestica mtTFB1
Rattus norvegicus mtTFB1
Mus musculus mtTFB1
Pan paniscus mtTFB1
Homo sapiens mtTFB1
Bos taurus mtTFB1
Canis familiaris mtTFB1
Trypanosoma brucei mtTFB
Leishmania major mtTFB
Schizosaccharomyces pombe mtTFB
Debaryomyces hansenii mtTFB
Candida albicans mtTFB
Eremothecium gossypii mtTFB
Kluyveromyces lactis mtTFB
Candida glabrata mtTFB
Saccharomyces kluyveri mtTFB
Saccharomyces douglasii mtTFB
Saccharomyces cerevisiae mtTFB
Drosophila melanogaster mtTFB2
Anopheles gambiae mtTFB2
Xenopus tropicalis mtTFB2
Takifugu rubripes mtTFB2
Tetraodon nigroviridis mttFB2
Danio rerio mtTFB2
Bos taurus mtTFB2
Canis familiaris mtTFB2
Rattus norvegicus mtTFB2
Mus musculus mtTFB2
Homo sapiens mtTFB2
Pan troglodytes mtTFB2
100
Archeal Dimethyltransferases
Eukaryottic Cytoplasmic Dimethyltransferases
Prokaryotic Dimethyltransferases
Eukaryotic Mitochondrial TFB1s
Fungal Mitochondrial TFBs
Eukaryotic Mitochondrial TFB2s
*
*

Figure 2 Phylogenetic analysis of the B family of mitochondrial transcription factors and rRNA adenine dimethyltransferases from bacteria, archaea, and eukaryotes.

Shown is a consensus parsimony tree generated from PROTPARS using a gamma model with $\alpha = 2.04$ and eight rate categories as estimated by Tree-Puzzle. Full genus and species names are listed at the ends of the branches they represent as is a designation of the type of protein: rRNA dimethyltransferase (DMT), prokaryotic rRNA dimethyltransferase (KsgA), mitochondrial transcription factor B1 (mtTFB1) ortholog, mitochondrial transcription factor B2 ortholog (mtTFB2), and a separate designation for organisms that contain only a single mitochondrial transcription factor B ortholog (mtTFB). Clades are defined by brackets to right with * denoting protist mtTFB homologs. Bootstrap values from 100 replicates are listed at each node. Arrows signify nodes discussed in text. The scale bar represents 100 substitutions per sequence.

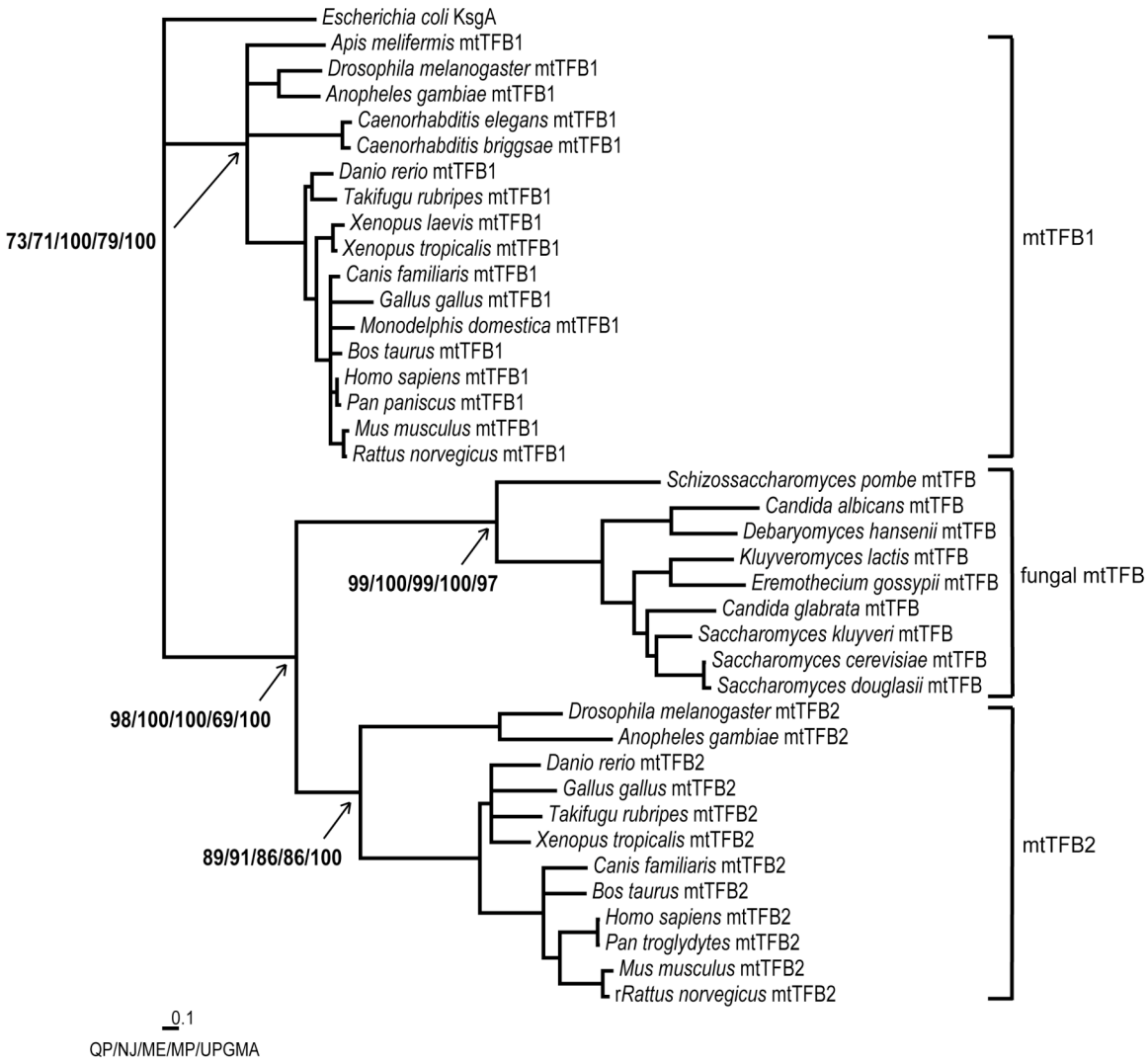

Escherichia coli KsgA
Apis melifermis mtTFB1
Drosophila melanogaster mtTFB1
Anopheles gambiae mtTFB1
Caenorhabditis elegans mtTFB1
Caenorhabditis briggsae mtTFB1
Danio rerio mtTFB1
Takifugu rubripes mtTFB1
Xenopus laevis mtTFB1
Xenopus tropicalis mtTFB1
Canis familiaris mtTFB1
Gallus gallus mtTFB1
Monodelphis domestica mtTFB1
Bos taurus mtTFB1
Homo sapiens mtTFB1
Pan paniscus mtTFB1
Mus musculus mtTFB1
Rattus norvegicus mtTFB1
mtTFB1
73/71/100/79/100
Schizossaccharomyces pombe mtTFB
Candida albicans mtTFB
Debaryomyces hansenii mtTFB
Kluyveromyces lactis mtTFB
Eremothecium gossypii mtTFB
Candida glabrata mtTFB
Saccharomyces kluyveri mtTFB
Saccharomyces cerevisiae mtTFB
Saccharomyces douglasii mtTFB
fungal mtTFB
99/100/99/100/97
98/100/100/69/100
Drosophila melanogaster mtTFB2
Anopheles gambiae mtTFB2
Danio rerio mtTFB2
Gallus gallus mtTFB2
Takifugu rubripes mtTFB2
Xenopus tropicalis mtTFB2
Canis familiaris mtTFB2
Bos taurus mtTFB2
Homo sapiens mtTFB2
Pan troglydytes mtTFB2
Mus musculus mtTFB2
rRattus norvegicus mtTFB2
mtTFB2
89/91/86/86/100
0.1
QP/NJ/ME/MP/UPGMA

Figure 3 Phylogenetic analysis of the mtTFB family of mitochondrial transcription factors.

Shown is a quartet-puzzling tree generated by Tree-Puzzle. Full genus and species names are listed at the ends of the branches they represent as is a designation of the type of protein: prokaryotic rRNA dimethyltransferase (KsgA), mitochondrial transcription factor B1 (mtTFB1), mitochondrial transcription factor B2 (mtTFB2), and a separate designation for fungal mitochondrial transcription factor B (mtTFB). Quartet puzzling (QP) values from Tree-Puzzle and bootstrap values of Neighbor-Joining (NJ), Maximum Evolution (ME), Maximum Parsimony (MP), and UPGMA analyses (gamma model, α = 3.18) from the Mega 3.1 software suite are listed in the order indicated at the four key branch points of the tree (arrows). The scale bar represents 0.1 substitutions per sequence.

4.1.3 Maintenance of rRNA methyltransferase activity

Dimethylation of two adenine residues in a 3'-terminal stem-loop of small subunit rRNAs is highly conserved, occurring on bacterial, eukaryotic cytoplasmic, and most mitochondrial rRNAs. In E. coli, KsgA methylates this stem-loop and loss of this activity results in resistance to the translational inhibitor kasugamycin (Helser et al. 1972; van Buul and van Knippenberg 1985). We showed previously that h-mtTFB1 is able to functionally complement an *E. coli* ksgA mutant by methylating this substrate, demonstrating that it has the predicted rRNA methyltransferase activity and implicating it as the enzyme responsible for modification of the homologous stem-loop in the human mitochondrial 12S rRNA subunit (Seidel-Rogol et al. 2003).

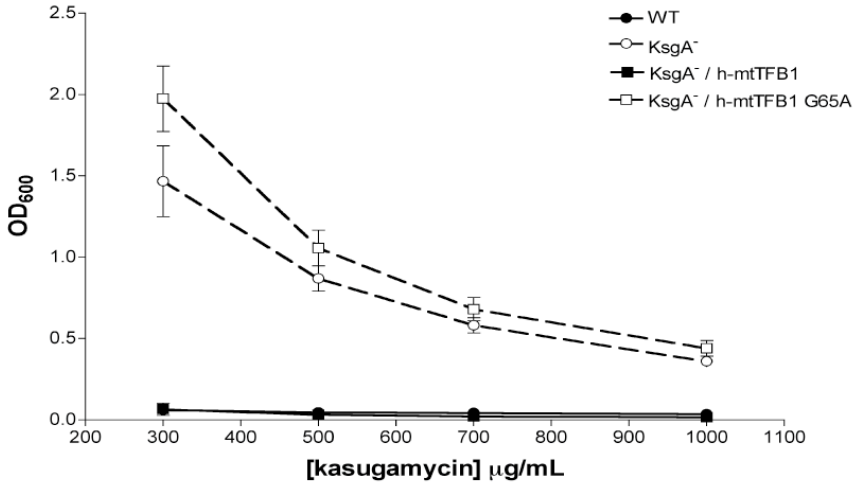
Using this same strategy, we determined whether h-mtTFB2 also has this enzymatic activity. As reported previously (Seidel-Rogol et al. 2003), expression of h-mtTFB1 completely restored sensitivity to the antibiotic in a ksgA mutant strain (Figure 4A). Likewise, expression of wild-type h-mtTFB2 in the same ksgA mutant background resulted in significantly increased sensitivity to kasugamycin compared to the corresponding kasugamycin-resistant control strain (Figure 4B), indicating that h-mtTFB2 also has rRNA methylation activity. However, h-mtTFB2 only partially restored this phenotype (Figure 4B), indicating that the activity of h-mtTFB2 was not as robust as h-mtTFB1 in this assay.

In S. cerevisiae, sc-mtTFB is the only obvious mitochondrial transcription factor B ortholog. Given that it too has structural similarity to KsgA rRNA methyltransferases (O'Farrell et al. 2004; Schubot et al. 2001), we tested whether it has this biological

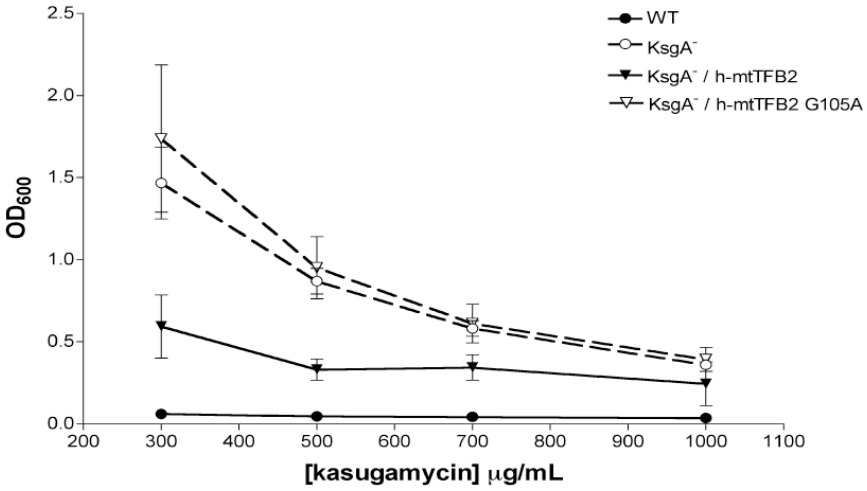
activity. Expression of sc-mtTFB in a ksgA mutant strain had no effect on kasugamycin sensitivity and, therefore, has no detectable activity in this assay (Figure 4C).

Given the apparent lower efficiency of h-mtTFB2 compared to h-mtTFB1, we wanted to confirm that its rRNA methyltransferase activity was co-factor dependent, as we showed previously for h-mtTFB1 (Seidel-Rogol et al. 2003). To examine this, we made a mutation in a conserved residue (G105A) in the predicted SAM-binding site of h-mtTFB2 by site-directed mutagenesis and tested its ability to restore antibiotic sensitivity in the ksgA mutant strain. The analogous mutation in h-mtTFB1 (G65A) was shown previously to inhibit SAM binding (McCulloch and Shadel 2003) and inactivates its enzymatic activity in this assay (Seidel-Rogol et al. 2003) and was used as a control in this experiment (Figure 4A). As was the case for the h-mtTFB1 G65A mutant, expression of the G105A h-mtTFB2 mutant had no effect on the sensitivity to kasugamycin (i.e. the strain was equally as resistant as the ksgA mutant strain alone; Figure 4B), indicating its rRNA methyltransferase activity was likewise SAM-dependent. We note that the final cell densities reached in the current study in the presence of drug differ from those we published previously using this assay (Seidel-Rogol et al. 2003). This difference is due to a modification of the original procedure used in this study that entails beginning the experiment with a greater number of cells (as detailed in Materials and Methods) and not a difference in the doubling times of the strains, which were virtually identical (data not shown).

A


WT
KsgA-
KsgA- / h-mtTFB1
KsgA- / h-mtTFB1 G65A
OD600
[kasugamycin] μg/mL

B


WT
KsgA-
KsgA- / h-mtTFB2
KsgA- / h-mtTFB2 G105A
OD600
[kasugamycin] μg/mL

C

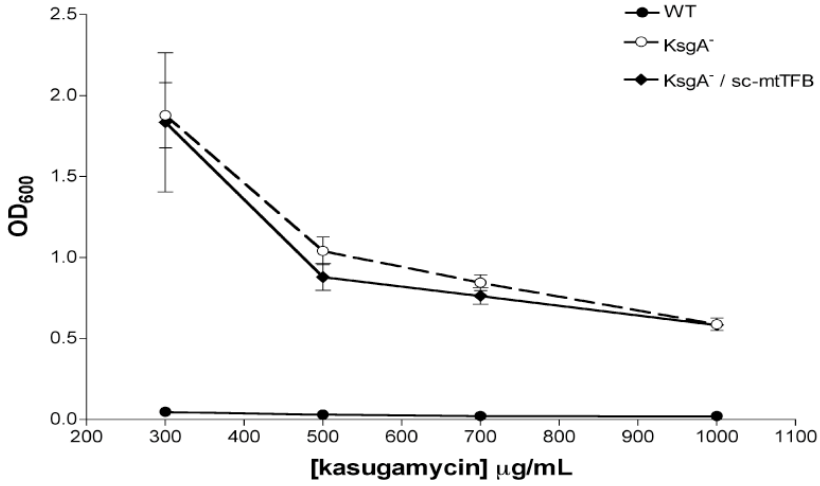
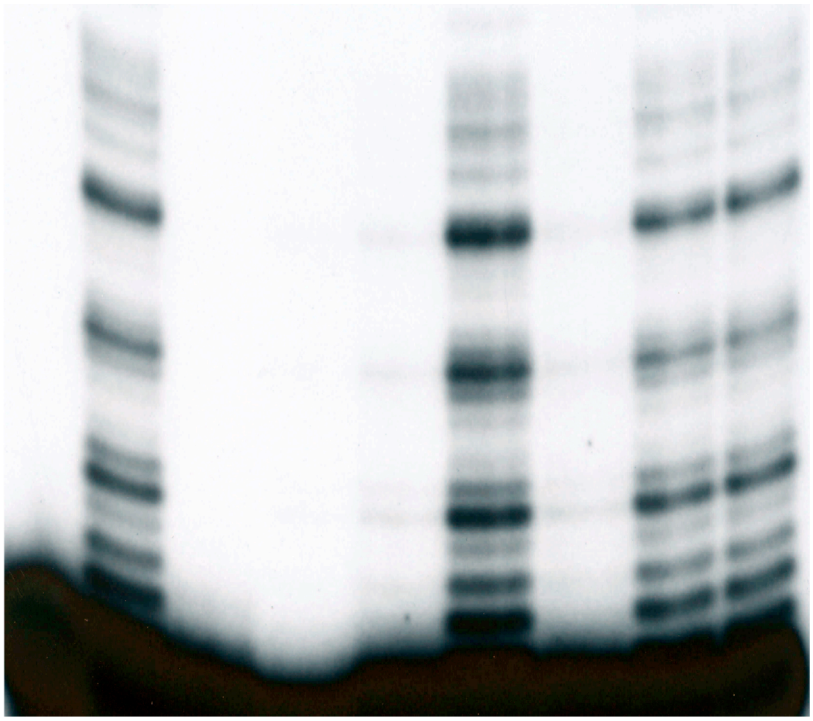

WT
KsgA-
KsgA- / sc-mtTFB
OD600
[kasugamycin] μg/mL

Figure 4 Functional complementation of *E. coli* KsgA rRNA methyltransferase activity by h-mtTFB2, but not its *S. cerevisiae* ortholog sc-mtTFB.

Results of a kasugamycin-sensitivity assay are shown. Plotted is the culture density (OD_{600}) after growth of cultures in the presence of 300, 500, 700, and 1000 μg/mL of kasugamycin. In all experiments two control strains were analyzed, the kasugamycin-resistant *ksgA* mutant strain with an empty-vector control plasmid (KsgA-) and its isogenic kasugamycin-sensitive KsgA+ JM101 parent (WT). A) Results of expressing h-mtTFB1 and the h-mtTFB1-G65A, SAM-binding mutant in the *ksgA* mutant are shown. Note that h-mtTFB1 expressing cells displayed the same sensitivity to kasugamycin as the parental wild-type strain, thus the plot of their data completely overlaps on the graph. B) Results of expressing h-mtTFB2 and the h-mtTFB1-G105A, SAM-binding mutant in the *ksgA* mutant are shown. C) Results of expressing sc-mtTFB in the *ksgA* mutant are shown. Error bars represent the standard deviation of six replicates.

We next determined if the restoration of sensitivity to kasugamycin by h-mtTFB2 was due to methylation of the E. coli 16S rRNA or an indirect effect. Dimethylation of the stem-loop has been shown by us (Seidel-Rogol et al. 2003) and others (Lafontaine et al. 1998) to impede extension by reverse transcriptase *in vitro*, providing a convenient primer extension assay for methylation at this site. Primer extension analysis using a radiolabeled primer that binds to the 3' terminus of 16S rRNA (Figure 5) revealed that the conserved stem-loop was indeed methylated in the ksgA- strain in the presence of h-mtTFB1 or h-mtTFB2 (demonstrated by the lack of longer products corresponding to extension of reverse transcriptase beyond the two methylated adenines; Figure 5, lanes 5 and 7). This methylation-induced block was virtually indistinguishable from that observed using rRNA from the wild type (KsgA+) strain (Figure 5, lane 4), in which the 16S rRNA is fully methylated. As was the case in the ksgA- strain that completely lacks methylation (Figure 5, lane 2), the block in primer extension (i.e. methylation) was significantly reduced in ksgA mutant strains expressing the h-mtTFB1 G65A (Figure 5, lane 6) or h-mtTFB2 G105A (Figure 5, lane 8) SAM-binding-site mutants, indicating loss of methyltransferase activity. Finally, strains expressing sc-mtTFB showed little or no methylation activity in this assay (Figure 5, lane 9). Altogether, these results largely confirm those of the kasugamycin-sensitivity assay (Figure 4) and demonstrate that h-mtTFB1 and h-mtTFB2 both have SAM-dependent, site-specific rRNA adenine methyltransferase activity, while the budding yeast ortholog, sc-mtTFB, does not.


primer alone
ksgA- / empty Vector
ksgA- / empty Vector (-dATP, dCTP)
KsgA+ / empty Vector
ksgA- / h-mtTFB1
ksgA- / h-mtTFB1 G65A
ksgA- / h-mtTFB2
ksgA- / h-mtTFB2 G105A
ksgA- / sc-mtTFB
← primer

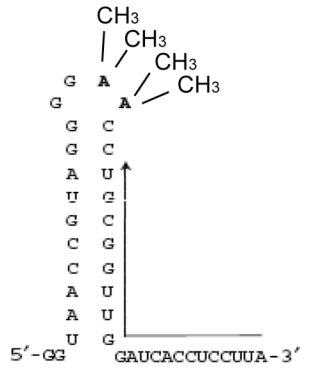

CH3
CH3
CH3
CH3
5′-GG
GAUCACCUCCUUA-3′
E. coli 16S

Figure 5 SAM-dependent methylation of an evolutionarily conserved stem-loop in the *E. coli* 16S rRNA by h-mtTFB2, but not its *S. cerevisiae* ortholog, sc-mtTFB.

An autoradiogram of primer-extension products from the *E. coli* 16S rRNA is shown. The sequence of the 3'-terminal stem-loop of *E. coli* 16S rRNA is shown to the right with the bent arrow representing the radiolabeled primer used in the assay and dimethylation of the two adenines in the loop shown. Reverse transcription from this primer is blocked by either dimethylation event, demonstrated by the lack of longer species that extend past the two adenine residues (i.e. that migrate slower than the primer alone in the gel, indicated by the arrow). Lane 1, radiolabeled primer alone; lane 2, *ksgA*$^-$ strain plus empty pBluescript KS+ (16S rRNA not methylated); lane 3: *ksgA*$^-$ strain plus empty pBluescript KS+ (control reaction lacking dATP, dCTP to mimic methylation block); lane 4: isogenic wild-type (*KsgA*+) strain plus empty pBluescript KS+ (16S rRNA fully methylated); lane 5: *ksgA*$^-$ strain + h-mtTFB1 plasmid; lane 6: *ksgA*$^-$ strain + h-mtTFB1-G65A plasmid; lane 7: *ksgA*$^-$ strain + h-mtTFB2 plasmid; lane 8: *ksgA*$^-$ strain + h-mtTFB2-G105A plasmid; lane 9: *ksgA*$^-$ strain + sc-mtTFB plasmid.

4.2 Relative levels of the human mitochondrial transcription machinery

4.2.1 Specifically identifying mtTFB1 and mtTFB2 via immunoblotting

To fully understand any transcription system a critical parameter to know is the relative amounts of the basal transcription machinery *in vivo*. The approach was to use immunoblotting to establish the levels of the mitochondrial transcription machinery on total cell and mitochondrial bases. This method requires antibodies that are specific for their target, especially for homologous proteins that are very similar in size. Since h-mtTFB1 and h-mtTFB2 share a large degree of similarity with each other and are predicted to be sizes that would not allow great separation on polyacrylamide gels, it was unclear if antibodies we had generated would cross react with the paralogous protein and confound this analysis. I determined the specificity of these antibodies with purified full-length recombinant h-mtTFB1 and h-mtTFB2. I found that the antibodies were indeed highly specific for their targets showing no reactivity with the paralogous protein (Figure 6).

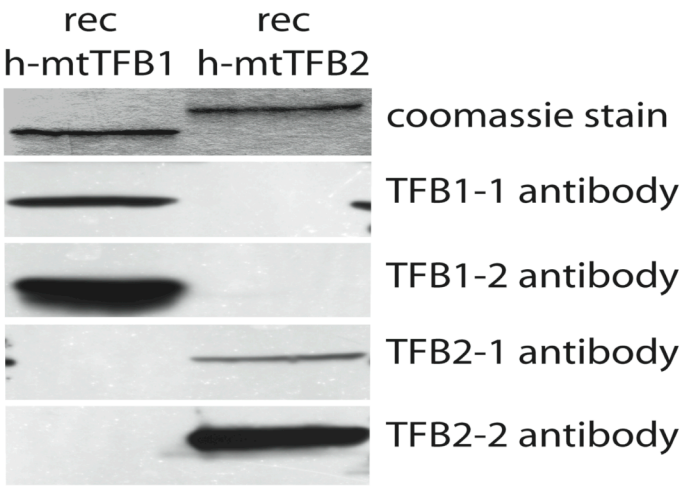


Figure 6 mtTFB antibody specificity

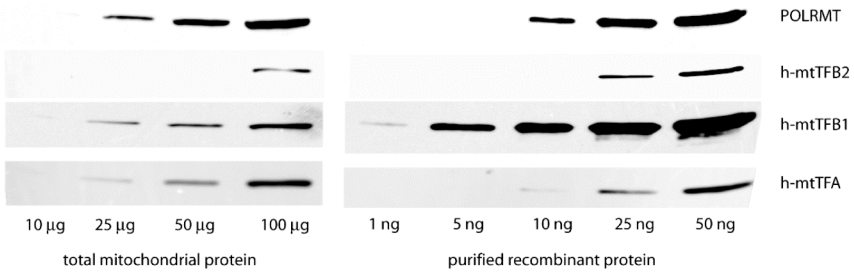
Shown are western blots on 200 ng of recombinant h-mtTFB1 and h-mtTFB2 proteins probed with four peptide antibodies: TFB1-1, TFB1-2, TFB2-1, and TFB2-2. Antibodies were found to specifically recognize their full-length recombinant peptide and not cross react with the paralogous protein. Coomassie staining of full-length recombinant h-mtTFB1 and h-mtTFB2 demonstrates their difference in molecular weight.

4.2.2 Estimation of levels of mitochondrial transcription machinery.

Having antibodies capable of specifically detecting each of the four human mitochondrial transcription proteins, as well as known amounts of corresponding recombinant proteins, allowed us to determine for the first time their relative abundance. We purified mitochondria from logarithmically growing HeLa cells and performed quantitative western blot analysis of POLRMT, h-mtTFA, h-mtTFB1 and h-mtTFB2 on known amounts of total mitochondrial protein. For each protein analyzed, multiple dilutions of total mitochondrial protein or total whole cell lysates were probed in parallel with a dilution series of known amounts of recombinant protein to allow the amount of each transcription component relative to the total amount of mitochondrial lysate to be determined. Representative western blots and standard curves are shown in Figure 7.

Based on this analysis, we calculated the following amounts of each transcription component/100 μg of total mitochondrial protein: 45.8 ± 9.7 ng of h-mtTFA, 4.03 ± .20 ng of h-mtTFB1, 14.9 ± .97 ng of h-mtTFB2, and 32.3 ± 2.1 ng of POLRMT. For whole cell lysates, we were only able to accurately measure h-mtTFA levels and found 11.1 ± 2.4 fg of protein per cell (data not shown). Next, we determined the number of molecules/cell based on the amount of total mitochondrial protein isolated from $4.5x10^7$ cells. We found there to be 25.2 ± 2.51 pg of mitochondrial protein per cell with recoveries of 75% to 80% of total mitochondria for three separate experiments based on HSP60 and h-mtTFA immunoblots (data not shown). These measurements result in $2.52x10^5$, $1.35x10^4$, $4.29x10^4$, and $3.12x10^4$ molecules per cell for h-mtTFA, h-mtTFB1, h-mtTFB2, and POLRMT respectively (Table 2).

A


POLRMT
h-mtTFB2
h-mtTFB1
h-mtTFA
10 µg
25 µg
50 µg
100 µg
1 ng
5 ng
10 ng
25 ng
50 ng
total mitochondrial protein
purified recombinant protein

B

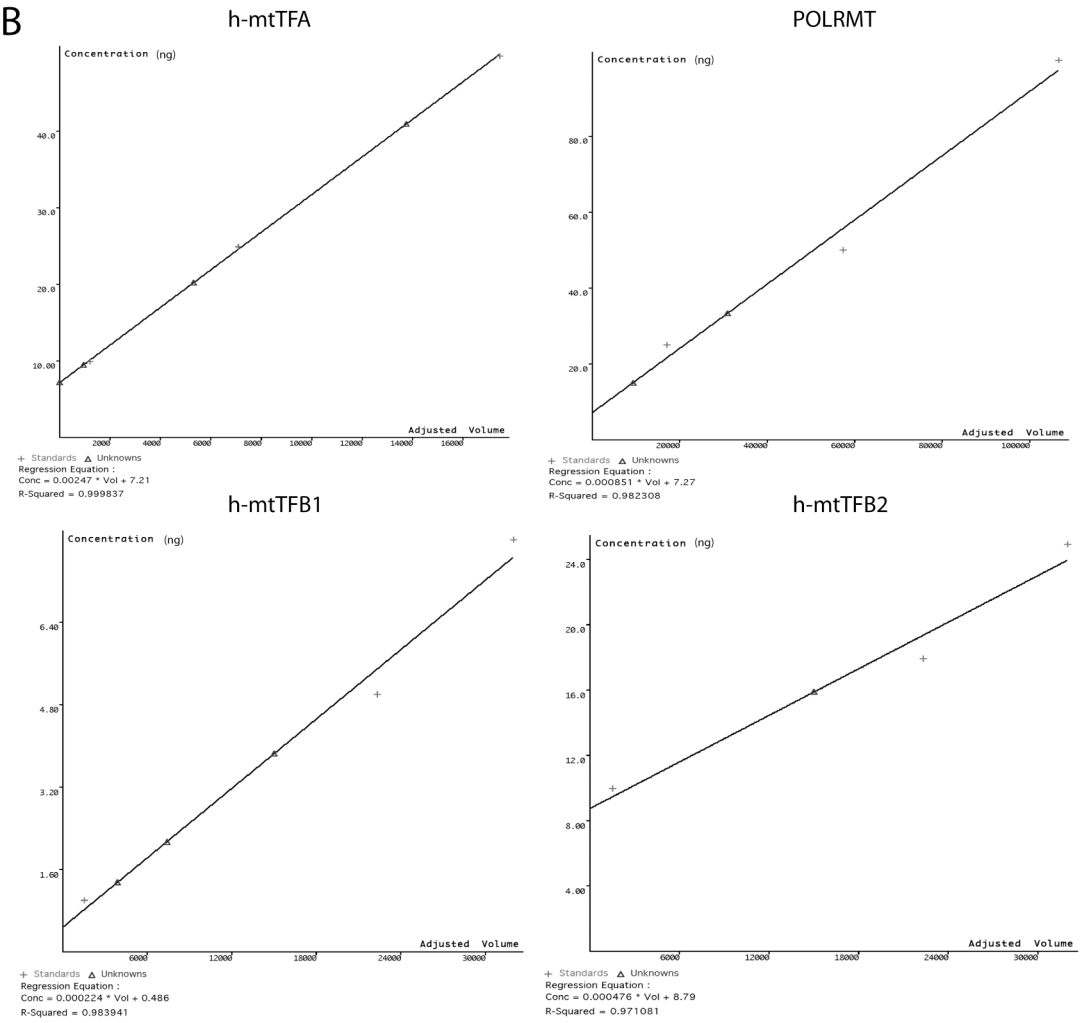

h-mtTFA
Concentration (ng)
Adjusted Volume
+ Standards △ Unknowns
Regression Equation :
Conc = 0.00247 * Vol + 7.21
R-Squared = 0.999837
POLRMT
Concentration (ng)
Adjusted Volume
+ Standards △ Unknowns
Regression Equation :
Conc = 0.000851 * Vol + 7.27
R-Squared = 0.982308
h-mtTFB1
Concentration (ng)
Adjusted Volume
+ Standards △ Unknowns
Regression Equation :
Conc = 0.000224 * Vol + 0.486
R-Squared = 0.983941
h-mtTFB2
Concentration (ng)
Adjusted Volume
+ Standards △ Unknowns
Regression Equation :
Conc = 0.000476 * Vol + 8.79
R-Squared = 0.971081

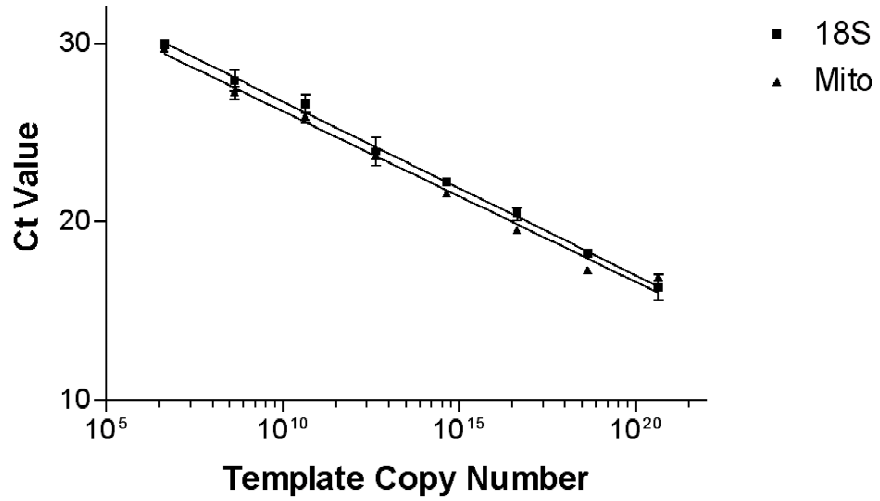
Figure 7 Estimation of levels of the mitochondrial transcription machinery in HeLa cells.

A) Shown is a representative immunoblot used to quantify individual components of the mitochondrial transcription machinery. Mitochondria were purified from untransfected HeLa cells. 10, 25, 50, or 100 μg of mitochondrial protein were loaded on a linear gradient polyacrylamide gel. On the same gel 1, 5, 10, 25 or 50 ng of each purified recombinant transcription component was combined and loaded into a single well. After blotting to a PVDF membrane blots were serially probed for the levels of POLRMT, h-mtTFA, h-mtTFB1, and h-mtTFB2. **B)** Signals obtained from recombinant proteins were used to generate standard curves for estimation of amounts of endogenous transcription components. Shown are representative standard curves from a single set of experiments performed to determine the linear range of detection for each noted protein. Each experiment was repeated three times in order to generate values for Table 1. Given are the equations and R-Squared value for each curve for standards (+) and unknowns (Δ).

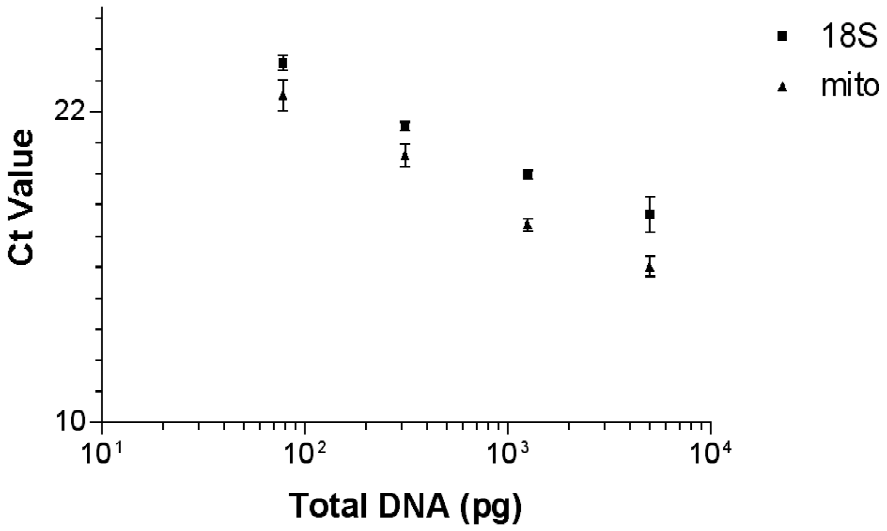
4.2.3 mtDNA copy number measurements

Finally, we also determined the mtDNA copy number in the cells to be 5010 +/- 386 using quantitative real-time PCR (Figure 8). From these data, we calculated the relative levels of the four transcription proteins on a molecule/cell basis and as function of the number of mtDNA molecules/cell (Table 2). Interestingly, there is ~3 fold more h-mtTFB2 than h-mtTFB1 in these cells, which results in a close to 1: 1 relationship between h-mtTFB2 and POLRMT (1.37: 1), but an excess of POLRMT to h-mtTFB1 (2.31: 1). Furthermore, my measurements place POLRMT at a ~6-fold molar excess over the number of mtDNA molecules and between 42 and 58 h-mtTFA molecules/mtDNA molecule in HeLa cells. This h-mtTFA to mtDNA ratio is consistent with that observed by Wiesner and colleagues (Wiesner et al. 2006), but more than 30 fold lower than that proposed by the Kang lab (Takamatsu et al. 2002).

A


30
20
10
Ct Value
10^5
10^{10}
10^{15}
10^{20}
Template Copy Number
18S
Mito

B


22
10
Ct Value
10^1
10^2
10^3
10^4
Total DNA (pg)
18S
mito

C

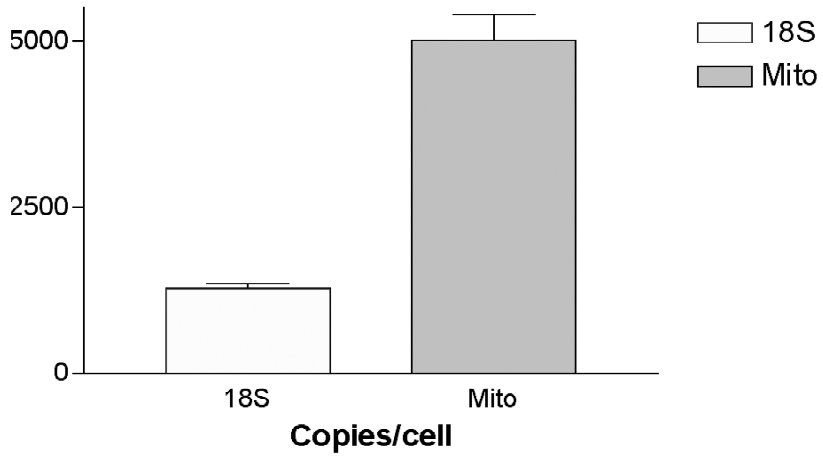

5000
2500
0
18S
Mito
Copies/cell
18S
Mito

Figure 8 Measurement of absolute mtDNA copy number in HeLa cells.

A) Real time PCR Ct value standard curve generated with known amounts of plasmid harboring either the COX3 or 18S target described in Materials and Methods. B) Real time PCR Ct values for both nuclear and mitochondrial targets for a range of dilutions of total cellular DNA used to determine absolute copy number. C). Shown is the quantification of three experiments as in B) averaging copy number from each dilution series. Mitochondrial DNA copy number was found to be 5010 ± 386, around four fold greater than the 18S rDNA locus at 1263 ± 88 copies per HeLa cell under growth conditions described in Materials and Methods.

Table 2 **Estimation of levels and ratios of the mitochondrial transcription components in HeLa cells.**

	molecules/cell	*molecules/mtDNA*	*molecules:h-mtTFA*	*molecules:h-mtTFB1*	*molecules:h-mtTFB2*	*molecules:POLRMT*
h-mtTFA	$2.52x10^5 \pm 4.86x10^4$	50.3 ± 8.8	-	$18.6 \pm 3.59 : 1$	$5.87 \pm 1.13 : 1$	$8.07 \pm 1.56 : 1$
h-mtTFB1	$1.35x10^4 \pm 4.17x10^2$	$2.7 \pm .08$	$1 : 18.6 \pm 3.59$	-	$1 : 3.17 \pm .21$	$1 : 2.31 \pm .15$
h-mtTFB2	$4.29x10^4 \pm 2.80x10^3$	$8.57 \pm .56$	$1 : 5.87 \pm 1.13$	$3.17 \pm .21 : 1$	-	$1.37 \pm .09 : 1$
POLRMT	$3.12x10^4 \pm 1.99x10^3$	$6.23 \pm .39$	$1 : 8.07 \pm 1.56$	$2.31 \pm .15 : 1$	$1 : 1.37 \pm .09$	-

4.3 Distinct biological roles of the mitochondrial transcription factor B family

4.3.1 Over-expression of wild type mtTFB1 and mtTFB2 in HeLa cells

With a determination of relative levels of the mitochondrial transcription machinery in hand, we next set out to examine the consequence of altering the relative amounts of these proteins on mitochondrial gene expression and to determine if their expression levels are coordinately regulated. To do this, we created stable HeLa cell lines that over-express either h-mtTFB1 or h-mtTFB2. From the initial analysis of these lines I immediately made two novel observations.

4.3.2 Processing and predicted cleavage site of mtTFB2

First, the mobility of h-mtTFB2 isolated from cells was significantly faster than that of the corresponding recombinant protein (Figure 9). Since many matrix-localized mitochondrial proteins have an N-terminal localization sequence (MLS) that is often removed upon import, I hypothesized that this was the case for h-mtTFB2. Consistent with this, using SignalP (a MLS and cleavage site prediction program), a single high-probability mitochondrial peptidase cleavage site for h-mtTFB2 was found between amino acids 30 and 31 (Figure 9). No such cleavage site was predicted for h-mtTFB1, consistent with no obvious change in mobility of this protein relative to its recombinant control (Figure 9). Furthermore, mapping amino acids 5-22 onto a helical-wheel diagram reveals a pattern consistent with a putative amphipathic helix, which is another common attribute of an MLS. Finally, using a peptide antibody (TFB2-1) that was made against amino acids 2-20 of h-mtTFB2 (which are predicted to be removed by SignalP), I was

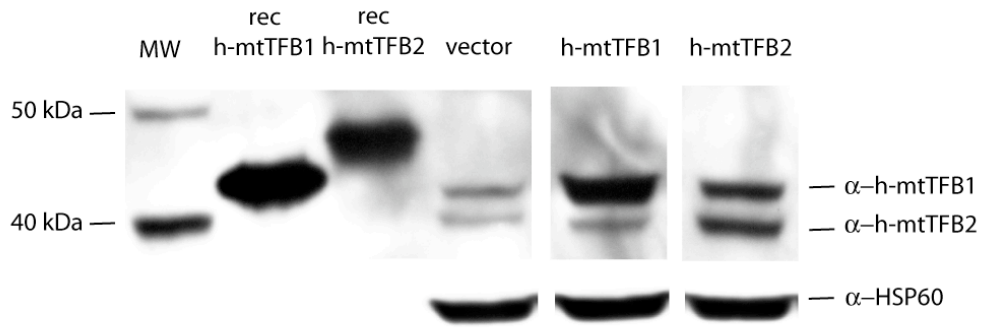
unable to successfully detect endogenous h-mtTFB2 by western blotting (data not shown). However using the peptide antibody (TFB2-2) generated against amino acids 49-70, which is used throughout this study, I readily detect endogenous h-mtTFB2 (Figure 6). Taken together, these data suggest that amino acids 1-30 likely comprise a significant portion of the MLS for h-mtTFB2 and are removed upon import, but the exact site of cleavage is unknown.

4.3.3 Over-expression of mtTFB2 induces mtTFB1 levels

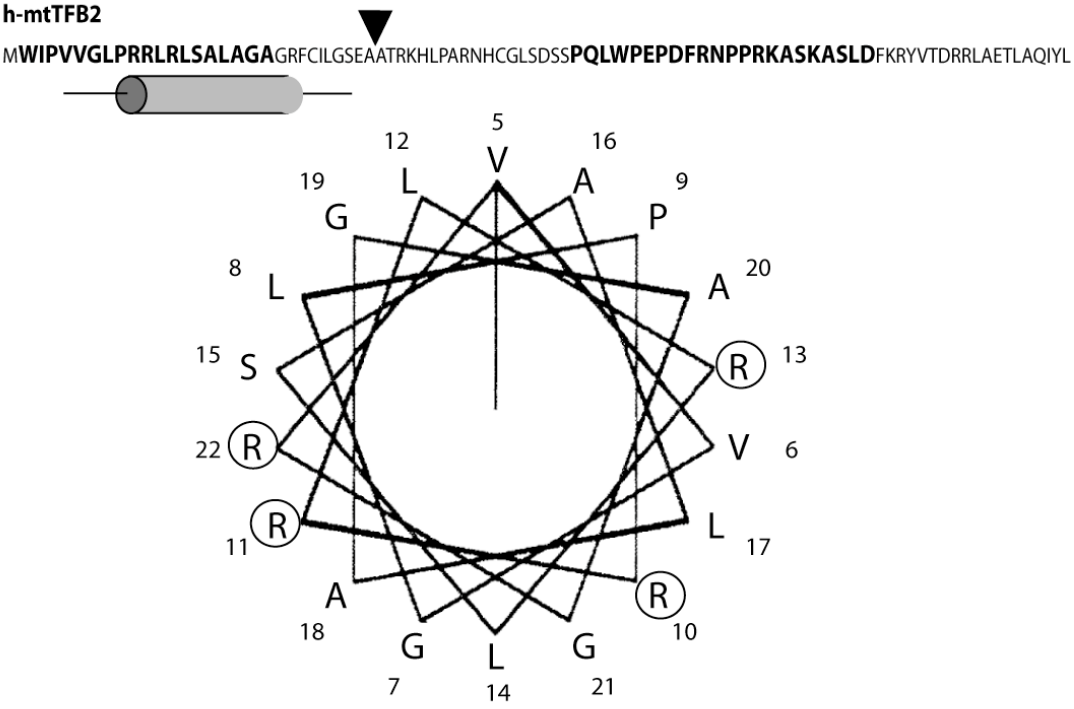
In addition to the increased mobility of h-mtTFB2 described above, in the cell lines that over-express h-mtTFB2 (~3-fold), there was a corresponding increase in the steady-state level of h-mtTFB1 (Figure 9, compare lanes 6, 7, and 8 to lane 4). However, the converse was not true. That is, there was no increase in h-mtTFB2 steady-state levels when h-mtTFB1 was over-expressed ~10-fold (Figure 9, compare lane 5 to lane 4). The steady-state levels of the remaining transcription machinery were also analyzed in the face of h-mtTFB1 or h-mtTFB2 over-expression. No obvious changes in the amounts of POLRMT or h-mtTFA per mitochondrion were observed in either case (Figure 9). To begin to determine the mechanism of increased h-mtTFB1 levels due to h-mtTFB2 over-expression we employed reverse transcriptase real-time PCR to measure amounts of h-mtTFB1 transcript in vector control cells, h-mtTFB1 over-expressors, and h-mtTFB2 over-expressors. We found for h-mtTFB1 over-expressors that h-mtTFB1 transcripts were greatly increased compared to empty vector control cells, but to an extent even greater than expected based on protein levels at approximately 64 fold. In h-mtTFB2 over-expressing samples, we found that h-mtTFB1 transcripts were increased 2.5 fold (Figure 9 and Figure 10). The increase in h-mtTFB1 transcripts coincided closely with

the increased h-mtTFB1 protein levels in the presence of h-mtTFB2 over-expression (Figure 9). Thus there is a form of regulation between h-mtTFB2 and h-mtTFB1 that increases transcription of the h-mtTFB1 gene, but does not result in increased levels of the rest of the transcription machinery per mitochondrial mass.

A


MW
rec h-mtTFB1
rec h-mtTFB2
vector
h-mtTFB1
h-mtTFB2
50 kDa
40 kDa
α–h-mtTFB1
α–h-mtTFB2
α–HSP60

B


h-mtTFB2
MWIPVVGLPRRLRLSALAGAGRFCILGSEAATRKHLPARNHCGLSDSSPQLWPEPDFRNPPRKASKASLDFKRYVTDRRLAETLAQIYL
5 V
16 A
9 P
20 A
13 R
6 V
17 L
10 R
21 G
14 L
7 G
18 A
11 R
22 R
15 S
8 L
19 G
12 L

C

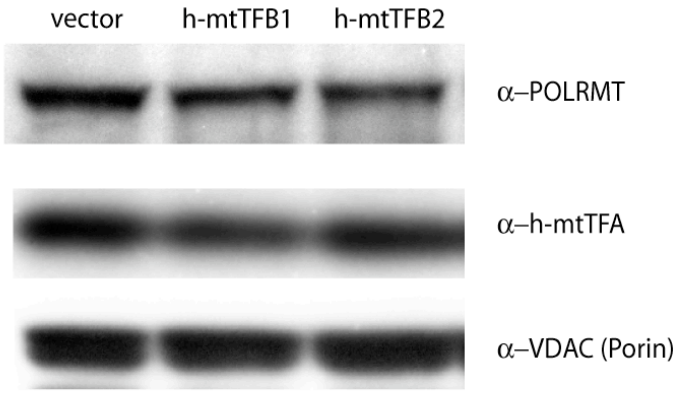

vector
h-mtTFB1
h-mtTFB2
α–POLRMT
α–h-mtTFA
α–VDAC (Porin)

Figure 9 h-mtTFB2 is processed *in vivo* and its over-expression in HeLa cells results in a coordinated increase of h-mtTFB1, but not POLRMT or h-mtTFA.

A) Western blot of mitochondrial extracts (100 μg protein) from HeLa cell lines over-expressing h-mtTFB1 or h-mtTFB2 used in this study in parallel with recombinant h-mtTFB1 and h-mtTFB2 run as controls. The blot was probed as indicated using peptide antibodies that distinguish h-mtTFB1 and h-mtTFB2 (α-h-mtTFB1 and α-h-mtTFB2) and an antibody that recognizes HSP60 (α-HSP60) that was used as a mitochondrial loading control. The lanes are loaded as follows: lane 1, molecular weight markers; lane 2, recombinant h-mtTFB1; lanes 3, recombinant h-mtTFB2; lanes 4-6, mitochondrial extracts from an empty pcDNA 3.1 zeo (+) vector-control, h-mtTFB1 over-expression, and h-mtTFB2 stable over-expression HeLa cell lines, respectively. **B)** Shown are the first 88 amino acids of h-mtTFB2 with the locations of peptides used to generate antibodies against h-mtTFB2 in bold. The triangle represents the location of a putative mitochondrial peptidase cleavage site predicted by SignalP. The cylinder below the text represents a α-helix predicted to be part of the mitochondrial localization sequence (MLS). Also shown is a helical-wheel diagram of amino acids 5-22 that are part of this predicted α-helix, with basic, positively charged residues circled. The arrangement demonstrates an amphipathic helix an often-typical feature of an MLS. **C)** Western blot of mitochondrial lysates from the same cell lines described in A) probed using antibodies that recognize h-mtTFA (α-h-mtTFA), POLRMT (α-POLRMT), and the outer mitochondrial membrane protein VDAC (α-VDAC/porin) as a mitochondrial loading control.

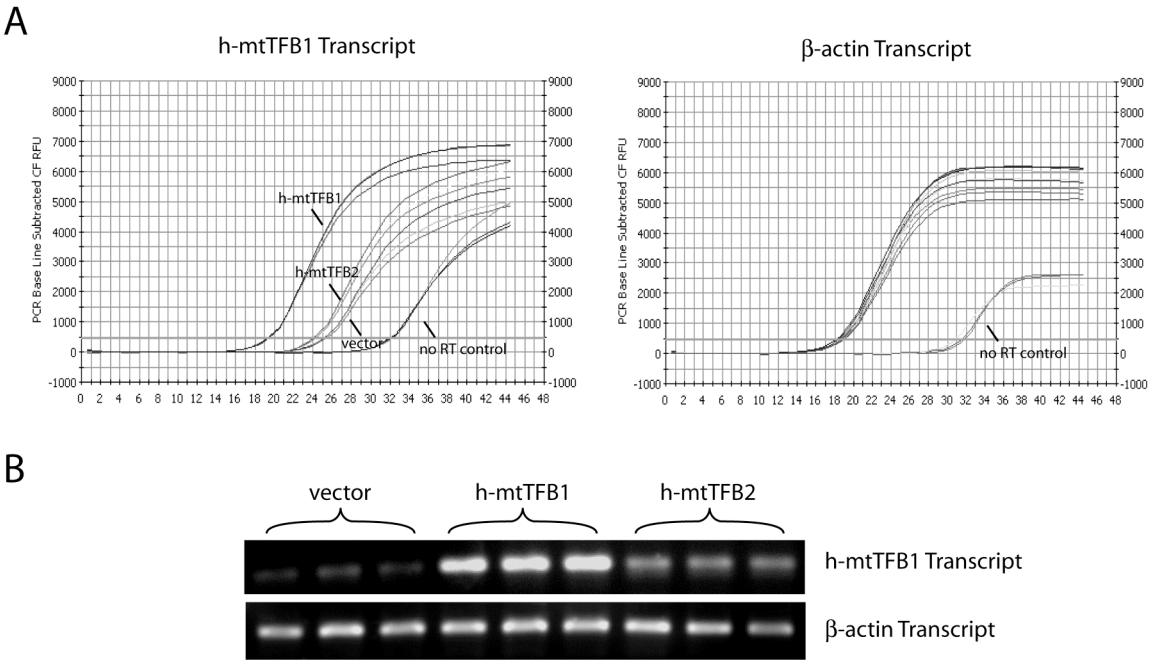


Figure 10 Detection of elevated h-mtTFB1 transcript in h-mtTFB2 over-expressing cells.

A) Shown are raw fluorescence curves obtained by the BioRad iCycler for the β-actin and h-mtTFB1 transcripts. An average Ct value of ~18.5 was found for the β-actin transcript for all samples verifying its use as a loading control. Average Ct values of 25.5, 24.3, and 19.2 were found for the h-mtTFB1 transcript for vector, h-mtTFB2 over-expressors, and h-mtTFB1 over-expressors respectively. B). After amplification 20% of each reaction was separated on a 1% agarose gel and visualized by ethidium bromide staining to qualitatively assay results and show that only a single product was present in each reaction. Shown are three replicates of both the h-mtTFB1 transcript and β-actin transcript reactions from vector, h-mtTFB1 over-expressors, and h-mtTFB2 over-expressors. Indeed h-mtTFB1 is highly over-expressed at an RNA level as to be expected, but the same transcript is also elevated consistently 2.5 fold in cells over-expressing h-mtTFB2 compared to empty vector.

4.3.4 Effects of over-expression of mtTFB1 and mtTFB2 on mitochondrial transcripts, mtDNA levels, and mitochondrial translation.

To address how altered levels of h-mtTFB1 and h-mtTFB2 affect mitochondrial gene expression, I next examined the h-mtTFB1 and h-mtTFB2 over-expression HeLa cell lines for changes in the steady-state levels of mtDNA-encoded transcripts and proteins, as well as for alterations in mtDNA copy number. Northern analysis of the mitochondrial 16S and 12S rRNAs and of ND2 and ND6 transcripts (representing mRNAs transcribed from each strand of mtDNA) revealed a ~2-fold increase in their steady-state levels in the h-mtTFB2 over-expression cell line, but no change in the h-mtTFB1 over-expression line (Figure 11A). Similar results were obtained when immunoblots of the mtDNA-encoded COX1 and COX2 proteins was performed and mtDNA copy number was measured. That is, over-expression of h-mtTFB2, but not h-mtTFB1, led to a significant increase in the steady-state levels of COX1 and COX2 proteins (Figure 11B) and a doubling of the mtDNA copy number (Figure 11C). Altogether, these results are consistent with a role for h-mtTFB2 in transcription and in transcription-primed mtDNA replication.

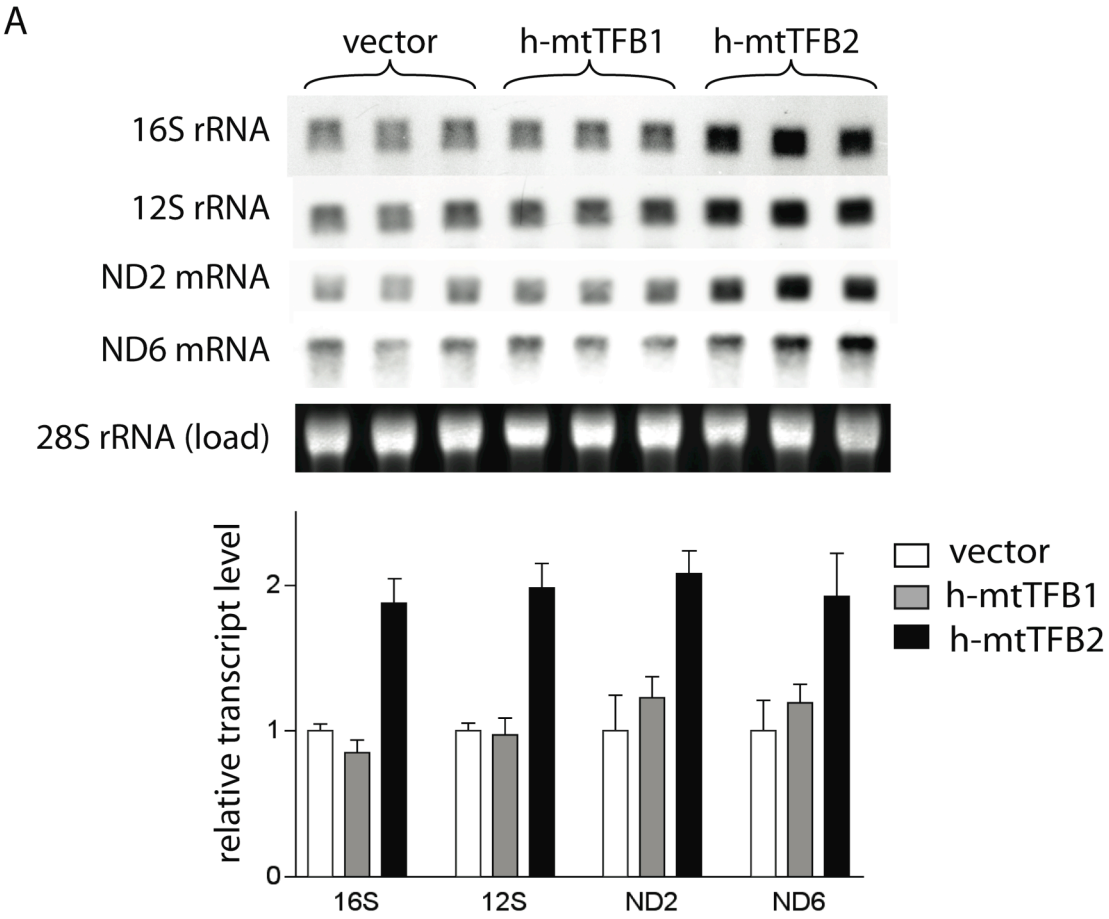

A
vector
h-mtTFB1
h-mtTFB2
16S rRNA
12S rRNA
ND2 mRNA
ND6 mRNA
28S rRNA (load)
relative transcript level
0
1
2
16S
12S
ND2
ND6
vector
h-mtTFB1
h-mtTFB2
B
vector
h-mtTFB1
h-mtTFB2
α–COX2
α–COX1
α–porin
C
relative mtDNA copy number
0
1
2
3
vector
h-mtTFB1
h-mtTFB2

Figure 11 Over-expression of h-mtTFB2 increases the steady-state levels of mtDNA-encoded transcripts and proteins, and doubles mtDNA copy number.

A) Northern analysis of the mtDNA-encoded 12S, 16S, ND2 and ND6 from the same cell lines described in Figure 1A. Total RNA (2 μg) from the indicated cell line was loaded in each lane and the analysis was performed in triplicate on samples from three independent cultures. Ethidium bromide staining of the cytoplasmic 28S rRNA is shown as a loading control. Results of a quantification of the blots are graphed below. The relative transcript level (ratio of the mitochondrial signal to that of the 28S control) is plotted with the ratios obtained in the h-mtTFB1 and h-mtTFB2 cell lines normalized to vector control, which was given a value of 1. The values are the mean +/- standard deviation (brackets). Only over-expression of h-mtTFB2 results in an increased level of mitochondrial transcripts. **B)** Western blot of mitochondrial extracts from the indicated cell lines as in Figure 1B, probed using antibodies that recognize the mtDNA-encoded COX1 and COX2 protein or VDAC as a loading control. This demonstrates that there is an increase per mitochondrial mass of these components unlike h-mTFA and POLRMT. **C)** Plotted is the relative mtDNA copy number (mtDNA relative to the nuclear 18S rDNA) normalized to that of the empty vector control, whose ratio was given a value if 1. The analysis was done in triplicate and values shown are the mean +/- standard deviation (brackets). Only over-expression of h-mtTFB2 results in a change in mtDNA levels.

Both h-mtTFB1 and h-mtTFB2 have retained rRNA methyltransferase activity that is postulated to modulate mitochondrial ribosome biogenesis and/or translation (Cotney and Shadel 2006). This, coupled to the fact that a function for *Drosophila* mtTFB1 in mitochondrial translation has been documented (Matsushima et al. 2005), prompted us to measure mitochondrial translation rates in the h-mtTFB1 and h-mtTFB2 over-expression HeLa cell lines using an *in vivo*-radiolabeling approach. Similar to the analysis of the other mitochondrial gene expression parameters (Figure 11), I again only saw differences in the h-mtTFB2 over-expression cell line, where a significant increase in the overall rate of translation was observed (Figure 12A). Interestingly, however, the labeling of specific products was not uniform (e.g. COX1, COX2, and ATP6), suggesting that increasing the amount of h-mtTFB2 and/or mitochondrial mRNAs is leading to increased translation of some mRNAs to the exclusion of others.

4.3.5 Over-expression of mtTFB1 and mtTFB2 increases sensitivity to kasugamycin

In all of the analyses described thus far no obvious consequences of over-expressing h-mtTFB1 alone were observed. I was somewhat surprised that over-expression of h-mtTFB1 did not influence mitochondrial translation rates given that it is more active as a rRNA methyltransferase than h-mtTFB2 (Cotney and Shadel 2006) and that its *Drosophila* ortholog has been implicated in translation efficiency (Matsushima et al. 2005). It remained a formal possibility that h-mtTFB1 is affecting mitochondrial ribosome function or biogenesis in manner that does no result in an increase in the overall rate of translation. In bacteria, methylation of the small subunit rRNA by the h-mtTFB1 ortholog KsgA results in sensitivity to the aminoglycoside antibiotic kasugamycin and I have shown that h-mtTFB1 and h-mtTFB2 can functionally replace KsgA in *E. coli*

(Cotney and Shadel 2006). Furthermore, the human mitochondrial 12S rRNA and the bacterial 16S rRNA are highly conserved at the site that is methylated (a stem-loop at the 3'end). I therefore determined whether over-expression of h-mtTFB1 alters the sensitivity of human mitochondrial ribosomes to kasugamycin. I found that over-expression of h-mtTFB1 resulted in a significant decrease in viability of HeLa cells grown in the presence of high concentrations of this drug (Figure 12B). Similar results were obtained in the h-mTFB2 over-expression line (Figure 12B). However, given that h-mtTFB1 is also upregulated when h-mtTFB2 is over-expressed (Figure 9), whether this effect is due to solely to h-mtTFB1 or h-mtTFB2 cannot be distinguished.

4.3.6 Effects of over-expression of mtTFB1 and mtTFB2 on mitochondrial biogenesis and membrane potential.

Given that over-expression of h-mtTFB2 (and as a result also h-mtTFB1, Figure 9) results in an increase in multiple mitochondrial gene expression parameters (Figure 11 and Figure 12), I next determined whether this signaled cells to increase overall mitochondrial biogenesis. In parallel, I also analyzed the h-mtTFB1 over-expression line. Surprisingly, I found that both cell lines exhibited a ~50% increase in mitochondrial membrane as measured by Mitotracker Green staining and FACS analysis (Figure 13A). Thus, despite the fact that over-expression of h-mtTFB1 does not increase mitochondrial gene expression in any manner examined thus far (Figure 11 and Figure 12), remarkably it does invoke a mitochondrial biogenesis response. The results of the Mitotracker Green staining were confirmed by immunoblots, where mitochondrial biogenesis was assayed as the ratio of the amount of the mitochondrial outer membrane

protein VDAC to that of tubulin (Figure 13B). I also analyzed the h-mtTFB1 and h-

mtTFB2 over-expression cell lines for mitochondrial membrane potential using Mitotracker Red staining and FACS analysis. Here I observed an ~80% increase in membrane potential in the h-mtTFB2 over-expression line, but no statistically significant change as the result of h-mtTFB1 over-expression alone (Figure 13A). These results suggest the interesting possibility that the simultaneous up-regulation of both h-mtTFB1 and h-mtTFB2 (as is the case in the h-mtTFB2 over-expression lines) is necessary to generate mitochondrial biogenesis and membrane potential with changes in mitochondrial gene expression.

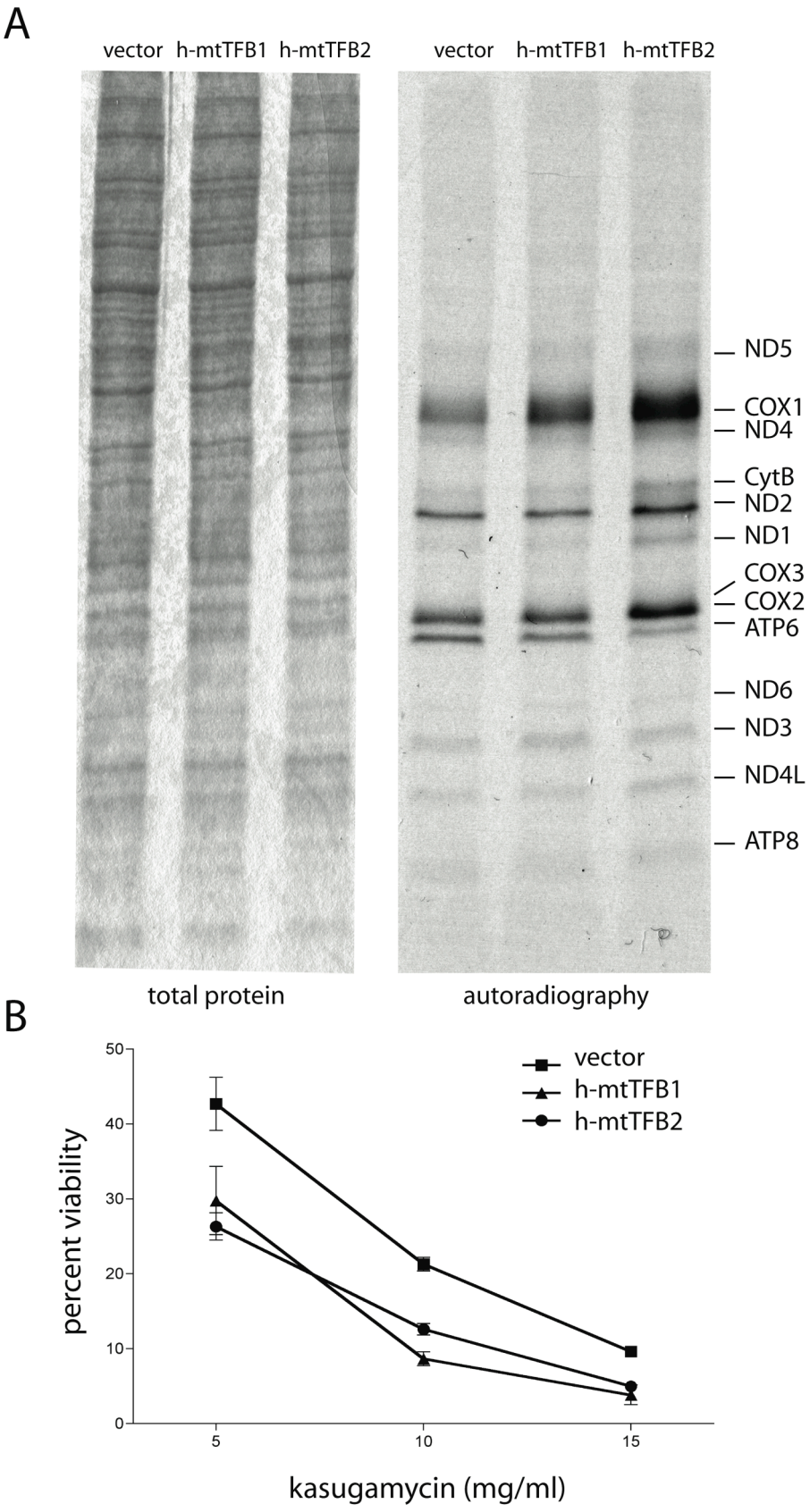
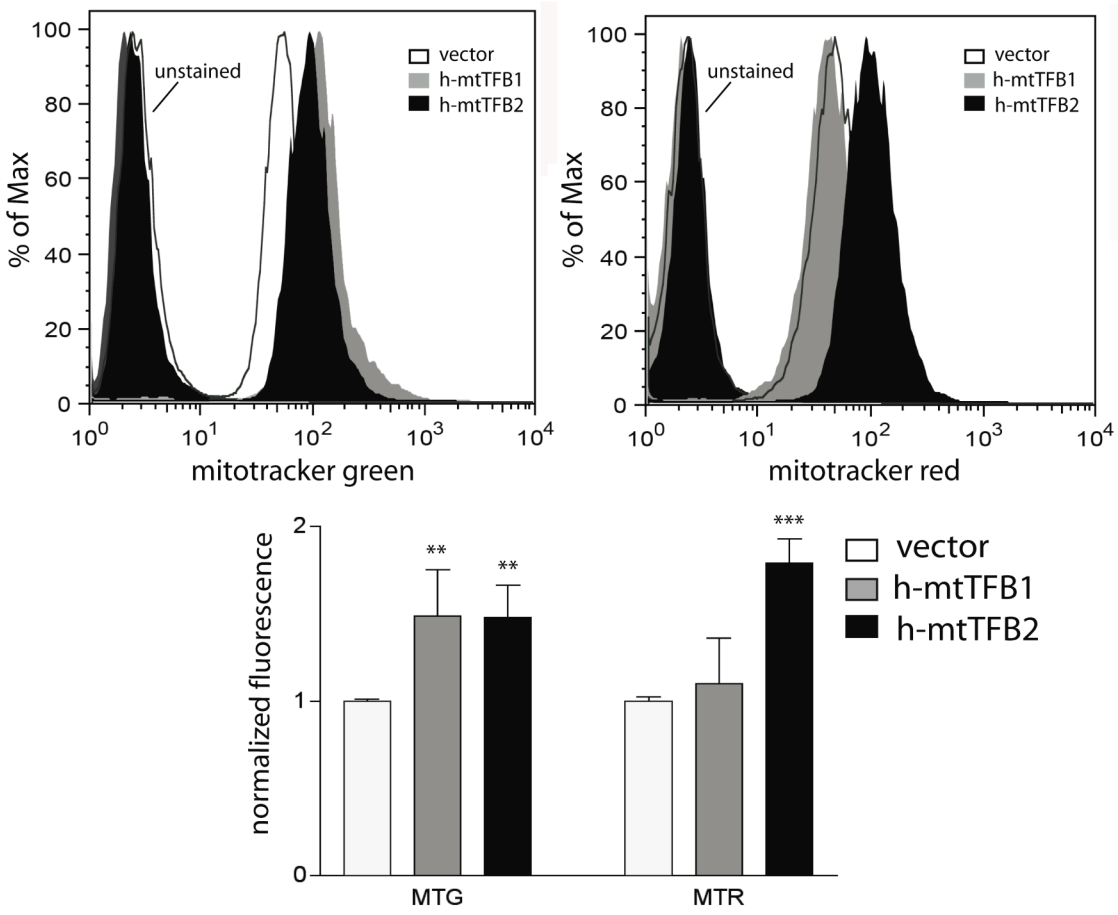

A
vector
h-mtTFB1
h-mtTFB2
vector
h-mtTFB1
h-mtTFB2
ND5
COX1
ND4
CytB
ND2
ND1
COX3
COX2
ATP6
ND6
ND3
ND4L
ATP8
total protein
autoradiography
B
vector
h-mtTFB1
h-mtTFB2
percent viability
50
40
30
20
10
0
5
10
15
kasugamycin (mg/ml)

Figure 12 Analysis of mitochondrial translation rates and kasugamycin sensitivity in h-mtTFB1 and h-mtTFB2 over-expression HeLa cell lines.

A) Mitochondrial translation products were labeled *in vivo* with ^{35}S-methionine in the presence of the cytoplasmic translation inhibitor emitine. Mitochondria were purified after labeling and 100 μg of mitochondrial protein were loaded on a linear gradient polyacrylamide gel. Shown on the left is a coomassie stain of the resulting gel demonstrating equal protein loading and on the right is an autoradiogram of the radiolabeled mitochondrial proteins (specific proteins are indicated on the right). Changes in the profile of labeled products only occur during over-expression of h-mtTFB2. **B)** Plotted are the percent viable cells of the indicated cell lines after growth for 72 hours in the indicated amount of the aminoglycoside kasugamycin. The analysis was done in triplicate and values shown are the mean +/- standard deviation (brackets). Over-expression of either h-mtTFB1 or h-mtTFB2 causes sensitivity to this aminoglycoside antibiotic.

A


unstained
vector
h-mtTFB1
h-mtTFB2
% of Max
mitotracker green
mitotracker red
normalized fluorescence
MTG
MTR
**
**
***

B

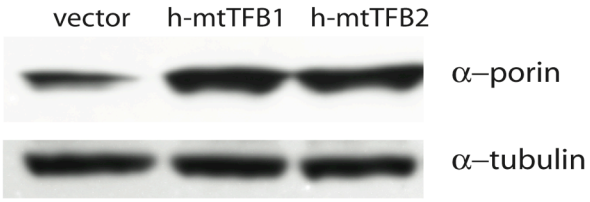

vector
h-mtTFB1
h-mtTFB2
α–porin
α–tubulin

Figure 13 Analysis of mitochondrial biogenesis and membrane potential in h-mtTFB1 and h-mtTFB2 over-expression HeLa cell lines.

A) Shown are representative results from FACS analysis of the indicated cell lines (see key in figure) stained with Mitotracker Green FM, as a measure of mitochondrial mass, or Mitotracker Red, as a measure of mitochondrial membrane potential. Plotted below is a quantification of a triplicate analysis of these parameters, with the mean fluorescence of the vector control cell line given a value of 1. Error bars represent standard deviation of three experiments and asterisks indicate statistically significant differences in mean fluorescence as determined by t-tests, **$p < 0.005$ and *** $p < 0.0005$. Over-expression of either h-mtTFB1 or h-mtTFB2 shows increased levels of Mitotracker green stain, but only h-mtTFB2 over-expression yields and increase in Mitotracker red staining. **B)** Western blot of whole-cell lysates (100 μg protein) from empty-vector control, h-mtTFB1 over-expression, and h-mtTFB2 over-expression cell lines with antibodies that recognize tubulin (α-tubulin) and VDAC (α -VDAC). The ratio of the VDAC signal to the tubulin signal indicates the relative amount of mitochondria per cell is increased in both h-mtTFB1 and h-mtTFB2 over-expressing cells.

4.4 Methylation status of human mitochondrial 12S rRNA is a component of a retrograde signaling pathway for mitochondrial biogenesis.

4.4.1 Over-expression of methylation deficient mutant h-mtTFB1 and h-mtTFB2

To uncover what contribution rRNA methylation activity of the mitochondrial transcription factors h-mtTFB1 and h-mtTFB2 makes on the phenotypes described above, I constructed point mutants of each protein that had demonstrated loss of methyltransferase activity in the previous bacterial methylation assay (Figure 4 and Figure 5). I obtained several HeLa cell lines stably over-expressing methyltransferase deficient forms of h-mtTFB1 or h-mtTFB2 (Figure 14). Similar to wild type h-mtTFB2, mutant h-mtTFB2 displayed increased levels of COX2 and no effects on POLRMT or h-mtTFA. Additionally, loss of methyltransferase activity had no effect on h-mtTFB2's ability to stimulate h-mtTFB1 levels, suggesting that methyltransferase activity is not responsible for this phenomenon. Over-expression of mutant h-mtTFB1 showed several unexpected effects. First, levels of h-mtTFB2 were increased over vector and wild type h-mtTFB1 levels. Second, levels of COX2 protein were elevated to levels similar to that of h-mtTFB2 over-expressors (Figure 14).

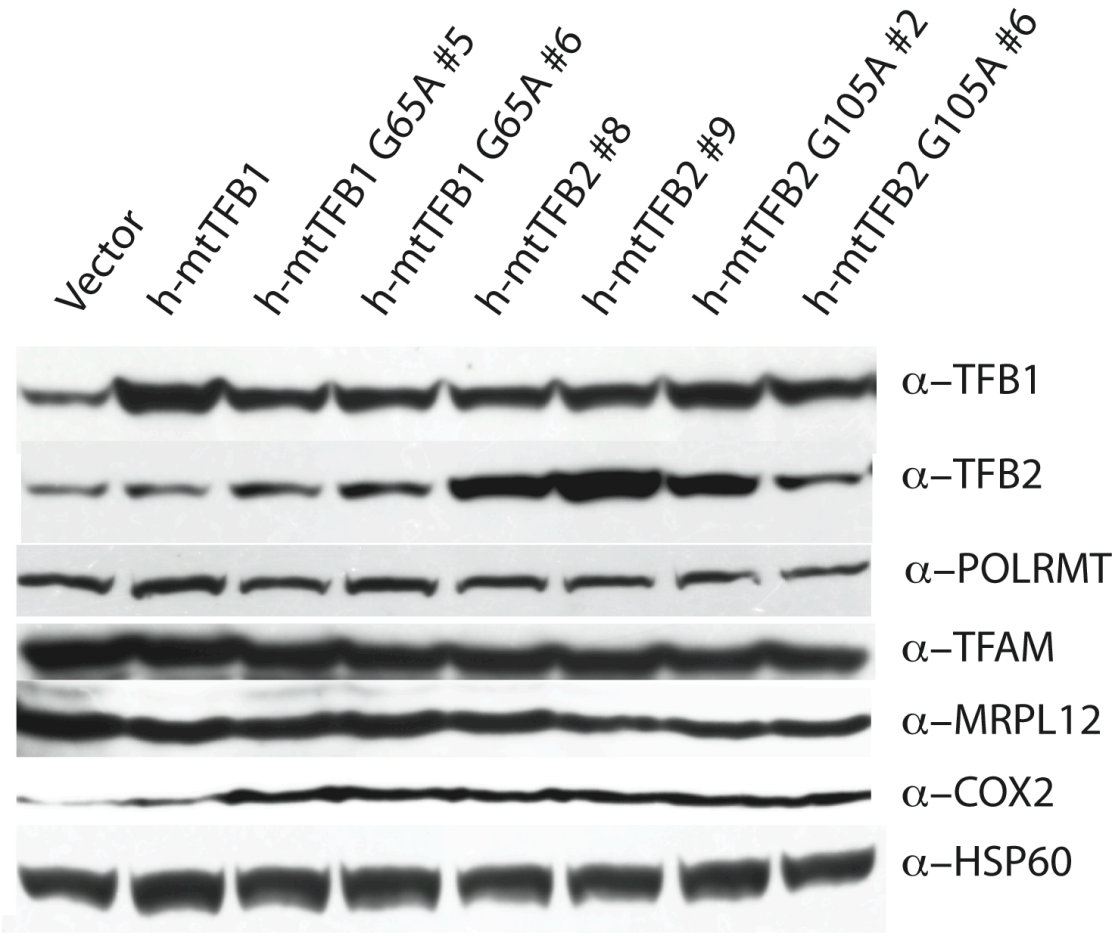
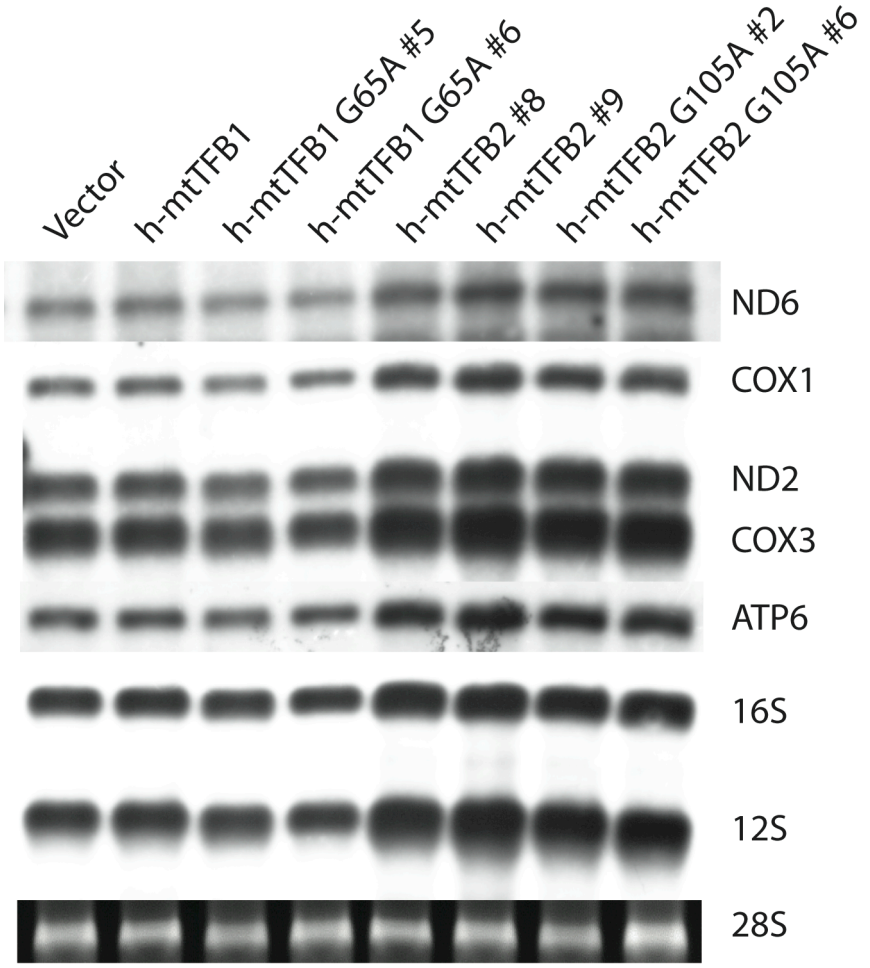


Figure 14 Over-expression of methyltransferase deficient forms of h-mtTFB1 and h-mtTFB2

Shown are immunoblots of mitochondrial extracts harvested from HeLa cell lines over-expressing the protein indicated. Blots were serially probed with the antibodies listed to the right of the figure and normalized for mitochondrial loading by signal obtained for HSP60. Clearly, over-expression of either wild type or mutant h-mtTFB2 elevates levels of h-mtTFB1. Over-expression of mutant h-mtTFB1, unlike wild type over-expression, causes a slight increase in h-mtTFB2 levels and dramatic changes in COX2 levels. No change in other transcription components was observed.

A


Vector
h-mtTFB1
h-mtTFB1 G65A #5
h-mtTFB1 G65A #6
h-mtTFB2 #8
h-mtTFB2 #9
h-mtTFB2 G105A #2
h-mtTFB2 G105A #6
ND6
COX1
ND2
COX3
ATP6
16S
12S
28S

B

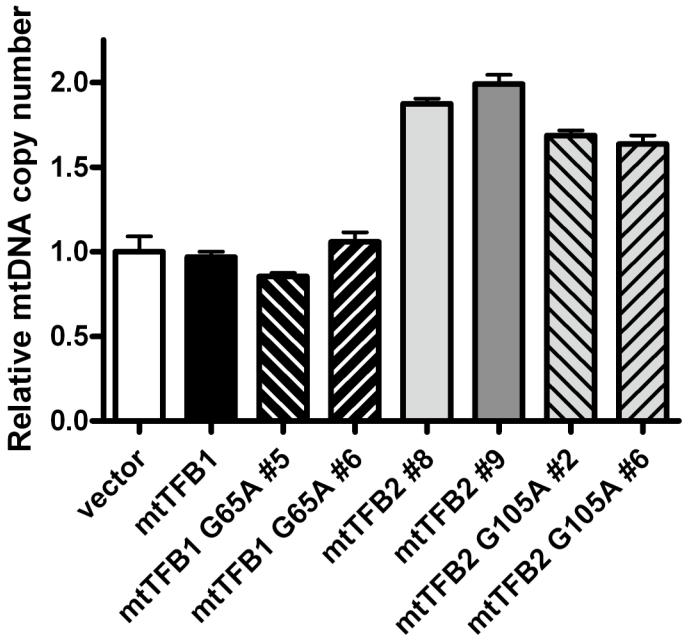

Relative mtDNA copy number
2.0
1.5
1.0
0.5
0.0
vector
mtTFB1
mtTFB1 G65A #5
mtTFB1 G65A #6
mtTFB2 #8
mtTFB2 #9
mtTFB2 G105A #2
mtTFB2 G105A #6

Figure 15 Methyltransferase activity is not required for transcript level increases induced by h-mtTFB2 over-expression

A.) Shown are northern blots of total RNA samples harvested from HeLa cell lines over-expressing the indicated protein. Blots were probed for the sequences indicated to the right of the figure. Over-expression of mutant h-mtTFB2 recapitulated increases in mitochondrial transcripts observed when over-expressing wild type h-mtTFB2. This indicates that this phenotype is not due to methyltransferase activity by h-mtTFB2 and that transcription and methylation may be separable in this protein. Over-expression of mutant h-mtTFB1 had little to no effect on levels of mitochondrial transcripts. B.) Results of quantitative PCR for measuring relative mtDNA levels in HeLa cells over-expressing the indicated proteins are presented. Mutant h-mtTFB2 is fully capable of increasing mtDNA levels, as it did transcript levels. Over-expression of mutant h-mtTFB1 did not cause any changes in mtDNA levels.

4.4.2 Methyltransferase deficient h-mtTFB2 maintains higher transcript levels and mtDNA copy number

To ascertain whether mutant h-mtTFB1 and h-mtTFB2 affected levels of mitochondrial transcripts, I repeated northern blots as described above (Figure 11). Mutant h-mtTFB2 over-expression shows transcript level increases for all messages probed similar to that of wild type h-mtTFB2 (Figure 15 A). These results suggest that the previous observation of ATP6 translation being decreased during h-mtTFB2 over-expression (Figure 12 A) is not due to defects in processing the transcript or levels of this mRNA. Again consistent with a transcription-coupled mode of replication, over-expression of mutant h-mtTFB2 results in higher levels of mtDNA. When methyltransferase deficient h-mtTFB1 was over-expressed, overall transcript levels appear to decrease slightly compared to both vector and wild type h-mtTFB1 over-expressors. There appeared to be little to no effect of mutant h-mtTFB1 over-expression on levels of mtDNA (Figure 15 B).

4.4.3 Methylation activity of mtTFB1 is required for stimulation of mitochondrial biogenesis and is a component of a mitochondrial retrograde signal.

Our previous analysis of mitochondrial 12S rRNA showed that a significant portion of transcripts contains unmethylated adenines in the 3′-terminal stem-loop at positions 1583 and 1584 (Seidel-Rogol et al. 2003). Over-expression of h-mtTFB1 would be predicted to reduce the portion of 12S rRNA that is unmethylated; therefore I performed primer extension analysis of 12S rRNA from mitochondria of cell lines

described above. A significant increase was detected in the level of fully methylated 12S rRNA transcripts in h-mtTFB1 over-expressing cells compared to vector control (Figure 16). h-mtTFB2 over-expression, wild type or mutant, did not affect the level of methylated transcripts. Consistent with a loss of methyltransferase activity, mutant h-mtTFB1 had no effect on the level of methylated 12S rRNA compared to vector control (Figure 16). These analyses demonstrate that indeed the 12S rRNA is an *in vivo* target of h-mtTFB1's methyltransferase activity and the G65A mutation disrupts this ability.

To investigate the possibility that the methylation activity of h-mtTFB1 was directly responsible for mitochondrial biogenesis previously observed, I measured the levels of Mitotracker Green staining in wild type and mutant h-mtTFB1 over-expressing cells via FACS. Mutant h-mtTFB1 over-expression led to a decrease in mitochondrial mass in one clone and no change in the other clone compared to vector control (Figure 17). Over-expression of mutant h-mtTFB2 maintained increases in both mitochondrial mass and mitochondrial membrane potential. In support of the idea that 12S rRNA methylation is a target of h-mtTFB1 and a signal for mitochondrial biogenesis, we detected elevated levels of mitochondrial mass in human cybrid cell lines harboring the A1555G mtDNA mutation (Figure 18 A). Decreases in total cellular ROS as measured by dyhidroethidium (DHE) staining were observed for both A1555G mutant cybrids and h-mtTFB1 over-expressors compared to the proper controls, showing another parallel between these two cell lines (Figure 18 B and C). This mutation has been demonstrated to be closely linked with antibiotic induced deafness in humans and h-mtTFB1 has been identified as a nuclear modifier locus of this phenotype (Bykhovskaya et al. 2004). When the levels of methylated to unmethylated 12S rRNA were measured via primer extension

in these cell lines compared to their appropriate controls, I found that the A1555G cybrids had a dramatic shift in the levels of methylated to unmethylated 12S rRNA without a change in the overall level of 12S rRNA (Figure 19). This change in methylation pattern could be due to altered h-mtTFB1 activity or changes in the level of h-mtTFB1, both of which remain to be determined.

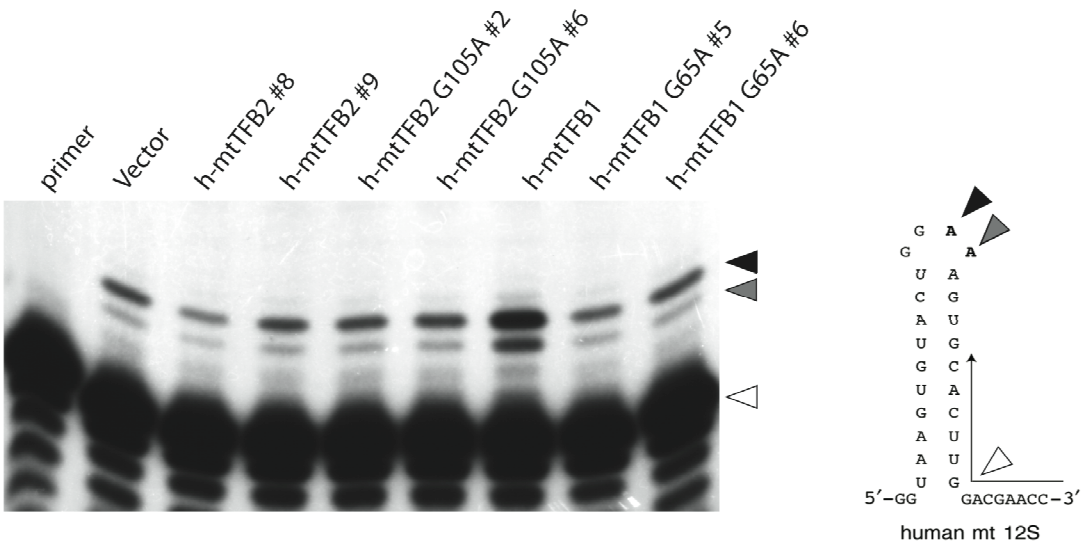


Figure 16 Over-expression of h-mtTFB1 leads to more methylated 12S rRNA

Shown here are the results of primer extension analysis of 12S rRNA obtained from purified mitochondria of HeLa cells over-expressing the indicated protein. Shown to the right are the predicted bands from either the primer alone (open triangle) or blocks of primer extension due to dimethylation of adenines (solid triangles). Only over-expression of h-mtTFB1 shows an increase in the overall levels of both dimethylation events. The G65A mutation of h-mtTFB1 fully prevents this activity.

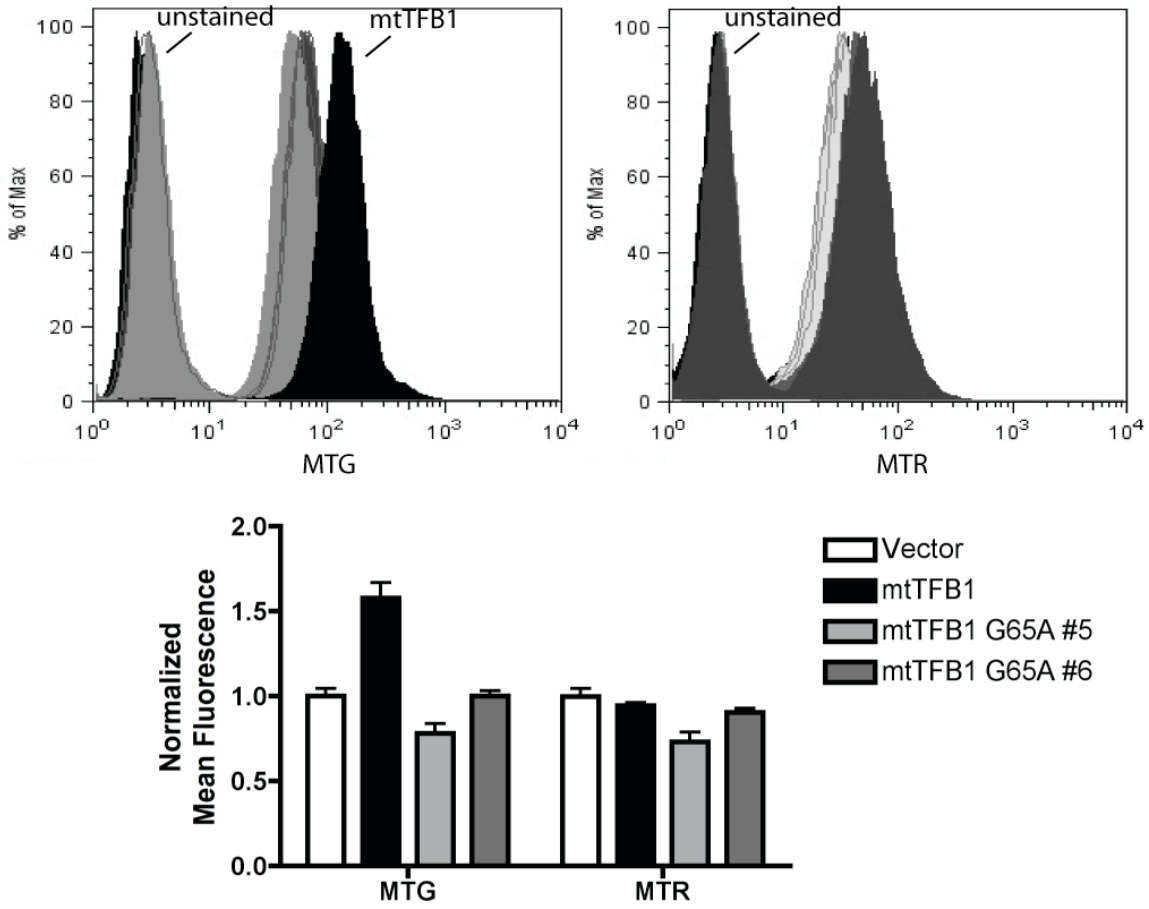


Figure 17 Methyltransferase activity is critical for mitochondrial biogenesis driven by h-mtTFB1 over-expression.

Depicted are representative FACS plots of HeLa cells stained with Mitotracker Green (MTG) and Mitotracker Red (MTR). h-mtTFB1 over-expression results in elevated levels of overall mitochondrial mass as indicated by increased MTG fluorescence, but no change in membrane potential was observed. Mutating the conserved G65A residue eliminates h-mtTFB1's ability to stimulate mitochondrial biogenesis.

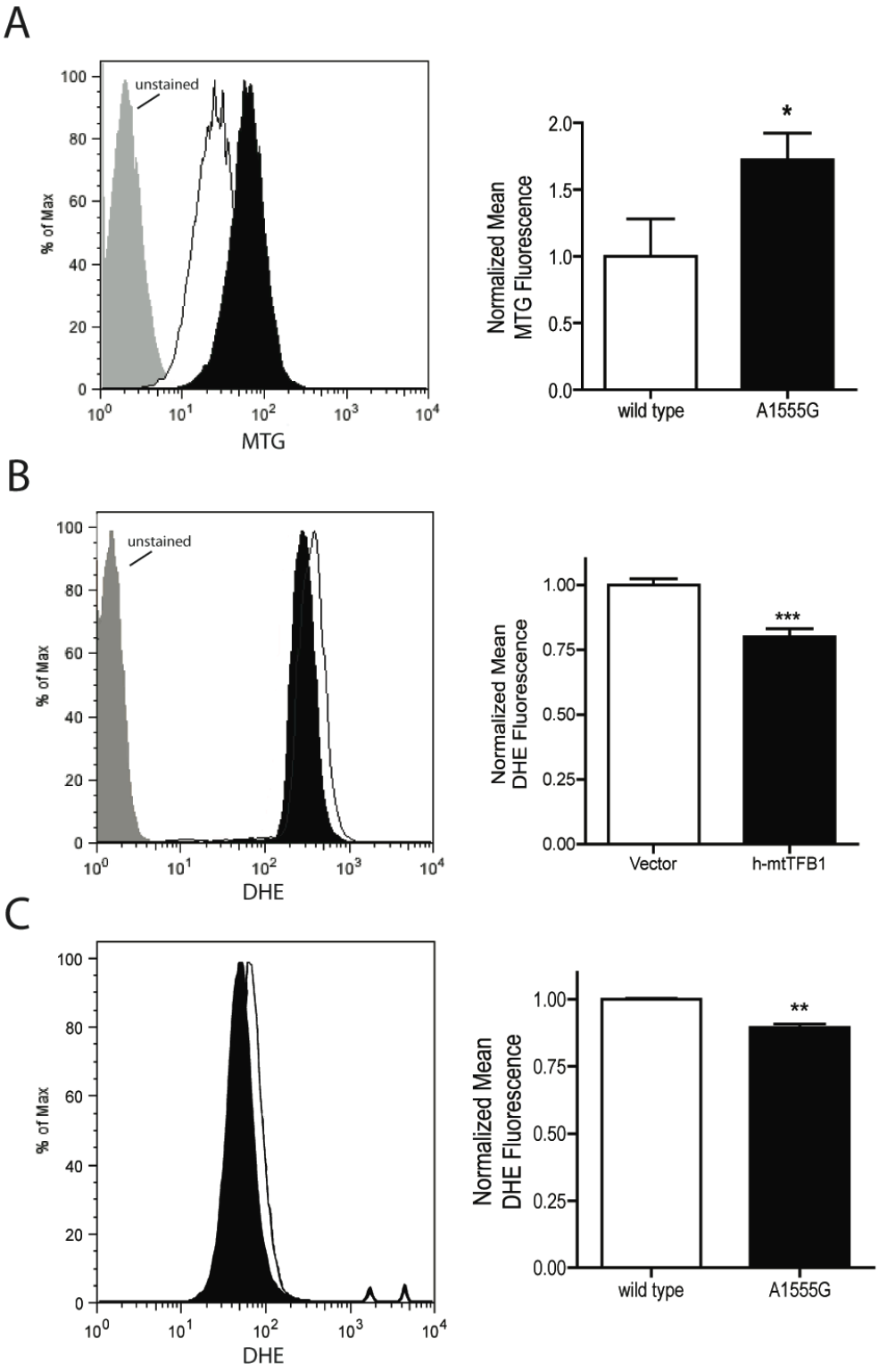

A
unstained
% of Max
MTG
Normalized Mean MTG Fluorescence
wild type
A1555G
*
B
unstained
% of Max
DHE
Normalized Mean DHE Fluorescence
Vector
h-mtTFB1
***
C
% of Max
DHE
Normalized Mean DHE Fluorescence
wild type
A1555G
**

Figure 18 Both A1555G mtDNA mutant cybrids and h-mtTFB1 over-expressing cell lines have elevated mitochondrial mass and decreased total ROS.

A.) Shown are representative FACS plots of hybrid cell lines containing wild type or A1555G mutant mtDNA stained with Mitotracker Green. The cybrid cell line harboring the mtDNA mutation shows significantly increased levels of mitochondrial mass according to Mitotracker Green staining. B.) Shown are representative FACS plots of either h-mtTFB1 an vector control cells or mutant and wild type cybrid cell lines stained with dihydroethidium (DHE). Both h-mtTFB1 over-expressing cell lines as well as cybrids harboring show decreased overall DHE fluorescence indicating lower levels of total cellular ROS. C.) A1555G mutants show decreases cellular ROS as measured by DHE fluorescence. (* = $p < 0.05$, ** = $p < 0.005$, *** = $p < 0.001$)

(These unpublished results provided with permission from Sharen McKay)

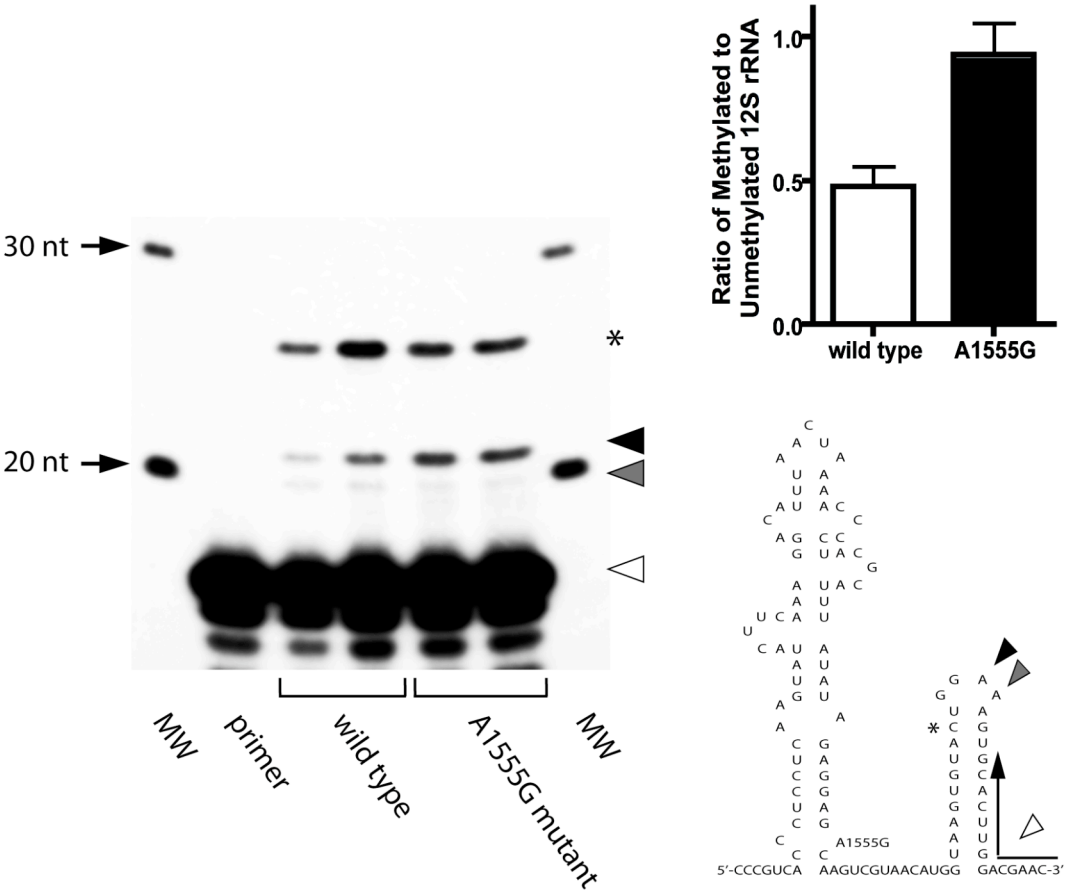


Figure 19 The ratio of methylated to unmethylated 12S rRNA is altered in A1555G mtDNA mutant cybrids.

Shown are the results of primer extension analysis of total RNA from cybrid cell lines containing wild type or mutant mtDNA. The predicted products are primer alone (open triangle), primer extension blocks by dimethylated adenines (filled triangles), and primer extension blocked by addition of ddGTP (*). Cell lines harboring wild type mtDNA show a nearly 2:1 ratio of unmethylated to methylated 12S rRNA. Cybrids containing the A1555G mutation show approximately equal levels of unmethylated to methylated 12S rRNA. The second adenine seems to show the greatest amount of change in methylation. Results are quantified in the bar graph shown with n=5 from two separate experiments.

Chapter V: Discussion

Portions of this chapter reproduced from Cotney and Shadel. J. Mol. Evol. (2006) and Cotney, Wang, and Shadel. Nucleic Acids Res. (2007).

5.1 Implications of divergent evolution and maintenance of dual functions in the h-mtTFB proteins

An important recent advance in our understanding of mitochondrial gene expression came with the discovery that humans have two paralogs of the well-characterized *S. cerevisiae* mitochondrial transcription factor B, h-mtTFB1 and h-mtTFB2. An additional unexpected finding was that this class of transcription factor is related to a family of site-specific rRNA methyltransferases at both the primary and tertiary structural level (Falkenberg et al. 2002; McCulloch et al. 2002; O'Farrell et al. 2004; Schubot et al. 2001). In fact, h-mtTFB1 was shown to possess this enzymatic activity, strongly suggesting it is a dual-function protein involved not only in transcription, but also some aspect of translation via its ability to methylate the mitochondrial 12S rRNA (Seidel-Rogol et al. 2003). However, whether h-mtTFB2 also possesses rRNA methyltransferase activity had not been addressed and the precise contribution of each factor to transcriptional activation and rRNA methylation has yet to be established. As I will discuss, the results of this study demonstrate that, like h-mtTFB1 (Seidel-Rogol et al. 2003), h-mtTFB2 possesses rRNA methyltransferase activity, which has novel implications regarding how these factors coordinately contribute to the regulation of mitochondrial gene expression in humans. In addition, my phylogenetic analysis lead me to conclude, that during the course of evolution, there has been differential selection of transcription and enzymatic activities of mtTFB orthologs within and between species in order to co-evolve the functions of these factors with changes in mtDNA structure and function in various organisms.

It has recently been pointed out that the mitochondrial transcription factors are mostly likely descendents of the dimethyltransferase of the mitochondrial endosymbiont (Shutt and Gray 2006). This idea suggests that these genes were the result of a translocation event from the mitochondrial genome to the nuclear genome and not duplication of the cytoplasmic ribosome dimethyltransferase gene. However the precise sequence of events that led to single mitochondrial transcription factors in some species and apparent paralogs in others has not been addressed. If indeed there are only single forms of these genes in some species (i.e. *C. elegans*) and that the apparent absence is not due to a large degree of divergence in the paralogous form, several modes of evolution could explain the trees presented (Figure 2 and Figure 3).

A duplication event appears to have occurred early in the evolution of eukaryotes due to the presence of both forms of the mitochondrial transcription factors before the appearance of a branch leading to the fungal mitochondrial transcription factors (Figure 2). The simplest explanation would be that a single translocation event occurred with subsequent duplication in the nuclear genome in an ancient eukaryotic ancestor. However, this analysis cannot discount the possibility of other types of duplication models including multiple translocations of the dimethyltransferase gene from the mitochondrial genome or possible lateral gene transfer from one eukaryote or endosymbiont relative to a eukaryote already harboring the translocated gene. While these are both plausible explanations, the conservation of exon-intron boundaries between human mtTFB1 and mtTFB2 (Shutt and Gray 2006) is most simply explained by a single translocation followed by nuclear duplication.

The divergence of the yeast mitochondrial transcription factors occurs from the same branch that yields the TFB2 homologs. This event occurred after the original duplication and would suggest either the loss of the mtTFB1 homolog in fungi or a high degree of sequence divergence such that the mtTFB1 homolog can no longer be identified (Figure 3). The loss of the mtTFB1 homolog is supported by the fact that methylation of the conserved stem-loop described throughout the text does not occur in yeast mitochondria (Klootwijk et al. 1975). This divergence pattern additionally suggests two things about the function of the two paralogs. One, the lack of methylation activity and presence of only one homolog in yeast would point to mtTFB1 being the original dimethyltransferase and mtTFB2 possibly being the duplicated form. Second, the fact that the fungal protein functions strictly as a transcription factor suggests that the mtTFB2 form may have more transcriptional activity than mtTFB1. This is supported by the higher overall transcriptional activity of mtTFB2 compared to mtTFB1 (Falkenberg et al. 2002) and the more dramatic effect upon mitochondrial transcription when mtTFB2 expression is reduced in Schneider cells (Matsushima et al. 2005).

The KsgA class of rRNA methyltransferases dimethylates two adjacent adenine residues in a stem-loop structure located at the extreme 3′-end of the small subunit rRNA (Helser et al. 1972; van Buul and van Knippenberg 1985). This stem-loop is highly conserved from bacteria to vertebrates, and is methylated in this fashion in both cytoplasmic and mitochondrial ribosomes in most eukaryotes. Lack of methylation at this site in bacteria results in resistance to the antibiotic kasugamycin, which normally inhibits bacterial translation (Helser et al. 1972; van Buul and van Knippenberg 1985). We reported previously that h-mtTFB1 functionally complements a *ksgA* mutation in *E.*

coli by restoring methylation at the heterologous stem-loop, demonstrating that h-mtTFB1 has rRNA methyltransferase activity (Seidel-Rogol et al. 2003). Utilizing this same strategy, I show here that h-mtTFB2 has rRNA methyltransferase activity as evidenced by its ability to restore sensitivity to kasugamycin in an *E. coli ksgA* mutant (Figure 4 B) and methylate the 16S bacterial rRNA (Figure 5). However, its activity is qualitatively lower than that of h-mtTFB1, which was analyzed in parallel (Figure 4 A).

The observed rRNA methyltransferase activity was eliminated in both h-mtTFB1 and h-mtTFB2 by a point mutation in a conserved residue required for binding SAM (McCulloch et al. 2002), (Figure 4 and Figure 5 and data not shown), demonstrating that the methyltransferase activity of each enzyme is co-factor dependent. However, unlike what would be predicted based on the drug-sensitivity assays (Figure 4), a lower amount of methylation by h-mtTFB2 compared to h-mtTFB1 in the primer extension assay was not observed. This is most likely explained by primer extension being blocked by partially methylated rRNA (i.e. either single adenine dimethylation event), and sensitivity to kasugamycin requiring dimethylation of both of the target adenines in the 16S stem-loop substrate (Figure 5). This would result in the primer extension assay overestimating the amount of dual adenine dimethylation and hence an inability to measure reduced amounts h-mtTFB2 activity effectively. I also note that the differences observed in these assays cannot be attributed to large differences in expression levels of in *E. coli* nor in differences in total RNA used in these experiments (data not shown). We therefore conclude from these results that both h-mtTFB1 and h-mtTFB2 have rRNA methyltransferase activity, but that h-mtTFB1 is likely a more active enzyme, at least on this heterologous substrate. However, more rigorous biochemical analysis is necessary to

address precisely how the rRNA methyltransferase activities differ between h-mtTFB1 and mtTFB2.

While the evolution of these two genes may suggest the separation of two functions between two genes, the findings in this work suggest that mtTFB2 has retained at least partial methyltransferase activity. Either this activity is relevant *in vivo* or I am observing intermediate steps in the loss of this activity and it does not play a biologically relevant role. The presence of transcription factor activity in mtTFB1 also points to a more complicated process taking place in species, notably higher eukaryotes thus far, that have retained both genes. The dual functions of these two genes and different predicted levels of each function could provide a scenario allowing more advanced eukaryotes the ability to fine-tune the energetic output of their mitochondria.

5.2 Model of impact of mtTFB levels and activities on mitochondrial transcription and translation

The presence of two interchangeable transcription factors with slightly different but overlapping functions could provide cells the mechanism to precisely modulate not only mitochondrial transcription, but also translation of those mitochondrial mRNAs. Depending on the energy requirements of the cell at a particular point in time, the expression of either protein could be controlled differently to tailor the ATP output of mitochondria. In cases where the mitochondrion needs to build up ribosomes for future translation, mtTFB1 could be employed such that rRNA production is immediately followed by proper modification. Once the correct levels of fully functional, stable ribosomes had been reached, expression could shift to an mtTFB2 driven mode. This

would allow for less processing of ribosomal rRNAs and overall increase in transcription to allow for accumulation of mitochondrial mRNAs.

Various combinations of levels of the two proteins could provide a fairly simple but robust mechanism that could modulate both transcription and translation rather precisely (Figure 20). Consistent with this idea, the two mtTFB proteins are differentially regulated during cell differentiation and growth (Gleyzer et al. 2005), yielding different mitochondrial transcriptional outcomes in cultured cells. Also consistent with the idea that these two factors could modulate mitochondrial transcription and translation, the actual levels of transcripts for each gene are present at different ratios in different cell types (Falkenberg et al. 2002). Those tissue types that are associated with the highest ATP requirements (i.e. heart, skeletal muscle, liver) have relatively equal expression levels of both genes. However, other cell types such as kidney and testes have higher levels of h-mtTFB2 or h-mtTFB1 expression, respectively. While these are only mRNA profiles and do not necessarily represent actual proteins levels, it is difficult to reconcile these expression levels with what is known about how mitochondrial transcription works. If each activity were separated to a specific protein (i.e. h-mtTFB1 is the methyltransferase, and h-mtTFB2 is the transcription factor) those tissues that have imbalanced expression of each protein would not be predicted to have fully functional mitochondria. However if both proteins have dual functions, albeit different levels of each, both transcription and translation could be occurring without the requirement of expression of the other protein.

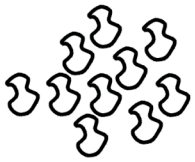
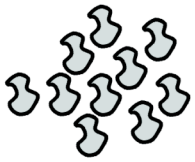
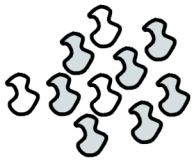
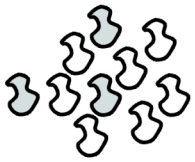
		LSP	HSP	12S methylation	replication
A		++	+	+	++
B		+	++	++	+
C		+	+	+	+
D		++	++	++	++

Figure 20 A putative mtDNA regulatory scheme based on h-mtTFB1 and h-mtTFB2 having partially overlapping, but non-identical transcription and methylation activities.

Shown are four regulatory scenarios (A-D), with hypothetical effects on transcription from the light-strand promoter (LSP) or heavy-strand promoter (HSP), methylation of the mitochondrial small subunit rRNA stem-loop (12S methylation), and LSP-dependent mtDNA replication (replication) indicated. One plus (+) denotes moderate or basal activity, while two pluses (++) indicates higher activity of the indicated process. The two orthologs of mtTFB are depicted as gray and white shapes, which can represent either h-mtTFB1 or h-mtTFB2. A situation where one paralog of h-mtTFB predominates significantly over the other (e.g. due to its higher relative abundance) is shown in scenarios "A" and "B". In "A" predominant activity of one form (white) is shown resulting in higher levels of LSP transcription and replication. While in "B", a situation where the other paralog (gray) predominates, HSP transcription and 12S methylation is preferentially increased. In scenarios "C" and "D," both forms are shown contributing significantly, but differentially (based on their relative abundance), such that basal ("C") or induced ("D") levels of all processes are attained. Depending on the precise relative abundance h-mtTFB1 and h-mtTFB2, a variety of outputs can be imagined that globally regulate the system according to changing cellular needs or in different tissues. A similar regulatory scenario could also pertain to differences in activity (as opposed to abundance) and the activities of h-mtTFB1 and h-mtTFB2 in transcription are almost certainly also a function of their relative levels compared to those of h-mtTFA and mtRNA polymerase.

5.3 Relative levels of mitochondrial transcription machinery and mtDNA

In order to test the model shown in Figure 20, I wanted to change the levels of the mtTFB levels *in vivo* and make predictions about possible mitochondrial outcomes. To better understand the impact of changing mitochondrial transcription factor levels, the first goal of this study was to define the relative amounts of the human mitochondrial transcription machinery *in vivo*. Previously, groups have estimated the amounts of h-mtTFA (Fisher et al. 1991; Shen and Bogenhagen 2001; Takamatsu et al. 2002; Wiesner et al. 2006), but a simultaneous assessment of the entire core transcription system has not been described. We now have antibodies that recognize the four core human mitochondrial transcription components and, importantly, we generated peptide antibodies that are capable of distinguishing the two related h-mtTFB paralogs, h-mtTFB1 and h-mtTFB2 (Figure 6). Using these antibodies, I was able to detect each component in mitochondrial extracts containing known amounts of total mitochondrial protein by western blotting and compare these to the signals obtained from known amounts of cognate recombinant protein analyzed in parallel. Since the total number of cells from which the extracts were derived and the copy number of mtDNA were also quantified, I was able to express the results for each individual component as the number of molecules/cell or the number of molecules/molecule of mtDNA (Table 2).

These results revealed a number of novel and salient points about the relative abundance of the mitochondrial transcription system in HeLa cells (see Table 2). First,

relative to the mtDNA copy number, which were found to be 5010 ± 386 per cell (Figure 8), POLRMT is in ~6-fold excess of the mtDNA on a per-molecule basis. Given that there are three known promoters in each mtDNA molecule, this reveals that POLRMT is at an ~2-fold excess with regard to the number of promoter binding sites. The second main conclusion that we reach is that there is roughly three times more h-mtTFB2 than h-mtTFB1 molecules/cell. This difference is perhaps most relevant when compared to the amount of POLRMT, from which it becomes clear that there is an excess of h-mtTFB2 to POLRMT (1.3: 1), but a limiting amount of h-mtTFB1 relative to POLRMT (0.43: 1). While the relevance of these differences is difficult to predict, it is noteworthy that for both h-mtTFB1 and h-mtTFB2 the levels relative to POLRMT are not too far removed from 1:1, which is consistent with the predicted optimal stoichiometry in the core transcription complexes for transcription *in vitro* (Falkenberg et al. 2002). Finally, I arrive at a value of ~25,000 molecules of h-mtTFA per cell, which places it in ~5-18 fold excess of the other core transcription components (Table 2).

Given that h-mtTFA has been postulated to have an mtDNA-packaging role in addition to its transcription factor function (Alam et al. 2003; Ekstrand et al. 2004; Kanki, Nakayama et al. 2004; Kanki et al. 2004; Takamatsu et al. 2002), its abundance relative to mtDNA is important to discuss. Based on my measurement of mtDNA copy number, the ratio of h-mtTFA: mtDNA molecules observed is 50 ± 8:1. This value is in good agreement with that of 35:1 reported by Wiesner and colleagues (Wiesner et al. 2006), but substantially lower than that proposed by Kang and colleagues, who suggest a ratio of ~1700:1 (Takamatsu et al. 2002). Kang and colleagues cite the ratio of mtTFA to mtDNA in *Xenopus* oocytes at 2000:1 (Shen and Bogenhagen 2001) to support their

findings (Takamatsu et al. 2002). However, ratios of *Xenopus laevis* mtTFA (xl-mtTFA):mtDNA are greatly upregulated during *Xenopus* oocyte maturation ranging from a resting immature oocyte level of ~200:1 (Antoshechkin and Bogenhagen 1995) to the noted ratio of ~2000: 1 (Shen and Bogenhagen 2001), which occurs only in mature oocytes. Given that *Xenopus* mtTFA (xl-mtTFA) binds mtDNA as a tetramer (Antoshechkin et al. 1997), there are effectively ~50 xl-mtTFA complexes per genome in immature oocytes, which we argue is a cell type that is more relevant for comparison to mammalian cell types than a mature oocyte, which has substantially more mitochondria and mtDNA in preparation for fertilization and development. Finally, if h-mtTFA levels were indeed high enough to completely coat the mtDNA genome as suggested by Kang and colleagues (Alam et al. 2003; Kanki et al. 2004; Takamatsu et al. 2002), this would seem incompatible with any significant transcriptional output based on *in vitro* transcription studies (Dairaghi et al. 1995; Falkenberg et al. 2002). Furthermore, it becomes difficult to explain why over-expression of h-mtTFA increases mitochondrial transcription and mtDNA copy number (Ekstrand et al. 2004; Garstka et al. 2003) if, in fact, mtDNA is already fully saturated with h-mtTFA (Alam et al. 2003). It is noteworthy that I observe less than one sixth the amount of h-mtTFA per cell and five times more mtDNA than Takamatsu et al measured in HeLa cells (Takamatsu et al. 2002). These differences likely begin to account for the apparent overestimation of the mtTFA: mtDNA ratio by Kang and colleagues compared to that reported herein and that of Wiesner and colleagues.

5.4 Outcomes of altering mitochondrial transcription machinery levels

With the new knowledge of the relative levels of the human mitochondrial transcription system in HeLa cells, I went on to analyze the consequences of over-expressing each of the h-mtTFB paralogs on mitochondrial gene expression and biogenesis *in vivo*. Stable HeLa cell lines were established that over-express h-mtTFB1 or h-mtTFB2 by ~10-fold or up to ~3-fold, respectively (Figure 9A). Characterization of these lines by western blotting immediately revealed a salient difference between these two factors; that h-mtTFB2 is processed *in vivo* (Figure 9A). Several additional lines of evidence strongly indicate that the processing of h-mtTFB2 is via cleavage by mitochondrial proteases upon import of the protein into mitochondria, including a strong predicted mitochondrial protease cleavage site between amino acids 30 and 31, the presence of a predicted amphipathic alpha helix (stereotypical of mitochondrial localization sequences) spanning amino acids 5-22, and an inability to detect h-mtTFB2 with a peptide antibody that was directed against amino acids 2-20, which would be removed by the predicted cleavage event (Figure 9B). No obvious mobility differences were observed between recombinant h-mtTFB1 and that isolated from cells. Furthermore, SignalP does not predict a mitochondrial cleavage site for h-mtTFB1 (data not shown). Thus, h-mtTFB2 and h-mtTFB1 are handled quite differently upon import, which is consistent with their distinct evolutionary history subsequent to the putative gene duplication event that created the two protein families early in eukaryotic lineage (Figure 2; Cotney and Shadel 2006; Shutt and Gray 2006). It is also tempting to speculate that the use of different modes of import for these two transcription factors

could provide a mechanism to control their relative levels in the organelle in response to different conditions.

A second important observation that came from the initial analysis of the h-mtTFB1 and h-mtTFB2 over-expression lines is that there is an increase in h-mtTFB1 when h-mtTFB2 is over-expressed (Figure 9A). However, the converse was not true. That is, in the h-mtTFB1 over-expression line there is no change in the steady-state amounts of h-mtTFB2 per mitochondrion (Figure 1B). Furthermore, there are no obvious changes in the levels of POLRMT or h-mtTFA per mitochondrion in either of the h-mtTFB factor over-expression lines (Figure 9C). We conclude that there is some form of one-way communication between h-mtTFB2 and h-mtTFB1. The fact that this upregulation occurs, at least in part, via increased levels of the h-mtTFB1 mRNA (Figure 10), suggests that this involved a retrograde signal transduction mechanism from the mitochondria to the nucleus (Butow and Avadhani 2004). It is tempting to speculate that this response is initiated via signals generated by alterations in mitochondrial transcriptional output or the amount of 12S rRNA methylation.

I next examined multiple mitochondrial parameters in the h-mtTFB1 and h-mtTFB2 over-expression cell lines. In the h-mtTFB2 over-expression cell lines (where it is important to keep in mind that there is also a compensatory increase in mtTFB1 as discussed above; Figure 9A), there is a ~2-fold increase in overall mitochondrial transcript levels as evidenced by Northern analysis of the two rRNAs (12S and 16S) and two mRNAs (ND2 and ND6) encoded by mtDNA and representing transcripts derived from both strands (Figure 11A). This was accompanied by a corresponding doubling of the mtDNA copy number (Figure 11C). Given the documented roles for the h-mtTFB

factors in directing transcription initiation efficiency (Falkenberg et al. 2002), these changes most likely represent an increase in the rate of mitochondrial transcription that is also driving a increase in transcription-primed mtDNA replication (Bonawitz et al. 2006). However, it remains a formal possibility that the increase in steady-state levels of mitochondrial transcripts is due to enhanced RNA stability and/or similar rates of transcription from the increased number of mtDNA templates. Also evident in the h-mtTFB2 over-expression lines was an increased rate of mitochondrial translation of most, but not all mitochondrial gene products (Figure 12A) that, at least in the case of COX1 and COX2, results in significantly increased steady-state amounts of protein (Figure 11B). However, the change in levels of mitochondrial translation were not uniform for all of the subunits. For example, there is apparently a reduced level of translation of ATP6 and no apparent change in the level of synthesis of ND3 or ND4L (Figure 12A). Thus, artificially raising the levels of h-mtTFB2 (and in response, also h-mtTFB1) does increase mitochondrial gene expression and mtDNA replication, but may come at the cost of imbalanced relative levels of mitochondrial protein synthesis. We suspect that most of the described effects on mitochondrial gene expression just listed are driven primarily by the increased levels of h-mtTFB2 since no major changes in mtDNA copy number, transcript levels, translation rates or steady-state levels of COX1 and COX 2 were observed in the cell line in which h-mtTFB1 alone was over-expressed. However, it is also just as likely that upregulation of both factors is necessary to coordinately mount all of the responses observed.

The lack of a "mitochondrial gene expression" response in the h-mtTFB1 over-expression lines is somewhat surprising given the documented ability of this protein to

activate transcription *in vitro* (Falkenberg et al. 2002; McCulloch et al. 2002). One potential reason for this may be explained via the relative levels of the transcription system I established in these cells. For example, the ratio of h-mtTFA to h-mtTFB1 required for optimal transcription *in vitro* is ~1: 2 (Falkenberg et al. 2002). Thus, the corresponding ratio determined *in vivo* of 18: 1 would likely not allow much of a contribution of h-mtTFB1 to the total transcriptional output. Furthermore, to reach the optimal ratio of 1: 2 would require nearly 40-fold over-expression of h-mtTFB1. Levels of h-mtTFB1 were only able to increase by ~10-fold, which likely explains why no alterations in mitochondrial transcripts were observed in the h-mtTFB1 over-expression cell lines (Figure 11A). In contrast, the ratio of h-mtTFA to h-mtTFB2 *in vivo* of 5: 1 (Table 1) is well within the range of 2: 1 - 40: 1 that is optimal for mitochondrial transcription *in vitro* (Falkenberg et al. 2002). Increasing this by ~3-fold in the h-mtTFB2 over-expression cell line would keep this ratio within this optimal range, entirely consistent with my results (Figure 3A). Whether the relative amounts of the transcription components vary in different cell or tissue types such that h-mtTFB1 also contributes to mitochondrial transcription remains an open question.

While over-expression of h-mtTFB1 does not show any major effects on the mitochondrial transcription and translation parameters measured, it did result in increased sensitivity of HeLa cells to the aminoglycoside antibiotic kasugamycin (Figure 12B). Sensitivity to this antibiotic in *E. coli* is modulated by dimethylation of two adenine residues in the 3' terminal stem-loop of the small subunit ribosomal RNA by KsgA (Helser et al. 1972; van Buul et al. 1983), the bacterial homolog of h-mtTFB1 and h-mtTFB2. We have shown previously that h-mtTFB1 and h-mtTFB2 are able to methylate

the homologous stem-loop in bacteria and restore sensitivity to kasugamycin (Figure 5; Cotney and Shadel 2006; Seidel-Rogol et al. 2003). Thus, we interpret the ability of increased levels of h-mtTFB1 to sensitize HeLa cells to kasugamycin to indicate that h-mtTFB1 is increasing the number of methylated ribosomes that are targets for inhibition by the drug. This provides the first *in vivo* confirmation that h-mtTFB1 is likely the 12S rRNA methyltransferase in human mitochondria as we predicted from our earlier studies (Figure 5; Cotney and Shadel 2006; Seidel-Rogol et al. 2003). Interestingly, I observed a similar sensitivity to kasugamycin in the h-mtTFB2 over-expression lines (Figure 12B). While this is most likely due to the fact that h-mtTFB1 is also upregulated in this cell line, the possibility that h-mtTFB2 is also contributing to the methylation of the 12S rRNA cannot be excluded since it too possesses rRNA methyltransferase activity, albeit at a much lower level (Figure 4; Cotney and Shadel 2006). In this regard it is interesting to note that others have shown that different methylation combinations on the *E. coli* 16S rRNA (adenines in the stem-loop being monomethylated, dimethylated, or mixed) result in different levels of aminoglycoside resistance (O'Farrell et al. 2006). Thus, the idea that h-mtTFB1 and h-mtTFB2 might orchestrate different states of 12S rRNA methylation with different functional outcomes is an intriguing possibility.

Finally, a new role for h-mtTFB1 and h-mtTFB2 in mitochondrial biogenesis has been elucidated. Over-expression of h-mtTFB1 alone results in a ~50% increase in mitochondrial mass as judged by Mitotracker Green staining and confirmed by western analysis of a mitochondrial marker protein, porin (Figure 13). Thus, even in the absence of a direct effect on mitochondrial transcription or translation, increased h-mtTFB1 somehow signals a change in mitochondrial biogenesis. It will be of interest to determine

the nature of this signal and whether it involves, for example, a sensing of the amount of ongoing mitochondrial ribosome assembly through the rRNA methylation status. A similar increase in mitochondrial biogenesis is observed in the h-mtTFB2 over-expression cell line (Figure 13), in which both h-mtTFB1 and h-mtTFB2 are elevated. However, here, unlike h-mtTFB1 over-expression alone, it was accompanied by a significant (~80%) increase in mitochondrial membrane potential (Figure 13A). Taken together, these results suggest that, while h-mtTFB1 can alone induce a mitochondrial biogenesis response, an increase in both h-mtTFB1 and h-mtTFB2 is needed to coordinate an increase in mitochondrial gene expression with the increase in mitochondrial mass. I propose that the lack of an increase in membrane potential in the h-mtTFB1 over-expression cell line (Figure 13) is a result of a defect in this response that leads to more mitochondria with fewer OXPHOS complexes per mass of organelle.

5.5 12S rRNA methylation is a component of a signaling pathway for mitochondrial biogenesis

Altering the levels of the two h-mtTFB factors had both predicted and unexpected effects on mitochondrial biology. Since these two proteins have been documented to have two functions I wanted to know what role one of these functions, methyltransferase activity, has on the phenotypes that I have previously observed. We have demonstrated that a conserved residue in h-mtTFB1 and h-mtTFB2, G65 and G105 respectively, is essential for SAM dependent methylation of rRNA (Figure 5; Cotney and Shadel 2006; Seidel-Rogol et al. 2003). To determine if the methyltransferase activity of either protein is contributing to my previous observations I over-expressed methyltransferase deficient forms of h-mtTFB1 (G65A) or h-mtTFB2 (G105A) in HeLa cells (Figure 14). I observed

that mutant h-mtTFB2 exhibited the same phenotypes as wild type h-mtTFB2. Mutant h-mtTFB2 was totally capable of increasing levels of mitochondrial transcripts and mtDNA copy number (Figure 15 A and B). These results suggest that the methyltransferase activity of h-mtTFB2 is not required for transcription. These two activities are known to be separable for h-mtTFB1 (McCulloch and Shadel 2003) and the same scenario seems to be present for h-mtTFB2. Mutant h-mtTFB2 expression was also sufficient to elevate levels of h-mtTFB1 to a similar degree seen by over-expression of wild type h-mtTFB2 (Figure 14). Therefore the pathway that generates h-mtTFB1 expression in response to h-mtTFB2 levels is not dependent on the methyltransferase activity of h-mtTFB2.

Methyltransferase deficient h-mtTFB1 had no effects on the expression of h-mtTFA or POLRMT, but surprisingly levels of h-mtTFB2 appear to be elevated and COX2 protein was dramatically increased (Figure 14). COX2 levels appear to be elevated whenever h-mtTFB2 levels are increased indicating an influence of this transcription factor on the stability or translation of COX2 as well as ATP6. Levels of mitochondrial transcripts also appear to be slightly depressed in response to mutant h-mtTFB1 expression, and, at least for one mutant, mtDNA copy number was also slightly decreased (Figure 15). These results could indicate that the G65A mutation of h-mtTFB1 is a dominant negative mutation and could block proper methylation of 12S rRNA. As you may recall Dim1p is involved in the processing and modification of rRNA and removal of this activity causes immature rRNA species to accumulate. In a similar scenario the h-mtTFB1 G65A mutant may be unable to quickly release the 12S rRNA stem loop in the absence of SAM binding and methylation. If so, this could be evidence that transcription and methylation of 12S rRNA are coupled and one cannot proceed efficiently without the

other. When such a dominant negative protein is present, the cell may compensate by increasing levels of h-mtTFB2 in hopes of elevating the levels of transcripts that can be produced to return to a relatively normal state of transcript levels and ribosome assembly.

It should be pointed out here that in these experiments I measured the levels of COX1, COX3, and ATP6 in addition to the previous set of probes. When I labeled mitochondrial translation products previously from cells over-expressing h-mtTFB2 the ATP6 protein was decreased specifically (Figure 12; Cotney et al. 2007). Since processing of the ATP6 and COX3 transcripts occurs through a mechanism not explained by the tRNA cleavage model I wondered whether this depletion of ATP6 was due to decreased levels of this transcript or processing defects. In both the wild type and mutant h-mtTFB2 over-expressing lines the processing and the levels of ATP6, and its adjoining gene COX3, are both increased (Figure 15). Therefore the ATP6 protein level decrease I observed must be due to either stability or decreased translation of the protein.

When I analyzed these cell lines for changes in mitochondrial mass and membrane potential I found that mutant and wild type h-mtTFB2 show similar changes in Mitotracker Green and Red staining (data not shown), but h-mtTFB1 G65A mutants were completely unable to increase levels of Mitotracker Green staining compared to their wild type counterpart (Figure 17). This indicates that the mitochondrial biogenesis response I observed previously during over-expression of h-mtTFB1 is entirely dependent on methyltransferase activity of this protein. To determine whether 12S rRNA methylation was actually altered in any of the cell lines generated I performed primer extension analysis of RNA from isolated mitochondria. Wild type and mutant h-mtTFB2 over-expression had no effect on the levels of methylated 12S rRNA (Figure 16 Lanes 3-6).

As expected, over-expression of h-mtTFB1 showed a major increase in the number of methylated 12S rRNA (Figure 16 lane 7) while the G65A mutant had no effect on levels of methylated 12S rRNA (Figure 16 Lanes 8 and 9). These results show for the first time that an *in vivo* target of h-mtTFB1 is indeed the 3'-terminal stem-loop of 12S rRNA and that levels of the protein have an effect on the methylation status of 12S rRNA. It is interesting to note that the second adenine in the loop appears to be the more abundantly dimethylated base (Figure 16 compare upper methylation stop to lower methylation stop) and that the ratio of the two seems to be maintained even when overall methylation is increased by h-mtTFB1. This could indicate a preference for methylation of this base, which has been observed for other members of this rRNA methyltransferase family (O'Farrell et al. 2006). Regardless of the preference of the base for methylation, it is apparent that the mutant h-mtTFB1 protein is unable to cause more methylation to occur. It is therefore clear that the methylation of the 12S rRNA is required for mitochondrial biogenesis by h-mtTFB1 expression in HeLa cells.

To determine if such a phenomenon exists in other cell lines naturally or during a disease state, we analyzed similar mitochondrial parameters in a human cell line containing the A1555G mtDNA mutation. This cell line was generated by transferring mitochondria from a lymphoblastoid cell line derived from a patient containing mutant mtDNA to a cell line lacking mitochondrial DNA ($\rho^0$260) (Guan et al. 2001). This mutation occurs near the conserved 3′-terminal stem-loop of human 12S rRNA and has been linked to aminoglycoside-induced deafness (Figure 19). Recently h-mtTFB1 has been implicated as a nuclear modifier locus of this non-syndromic deafness (Bykhovskaya et al. 2004). The exact mechanism by which h-mtTFB1 could alter the

occurrence of deafness in individuals with this mtDNA mutation is not completely clear, but is postulated to be due to decreased h-mtTFB1 levels or activity (Bykhovskaya et al. 2004). If this idea is correct, cells harboring the A1555G mutation could mimic h-mtTFB1 over-expression by either increasing h-mtTFB1 activity or allowing the 3′-terminal stem-loop of 12S rRNA to be more readily available for methylation. Comparisons of cybrid cell lines containing either wild type or A1555G mutant mtDNA revealed several phenotypes similar to h-mtTFB1 over-expression in HeLa cells. First, the A1555G cybrid cells exhibit elevated levels of mitochondrial mass compared to wild type control cybrid cells as measured by Mitotracker Green staining (Figure 18A). Second, both h-mtTFB1 over-expressing cell lines and A1555G mutant cybrids exhibit lower levels of total ROS measured by DHE staining compared to their respective controls (Figure 18 B and C). Lastly, in agreement with that there is more h-mtTFB1 activity in A1555G mutant cell lines, primer extension analysis revealed a significant shift in the ratio of methylated to unmethylated 12S rRNAs (Figure 19 compare uppermost band to two lower methylation stop bands). In this experiment, as well as the measurement of methylation in h-mtTFB1 over-expressing HeLa cells, the second dimethylation event appears to be the more predominant modification. Whether these results are directly attributable to increased h-mtTFB1 activity and not protein levels has not been determined, but either would be in response to the A1555G mutation and might provide a link to methylation and mitochondrial biogenesis.

In this study I have explored the possibility that h-mtTFB1 and h-mtTFB2 have direct roles in mitochondrial biogenesis. Increased mitochondrial mass seems to follow levels of these factors and, at least for the case of h-mtTFB1 depends entirely upon the

methyltransferase activity of the transcription factor. By what mechanism could h-mtTFB levels or activity control mitochondrial biogenesis and are the h-mtTFB proteins the prime downstream mediators of known mitochondrial biogenesis pathways? Take for instance, a normal cellular response generated by an external stimulus such as chronic exercise. PGC-1α and the nuclear respiratory factors, NRF-1 and NRF-2, stimulate the expression of the mitochondrial transcription factors mtTFA, mtTFB1, and mtTFB2. Upregulation of the mitochondrial transcription machinery could drive the expansion of mtDNA numbers and OXPHOS complexes. How altering levels of these factors might also generate mitochondrial mass is unknown. mtDNA is often thought to be a measure of mitochondrial mass and could signal for mitochondrial biogenesis itself. mtTFA has been the most closely linked of the mitochondrial transcription machinery to mtDNA levels and is typically thought to be the major downstream mediator of the PGC-1α pathway to increase mtDNA and ultimately mitochondrial mass. However, ectopic expression of mtTFA, while it changes mitochondrial transcription and may increase mtDNA levels in some systems, has not been documented to cause major changes in mitochondrial mass (Ekstrand et al. 2004). Over-expression of h-mtTFB2 generates increased mitochondrial transcripts, mtDNA copy number, mitochondrial mass, membrane potential, and h-mtTFB1 levels (Figure 10; Figure 11; Figure 13; Cotney et al. 2007). Over-expression of h-mtTFB1 on the other hand only generates more mitochondrial mass and this phenomenon requires 12S rRNA methylation. So the common thread between these two scenarios is that changes in levels of h-mtTFB1 occur along with changes in mitochondrial mass. Since 12S rRNA methylation is critical for the mitochondrial mass increase by h-mtTFB1, I postulate that the methylation status of

the 3′-terminal stem-loop of this rRNA is measured and generates a retrograde signal for mitochondrial mass. To integrate what we have learned from the ectopic expression of both h-mtTFB1 and h-mtTFB2, I present a model of regulation of mitochondrial mass and membrane potential (Figure 21). This model has both anterograde and retrograde components with the methylation status of the 12S rRNA being a central player for coordinating mitochondrial biogenesis with mitochondrial membrane potential. I will discuss below how such a system might function.

When h-mtTFB1 is over-expressed, the 12S rRNA is hypermethylated and this could be a signal for good mitochondrial ribosomal health and possibly a meter for future translation. The cell increases mitochondrial mass in preparation of more OXPHOS complexes, but because h-mtTFB1 does not stimulate mitochondrial transcription there are no more mRNAs to be translated and the mitochondrial membranes are not filled with new complexes. This would result in higher mitochondrial mass, but less membrane potential overall which I do observe (Figure 18B). Expression of h-mtTFB2 on the other hand produces more mitochondrial transcripts, but if indeed its methyltransferase activity is much lower than h-mtTFB1, most of the 12S rRNAs produced would initially be hypomethylated or totally unmethylated. These under methylated 12S rRNAs generate a signal that drives expression of h-mtTFB1 (Figure 10). Increased levels of h-mtTFB1 then enter mitochondria and methylate their targets. Once 12S methylation status has been restored to normal or above normal levels it becomes a signal again, like the case of h-mtTFB1 expression alone, and drives mitochondrial membrane production. This time around, however, there are more ribosomes and mRNAs in the organelle and more OXPHOS complexes are produced and placed in the inner membrane. With more mass

and more OXPHOS complexes more membrane potential can be generated which I do in fact see (Figure 13). These two scenarios effectively create two arms of a mitochondrial biogenesis pathway. One arm is responsible for measuring 12S rRNA methylation status and possibly mitochondrial ribosome formation resulting in changes in mitochondrial mass. The other arm is responsible for generation of mtDNA and mitochondrial transcripts. The two arms are then connected by the 12S rRNA itself to generate both more mitochondrial mass and more functional organelle (Figure 21).

How the 12S rRNA could perform double duty in such a scenario and the nature of such a signal in this pathway are unknown, but are open to speculation. Bacterial ribosomes that contain unmethylated small subunit rRNA are known to have lower levels of certain ribosomal proteins and require higher levels of translation initiation factors to achieve normal translation (Poldermans et al. 1979). In human mitochondria, 12S rRNA methylation status could be monitored by the level of some free mitochondrial ribosomal proteins. If low levels of free protein are detected, due to fully methylated 12S rRNA, mitochondrial mass might be induced to match these fully active ribosomes. If large amounts of proteins are free, due to hypomethylation, there could be an interaction with the mitochondrial gene expression systems or processing events that allow these proteins to exit from mitochondria and stimulate pathways in the cytoplasm. In fact such a protein and scenario may have already been identified. A free mitochondrial ribosomal protein, MRPL12, has been shown to interact with POLRMT and increase the steady state level of mitochondrial transcripts (Appendix A; Wang et al. 2007). Changing levels of mitochondrial transcripts could have numerous effects on mitochondrial function and could generate many different types of signals to control expression of nuclear genes. In

bacteria the homologous protein, RPL12, has been shown to regulate the expression of itself and other genes within the same operon setting a precedent for a ribosomal protein functioning in a regulatory pathway (Little et al. 1981). An alternative hypothesis could include alterations in translational fidelity producing mutated OXPHOS complexes that produce more ROS or even the possibility of mitochondrial ribosomes consuming less GTP during translation, but until putative pathways controlling this type of membrane biogenesis are identified any number of possible signaling mechanisms are viable.

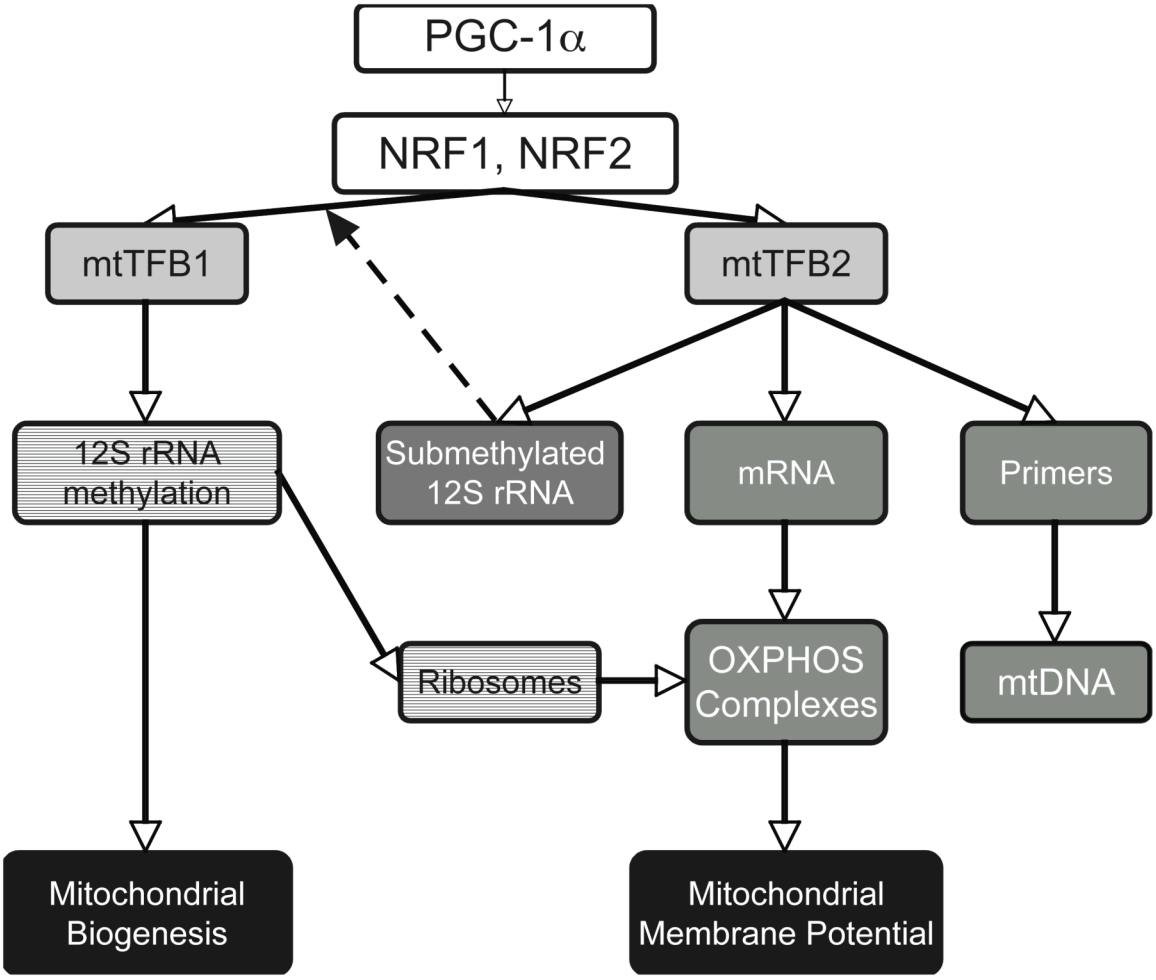


Figure 21 Mitochondrial biogenesis and membrane potential control by h-mtTFB1 and h-mtTFB2

Illustrated are the possible outcomes and participants in a pathway that controls complete mitochondrial biogenesis. Here, h-mtTFB1 and h-mtTFB2 are the major downstream effectors of general mitochondrial biogenesis pathways and their activities alone are sufficient to generate mitochondrial alterations. See text for further details.

5.6 Closing remarks

This research has illuminated how the mitochondrial transcription factors Bs may have evolved in many different organisms and identified them as major contributors to variation between species in mitochondrial gene expression. My results show that, at least for the human mtTFBs, these proteins have maintained the methyltransferase activity of their ancestors in addition to the gain of transcription factor function. *In vivo*, the roles of these proteins appear to differ substantially with h-mtTFB1 being the primary methyltransferase and h-mtTFB2 being the primary transcription factor. This research has also identified these factors as major players in the mitochondrial biogenesis response. Combining their activities results in a robust remodeling of the mitochondrial compartment generating a greater organelle mass along with more capacity for respiration. This response is dependent upon the methylation activity of h-mtTFB1, and uses the methylation status of 3′-terminal stem-loop of 12S rRNA as a metric for mitochondrial function. Inappropriate modulation of these proteins could be detrimental to overall cellular health by producing mitochondria that do not function properly or generate excessive amounts of ROS. Whether these two factors control such mitochondrial outcomes in all cells is of great interest and could provide new insight into how mitochondria can communicate with the rest of the cell. Finally, this research has added another piece to the puzzle of mitochondrial biology and provides a glimpse of the complex systems and communication that must be in place to regulate such an important organelle.

Appendix A: POLRMT and MRPL12

In bacteria, transcription and translation are known to be coupled. Transcription is initiated and elongated, but before the transcript is completed translation of the message has already begun. In yeast mitochondria a similar situation has been identified. The yeast mitochondrial RNA polymerase, through its ATD, associates with specific mitochondrial translation factors that serve to ferry the transcribing RNA polymerase to the mitochondrial ribosome (Rodeheffer and Shadel 2003). The events surrounding the handoff of the transcript from the RNA polymerase to the ribosome are unknown, but these binding activities are essential for proper mitochondrial translation in yeast. In humans the mtDNA and the translation machinery are both located in close proximity to one another at the inner mitochondrial membrane. This led us to speculate that mitochondrial transcription and translation in humans may also be coupled, possibly through the mitochondrial RNA polymerase.

The human mitochondrial RNA polymerase also contains an amino terminal domain extension like the Rpo41p protein, but to date has no documented function. We hypothesized that proteins may associate with the mitochondrial RNA polymerase other than those known to be specifically involved in transcription and serve to either regulate mitochondrial transcription or couple transcription to other processes such as translation. To begin to identify proteins interacting with POLRMT, we produced a clone of the protein with a 6x His tag at its amino terminus. This tagged protein was expressed and purified from *E. coli* and bound to a column with high affinity for his-tagged proteins. Soluble mitochondrial extracts from HeLa cells were then passed over this POLRMT

"bait" column. Subsequent washing, elution, and separation of proteins by SDS-PAGE revealed very few bands that specifically bound to POLRMT. One prominent band was excised from the gel and submitted for protein identification by mass spectrometry. This putative POLRMT interactor was identified as the mitochondrial ribosomal protein L12 (MRPL12). MRPL12 is a small protein located on the stalk of the large subunit of the mitochondrial ribosome. This protein is necessary for recruitment of translation initiation factors to bacterial ribosomes and affects translational output. To determine if this was a direct interaction, GST-tagged and purified MRPL12 was incubated with POLRMT. Using the same His-binding beads, POLRMT was pulled down and samples were assayed for presence of MRPL12 via immunoblotting against the MRPL12 protein. We confirmed that MRPL12 and POLRMT form a direct interaction *in vitro*. To begin to understand the biological relevance of this interaction we tested whether MRPL12 could affect levels of transcription by POLRMT *in vitro*. Addition of MRPL12 to a transcriptionally competent mitochondrial extract resulted in an increase of transcripts from the LSP (Wang et al. 2007). To identify if this activity is present *in vivo*, we over-expressed MRPL12 in HeLa cells and assayed for levels of mitochondrial transcripts. Over-expression of this protein stimulated the levels of steady state transcripts of mitochondrial transcripts from both the heavy and light strands *in vivo* (Figure 22).

This study identified the first protein to interact with the mitochondrial RNA polymerase in humans not known to be essential for basal mitochondrial transcription. Whether the ATD of POLRMT is essential for this interaction and the stimulation of transcript levels is due to stability of transcripts or increased initiation has yet to be determined. It is tempting to speculate that the interaction of these two proteins occurs at

the ATD, through which a direct interaction with free MRPL12 or protein bound to the mitochondrial ribosome could provide a mechanism of coupling transcription to translation as well as regulating transcription based on the levels or assembly of mitochondrial ribosomes. Further study will aim at identifying the regions required for this interaction and characterizing the nature of the elevated transcript levels. See Wang, Cotney, and Shadel, JBC, 2007 vol. 282 (17) pp. 12610-8 for more complete details of this work.

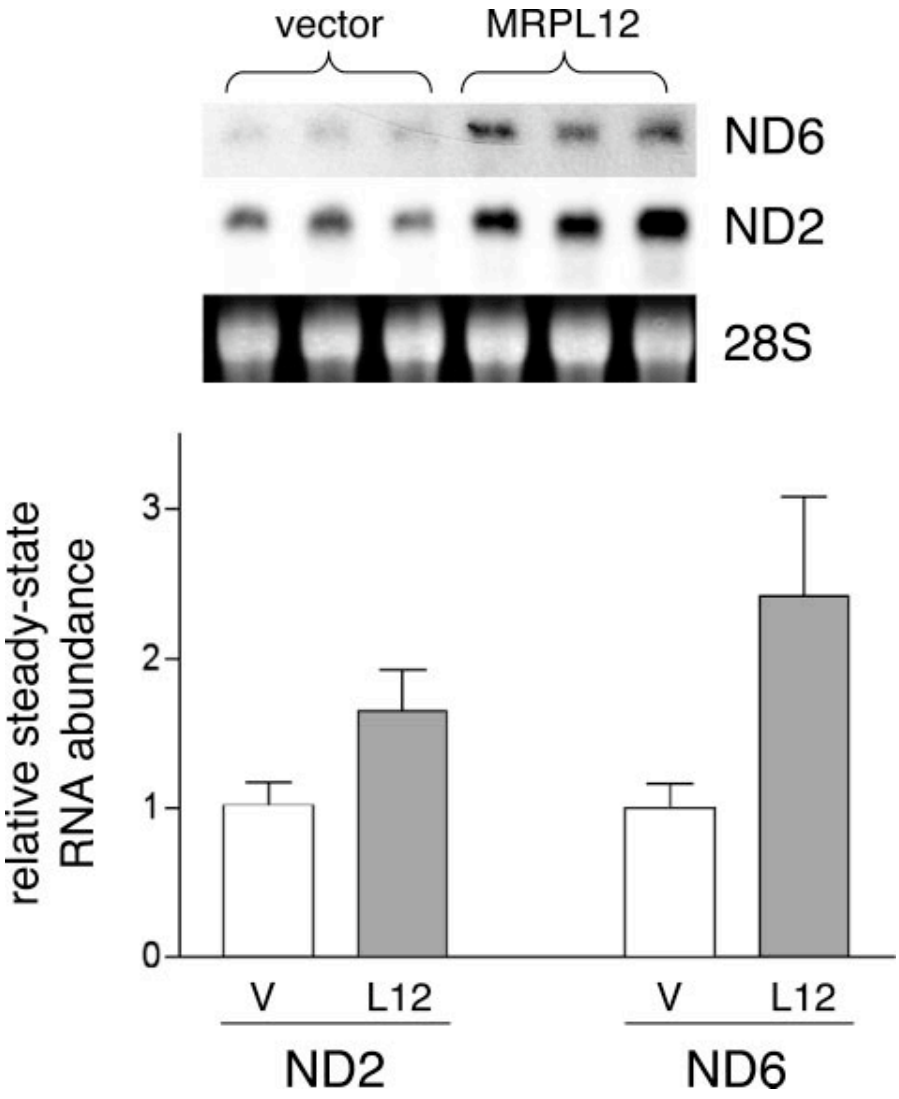

vector
MRPL12
ND6
ND2
28S
relative steady-state RNA abundance
3
2
1
0
V
L12
V
L12
ND2
ND6

Figure 22 Over-expression of MRPL12 in HeLa cells enhances the steady state level of mtDNA-encoded transcripts.

Shown is a Northern analysis of the mtDNA-encoded mRNAs from HeLa cells. The *top* is an autoradiogram of the blots probed for either the ND2 or the ND6 mRNA in three biological replicates of RNA isolated from empty vector control (*vector*) or MRPL12 over-expression HeLa cells. The 28 S cytoplasmic rRNA from the ethidium-stained gel is also shown and was used as the loading control. The relative abundance of the ND2 and ND6 RNA was quantified based on the 28 S loading control and graphed (*bottom*). The ratio of the ND2/28S or ND6/28S in the vector control cell lines (*V*) was given a value of 1.0 (*white bars*) and used to normalize the signals from the MRPL12 over-expression lines (*L12*). Thus, the *gray bars* represent the fold up-regulation ofND2 and ND6 mRNA in the MRPL12 over-expression cell line. All data represent the average of six values S.D. (*brackets*), that is, two experiments done in triplicate (one representative experiment is shown).

Appendix B: mtDNA Levels and Effects on Abundance of Transcription and Translation Machinery

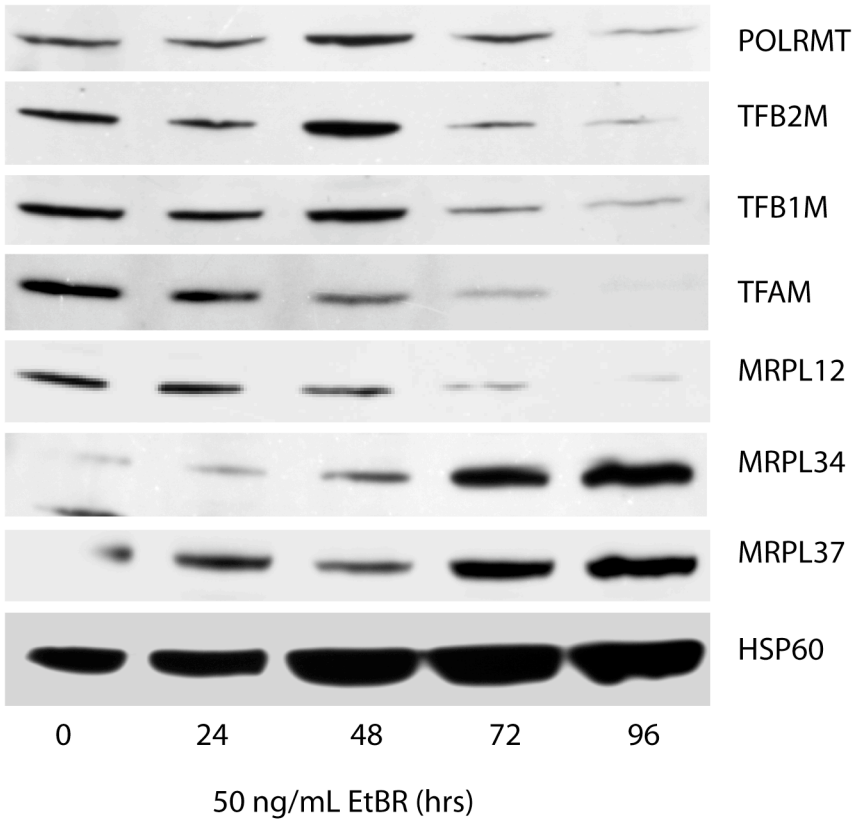
mtDNA has long been considered a marker for general mitochondrial mass. We have previously shown that two components of the mitochondrial transcription machinery, TFAM and POLRMT, are highly dependent on the level of mtDNA in a cell (Seidel-Rogol and Shadel 2002). If treated with an agent that causes depletion of mtDNA, levels of both TFAM and POLRMT decrease. The kinetics and degree of protein levels roughly mirrors the decline of mtDNA caused by the presence of ethidium bromide. We originally proposed that this phenomenon suggested a mutual dependence of mtDNA and the mitochondrial transcription machinery. In the absence of mtDNA the mitochondrial transcription machinery is no longer stable and levels decrease.

Excess TFB1M levels have no effect on mitochondrial transcription or mtDNA copy number in Drosophila or human cells. However, levels of the TFB2M protein greatly influence mitochondrial transcription and abundance of mtDNA in these two systems. From this data, it can be postulated that the TFB1M protein is not part of the mitochondrial transcription machinery. If this is the case, TFB2M should show depletion kinetics similar to TFAM and POLRMT while TFB1M will not be affected by loss of mtDNA. To address this issue we treated cells with ethidium bromide for four days and assayed the levels of each of the proteins mentioned above by western blot.

Over the course of four days, TFAM and POLRMT levels decrease dramatically as previously shown. As predicted, TFB2M levels also follow this general pattern of transcription machinery depletion (Figure 23). When TFB1M abundance was measured, I saw that this protein also followed a general profile of depletion. This suggests that

while TFB1M may not have direct effects on levels of transcription it is intimately dependent on either the levels of mtDNA or the other compoents of the mitochondrial transcription machinery. We have also recently idientified a mitochondrial ribosomal protein, MRPL12, which interacts with POLRMT and effects levels of mitochondrial transcription (Wang et al. 2007). If it is indeed involved with the mitochondrial transcription machinery it could also show depletion kinetics observed above. Indeed as mtDNA levels were depleted and the transcription machinery decreased, the levels of MRPL12 also declined (Figure 23). Other components of the large subunit of the mitochondrial ribosome did not show a depletion profile and instead increased during this treatment (Figure 23 B). This observation bolsters the argument that MRPL12 is involved in mitochondrial transcription and that depletion of mtDNA does not cause a general decline in all mitochondrial proteins.

A


POLRMT
TFB2M
TFB1M
TFAM
MRPL12
MRPL34
MRPL37
HSP60
0
24
48
72
96
50 ng/mL EtBR (hrs)

B

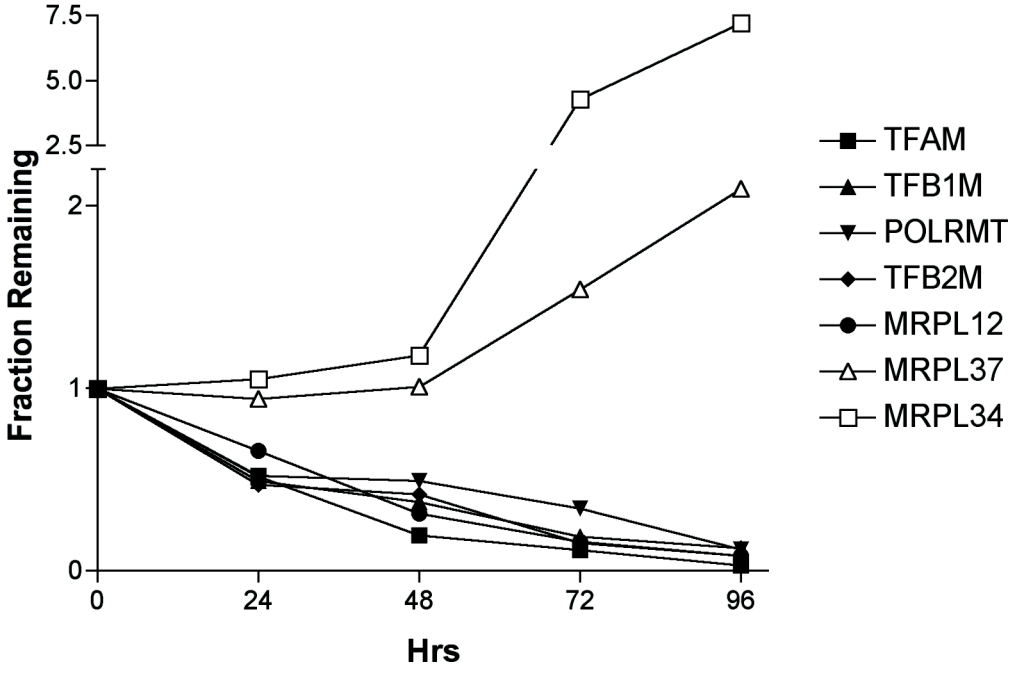

Fraction Remaining
7.5
5.0
2.5
2
1
0
0
24
48
72
96
Hrs
TFAM
TFB1M
POLRMT
TFB2M
MRPL12
MRPL37
MRPL34

Figure 23 Ethidium Bromide depletion of mtDNA reduces levels of mitochondrial transcription machinery.

A.) Immunoblot analysis of mitochondrial extracts obtained from HeLa cells treated with ethidium bromide. HeLa cells were incubated with 50 ng/mL ethidium bromide for up to four days. After each 24-hour period cells were harvested and mitochondria were extracted. 100 μg mitochondrial extract from each time point were electrophoresed and then probed for the indicated proteins. HSP60 serves as a loading control for mitochondrial protein. All identified components of the mitochondrial transcription machinery are depleted during ethidium bromide treatment. Additionally, MRPL12 shows similar kinetics. Two other proteins of the large subunit of the mitochondrial ribosome do not show depletion and instead increase substantially. B.) A graphical representation of above immunoblot. Values are plotted as a fraction remaining from the time zero level, normalized to HSP60 levels.

Appendix C: POLRMT and MYC

The MYC gene has been identified as a major activator of cell cycle progression by influencing the expression of a wide array of genes necessary for cell growth and division. Many different cancers have utilized this factor to initiate rapid proliferation and elevated cellular metabolism. How Myc directly influences cellular metabolism is not fully understood, but some of its targets include NRF-1 and cytochrome c (Morrish et al. 2003). More recently Myc expression has been shown to stimulate mtDNA levels, production of ATP, and total oxygen consumption (Zhang et al. 2007). These data would suggest a more direct role of Myc in changing cellular metabolism, possibly by influencing mitochondrial gene expression. Gene expression profiling has identified POLRMT to be a direct target of Myc (Zhang et al. 2006). Myc has also been documented to bind to the promoter of mtTFA and stimulate its expression (Li et al. 2005). Increasing the levels of these two factors could provide a mechanism for robust change in mitochondrial gene expression an overall mitochondrial respiration.

To begin to address whether POLRMT is a direct target of Myc *in vivo*, ChIP experiments were performed to identify DNA sequences bound by Myc. Binding to the POLRMT promoter was confirmed, with two sites showing significant binding affinity. To determine if these binding events can change POLRMT levels and influence relevant changes in mitochondrial gene expression, Myc levels were knocked down by RNAi. Depletion of the Myc protein caused a dramatic decrease in levels of POLRMT as well as the levels of mtTFA (Figure 24 A). This resulted in mtDNA copy number depletion and decreased steady state levels of all mitochondrial mRNAs assayed. Knockdown of Myc

also had major impacts on the organization of the mitochondrial network in H1299 cells. Normal cells show a reticulated network, but cells with Myc depletion display a highly fragmented mitochondrial compartment (data not shown). All of these effects show a major contribution of Myc to overall mitochondrial function and morphology, but whether these effects are mediated by many different genes or through the expression levels of POLRMT alone is unknown.

To address this issue, POLRMT was ectopically expressed in conjunction with Myc knock down. POLRMT levels were restored to nearly normal levels, but levels of mtTFA were unable to recover (Figure 24 A). This ectopic expression was sufficient to rescue both mitochondrial DNA copy number and mitochondrial transcript levels (Figure 24 B). These results suggest that POLRMT is indeed a direct target of Myc and changes in POLRMT alone are sufficient to influence major changes in mitochondrial gene expression. Since Myc is considered a major oncogene, many of its downstream effectors could be relevant targets for cancer therapy. These results could suggest a link between mitochondrial gene expression, specifically POLRMT, and tumor progression. It will be interesting to see if changes in POLRMT levels occur in other cancers and if targeting this RNA polymerase could slow tumor growth.

This work was performed in collaboration with Xiaoyong Zhang and Steve McMahon at Thomas Jefferson University.

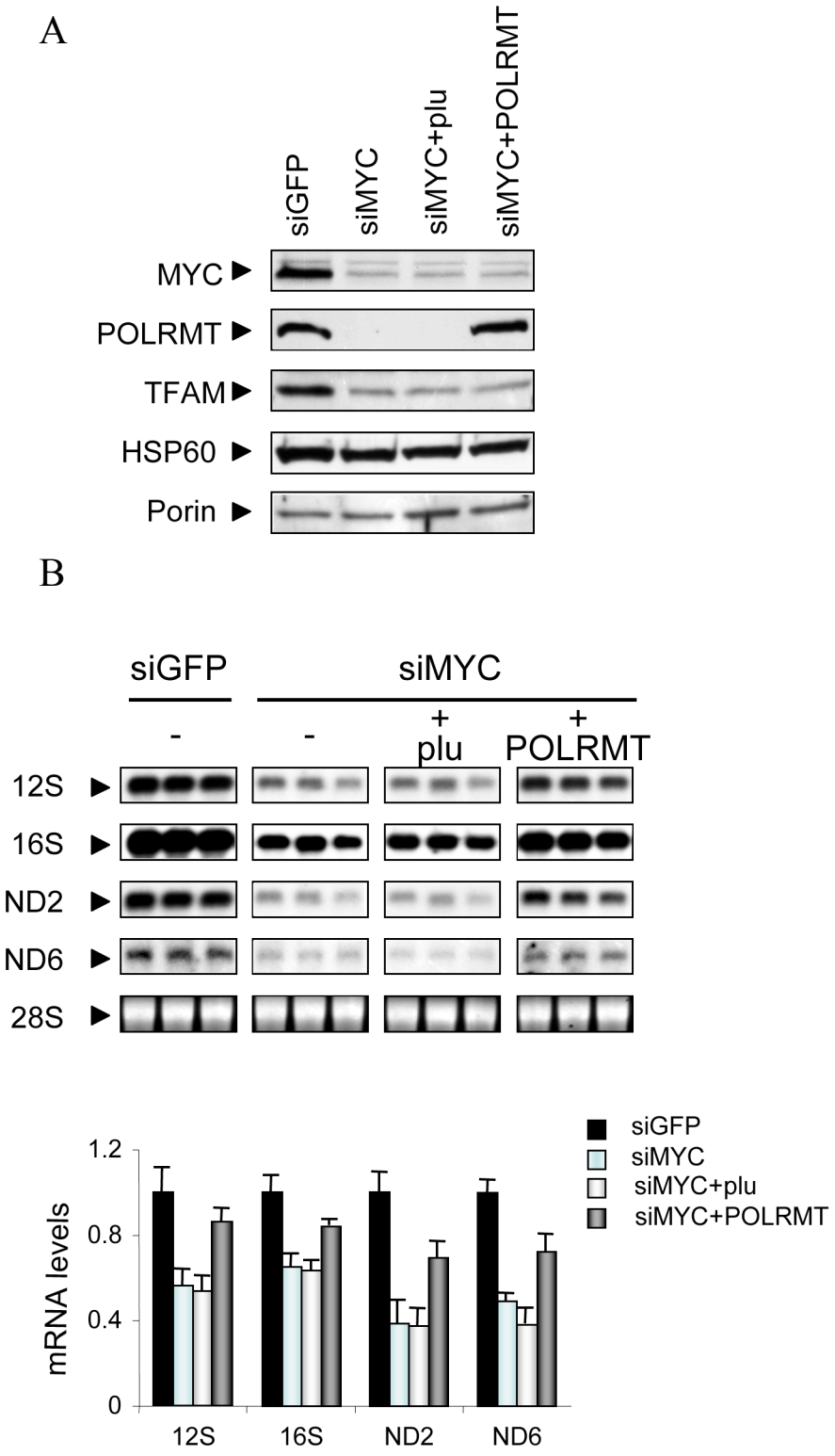

A
siGFP
siMYC
siMYC+plu
siMYC+POLRMT
MYC
POLRMT
TFAM
HSP60
Porin
B
siGFP
siMYC
-
-
+ plu
+ POLRMT
12S
16S
ND2
ND6
28S
siGFP
siMYC
siMYC+plu
siMYC+POLRMT
mRNA levels
1.2
0.8
0.4
0
12S
16S
ND2
ND6

Figure 24 Regulation of POLRMT by MYC controls transcription from the mitochondrial genome.

(A) H1299 cells were transfected with MYC siRNA oligos (Dharmacon) or GFP siRNA as a control. Twenty-fours hours post-transfection, one group of MYC depleted cells was infected with a lentivirus expressing human POLRMT or with the empty vector control (plu). Three days post-transfection, cells were harvested and mitochondrial proteins POLRMT, TFAM, HSP60 and Porin were measured by western blotting. The level of MYC depletion was also evaluated by western blotting. (B) Total RNA was isolated from the MYC-depleted cells described in (A). Levels of transcript for the mitochondrial genes 12S, 16S, ND2 and ND6 were determined by northern blotting (top panel). Levels of the 28S transcript were used as a loading control. Signals for northern analysis were quantified and expressed as fold change relative to the GFPi control (bottom panel). Averages from three different samples ± standard errors are displayed.

Appendix D: Materials and Methods

D.1 Plasmids

The plasmids used in this study to express h-mtTFB1 and h-mtTFB2 in *E. coli* are derivatives of pBluescript II KS+ (Stratagene), which harbors an IPTG-regulated *lac* promoter. The h-mtTFB1 cDNA was amplified by PCR with primers that introduced an *Nde*I restriction site (that included the ATG start codon) and a *Bam*H1 restriction site downstream of its natural stop codon. This product was inserted into the plasmid pT7-7 using these enzymes, which placed a consensus ribosome binding site (rbs) upstream of the ATG for optimal bacterial expression. Using an engineered *Eco*R1 site located upstream of the rbs, an expression cassette was inserted into pBluescript II KS+ as an *Eco*R1-*Bam*H1 fragment, to create the plasmid pBS-HB1. To construct the h-mtTFB2 expression vector (pBS-HB2), the h-mtTFB1 cDNA in pBS-HB1 was replaced by an analogous h-mtTFB2 PCR product with an *Nde*I site overlapping its start codon and a *Spe*I site downstream of its natural stop codon. The h-mtTFB1-G65A and h-mtTFB2-G105SA point mutations were created by a whole-circle PCR mutagenesis. The sc-mtTFB expressing plasmid was essentially the same as the pBS-HB1 and pBS-HB2 plasmids, except a *Xba*I-*Hind*III fragment from pGS344, a pT7-7 plasmid containing the *MTF1* gene engineered with a rbs for bacterial expression, was cloned in pTZ18R instead of pBluescript II KS+.

Bacterial expression vectors for production of h-mtTFB1 and h-mtTFB2 in *E. coli* were created using pET21b (Promega). The h-mtTFB1 and h-mtTFB2 cDNAs were amplified using primers

5′-AACATATGGCTGCCTCCGG-3′ and

5′-AAGAGCTCGAGTCTGTAATTCTC-3′

or primers

5′-AACATATGTGGATCCCAGTGG-3′ and

5′-GCGGCCGCCCTATCTTCCAGGGTTTCATC-3′,

respectively, and the resulting PCR products were ligated into pGEMT-Easy (Promega). These plasmids were then digested with appropriate restriction enzymes for ligation in pET21b (NdeI and XhoI for h-mtTFB1 or NdeI and NotI for h-mtTFB2). To generated plasmids for over-expression in human cells, the h-mtTFB1 and h-mtTFB2 cDNAs were both cut from previously described vectors (Cotney and Shadel 2006) using EcoRV and NotI and ligated into pcDNA3.1 zeo (+) (Invitrogen) cut with the same enzymes. The vector used to express POLRMT in bacteria was pProEX-Htb (Invitrogen). A portion of the human cDNA encoding amino acids 41-1250 and the stop codon was cloned into the *Bam*H1 and *Xho*I of this vector via a *Bam*H1-*Sal*1 restriction fragment. Amino acids 1-40 were deleted since they compose the mitochondrial localization sequence (MLS) that is predicted to be removed by proteases during import into mitochondria (Tiranti et al. 1997). However, in place of the MLS, there are 29 unnatural amino acids fused to POLRMT that include a initiator methionine, a His6 tag, a spacer of seven amino acids, a TEV protease cleavage site, and another spacer of six amino acids. The vector has an

intact *E. coli* lacI^q gene allowing POLRMT expression from the *trc* promoter to be regulated by addition of isopropylthiogalactoside (IPTG; Sigma).

D.2 Bacterial kasugamycin-resistance assays

Drug inhibitory assay was performed essentially as described (Huntington et al. 2000). Briefly, bacterial strains were grown overnight at 37°C in LB media + 100 μg/mL ampicillin (LB/amp). Cultures were diluted 100 fold in fresh LB/amp and grown to an OD_{600} of 0.2. The cultures were diluted again to an OD_{600} of 0.01 in fresh LB/amp + kasugamycin (BIOMOL) at various concentrations (see legend to Fig. 1). Cultures were then grown for 18 hours at 37°C and growth assessed by measuring absorbance of the cultures at OD_{600}. The bacterial strains used are JM101 (KsgA+ wild-type parent) and the corresponding *ksgA* mutant (Vila-Sanjurjo et al. 1999), which were kindly provided by Dr. A. Vila-Sanjurgo.

D.3 Primer extension analysis of bacterial 16S rRNA adenine methylation

Bacterial strains used for RNA isolation were grown overnight in LB/amp medium. The next day, cultures were diluted 20-fold into fresh LB/amp and grown for four hours. Approximately $1x10^9$ cells were then harvested by centrifugation and total RNA was obtained using an RNeasy Mini Kit (Qiagen) according to manufacturer's guidelines. Total RNA (1 μg) was incubated with ^{32}P-labeled EC16S primer (5'-GATATTTATAACCCCACCGCA-3') in 10 μL of 1x reverse transcriptase buffer (50 mM Tris-HCl, 8 mM $MgCl_2$, 30 mM KCl, 1 mM DTT, 1.2 mM dNTPs, pH 8.5) at 55°C for 20 minutes, at room temperature for 5 minutes, and then placed on ice. Annealed samples then were added to 10 μL of extension reaction mix (20 mM Tris-HCl, 10 mM $MgCl_2$, 1.2 mM dNTPs, 0.1 unit/μL RNAsin (Roche), 0.5 unit/μL AMV Reverse

Transcriptase (Roche) and incubated at 42°C for 30 minutes. 20 μL of formamide loading buffer (80% deionized formamide, 10 mM EDTA, 1 mg/mL xylene cyanol, 1 mg/mL bromophenol blue) was added to each reaction and samples were denatured at 95°C for three minutes and then placed on ice. Samples (15 μL) were loaded onto a 10% polyacrylamide/7M urea gel and separated by electrophoresis at 300 volts. The gel was then fixed in 10% acetic acid, dried using a vacuum gel dryer, and developed by autoradiography using x-ray film.

D.4 Phylogenetic analysis

Sequences listed in Table 1 were compiled and aligned using ClustalX (version 1.83.1) (Thompson et al. 1997) with default parameters. Alignments were then visually inspected and manually aligned using BioEdit (version 7.0.5) (Hall 2005). Regions of ambiguous sequence, extreme N-terminal and C-terminal sequences as well as regions where one sequence caused gaps to form in all other sequences were removed and this process was repeated to ensure the best possible alignment. Sequences were then analyzed using the PHYLIP package (version 3.65) (Felsenstein 2005). The alignment file was bootstrapped 100 times with SEQBOOT. Trees were generated with PROTPARS using estimated values from TreePuzzle and jumbling of sequences ten times. A consensus tree was then generated with CONSENSE and plotted using TreeView (version 1.6.6) (Page 1996).

D.5 Purification of recombinant proteins

BL21-Codonplus® *E. coli* (Stratagene) were transformed with pET21b vectors containing either mtTFB1 or mtTFB2. For purification of mtTFB1, an overnight culture of bacteria was diluted into 1L of LB media and grown at 37°C until culture reached an

OD_{600} of 0.8. IPTG was added to 1 mM for induction of expression and the culture was allowed to incubate at room temperature with shaking for 24 hours. Cells were harvested by centrifugation, then pellets were resuspended in 80 mLs of lysis buffer (50 mM Sodium Phosphate, pH8.0; 0.3 M NaCl, 10% glycerol, 0.1% Tween-20, and 10 mM Imidazole). Cells were lysed by sonication and the suspension was cleared by centrifugation. The clarified lysate was applied to a BD-Talon column (BD Biosciences). The column was washed with 10 column volumes of lysis buffer and eluted with three column volumes of lysis buffer plus 150 mM Imidazole. Fractions were assayed for presence of mtTFB1 by SDS-PAGE followed by either coomassie staining or western blot. Peak fractions were combined, concentrated, and dialyzed into storage buffer (20 mM Tris-HCl, pH 8.0; 0.5 mM EDTA, 0.25 M Sucrose, 15% glycerol, 1 mM DTT, and 1mM PMSF). Protein concentration was determined with a protein assay kit from Bio-Rad and confirmed by SDS-PAGE and coomassie staining with comparison against a BSA standard curve. For purification of mtTFB2, the same procedure was followed for mtTFB1, except that mtTFB2 was found to be highly insoluble and present in the pellet after sonication. Pellets after sonication were resuspend in lysis buffer with the addition of 0.1% Triton, 5 mM BME, 1 mM PMSF, and 6M guanidine-HCl. The solution was stirred overnight at 4°C then clarified by centrifugation. The protein solution was added to a Talon column as before then washed with 10 column volumes of the pellet resuspension buffer. The column was eluted with three column volumes of the same buffer plus 200 mM Imidazole. Fractions were analyzed as above and peak fractions were collected. Concentrated samples were dialyzed extensively against storage buffer and protein concentration was determined. Purification of mtTFA was performed as

described (Dairaghi et al. 1995), except the process was stopped after the BioRex-70 column (Bio-Rad). The expression and purification of POLRMT was performed as described (Wang et al. 2007).

D.6 HeLa cell growth and transfection

For transfection of HeLa cells (Clonetech) with h-mtTFB1 or h-mtTFB2 over-expression plasmids, cell were seeded at $5x10^5$ per 10 cm dish in DMEM (Sigma) plus 10% Bovine Growth Serum (Hyclone) and allowed to grow for 24 hours. Cells were transfected with empty pcDNA3.1 vector, pcDNA3.1-mtTFB1, or pcDNA3.1-mtTFB2 using Effectene (Qiagen) according to manufacturer's suggestions. Cells were allowed to incubate with plasmid/reagent complexes for 24 hours. Transfected cells were subcultured by diluting cells 1: 50 or 1: 100 in growth media with 400 μg/ml Zeocin (Invitrogen) and plated in 10 cm dishes. Cells were grown until individual colonies were visualized and 10-15 clones were individually selected with glass cylinders and transferred to 24-well plates. Cells were grown to confluence, subcultured into 6-cm dishes in growth media plus 100 μg/ml Zeocin and grown for 48 hours. One dish for each clone was harvested and assayed for protein expression via western blot. Remaining plates were harvested and stored in growth media plus 10% DMSO at -80°C. Stably transfected HeLa lines were consistently plated at 5000 cells/cm^2, grown for 72 hours at 37°C and 5% CO_2 before subculturing again. All experiments were performed on cells passaged at most 10 times.

D.7 Antibody production and western analysis of whole-cell and mitochondrial proteins

Four polyclonal peptide antibodies (two for h-mtTFB1 and two for h-mtTFB2) were generated for us by Multiple Peptide Systems. The peptides used as antigens were as follows:

TFB1-1 H-CVPKPEVDVGVVHFTPLIQPKIE-NH_2,

TFB1-2 H-CREELKRRKSKNEEKEEDDAENYRL-NH_2,

TFB2-1 H-CWIPVVGLPRRLRLSALAGA-NH_2,

TFB2-2 H-CPQLWPEPDFRNPPRKASKASLD-NH_2.

Multiple Peptide Systems performed the synthesis of peptides, injection of rabbits, and collection of serum, and also provided a small batch of peptide-affinity purified antibody. We also used antibodies that were purified from provided crude serum using protein-A sepharose (Amersham). Specificity of each antibody was determined by immunoblots of 200 ng of recombinant full-length h-mtTF1 and h-mtTFB2 proteins run alongside one another on the same gel. Antiserum used to detect human POLRMT was the same as that described previously (Seidel-Rogol and Shadel 2002) and polyclonal antibodies for detection of h-mtTFA were generously provided by Dr. David Clayton.

For whole cell extracts, $1x10^6$ cells were suspended in 100 μL of cold lysis buffer (50 mM Tris-HCl pH 8.8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 2 mM EDTA, 10% glycerol, 5 mM DTT, 1 mM PMSF) and incubated at 4°C with rotation for 30 minutes. Protein concentration was determined using the BioRad Protein Assay Kit and indicated amounts of total protein were loaded on polyacrylamide gels for

analysis. For mitochondrial extracts, mitochondria were harvested by differential centrifugation. Briefly, cells were resuspended in 10 pellet volumes of RSB buffer (10 mM NaCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl pH 7.5), swelled on ice for 10 minutes, homogenized with a motorized Teflon pestle, then 2.5x MS Buffer (525 mM Mannitol, 175 mM Sucrose, 125 mM Tris-HCl pH 7.5, 2.5 mM EDTA) was added to 1x. The homogenate was centrifuged at 980g for 10 minutes twice to pellet nuclei and unbroken cells. The supernatant was transferred to a fresh tube and spun at 17,000g for 30 minutes to pellet mitochondria. The mitochondrial pellet was washed three times with 1x MS buffer then stored at -80°C until further use. Mitochondrial pellets were resuspended in lysis buffer equal to one volume of the original cell pellet. The protein extract was then treated and quantified as described above.

Protein extracts were separated on SDS-polyacrylamide gels and then transferred to PVDF membranes. Membranes were blocked with 5% Milk/TBST (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.05% Tween-20) for 30 minutes at room temp then probed with primary antibody in 5% Milk/TBST overnight at 4°C. The primary antibody was removed and the blot was washed three times for 10 minutes with TBST. The appropriate secondary antibody was applied in 5% Milk/TBST and allowed to incubate at room temperature for one hour. The blot was washed again, ECL reagent was added and blots were exposed to film. Several exposures were obtained then films were photographed and analyzed with a BioRad VersaDoc and Quantity One software (v4.6.1). For serial western blots, membranes were stripped with stripping buffer (65 mM Tris-HCl pH 6.8, 2% SDS, 100 mM BME) at 50°C for 30 minutes then reblocked with 5% Milk/TBST.

D.8 Nucleic acid extraction, Northern blotting, Reverse transcriptase Real-Time PCR, and mtDNA copy number analysis

RNA for northern blot analysis was isolated from $1x10^6$ cells from the indicated cell lines using the RNEasy Kit (Qiagen) according to manufacturer's instructions. RNA was eluted in the final step using RNase-free dH_2O, quantified by absorbance at 260 nm, and stored frozen at -80°C until used. RNA (2 μg) was separated by size on 1.2% agarose/formaldehyde gels and transferred to uncharged nylon membranes (Osmonics) via upward capillary flow. RNA was crosslinked to the blots by UV irradiation using a Stratalinker (Stratagene), stained with 0.1 μg/mL ethidium bromide/100 mM ammonium acetate and destained with 100 mM ammonium acetate. The stained membrane was photographed using the BioRad VersaDoc and load was quantified using Quantity One software. The ethidium-stained 28S rRNA band was used as a loading control for all experiments as described by others (Correa-Rotter et al. 1992; Duhl et al. 1992; Eykholt et al. 2000). Body-labeled DNA hybridization probes for mtDNA-encoded 12S and 16S rRNAs and the ND2 and ND6 mRNAs were generated by PCR with 32P-dCTP using the following gene-specific primers:

h16S-5′, 5′-CCCTCAACTGTCAACCCAACACAGG-3′;

h16S-3′, 5′-CCGGGCTCTGCCATCTTAACAAAC-3′;

h12S-5′, 5′-GACCCAAACTGGGATTAGATACCCCAC-3′;

h12S-3′, 5′-GACCCAAACTGGGATTAGATACCCCAC-3′;

ND2-5′, 5′-GGCCCAACCCGTCATCTAC-3′;

ND2-3′, 5′-GAGTGTGGGGAGGAATGGGG-3′;

ND6-5′, 5′-GGGGTTTTCTTCTAAGCCTTCTCC-3′;

ND6-3′, 5′-CTAATCAACGCCCATAATCATAC-3′;

COX1-5′, 5′-CGGCGCATGAGCTGGAGTCC-3′;

COX1-3′, 5′-GAGAGATAGGAGAAGTAGG-3′;

COX3-5′, 5′-CCTAATGACCTCCGGCCTAG-3′;

COX3-3′, 5′-GAGCCGTAGATGCCGTCGG-3′;

ATP6-5′, 5′-CCACAATCCTAGGCCTACC-3′, and

ATP6-3′, 5′-GCATGAGTAGGTGGCCTGC-3′.

PCR products were purified using the Qiagen PCR Purification Kit, denatured at 95°C, then placed on ice until hybridized to blots. RNA blots were pre-hybridized with 20 mL of Rapid-Hyb Buffer (GE Healthcare) for one hour at 68°C in a hybridizing oven with horizontal rotating cylindrical jars (Techne). After prehybridization, 20 mL of fresh Rapid-Hyb containing the desired radiolabeled probe was added and incubation was carried out at 68°C overnight. Probe solution was then removed and hybridized blots were washed once with 2x SSC, 0.1% SDS at room temp for 10 min, then three times with 1x SSC, 0.1% SDS at 68°C for 10 min. The blots were then wrapped in plastic wrap and exposed to x-ray film with intensifying screens at -80°C. Films were photographed using a BioRad VersaDoc and quantified using Quantity One software. When blots were analyzed serially for multiple transcripts, they were first stripped with 10 mM Tris·HCl [pH 7.4], 0.2% SDS at 72°C for two hours to remove previously hybridized probe.

For analysis of h-mtTFB1 transcript levels, first strand cDNA synthesis was performed by combining 2 μg of total RNA with 1 mM dNTPs and 8 μM oligodT_{15} in a 20 μL reaction volume. This mixture was heated to 70°C for 10 minutes then allowed to cool to 4°C over 10 minutes to anneal the oligodT primer to polyA tails of mRNAs. An equal volume of reverse transcriptase mix (200 units M-MuLV Reverse Transcriptase (NEB), 2X M-MuLV RT Buffer, and 25 units of RNAseOUT (Invitrogen)) was added to

each primer-RNA mix and incubated at 42°C for 1 hour. The reactions were then heated to 70°C for 10 minutes to deactivate the reverse transcriptase. Finally an equal volume of 3 mM Tris-HCl pH 8.5, 0.3 mM EDTA was added to each reaction and samples were stored at -20°C until further use. cDNA samples were thawed and diluted from 16 to 32 fold in dH_2O. 10 μL of diluted cDNA sample was added to 14 μL of SYBR Green reaction mix (Taylor et al. 2005) to each well along with 0.5 μL of both appropriate 25 μM primers. β-actin primers,

β-actin RT F, 5′-TGGCACCACACCTTCTACAATGAGC-3′ and

β-actin RT R, 5′-GCACAGCTTCTCCTTAATGTCACGC-3′,

were used as controls for total cDNA in each reaction. To amplify the h-mtTFB1 transcript, primers

TFB1 RT F, 5′-TCTGCAATGTTCGACACATC-3′ and

TFB1 RT R, 5′-ACCTATATAAGAAGCTCCAC-3′ were used.

Reaction conditions for the BioRad iCycler were as follows: 95°C, 10 minutes, 1 cycle; 95°C, 30 seconds, 56°C, 30 seconds, 72°C, 30 seconds, 40 cycles. Fluorescence was measured after each 56°C step. Melt curve analysis to ensure single products was performed immediately after completion of steps above by using a temperature step gradient from 55°C to 80°C in 0.5°C increments with fluorescence measured after a 10 second incubation at each temperature. C_t values for each reaction were obtained through the iCycler iQ software. h-mtTFB1 transcript levels were then normalized by β-actin transcript levels and compared to empty vector controls.

Total DNA was extracted from $1x10^6$ cells by addition of 500 μL of extraction buffer (50 mM Tris-HCl pH 8.5, 0.25% SDS, 1 mM EDTA, 5 mM DTT) and boiling for

10 minutes. After cell lysis, tubes were allowed to cool to room temperature, 100 μg of RNAse A were added, and tubes were allowed to incubate at 37°C. Following a three-hour incubation, 100 μg of Proteinase K were added and samples were placed at 55°C overnight. Samples were heated to 95°C for five minutes and allowed to cool to room temperature. Total DNA concentration was measured and samples were stored at -20°C until further use.

Relative and absolute mtDNA copy numbers were measured by SYBR Green fluorescence using a BioRad iCycler and accompanying software (v3.1). Total DNA samples were diluted to a range of DNA concentrations from 500 pg/μL to 7.8 pg/μL and 10 μL were dispensed to appropriate wells of a 96 well PCR plate. 14 μL of SYBR Green reaction mix (Taylor et al. 2005) were added to each well along with 0.5 μL of both appropriate 25 μM primers. For detecting relative levels of mtDNA, a region of the mitochondrial genome encompassing a portion of COX3 and a region of the multicopy nuclear 18S rDNA locus were used. Primers are as follows:

RTQ COX3 F, 5′ -CACCCAAGAACAGGGTTTGT-3′;

RTQ COX3 R, 5′-TGGCCATGGGTATGTTGTTAA-3′;

RTQ 18S F, 5′-TAGAGGGACAAGTGGCGTTC-3′; and

RTG 18S R, 5′-CGCTGAGCCAGTCAGTGT-3′.

Reaction conditions for the iCycler were as follows: 95°C, 10 minutes, 1 cycle; 95°C, 15 seconds, 60°C, 1 minute, 40 cycles. Fluorescence was measured after each 60°C step. Melt curve analysis to ensure single products was performed immediately after completion of steps above by using a temperature step gradient from 55°C to 80°C in 0.5°C increments with fluorescence measured after a 10 second incubation at each

temperature. Standard curves to determine absolute copy number were constructed with known amounts (1638400 to 100 templates) of the plasmid pGEMT Easy (Promega) containing either one of the PCR products obtained using the mitochondrial and nuclear primers described above (COX3 or 18S).

D.9 Labeling of mitochondrial translation products *in vivo*

Labeling was performed as described (Chomyn 1996) with modifications. HeLa cells were seeded and grown as described above in 10 cm dishes. After 72 hours of growth, cells were washed three times with five mLs of sulfur free media (Gibco) without serum. Cells were allowed to incubate for five minutes at 37°C, 5% CO_2 before removing media. After the last wash, five mLs of sulfur free media with 10% BGS plus 100 μg/mL emitine were added to each plate and allowed to incubate at 37°C, 5% CO_2 for five minutes. After this incubation step, Expre^{35}S^{35}S Protein Labeling Mix (Perkin Elmer) was added to each plate to a concentration of 125 μCi/mL and labeling was allowed to proceed at 37°C and 5% CO_2. After one hour the labeling media was removed and the plates were rinsed once with five mLs of normal DMEM with 10% BGS. Plates were then washed twice with 10 mLs of TD buffer (25 mM Tris pH 7.5, 137 mM NaCl, 10 mM KCl, 0.7 mM Na_2HPO_4), trypsinized, and harvested by centrifugation at 1000g. Cell pellets were washed once with 10 mLs of TD buffer then transferred to a 1.5 mL tube with one mL of TD buffer and harvested again by centrifugation. Mitochondria were harvested as above and protein was quantified. 50 μg of total protein were added to each lane of a 10% to 20% linear gradient SDS-polyacrylamide gel. Samples were separated then the gel was dried and exposed to film at -80°C.

D.10 Kasugamycin sensitivity assays

HeLa cells were plated and grown as described above. After 48 hours of growth, kasugamycin was added at indicated concentrations and cells were allowed to incubate for 72 hours. Cells were washed, trypsinized, and treated with 0.08% Trypan Blue in PBS. The total number of viable cells at each drug concentration was compared to that of zero-drug controls to yield a percentage of viable cells that is plotted as a function of the drug concentration.

D.11 FACS analysis

HeLa cultures were plated and grown as described above. After 72 hours of growth cells were stained in culture media with 70 nM Mitotracker Green and 90 nM Mitotracker Red CMXRos, both from Molecular Probes, for 30 minutes at 37°C, 5% CO_2. After staining, dyes were removed and cells were washed three times with PBS. Cells were trypsinized and collected by centrifugation. Cell pellets were resuspended in 1 mL of PBS and analyzed on a BD FACScalibur instrument with accompanying software. Histograms and means for these data were obtained using FlowJo (v 8.0.1)

D.12 Mitochondrial localization sequence and cleavage prediction

Full-length sequences of h-mtTFB1, h-mtTFB1, and sc-mtTFB were obtained from NCBI. These sequences were submitted to cleavage site prediction by the SignalP program (v3.0, http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen et al. 2004). Hidden Markov Modeling option was selected and the first 70 amino acids for each sequence were used for signal and cleavage site prediction (Nielsen and Krogh 1998).

D.13 Primer extension analysis of human 12S rRNA

Total RNA or mitochondrial RNA from purified mitochondria was harvested using the RNeasy kit according to manufacturer's directions. A primer designed to bind the 3'-terminus of the 12S rRNA was labeled with 32P by T4 polynucleotide kinase as previously described (Seidel-Rogol et al. 2003). 10 μg of total RNA or 2 μg of mitochondrial RNA were incubated with 200 pmol of labeled primer in 1x AMV RT buffer (Roche) with either 1mM dNTPs or 1x dNTP mix (1mM dATP, dTTP, dCTP and 250 uM ddGTP) in a 25 μL reaction volume for 20 minutes at 55°C. The tubes were allowed to cool to room temp for five minutes then 10 μl of AMV RT mix was added (1x AMV RT buffer, 20 units AMV RT, 1 mM dNTPs of 1x dNTP mix) and samples were incubated at 42°C for 30 minutes. After elongation had occurred an equal volume of formamide loading buffer was added to each sample and loaded on a 15% acrylamide:7M Urea sequencing gel. Samples were electrophoresed at 25 mA for approximately 2.5 hrs. The gel was then place on filter paper, dried, and exposed to film.

Appendix E: Sequence Alignments

Supplemental Figure 1 Sequence alignment used to generate tree shown in Figure 2.

```
Phylip 3.4 Format

    94    251
mmKsgA      GGT-FDQHFL IDAGYLDRIV AAAELSPQDT VLEIGAGIGN LTERLAR-RA
paKsgA      KRRRLAQHFL RDPSVAEYIA G---LVPSGL VIEVGPGAGA LTIPLAK-RS
pyaKsgA     PNDRIGQHFL IVKDVIDKAI EVAEVSKSDV VLEVGPGLGF LTDELSK-RA
mgKsgA      PKKKLGQCFL IDKNFVNKAV ESANLTKDDV VLEIGLGKGI LTEELAK-NA
mkKsgA      PRRRLGQHFM VDDNILEFMV EAAEVREDDI VLEIGPGPGL LTRYLMT-RA
mhKsgA      MRAYRDQHFL TDPRIVARIA DILDISGRIV -LEIGPGEGI LTEALLE-RG
afKsgA      LRKSLGQHML VDRRVISRIV GYAELSEDDV VLEVGCGTGN LTSALLR-KC
hmKsgA      ADTRQDQHFL VDDRVLDRIP EYATDADIDL VLEIGAGPGA LTDRLLA-TA
faKsgA      FAKKYGQVFL NDKNIAAKEV RLLGIEPGDH VLEIGPGHGI LTGILLS-EP
ecKsgA      ARKRFGQNFL NDQFVIDSIV SAINPQKGQA MVEIGPGLAA LTEPVGE-RL
stKsgA      ARKRFGQNFL NDRFVIDSIV SAINPQKGQA MVEIGPGLAA LTEPVGE-RL
ypKsgA      ARKRFGQNFL NDQFVIDSIV SAIHPVPGEA VVEIGPGLGA LTEPVAA-RM
spKsgA      ARKRFGQNFL TDGNVINRIV GAIAPDNNHV MVEIGPGLGA LTEPVAM-AV
bmKsgA      ARKRFGQNFL VDHGVIDAIV AAIRPERGER MVEIGPGLGA LTGPVIA-RL
tdKsgA      PRKRFGQNFL IDDGIVHAIV NAIHPQAGET VVEIGPGLGA LTRPLLE-RL
ppKsgA      ARKRFGQNFL HDAGIIDRIL RAINAKAGEH LLEIGPGQGA LTEGLLG-SG
nmeKsgA     ARKRFGQNFL QDTRIISDIV NAVRPQADDV VIEIGPGLAA ITEPLAK-KL
lpKsgA      PRKRFGQNFL QDKYIINEIL RAINPLADDN MLEIGPGLGA LTQPLLQ-KL
bpKsgA      ARKRFGQHFL TDESVVESIV RAIGPARDDR VVEIGPGLSA LTRPLLD-RI
xfKsgA      AKKAFGQHFL VDRYYIDRII HAITPQPNDH IVEIGPGQGA ITLPLLK-CC
nmKsgA      ARRRFGQNFL HDPSILHRMV DSIDPRPGQC CIEIGSGLGA LTRPLLE-RA
neKsgA      PRRRFGQHFL VDHHIIAEII HIICPLPGDR MIEIGPGLGA LTQPLLN-NL
ccKsgA      AKKKLGQNFL TDRNITRKTV LLSGAKPDDQ VVEIGPGFGA LTRELVE-EC
bcKsgA      FKKSLGQNFL IDTNVLNRIV DHAEIGSESG AIEIGPGIGA LTEQLAK-RA
teKsgA      ARKRFGQHWL RSEAILAQII AAAELHPGDR VLEIGPGRGA LTRPLLV-SG
cdKsgA      FSKSLGQNFL IDSNIIDKIL SGARITRGDN IIEVGPGIGT LTREMGK-IA
cbyaKsgA    PRKRFGQHWL KDPAVHEAIV RAAQLPPPQR VLEIGPGTGQ LTQRLLA-QG
ttKsgA      ADKRFGQNFL VSEVHLRRIV EAARPFTGP- VFEVGPGLGA LTRALLE-AG
scDIM1      FNTDLGQHIL KNPLVAQGIV DKAQIRPSDV VLEVGPGTGN LTVRILE-QA
klDIM1      FNTDLGQHIL KNPLVAQGIV DKAQIKPSDI VLEIGPGTGN LTVRILE-QA
skDIM1      FNTDLGQHIL KNPLVAQGIV DKAQIKPSDV VLEVGPGTGN LTVRILE-QA
spDIM1      FNKDFGQHIL KNPLVAQGIV DKADLKQSDT VLEVGPGTGN LTVRMLE-KA
ptDIM1      FNTGIGQHIL KNPLIINSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
hsDIM1      FNTGIGQHIL KNPLIINSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
mmDIM1      FNTGIGQHIL KNPLIVNSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
ceDIM1      FNTDKGQHIL KNPGVVNAIV EKSALKATDT VLEVGPGTGN LTVKMLE-VA
rnDIM1      FNTGIGQHIL KNPLIVNSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
dmDIM1      FNKDFGQHIL KNPLVITTML EKAALRATDV VLEIGPGTGN MTVRMLE-RA
mdDIM1      FNTGLGQHIL KNPLIVNSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
xtDIM1      FNTGLGQHIL KNPLIVNSII DKAAIRPTDV VLEVGPGTGN MTMKLLE-KA
ciDIM1      FNTGVGQHIL KNPLIINGMV EKAALKATDT VLEIGPGTGN MTVKMLE-KV
trDIM1      FNTGIGQHIL KNPLIVNSII EKAALRPTDV VLEVGPGTGN MTVKLLE-KA
cfDIM1      FNTGIGQHIL KNPLVVNSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
```

```
agDIM1      FNKDFGQHIL KNPLVITSML EKAALRPTDV VLEIGPGTGN MTVKILE-KV
tnDIM1      FNTGIGQHIL KNPLIVNSVI EKAALRPTDV VLEVGPGTGN MTVKLLE-KA
cbDIM1      FNTDKGQHIL KNPGVVNAIV EKSALKSTDT VLEVGPGTGN LTVKMLE-VA
drDIM1      FNTGIGQHIL KNPLVVNGII EKAALRPTDV VLEVGPGTGN MTVKLLE-KA
btDIM1      FNTGIGQHIL KNPLIVNSII DKAALRPTDV VLEVGPGTGN MTVKLLE-KA
xlDIM1      FNTGLGQHIL KNPLIVNSII DKAAIRPTDV VLEVGPGTGN MTMKLLE-KS
amDIM1      FNTDIGQHIL KNPLVIQSMV EKAAVRPTDV VLEIGPGTGN MTIKLLD-KA
tbDIM1      FNKGFGQHIL KNPLVIAAIV EKAAVKPTDI VLEIGPGTGN LTEKLLQ-AA
ddDIM1      MNKSYGQHLL KNPLIIDAIV DKSQLKSTDT VLEIGPGTGN LTMKLLE-NC
cemtTFB1    AKKILSQNYL MDMNITRKIA KHAKVIE-KD VIEIGPGPGG ITRAILEAGA
hsmtTFB1    AAKQLSQNFL LDLRLTDKIV RKAGNLT-NA VYEVGPGPGG ITRSILNADV
mmmtTFB1    AVKQLSQNFL LDLRLTDKIV RKAGSLA-DV VYEVGPGPGG ITRSILNANV
drmtTFB1    AQKQLSQNFL LDTRLTDKIV RQAGNLN-NA VCEVGPGPGG LTRSILKAGA
dmmtTFB1    ARKQLSQNFL MDERLTDKIV KSAGRIDPRD VLEVGPGPGG ITRSILRRHP
rnmtTFB1    AAKQLSQNFL LDLRLTDKIV RKAGSLA-DV VYEVGPGPGG ITRSILNADI
agmtTFB1    AIKQLSQNFL MDERLTDKIV RAAGNIR-DH VLEVGPGPGG ITRSIIRQNP
ppmtTFB1    AAKQLSQNFL LDLRLTDKIV RKAGNLT-NA VYEVGPGPGG ITRSILNADV
ammtTFB1    AMKELSQNFI LNQNLADKII KKTGNLN-DC VLEIGPGPGA LTRSILKCQP
cbmtTFB1    AKKILSQNYL MDMNINRKIA KHAKVNE-ND VIEIGSGPGG ITRGILEAGA
tnmtTFB1    AEKQLSQNFL LDLKLTDKIV RQAGCLK-GA VCEVGPGPGG LTRSILNAGA
cfmtTFB1    AAKQLSQNFL LDLRLTDKIV RKAGNLT-NA VYEVGPGPGG ITRSILNANV
xlmtTFB1    AEKQLSQNFL LDLKLTDKIV RRAGNLQ-NA VCEVGPGPGG ITRSILNAGV
xtmtTFB1    AEKQLSQNFL LDLKLTDKIV RKAGNLQ-NA VCEVGPGPGG ITRSILNAGV
trmtTFB1    AEKQLSQNFL LDLKLTDKIV RQAGCLK-DA VCEVGPGPGG LTRSILNAGA
btmtTFB1    AVKQLSQNFL LDLRLTDKIV RKAGNLT-NA VYEVGPGPGG ITRSILNAGV
mdmtTFB1    AQKQLSQNFL LDLRLTDKIV RKTGDLK-NA VYEVGPGPGA FTRSILNAQV
cimtTFB1    ARKQLSQNFL LDPLITNRFV LCGAKDLAGH VCEVGPGPGP ITRSILQRKP
hsmtTFB2    SKASLDFKRY VTDRRLAETL AQIYLGKPSR LLECNPGPGI LTQALLE-AG
mmmtTFB2    SRNRSEPTRH IACKKAARNL VRDLLEHQNP ILECNPGPGI LTGALLK-AG
rnmtTFB2    HKSRSEPTRH IACKKSARNL VRDLLEHQNP ILECNPGPGI LTGALLK-AG
dmmtTFB2    KQKVPTHMYI ANSEAAARIN QYLEPHFQSS VMELNSGAGY FTRHLLD-RE
cfmtTFB2    PRCRSEPRSY ITSPKLATTL AHILENERKT LLECNPGPGI LTEALLK-SK
ptmtTFB2    -KASLDFKRY VTDRRLAETL AQIYLGKPSR LLECNPGPGI LTQALLE-AG
drmtTFB2    ALVCKNLRRF IVDPALATIV TDHLSRDIDD IFECNPGPGV LTRALLN-RG
btmtTFB2    YKSRSETKRY VTSPRVAETV VRVLRGKRK- ILECNPGPGV LTRALLE-SG
trmtTFB2    ------MRHF IVNPDLAKLI TQHLL--PEN IFECNPGPGV LTRTLLN-SG
xtmtTFB2    ---------- ---------- ---------- -----PGPGI LTQTLLA-AG
agmtTFB2    SSLSTERFYL ANRTTAARIA DVLTQDLPAD LVEVNPGPGL LTEQLLQ-RN
tnmttFB2    AQACKKIRHF IVDPDLAKLV TQHLQ--PDN IFDCNPGPGV LTRTLLN-SG
scmtTFB     LKFFYGFKYL WNPTVYNKIF DKLDLTKTYK VLDLYPGVG- IQSAIFYNKY
spmtTFB     GRSSYKSSYL VNQNLMDEAL VSNLLKEYNS ILEMAPGPGV TTTSLFN--Y
skmtTFB     IKHYYGFKYL LNSSVHTQIY NKLQLQSTYK VLDLYPGPS- QHSAIFRNIF
sdmtTFB     LKFFYGFKYL WNPTVYNRIF DKLDLTKTYK VLDLYPGVG- VQSAIFYNKY
klmtTFB     IHTSYGSQFE KNPKVINQIL DKLNLESYYK ILDIYAGPL- IQSVILNERL
egmtTFB     VKFSYGRTLL HSPTAIELAY KRLNLQEHYD VLELYPGTG- LPSYIFHDIY
cgmtTFB     IKSYYGFRYL LDPDWHQKVF EKLGLSS--- VLDLYPGPG- IHSAVLYNKV
cdmtTFB     IPYYYSRAHM ADPEACQKIL DKLDLKSKYD IVDVNPGYG- LFSTMLNYEL
dhmtTFB     PKYQYGYRHV VRPQPCQDII DKLNLKKKYP IIDVFSGYG- LFSSMINYEL
ddmtTFB     AKQQLSQNFL IDKNITDKIC KKSGGFD-DC VIEVGAGPGG LTRSLLTSGA
lmmtTFB     FLAKYDQRFM LNLKLTHQLV SRTTLTTPDK LVELGPGVGS LTRSLLTRPC
tbmtTFB     YLAKYGQRFI LNLKLTHQVA ARTTLRTPDK LLELGPGAGA LTRSLLTRPC

            KKVIAVELDP ALVSVLHDRF -D---AAENI EIIAGDALKV DFP--EFDKV
```

```
KTVYAIEIDK ALAERLRG-- ----IAPPNV VIIVGDALEV E--WPRADFF
KKVFTIELDR RIIEILRN-E -Y---SWNNV EIIQGDAVKV EWP--SFNKV
KKVYVIEIDK SLEPYANKLK -E---LYNNI EIIWGDALKV DLNKLDFNKV
GQVIAVELDG RMVEILKREL -G---EAPNL EIVRADFLEY DVP-DDVNKV
ARVISVELDR TLIERLSRRF -ASEIADGSL TLLQGDAVKV P--LPPFEIV
S-VVGIEKDP LMVKRLRERF -SDFIGKGRF RLIQGDALKV D--FPYFTKF
ERVTAVERDP DFAAHLREEF -TEEVAADRL TIVEGDALEV DLP--DFTAS
VKLTAIEPDH RFYESLKISY -HDHIVSGKF NIIKESFLDT EP--SYFDHI
DQ----LTVI ELDRDLAARL QTHPFLGPKL TIYQQDAMTF NFGELAEKRV

DK----LTVI ELDRDLAARL QTHPFLGPKL TIYQQDAMTM NFGELSAQRV
DH----MTVI ELDRDLAARL ASHPQLKDKL TIHQQDAMKV NFSELSEQRV
DN----LTVV ELDRDLVERL HKHPVLKDKL TIHQGDALQF DFSQLVVPKV
ATPGSPLHAV ELDRDLIGRL EQ-RFG-ELL ELHAGDALTF DFGSIARPRI
PH----LHAV ELDRDIIARL RR-AWPPERL TLHAGDALKF DFGSLG--RI
AQ----LDVV ELDKDLVPIL HHKFADRSNF RLHQGDALKF DFNQLGVPKV
NR----LHVV EIDRDIVCRL KTLP-FADKL VIHEGDVLQF DFNGIAGKKI
NR----LTAI EIDTDLQSYL TCLPVSQGKL NLIPADALTV DFCQFGP-RV
DH----LTAV EIDRDLAARL RR-QYPAERL TVVEADALTV DFAQFG--RV
GS----LTAI ELDRDLIAPL TAAATPLGKL DIIHRDVLTV DLSILAKPRL
RA----LVAI ELDRDLIEPL RRCCDGAGEL EIIQADALGL DFACFRQGRV
DT----LQAI ELDRDIVDYL SR-NYA-EKL VIHNVDALKF DFSALG--RI
HNLTVIEKDP TLATFIRNEY PQIKVIEGDV LTINFSAMAQ AG--KPL-QI
KK----VVAF EIDQRLLPIL DETLAPYGNV TVINKDVLKA DVHEVFSEMV
AE----VVAV ELDRDLCGQL RRQFDSE-RF QLIEGDILRL DLAPLG--KV
EK----VVAI EIDRNLIPIL KDTLSDLDNT EVVNQDILKV DIQELVKDKL
VH----VVAV EIDRDLCRLL QKRFADQPRF HLVEGDFLRL PLPPQ---LL
AE----VTAI EKDLRLRPVL EETLSGLP-V RLVFQDALLY PWEEVPQG-L
KNVVAVEMDP RMAAELTKRV -RGTPVEKKL EIMLGDFMKT E--LPYFDIC
RKVVRRGVRS SYGSGRTKRV -HGTPVEKKL EILLGDFMKT E--LPYFDVC
RRVVAVEMDP RMAAELTKRV -HGTAAEKKL EILLGDFMKT E--LPYFDIC
RKVIAVEMDP RMAAEITKRV -QGTPKEKKL QVVLGDVIKT D--LPYFDVC
KKVVACELDP RLVAELHKRV -QGTPVASKL QVLVGDVLKT D--LPFFDTC
KKVVACELDP RLVAELHKRV -QGTPVASKL QVLVGDVLKT D--LPFFDTC
KKVVACELDP RLVAELHKRV -QGTPLASKL QVLVGDVLKS D--LPFFDAC
KTVIACEIDP RMIAEVKKRV -MGTPLQNKL QVNGGDVMKM E--WPFFDVC
KKVVACELDP RLVAELHKRV -QGTPLASKL QVLVGDVLKS D--LPFFDAC
KKVIACEIDT RLAAELQKRV -QATPLQPKL QVLIGDFLKA E--LPFFDLC
KKVVACELDP RLVAELHKRV -QGTPQASKL QVMVGDVLKT D--LPFFDAC
KRVVACELDT RLVAELQKRV -QGSAVASKL QVMVGDVLKT D--LPFFDLC
NKVIACEIDP RMSAELQKRV -CGTPLQKKL HLMVGDVLKL ES-LPFFNVC
KKVVACELDC RLVAELQKRV -QCTPMQNKL QILVGDVLKT D--LPFFDIC
KKVIACELDP RLVAELHKRV -QGTPLASKL QVLVGDVLKT D--LPFFDAC
KKVVACEIDT RLVAELQKRV -QGTHMQPKL QILIGDVLKT D--LPFFDIC
KKVVACELDC RLVAELQKRV -QCTPMQNKL QILVGDVLKT D--LPFFDVC
KTVIACEIDP RMIAEVKKRV -MGTPLQTKL QVNGGDVMKQ E--WPFFDVC
KKVVACELDT RLVAELQKRV -QCTPMQNKL QILIGDVLKT E--LPFFDVC
KKVIACELDP RLVAELHKRV -QGTPLASKL QVMVGDVLKA D--LPFFDAC
KRVIACELDT RLVAELQKRV -QGSAVASKL QVMVGDVLKT D--LPFFDLC
KKVIACELDT RMVAELQKRV -QGTAYQSKL QIMIGDVLKT D--LPFFDLC
KKVIAFEVDP RMVAELNKRF -QNSPLAPKL QVIRGNCLDH E--FPYFDKC
KKVIAIEVDP RMAAELQKRV -AASPYAQHL QIILGDFLKV D--LPYFDVC
SRLDVVEIDN RFIPPLQHLA EAADSR---M FIHHQDALRT EIGDIWKNHV
```

```
AELLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTF KVEKAFSEHI
AELLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTY KIEKAFPGHI
ADLLVVEKDM RFIPGLQLLS EAAPGR---I RIAQGDILAY KLERRFPAHI
QRLLLVEKDP RFGETLQLLK ECASPLNIQF DIHYDDILRF NIEQHIPDHL
AELLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTY KIEKAFPDHI
RHLVVVEKDR RFMPTMEMLA EVAQPF-MRM DIVQGDILDY RVAEAFPDHL
AELLVVEKDT RFVPGLQMLS DAAPGK---L RIVHGDVLTF KVEKAFSEHI
KKLIVVEKDK RFEPTLEMLA DAFETINGKM EIIFDDIMKI NMSNLFPSKL
SRLDVVEIDN RFIPPLQHLA EAADSR---M FIHHKDALRT EIGDIWKKHV
ADLLVVEKDS RFIPGLQLLS EAAPGR---L RIVHGDILTY RMDRGFLGHV
AELLVVEKDS RFIPGLQMLS DAAPGK---L RIVHGDVLTF KIERAFPEHI
EELLVVEKDT RFIPGLKMLN EASGGK---V RTVHGDILTY RMDRAFPKHI
EELLVVEKDT RFIPGLKMLN EASSGK---V QIVHGDILTY RMDRAFPKHI
ADLLVVEKDS RFIPGLKLLS EAAPGR---V RIVHGDILTY RMDRGFPRHI
AELLVVEKDS RFIPGLQMLS DAAPGK---L RIVHGDVLTF KIERAFPEHI
ADLLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTF KVERMFPEYI
ERLTVVEKDH RFLPMLKYVA DVSNDR---M TIVHGDILKY DLSQCFPQIV
AKVVALESDK TFIPHLESLG KNLDGKLRVI HCDFFKLDPR SGGVIPWTAD
ARVVAFESEK TFIPHLEPLQ RNMDGELQVV HCDFFKMDPR YQEVVPWSAG
ARVVAFESEK MFIPHLESLR KNADGELQVV HCDFFKIDPR YQELVPWSAG
SQIILLESMD HFMPKIQELH TLYPERVKVR QGDFVNLWKL VYMDKAFTDD
AKVIALESNR NFLPHLQSLR KKVDGELEVI YCDFFKMDPR NFGIVPWSED
AKVVALESDK TFIPHLESLG KNLDGKLRVI HCDFFKLDPR SGGVIPWTAD
AQVVALESDA NFLPELLELE SRLEGQLDVV HCDFFKLDPI GNGIMPWTAD
ARVIALESDK NFIPELKSLG NSVNGRLEVI YCDFFKLDPR NHGMVPWKKG
VQVVALEGDK VFLSELQALE VQLDGQLEVV NCDFFKLDPI GSGNLSWTDD
ARVVALESNK DFLPSLQLLE NNMDGQLEVV HCDFFKLDPL GHGTMPWIKD
VQLRLYETDA SFEARLNATF NLPKD--ALR IGDFNGLWRL SYLDGKWQEE
IQVVALEGEK FFLPELQDLE IQLDGQLEVV HCDFFKLDPI GSGNLSWTDD
CPRQLLEKR- ---LYKFLNA KFEGSP-LQI LKRDPYDWST YSNLIDLTVA
FQPKVVLESR EVFSKPLQKL CTLSDGRIKW VHQDGYYWQT YEDVYVLFFA
NPKQLMDSR- ---FVKFIQD NFAGTS-MEL YQRDPYEWSS YTDMIELVMA
CPKQLIEKR- ---LYKFLNA KFEESP-LQI LKKDPYDWST YSSLIDLTVA
KPKKLLEDR- ---FVELYQA TLKDHPSMVN YNKNPYKWET FLEMTNLIAA
KPKLLMESK- ---YAKVIEQ HLTLLDNIKL HKEDPYMWES FVSLIDIVMG
QPQQMMECR- ---FLSNLRN IYTDGP-MEL VKKDPYKWES YTELIDYIMA
KPRNLIENKE RCVIINKLVE ETGHNSNFTL YKKDSFIWET YNDLIDLILA
KPRKIIEDSK ENVRIKYLKE KTNNAENFEL YPHNGYNWST FESLIKLIIG
KKVIAVEMDP RFYPALKMLE ESSGGR---M SLIMANMMDV DEAKLLRDKI
VGVLGIEVDE RFNPHLEQIR ---NYTNNKF QWVTADVLKV DELELLKSEV
VGVLGIEQDE RFNGHLEQIR ---QYTSGKF QWTNGDVLRI NELEIVESEV

VSNLPYSISS EITFKLLR-- HKFKLGVLMY QYEFAVRMVS PPGC-KDYSR
VSNVPYSITS PLLFKLIR-- -HRLPAVLTI QREVAERLVA RPGS-EDYGR
VSNIPYQISS PFTFKLLK-- MEFERAVVMY QLEFALRMTA KPGD-RNYSR
VANLPYQISS PITFKLIK-- RGFDLAVLMY QYEFAKRMVA KEGT-KDYGR
VANIPYNISS PITFKLLE-- LDIDVAVLTY QREFAERMVA EPGS-KKYSR
MANLPYSISS PITFRLLD-- IGFEAAILMY QKEFADRMMA HPGT-RDCGR
VANIPYKISS PLTFKLLK-- TDFRLAVVMY QREFAERLCG EDN------R
ISNLPYGASS EIAFRLLP-- EQRPL-LLMF QQEFAERMAA DPAT-DDYGR
IGNIPYNISS PVLFKILDFN --FKSSILMV QKEFARRLVA RPGT-KEYSR
FGNLPYNIST PLMFHLFSYT DAIAD-HFML QKEVVNRLVA GPNS-KAYGR
FGNLPYNIST PLMFHLFSYT DAIAD-HFML QKEVVNRLVA GPNS-KEYGR
```

```
FGNLPYNIST PLMFHLFSYT DAIRD-HFML QKEVVNRLVA GPNS-KTYGR
FGNLPYNIST PLMFHLFEFA EQIET-HFML QKEVVLRLSA SPGC-KAYGR
IGNLPYNISS PLLFHLMSFA PVVID-HFML QNEVVERMVA EPGT-KAFSR
VGNLPYNIST PLLFHLLEFA PRIRD-HFML QKEVVERMVA SPAT-ADYGR
VGNLPYNIST PLIFHLLSHA GLIRD-HFML QKEVVERMAA GPGG-GDWGR
VGNLPYNIST PLLFKLAEVA DDVVD-HFML QKEVVERMVA APKS-NDYGR
VGNLPYNIST PLLIYLLKFI TCIDD-HFML QKEVVERIAA AHGT-KAYGR
VGNLPYNISS PLLFHLMGAA EQVRD-HFML QREVIDRMVA EPGS-GDYSR
VGNLPYNISS PILFHVLQQA AIIAD-HFML QKEVVDRMAA PPGS-KVYGR
IGNLPYNIAT PLLFHVTGFA EHLED-HFLL QKEVVERMAA GAGQ-ASYGR
VGNLPYNIST PLLFHLSRFS NLIID-HFML QLEVVERMVA QPST-PDYGR
LGNIPYSITS PILFHLLEHR RAFRSATLMM QHEVALRLAA KPAT-KEYGI
VANLPYYITT PILFKLLEEK LPVRG-VVMM QKEVGDRLAA KPGT-KEYGS
VANIPYNITG PLLGHLLGSI ARPRRPILLV QKEIGDRLMA SPGS-KAYGA
VANLPYYITT PIVMKFLEED IPVTD-VVMV QKEVADRMNA IPGT-KDYGA
VANIPYNLTG PILEKVLGSP AQPVR-VLMV QKELAERLQA GPGS-KAYGA
VANLPYHIAT PLVTRLLKTG RFAR--VFLV QKEVAERMTA RPKT-PAYGV
ISNTPYQISS PLVFKLINQP RPPRVSILMF QREFALRLLA RPGD-SLYCR
ISNTPYQISS PLVFKLINQP KPPRVSILMF QREFAMRLLA RPGD-SLYCR
ISNTPYQISS PLVFKLINQP RPPRVSILMF QREFAMRLLA RPGD-ALYCR
VSNTPYQISS PLVFKLLQQR PAPRGAILMF QREFALRLVA RPGD-PLYCR
VANLPYQISS PFVFKLLLHR PFFRCAILMF QREFALRLVA KPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAILMF QREFALRLVA KPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAILMF QREFALRLVA KPGD-KLYCR
VANLPYQISS PFVQKLLLHR PLPRYAVLMF QKEFADRLVA RPGD-KDYSR
VANLPYQISS PFVFKLLLHR PFFRCAILMF QREFALRLVA KPGD-KLYCR
IANVPYQISS PLIFKLLLHR PLFRCAVLMF QREFAERLVA KPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAILMF QREFALRLVA KPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAVLMF QREFALRLVA KPGD-KLYCR
VANLPYQISS PIVFKLLLHR PLFRCAVLMF QREFAQRLVA KPGE-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAVLMF QREFAMRLVA PPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAVLMF QREFALRLVA KPGD-KLYCR
VANMPYQISS PFVFKLLLHR PFFRCAVLMF QREFAQRLVA KPGD-KLYCR
VANLPYQVDT RIVFKLLLHR PFFRCAVLMF QREFPMRLVA PPGD-KLYCR
VANLPYQISS PFVQKLLLHR PLPRYAVLMF QKEFADRLVA RPGD-KDYSR
VANLPYQISS PFVFKLLLHR PFFRCAVLMF QREFAMRLVA KPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAVLMF QREFALRLVA KPGD-KLYCR
VANLPYQISS PFVFKLLLHR PFFRCAVLMF QREFAMRLVA KPGD-KLYCR
VANIPYQISS PLVFKLLLHR PMFRCAILMF QREFAERLVA KPGD-KLYCR
VANVPYAISS ALVFKLLK-R PNFKCAVLMF QREFALRVCA QPGS-EAYCR
VANVPYQISS PLTFKLLAHR PIFRTAVLMF QKEFALRLGA KPGD-SLYCR
IGNLPFNIAS PLIIKYLRDM SYRRGVTLTF QLEVAKRLCS PIAC-DTRSR
IGNLPFSVST PLIIKWLENI SCRDGPTLTF QKEVAERLAA NTGS-KQRSR
IGNLPFSVST PLIIKWLENI SLKDGPTLTF QKEVAERLVA TTGS-KQHSR
IGNLPFNVST PLIIKWLEQM SNRTGITLTF QKEVAERLTA STGS-RQRSR
IGNLPFAIST RLLINWLDDL AARRGATLTF QQEVAERICA PVGG-EQRCR
IGNLPFSVST PLIIKWLENI SLKNGPTLTF QKEVAERLVA TTGS-KQRSR
IGNLPFAIST RLLINWLRDM SLRTGATLTF QKEVAERIVA PILS-DQRCR
IGNLPFSVST PLIIKWLENI SCRDGPTLTF QKEVAERLAA NTGS-KQRSR
IGNLPFNVST PLIIKLLHAI SEKRDATLTF QKEVAERLIA QPLD-VQRCR
IGNLPFNIAS PLIIKYLRDM SYRRGVTLTF QLEVAKRLCS PIAC-DTRSR
IGNLPFNVST PLIIKWLENI ANQSGPPHAD IPEGGGREAD SQYG-QQTEE
IGNLPFSVST PLIIKWLENV SCRDGPTLTF QKEVAERLTA NTGS-KQRSR
```

```
IGNLPFSVST PLIIKWLEQV ADRTGPTLTF QQEVAERLTA STKN-KQRSR
IGNLPFSVST PLIIKWLEQL ADRTGPTLTF QKEVAERLTA STSS-KQRSR
IGNLPFSVST PLIIKWLENI ANQSGPTLTF QKEVAERLTA STGS-RQRSR
IGNLPFSVST PLIIKWLENV SQRNGPMLTF QKEVAERLTA TTGS-KQRSR
IGNLPFSVST PLIIKWLENI SKRDGPTLTF QKEVAERLTA GTGN-KQRSR
FGNLPFNVSL PLIFKWFEQI SRKDGMVLTF QREVVERFLA QTGD-KQRCR
IPLKVVGRAL WKLAYDLYSC TSINMFIGEK EFQKLMADPG N-PDLYHVLS
VPIKVFGRIL WKILFDLYSC ESINMFVSEK EFRKLIATPK R-PDLYQVMA
VPIKVFGRLL WKILFDLYSC ESINMFISEK EFRKLIATPK R-PDLYQVLG
INMLVFGPFF KHLINSLIFQ TSLILAMPPP IYIHLTCNNE IGYLIYRSTS
TPLRVVGKIL WKLLYDLYSS TSVNMFITER EYEKLVASPE T-PHLYQVLS
IPLKVIGRAL WKLAYDLYSC TSINMFIGEK EFQKLMADPG N-PDLYHVLS
VPVKIVGNLM WKMIYNLFER RSIIMFISQK EYTKLVTRPR D-YKNYQAFS
FPLKVVGNTL WKILHDLYSC SSVNLFISEK ECRKLTANPQ T-PALYQSLS
IPVKVVGGML LKMVYALFER LSINLFMSEK EYLKLSSQPG D-MMNYRASS
VPFKVFGTFL WKQIYNLFER NSINVFISEK QYTKLVSQPG D-MRNYQALS
VSFRLFSKYL RYLMNSITQQ SELFLVMSPL LFSHIASTKD AGYKLYRGGT
IPVKVVGGML LKMVYALFER LSFNLFISEK EYLKLTSRPG D-MMNYRAFS
NVTGEGS-GL IMQWLSCIGN KNWLLLWMPS TTARKLLARP GM-HSRSKCS
HLPHGYAGLF VSQILDFLSA RDWLLLWLPC SPTVTLLGSR GF-SKRSKTS
NLTGMIG-GL FMQWLSCIGN KNWLLVWVPE ATAHKVLARP GS-LIRAKCS
NVTGEGS-GL IMQWLSCIGN KNWLLLWMPS TTAKKLLARP ST-HARSKCS
NLTNKKG-QL YVQYLQCIAN QNWMLVWIPQ QTARKLFAPF SN-KDRNRLT
NLTDKRG-QL YMQYLQCIAN KNWMLFWVPQ TTAIKLLSPC GF-KSRSRCS
NLTEKKH-GL LMQWMNCVGN RNWLLIWMPT PTASKLLADS GD-HSRHKCS
NWTGNKE-SV LAQWIKCCGH RNWLVIFAPS VSAMKFLGEP GF-KKRRRTG
NLTPQKFGSL LAQWIMCSVY QNWLVCFVPD STAQKFLSGP RF-PKRNKSA
IGNLPFNVGT HLMLKWIRQI APRQGLYLMF QKELSDRICA QVGS-EEYSR
IANLPFEIIT ELLMRYAADC SQHRGLHVFT QREVAERILA PAGS-VQFSR
IANLPFNIIT ELLMRYAVDC SRKQNLHVFT QQEVAECIIA PAGS-IHFSR

LTIDT-CYFA DASIVMKVPK GAFQPAPEVD SAVIKLIPR- -PAPFEVRFL
LTVAV-QCFY DVEILRVLPP YVFDPPPKVY SAVVRLMPKA PCVDNFDEFE
LSLMT-QALA DVEIVMRIGK GAFYPKPKVD SALVLITPK- -KDRIELN-E
LSVAV-QSRA DVEIVAKVPP SAFYPKPKVY SAIVKIKPN- -KGKYHIEFD
LTVMV-NLLA DVELLRGVPR RAFIPPPRVG SSVVRLTPKS EEERPDVDLE
LSIML-QTYA RANRCFDLPP GAFSPPPAVR STVMWIEPRE PLFPIHDRYE
LGVIS-KTYC KAEILEIVKP SSFNPPPKVE SAIVRIVPEP EVFVENREFE
LSVTA-GHYA DVEVVETVPP EAFDPQPRVT SALVRTMPR- -TPDYTVPFM
ITINT-SVRS TIKILFNVTR KVFSPVPDVD SAVISIMKKD VD--IDLAFD
LSVMA-QYYC NVIPVLEVPP SAFTPPPKVD SAVVRLVPHA -TMPHPVKLS
LSVMA-QYYC QVIPVLEVPP SAFTPPPKVD SAVVRLVPHA -TMPYPVKLS
LTVMA-QYYC NVIPVLEVPP TAFTPAPKVD SAVVRLIPHV -QMPHPVGLS
LTVMA-QYFC QVVPVLEVPP HSFTPAPKVD SAVVRLLPYA -EKPFPCKLR
LSVML-QYRY VMDKLIDVPP ESFQPPPKVD SAIVRMIPHA -PHELPAVLG
LSIML-QRRF HMEWLLDVPP TAFDPPPKVE SAVVRLIPKS -TAEVPSVFA
LSIMV-QYHC RVEHLFNVGP GAFNPPPKVD SAIVRLVPHE -VLPFPAKLE
LGVML-QYFF DMEMLIDVPP ESFDPAPKVD SAVVRMIPVK -HRIGKADFA
LSVML-QYHC EVEYLFDVPP EAFEPRPKVD SAIVRLTPHR -VSPFESVLE
LSVML-QARY RMEKLFDVAP EAFDPPPRVV SAVVRMAPLP -ADRLRPAFE
LSVML-QAWC EVTTMFVVPP DAFQPPPKVN SAITRLVPRD -PTTIRIAFS
LSVMI-QYRC RVEPLFDVLP NAFRPVPKVT SSWVRLTPLS RPPRGTWDLA
LSLML-QNRF EMEQMLIVPA EAFNPPPRVQ SAIVCMRPRV -VPVIPFGFG
```

```
LAVQM-QAFC KVEYLFKVSR KVFKPQPKVE SAVIKLTPHA TDPALDADFR
LSIAI-QYYT EVETVMTVPR TVFVPQPNVD SAIIRLLKRP -KPVVEVTFF
LSVRV-QFLA TCEKVCAVPP RAFQPPPKVD SVVVCLRPHR -TLPR-VGLE
LSIAV-QYYC DTEIVAKAPR HMFIPQPNVD STVIGLHVRD -KRKYDVHFF
LSVRV-RYLA ECELICRVPP SAFRPPPQVE SAVVRLTPRP -APTP-ARFS
LTLRV-AHHA VAERLFDLPP GAFFPPPKVW SSLVRLTPTG -AP-----LF
LSANV-QMWA NVTHIMKVGK NNFRPPPQVE SSVVRLEIKN PRPQVDYNWD
LSANV-QMWA NVTHIMKVGK NNFRPPPKVE SSVVRIEIKN PRPQVDFNWD
LSANV-QMWA NVTHIMKVGR NNFRPPPQVE SSVVRIEIKN PRPKVDFNWD
LPANV-QMWA HVKHIMKVGK NNFRPPPLVE SSVVRIEPKN PPPPLAFEWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSVNV-QLLA KVEMLMKVKR TEFRPPPKVD SAVVRIAPKN PPPPVNFVWE
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDMLMKVGK NNFRPPPKVE SSVVRLEPKN PPPPVNFTWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDHIMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDLLMKVGK NNFKPPPKVE SSVVRIEPRN PPPPINFTWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDMLMKVGK NNFKPPPKVE SSVVRIEPRN PPPPINYTWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSVNV-QLLA KVEMLMKVKR TEFRPPPKVD SAVVRIAPKN PPPPVNFVWE
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPVNFQWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDHLMKVGK NNFRPPPKVE SSVVRIEPKN PPPPINFQWD
LSINT-QLLA RVDLLMKVGK NNFRPPPKVE SNVVRIEPRN PPPPINYQWD
LSVNS-QLLA RCSHLMKISK NSFNPPPKVE SSVIRLDPKH PAPSVDFEWD
LSVNT-QLLS KVTHLMKVGK NNFLPPPKVE SAVVRIEPFN PPPPINFVWD
ISIMS-QYVA EPKMVFQISG SCFVPRPQVD VGVVRFVPRK TPLVNTSFLE
LSVMA-QYLC NVRHIFTIPG QAFVPKPEVD VGVVHFTPLI QPKIEQPFVE
LSIMA-QYLC NVEHLFTIPG KAFVPKPKVD VGVVHLTPLI EPKIKQPFVE
LSVMA-QYLT TVKSCFTIPG QAFVPKPNVD VGVVHFTPLA QPQIQQPFVE
LSVMS-QVWT EPVMKFTIPG KAFVPKPQVD VGVVKLIPLK RPKTQLPFVE
LSIMA-QYLC NVEHLFTIPG KAFVPKPEVD VGVVHLMPLV QPKIKQPFVE
LSVMN-QIWS TPELRFMISG RAFVPKPEVD VGVVTIVPLQ TPLTQVHFVE
LSVMA-QYLC NVRHIFTIPG RAFVPKPEVD VGVVHFTPLI QPKIEQPFVE
LSVMA-QAWT HPVLHFIIPG TAFIPKPKVD VGLVTFVPLT IPRTKHEFFE
ISIMA-QYVA EPKLVFQISG SCFVPRPQVD VGVVRFVPRK TPLVKTSFLE
PSLHHGSVPL HGPQLLHHPR KAFVPKP--- ---------- --------QQ
LSIMA-QYLC NVQHILTIPG QAFVPKPEVD VGVVHFTPLV RPRIEQPFVE
LSIMS-QYLC NVKNCFTIPG RAFIPKPKVD VGVVHLTPFV QPKIEQPFVE
LSIMA-QYLC NVKNCFTIPG RAFVPKPEVD VGVVHFTPFI QPKIEQPFVE
LSIMA-QYLC TIHSCFTIPG RAFVPKPEVD VGVVHFTPLV KAQIQQPFVE
LSIMA-QYLC DVQHILTIPG QAFVPKPEVD SGVVHFTPLT RPRIKQPFVE
LSIMA-QYLC KVDNSFIIPG RAFVPKPEVD VGVVHFTPLV QPQINQPFVE
LSVSA-QNFC DIDYKFIIAG GSFVPPPKVE VGVVKIVPKK TFDINLPFID
VIWQLACEIK VLHMEPWSSF DIYTRKKLYL IQMIPRQNLF TKNLTFFHLL
VLWQVACDVK FLHMEPWSSF SVHTENNLYL VRMTPRRTLF TENLSFFHLV
VLWQVACEIK FLHMEPWSSF SVHAENNLYL VRMTPRRTLF TENLSFFHLV
VLFQILFEHK FIAKVPREDF LPQQMAYLYL VKFTPRRNLH ELCQSLWFFI
VLWQVACKIK LLHVEPWSSF GVYGQRNMCL IQLTPHQNLF TENLTFFLML
```

```
VIWQLACEIK VLHMEPWSSF DIYTRKKLYL IQMTPRQNLF TKNLTFFHLL
ALAQMAFDIE LLHEEPLSSF LTTTNN-LCL VRITPREDLF SSHLTLVLMV
VLGQTACGIK VLCTEPSSLF DTYAIK-LCF VQLTPHRNLF TGTLTFFHML
VLWQMACDIE LLHKESWDSF VMSSRP-LCL VRLRPHANLF SAGLTLLMMV
ALFQASCDIQ LLHMEPWSSF LTPLRF-LCL VRITPRRDLF TDSLTFIVMV
IVFQLYFEHE FLGKVPSKHF LPWCTAEWYL VRIVPRRNLF DHLLPFASFV
VLWQMACDIE LLHKESWESF VTSSRR-LCL VRLRPRADLF SAGLTLLMMV
VVR-EAFTDT KLIAISDANE LKG----SQC IEEWDPILFS AAEIWPALVE
VFR-EAFTDS RVLAASESTL QKLCMGYENY QISPNPLLVS PTPITSLTLV
VVT-EAFTDT KLVATSDSST LQK----SSL LEGHDPIIFS TRDTWLSLLE
VVR-EAFTDT KLIAISDANE LKG----SHC IEEWDPVIFS AADIWPALVE
LLS-ELATNT KLVATS-ENS VKK----PDC IEKFDPVIIP SDNKSPSLVE
VIT-EAVTDT RLIATT-PDN LAS----PGV LDKHDPLILP -ENKT-ALLE
LVR-EAFTDT KLVALSQEED MKS----SVC LDKSNPLVIP DGTDYGALVE
LKR-DLYTDS RLIGVVNSEK ----APGARV LVRDQPVLLE PSSAHRSVIE
IKR-ELFTDS KLIGLVESSG DHSIPDGPNL LVKDQPCLIP SRSILPAVVE
LSVMV-QQMC QPSIVYSIPG TAFVPPPKVD ASVVAIEPRI SPLGDEPVFE
LSVLC-QCFF HVRLKQTFVD QTYYPRTEVE GAMLTLEPRS VPLAHGLSLI

LSVLC-QCFF HTQLLRTFRE MTYYPKTAVL GALITLQPRA VPLLPGLDLI

QFVAAVFSQR RKKLRNAILN KEIVSQLPED --FMNKRAED LTPEELASVA
KFSAWLFSAR RKTLRRLKLA ---------- ----TKRVYQ LTLEELVELF
SLVKALFQHR RKVVSKALRE KDILSSVPHS ----NKRVFH LTPEEVK---
DFLRAIFQHR NKSVRKALID EDFLNTNSEI KNLINEKVFK LSVKDIVNLS
SVCRALFQHK NKTVRNALLL REVLEELPED --LLSERPLH LPPERVAELA
DLVRELFTRR RKTVQSTLKA ---------- -EILSSRPEA LYLEDFATIS
KFVTFAFSMR RK------RM ---------- -NLAEKRPEE LGARKFAEIV
DFLKAVFTQR RKTMRNAVRN DAVVEAADEG --LMSARAGK LTPADFATLA
KFIAKIFSMR RKKLSTILN- ---------- ------RPEE LDLSQLITLY
RITTEAFNQR RKTIRNSLGN LFSVEVLTGM GIDPAMRAEN ISVAQYCQMA
RITTEAFNQR RKTIRNSLGN LFSVETLTEM GIDPAMRAEN ISVAQYCQMA
RITTQAFNQR RKTVRNSLGD LFTSEQLIEL GIDPILRAEN ISVAQYCKLA
QLCTTAFNMR RKTLRNNLKH MLSDAEFEQL GIDQSQRPEQ ISVEQYVAMA
EVVTAAFSQR RKMLRNTLGG YRDLVDFDAL GFDLARRAED IGVDEYVRVA
RVVAAAFAQR RKTLRNTLSA LMRPEDFVAL GIDPGLRAEA LHVADYEAIT
RVVREAFNQR RKTLRNTMKG LLDSAAIEAA GVDGSLRPEQ LDLAAFVRLA
KLVKLAFHQR RKTIRNNLKE LAGDDDLQAV GINPQDRAEH IAPEKYVALS
NIVAKAFAMR RKTLTNNLKG IISLSQLNDL GIDGGKRPEQ ISVAEYVQLA
TVVARAFSQR RKMLRRVLDD WAALTPWDEL GIAPTARAEE VGVAQFIGLA
DIVRAAFGQR RKTLRNSLAD ICTPAHFEHA GIRTNARAEQ LEVTEFIALA
EVVARAFGQR RKTLRNALRG MISEQQIKAA GIEPSARAET IDLDHYLRLA
EMVLAAFSQR RKTLRNTLRH YLTIKDFDQL RIDPGLRAEN LSLEQYSAIT
RFVRIAFHQR RKTLLNNLKE ---------- ---LQLRAEA LSIEELLELF
EVVRASFAQR RKTLMNNLSN ELLDRILTEV GIDPKRRGET LSIEEFATLS
TLLKQGFATR RKMLANALKS EQVRQALLQL GRDANSRAEA LSLEDWLALS
KTVKASFGQR RKTLLNSLGG DEIREILKEA NIDEKRRGET LSIEEFSVLS
QLVCQGFSAR RKQLVNALGG QTVAAALAQL RLSPTARAEE LDLPDWLALS
RLVEAAFGKR RKTLLNALAA ARVEEALRAL GLPPRVRAEE LDLEAFRRLR
GLLRIVFVRK NRTISAGFKS KEKIDTVLKE TDLGDKRAGK CDQNDFLRLL
GLLRIVFVRK NRTIAAGFKS KEKIEQVLSE TGLAEKRAGK CDQTDFLKLL
GLLRIVFVRK NRTIAAGFKS KQKIETVLKE TGLSDKRAGK CDQTDFLRLL
GLLRIVFLRK NKTIGACFKT KSLIDGVLQQ CNLQDARASK CGQ-QSLSLL
```

```
GLVRITFVRK NKTLSAAFKS ADKIQQILTS TGFSDKRARS MDIDDFIRLL
GLVRITFVRK NKTLSAAFKS ADKIQQILTS TGFSDKRARS MDIDDFIRLL
GLIRSLFSRL ALNLQFPRLS ADKIQQILTS TGFSDKRARS MDIDDFIRLL
GLLRLCFMRK NKTLMAIFRL KKVIEETLTA SGYGESRARK MRVEDFLALL
GLVRITFVRK NKTLSAAFKS ADKIQQILTN TGFSDKRARS MDIDDFIRLL
GLTRIAFLRK NKTLAATFKV QDKVISILEE QDMAAKRARS MDIDDFMRLL
GLVRITFVRK NKTLSAAFKS ADKIQKILMD TGFSDKRARS MDIDDFIRLL
GLVRIAFVRK NKTLAAAFKS AEKIEGVLTE TGFSEKRARS MDIDDFMALL
GLVRIAFLRK NKTLSGLFRT KTKVCSILEE NNFDSMRART MDIDDFLRLL
GLVRIAFVRK NKTLGAAFKS SKKIEGVLQE ASFSEKRARS MDIDDFMVLL
GLVRITFVRK NKTLSAAFKS ADKIQQILTS TGFSDKRARS MDIDDFIRLL
GLTRIAFLRK NKTLAAAFKQ KEMVEKILEK ADASDKRARS MDIDDFMAVL
GLVRIAFVRK NKTLSAAFKS AKKIEGVLQE ASFCEKRARS MDIDDFMVLL
GLLRLCFMRK NKTLLAIFRL KKLIEETLTA SGYAENRARK MRVEDFLALL
GLVRIAFVRK NKMLSAAFKS APKIESILQE SKFSDKRARS MDIDDFMVLL
GLVRITFVRK NKTLSAAFKS ADKIQQILTS TGFSDKRARS MDIDDFIRLL
GLVRIAFVRK NKTLAASFKS AEKIEGILTE TGFSEKRARS MDIDDFMALL
GLTRIAFVRK NKTLSAAFKQ KDMISDILQK ADAENKRART MDIDDFISLL
GLVKHIFNRK NKKVSSIFRT RQHLESILQE P-IFDKRARV LDQESIMELL
GLVKLCFSRK NKTLSGIFRV KELIIKTLTD NDFLDSRSSK LDINDFLKLL
KVCRQVFHYR QK-LKTLYPE RLGIEQFADL AEGYNEQCIR YPGLFLYDYT
KVVQNVFQFR RK-LRMLFPE QLSISHFKSL CDVYRKMCDE DPQLFAYNFR
KVVQNAFQFR RK-LGMLFPE HLSLMHFKSL CDVYRKMCDE DPQLFTYNFR
KIVKNAFQFR RK-LGMLFPE EISISEFRAL ADAYSKLCNE NQGLFDYDFR
RVVRHIFSMR QK-YGTLLPP ELTVEQCLRL AEVYSEHLVT RPEVAAYDYR
KVVQNVFQFR RK-LGMLFPE HLSLMHFKSL CDVYRKMCDE DPQLFAYNFR
KVVRHIFSMR QK-VANLYPP QLSVAECLRI VEAYDGLVRE RPEIAAYDYR
KVVQNVFQFR RK-LRMLFPE QPSISHFKSL CDVYRKMCDE DPQLFAYNFR
KVTRHIFSFR QK-VETLFPL ELTIEN---- ---------- IDNLISYHYL
KLCRQVFHYR QK-LKTLYPP KLGMEQFADL AEGYNEQCIR HPGLFLYDYT
TVVRALPSTP LS-LQMFPVY ---------- ---------- ----------
KVVQNTFQFR RK-LGMLFPQ QLSVSHFKSL CDVYRKMCDE DPHLFAYNFR
KVVRCIFQFR RK-VSILFPE ELTMTHFKKL CNVYREMCDQ NPHLFSYNYR
KVVRSVFQFR RK-VSILFPE ELTMTHFKKL CNVYREMCDQ NPNLFAYNFR
KVVRNVFQFR RK-IEKLFPE ELTIPQIRAL ADAYAHLCTL EPDLQSYEFR
KVVQNAFQFR RK-LGMLFPE QLTVSHFKSL CDVYRKMCDE DPHLFAYNFR
KVVQSTFQFR RK-IGILFPQ QLSILHFKNL CDVYRKMCDE NPNLFAYNFR
YVVKHTMHRK SK-VKTMFP- ELENLHFRDL CYAYEEIADR IPNLRKVDTY
KHCFGRRSAT VIDHLRSEKV VNMHPQDFKT LFETIERSKD CAYKWLYDET
KHCFGKRNAP IIRHLRSDTA ARMYPHDFKK LFETIEQSED SVFKWIYDYC
KHCFGKRNAP IIHHLRSDTA AKMYPHDFKR LFETIERSED SVFKWIYDYC
KQNYVSRRNR IIPNLEKTQF GDLLPSQILT LFSQFRQWPE YGESSFLASL
KQCFMKRNFK LMDHLQLMKI TDLYPQDFKQ LFEAIEYFR- ESCRWLLDDC
KHCFGRRSAT VIDHLRSEKV VNMHPQDFKT LFETIERSKD CAYKWLYDET
KQCLAKRKGK LIQQINSTLT GDVYPDEYKR LFELMEQSGS FAESWLYQET
RQCFMKRNAK LIDHLPSVRV VDMYPKDFLR LFETIECSKD DTCKWLYDEF
KQCLAKRKVK LIDRLNLILT GHVHPEEYLQ LFQMMDKSQE FTQSWLYEEI
KQCLGKRKAK LVDRLNSIQT GNVSPEQYKQ LFEIMECSEE FNKSWIFDET
TQHYVSRRNR IIPSLEHTEF GELTPAQVLT LFNEFINWSE FHQSPFMQAV
KQCLAKRKVK LIDRLNLILT GHVYPEEYLR LFQLMDKSQE FTQSWLYEEI
MDP-IDFDFD VILKRTPLKK CPIDLTNDEF IYLTKLFMEW PFKPDILMDF
EMCSKPQDKQ LFESIVRIET LLPSFTQCGI NIDMPVGLLS AADFLTISKI
VNP-IDHDID LILKSTPLDK CPKDLTNKEF VYLTSIFNNW PFKPDIYMDF
```

```
MDP-IDFDFD VILKRTPLKK CPIDLTNDEF IYLTKLFMEW PFKPDILMDF
INP-RDHSID LILKSKPVKK KPYELTYLEI DEIAKVFALW PFKPSLLVDF
VLP-LNHNMK LVCKSTPLKK KPMQLTNQEF TKIASLYALW PFKPS-IYDF
FTP-KKHDID LVLYKTPLKK CPQDFTNNDF VYLTNLFHLW PFKPDVYMDF
ISPGKYTASQ IYLTNKKLQK SSYEFTCEDI IKLSEAYENW PFKPS-VADL
IDPKNLLEID FYKSTGKLNK CPRDLLADEI WEVFHVFDKW AFKPA-LTDT
FVCRELFSQR RKKAESLLGD EQREKEKQEL KENFEKQEIE IQNLTKFDLE
HFTNLLMKPG LRAVHKSLSR MDGAMTVLDL SVVEVTRLAC LWQQFVTMSQ
HFTDLLMRPG QRGVYKALQQ TDGALTVLDL TTEEVCKLAT LWHRFLEAXS

N
K
-
N
A
N
V
T
S
N
N
N
N
Q
A
D
N
K
D
N
E
R
S
N
E
N
D
E
Y
Y
Y
H
H
H
H
L
H
L
H
H
H
H
H
```

Q
H
L
H
H
H
H
C
N
N
E
E
E
A
E
A
E
I
N
-
E
E
E
E
E
E
G
L
P
S
E
M
L
L
M
L
L
D
L
V
I
I
V
Y
Y
L
Y
I
D
Q
Q

Supplemental Figure 2 Sequence alignment used to generate tree in Figure 3.

```
Phylip 3.4 Format


    39    309
hsmtTFB2   LDFKRYVTDR RLAAQIYLGK PSRPP-HLLL ECNPGPGILT QALLEAGAK-
ptmtTFB2   LDFKRYVTDR RLAAQIYLGK PSRPP-HLLL ECNPGPGILT QALLEAGAK-
cfmtTFB2   SEPRSYITSP KLAAHILENE RKTPD-KLLL ECNPGPGILT EALLKSKAK-
btmtTFB2   SETKRYVTSP RVAVRVLRGK RK-AG-QLIL ECNPGPGVLT RALLESGAR-
mmmtTFB2   SEPTRHIACK KAAVRDLLEH QNPSR-QIIL ECNPGPGILT GALLKAGAR-
rnmtTFB2   SEPTRHIACK KSAVRDLLEH QNPSH-QLIL ECNPGPGILT GALLKAGAR-
drmtTFB2   KNLRRFIVDP ALATDHLSRD IDDGK-AVIF ECNPGPGVLT RALLNRGAQR
xtmtTFB2   ---------- ---------- ---------- ---PGPGILT QTLLAAGAR-
trmtTFB2   --MRHFIVNP DLATQHLLP- -ENAA-TIIF ECNPGPGVLT RTLLNSGVQK
ggmtTFB2   -PWGRFIACP QLAQRCLQAG -SGPQ-PVVL ECAPGPGVLT RTLLNAGVR-
dmmtTFB2   VPTHMYIANS EAANQYLEPH FQSSGCDTVM ELNSGAGYFT RHLLDRESQF
agmtTFB2   STERFYLANR TTAADVLTQD LPAD--RLLV EVNPGPGLLT EQLLQRNVQN
xlmtTFB1   QLSQNFLLDL KLTDKIVRRA GNLQ-NAYVC EVGPGPGGIT RSILNAGVEE
xtmtTFB1   QLSQNFLLDL KLTDKIVRKA GNLQ-NAYVC EVGPGPGGIT RSILNAGVEE
hsmtTFB1   QLSQNFLLDL RLTDKIVRKA GNLT-NAYVY EVGPGPGGIT RSILNADVAE
ppmtTFB1   QLSQNFLLDL RLTDKIVRKA GNLT-NAYVY EVGPGPGGIT RSILNADVAE
cfmtTFB1   QLSQNFLLDL RLTDKIVRKA GNLT-NAYVY EVGPGPGGIT RSILNANVAE
btmtTFB1   QLSQNFLLDL RLTDKIVRKA GNLT-NAYVY EVGPGPGGIT RSILNAGVAE
mmmtTFB1   QLSQNFLLDL RLTDKIVRKA GSLA-DVYVY EVGPGPGGIT RSILNANVAE
rnmtTFB1   QLSQNFLLDL RLTDKIVRKA GSLA-DVYVY EVGPGPGGIT RSILNADIAE
mdmtTFB1   QLSQNFLLDL RLTDKIVRKT GDLK-NAHVY EVGPGPGAFT RSILNAQVAD
ggmtTFB1   ELSQNFLLDL RLTDKIVRQA GKLK-NAYVC EVGPGPGGIT RSILNAGVEQ
drmtTFB1   QLSQNFLLDT RLTDKIVRQA GNLN-NAHVC EVGPGPGGLT RSILKAGAAD
trmtTFB1   QLSQNFLLDL KLTDKIVRQA GCLK-DAHVC EVGPGPGGLT RSILNAGAAD
dmmtTFB1   QLSQNFLMDE RLTDKIVKSA GRIDPRDLVL EVGPGPGGIT RSILRRHPQR
agmtTFB1   QLSQNFLMDE RLTDKIVRAA GNIR-DHYVL EVGPGPGGIT RSIIRQNPRH
ammtTFB1   ELSQNFILNQ NLADKIIKKT GNLN-DCHVL EIGPGPGALT RSILKCQPKK
cemtTFB1   ILSQNYLMDM NITRKIAKHA KVIE-KDWVI EIGPGPGGIT RAILEAGASR
cbmtTFB1   ILSQNYLMDM NINRKIAKHA KVNE-NDWVI EIGSGPGGIT RGILEAGASR
scmtTFB    NPTVYNKIFD KLD--LTKTY KHP-EELKVL DLYPGVGIQS AIFYNKYCPR
sdmtTFB    NPTVYNRIFD KLD--LTKTY KHP-QELKVL DLYPGVGVQS AIFYNKYCPK
skmtTFB    NSSVHTQIYN KLQ--LQSTY KM--DELKVL DLYPGPSQHS AIFRNIFNPK
cgmtTFB    DPDWHQKVFE KLG--LSS-- --N-PNMKVL DLYPGPGIHS AVLYNKVQPQ
klmtTFB    NPKVINQILD KLN--LESYY KS--ESLQIL DIYAGPLIQS VILNERLKPK
egmtTFB    SPTAIELAYK RLN--LQEHY DM--SKVQVL ELYPGTGLPS YIFHDIYKPK
cdmtTFB    DPEACQKILD KLD--LKSKY DGS--KLDIV DVNPGYGLFS TMLNYELKPR
dhmtTFB    RPQPCQDIID KLN--LKKKY PNSHQNLDII DVFSGYGLFS SMINYELKPR
ecKsgA     -FGQNFLNDQ FVI---IVSA INPQKGQAMV EIGPGLAALT EPVGERLDQ-
spmtTFB    YLVNQNLMDE ALV-NLLKEY NSEK--MTIL EMAPGPGVTT TSLFNYFQPK

           --VVALESDK TFIPHLESLG KNLDGKLRVI HCDFFKLDPR SGGVIKPPAM
           --VVALESDK TFIPHLESLG KNLDGKLRVI HCDFFKLDPR SGGVIKPPAM
           --VIALESNR NFLPHLQSLR KKVDGELEVI YCDFFKMDPR NFGIVKPPIM
           --VIALESDK NFIPELKSLG NSVNGRLEVI YCDFFKLDPR NHGMVTPPVM
           --VVAFESEK TFIPHLEPLQ RNMDGELQVV HCDFFKMDPR YQEVVRPDVS
           --VVAFESEK MFIPHLESLR KNADGELQVV HCDFFKIDPR YQELVRPDVN
           --VVALESDA NFLPELLELE SRLEGQLDVV HCDFFKLDPI GNGIMKPPVM
```

```
--VVALESNK DFLPSLQLLE NNMDGQLEVV HCDFFKLDPL GHGTMQPPVM
--VVALEGDK VFLSELQALE VQLDGQLEVV NCDFFKLDPI GSGNLKPPAM
--VVALESHP AFLSKLQSLE NSLDGQLKVI YGDFFRLDPL VTGAVKPPAV
RRIILLESMD HFMPKIQELH TLYPERVKVR QGDFVNLWKL VYMDKMDGGS
--LRLYETDA SFEARLNATF NLPKD--ALR IGDFNGLWRL SYLDGFDNGQ
--LLVVEKDT RFIPGLKMLN EASGGK---V RTVHGDILTY RMDRAFPKHL
--LLVVEKDT RFIPGLKMLN EASSGK---V QIVHGDILTY RMDRAFPKHL
--LLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTF KVEKAFSESL
--LLVVEKDT RFVPGLQMLS DAAPGK---L RIVHGDVLTF KVEKAFSESL
--LLVVEKDS RFIPGLQMLS DAAPGK---L RIVHGDVLTF KIERAFPEIL
--LLVVEKDS RFIPGLQMLS DAAPGK---L RIVHGDVLTF KIERAFPESL
--LLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTY KIEKAFPGNI
--LLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTY KIEKAFPDNI
--LLVVEKDT RFIPGLQMLS DAAPGK---L RIVHGDVLTF KVERMFPEHL
--LLLVEKDT RFIPGLQIN- ----CR---M ILIN------ -----VCVNI
--LLVVEKDM RFIPGLQLLS EAAPGR---I RIAQGDILAY KLERRFPANI
--LLVVEKDS RFIPGLKLLS EAAPGR---V RIVHGDILTY RMDRGFPRDI
--LLLVEKDP RFGETLQLLK ECASPLNIQF DIHYDDILRF NIEQHIPDTS
--LVVVEKDR RFMPTMEMLA EVAQPF-MRM DIVQGDILDY RVAEAFPDCP
--LIVVEKDK RFEPTLEMLA DAFETINGKM EIIFDDIMKI NMSNLFPSTE
--LDVVEIDN RFIPPLQHLA EAADSR---M FIHHQDALRT EIGDIWKNET
--LDVVEIDN RFIPPLQHLA EAADSR---M FIHHKDALRT EIGDIWKKEP
QYSLLEKR-- -----SSLYK FLNAKFEGSP -LQILKRDPY DWSTYSNLID
QYSLIEKR-- -----SSLYK FLNAKFEESP -LQILKKDPY DWSTYSSLID
QYVLMDSR-- -----PDFVK FIQDNFAGTS -MELYQRDPY EWSSYTDMIE
QYTMMECR-- -----RDFLS NLRNIYTDGP -MELVKKDPY KWESYTELID
KHVLLEDR-- -----LKFVE LYQATLKDHP SMVNYNKNPY KWETFLEMTN
LQVLMESK-- -----PAYAK VIEQHLTLLD NIKLHKEDPY MWESFVSLID
NHILIENKER CVTSLSSIIN KLVEETGHNS NFTLYKKDSF IWETYNDLID
KHIIIEDSKE NVNIWEERIK YLKEKTNNAE NFELYPHNGY NWSTFESLIK
--LTVIELDR DLAARLQTHP FLGPKLT--- -IYQQDAMTF NFGELAEKMG
S-HVVLESRE VFSKPLQKLC TLSDGRIKWV HQDGYYWQTY EDVYVSKVLD

SSRGLFKNLG IEAVPWTAD- -IPLKVVGMF PSRGEKRALW KLAYDLYSCT
SSRGLFKNLG IEAVPWTAD- -IPLKVIGMF PSRGEKRALW KLAYDLYSCT
ISETLFQHLG IAAVPWSED- -TPLRVVGIF PAKNEKKILW KLLYDLYSST
TSDMLFQYLG VKAHPWKKG- -FPLKVVGIL PAKTERNTLW KILHDLYSCS
-SQAIFQNLG IKAVPWSAG- -VPIKVFGIL PYKHERRILW KILFDLYSCE
-SHTIFQNLG IKAVPWSAG- -VPIKVFGIL PNKHERRLLW KILFDLYSCE
YSEKLFSDLA ISEVPWTAD- -VPVKIVGLF TQRNERNLMW KMIYNLFERR
YSSVLFNSLC IPEAPWIKD- -VPFKVFGIL PQKNESTFLW KQIYNLFERN
FTDKLFTDLG ISEASWTDD- -IPVKVVGVL PMSNERGMLL KMVYALFERL
CSDKLFEAMG IAAVPWRAD- -VPVKIFGIF PQRKERNTLW RLLFILYECS
RVADLLS--D VPQKAFTDD- -INMLVFGAV G---SYPFFK HLINSLIFQT
RVCGMLS--G IPHRKWQEE- -VSFRLFSVI G---TVKYLR YLMNSITQQS
IKS------- -----WDDE- PPNVHIIGNL PFSVSTPLII KWLEQVADRT
KKP------- -----WDDD- PPNVHIIGNL PFSVSTPLII KWLEQLADRT
KRP------- -----WEDD- PPNVHIIGNL PFSVSTPLII KWLENISCRD
KRP------- -----WEDD- PPDVHIIGNL PFSVSTPLII KWLENISCRD
KRP------- -----WEDD- PPNVHIIGNL PFSVSTPLII KWLENVSCRD
KRQ------- -----WEDD- PPNVHIIGNL PFSVSTPLII KWLENVSQRN
RRQ------- -----WEDD- PPNVHIIGNL PFSVSTPLII KWLENISLKD
RRQ------- -----WEDD- PPNVHIIGNL PFSVSTPLII KWLENISLKN
```

```
KRR------- -----WEDD- PPNVYIIGNL PFSVSTPLII KWLENISKRD
LRT------- -----WKE-- PPNIHIIGNL PFSVSTPLIV KWLENVSKRD
TKT------- -----WEDD- PPNLHIIGNL PFNVSTPLII KWLEQMSNRT
SKK------- -----WHED- PPNLHIIGNL PFSVSTPLII KWLENIANQS
QR-------- ---------- ---IHLIGNL PFAISTRLLI NWLDDLAARR
PHD------- -----WMDRK RAPVHLIGNL PFAISTRLLI NWLRDMSLRT
IKA------- -----WTEK- CPRIKLIGNL PFNVSTPLII KLLHAISEKR
ARPESVD--- -----WHDSN LPAMHVIGNL PFNIASPLII KYLRDMSYRR
NRPSSVP--- -----WHDSK LPQMHVIGNL PFNIASPLII KYLRDMSYRR
EERIFVP--- -EVQSSDHIN DKFLTVANVT GEGS-EGLIM QWLSCIGNKN
EERIFVP--- -EVQPSDHIN DKFLTVANVT GEGS-EGLIM QWLSCIGNKN
KEKRFVP--- -NRQSRDKIH NQFLVMANLT GMIG-EGLFM QWLSCIGNKN
QEKLFVP--- -KKADYNVVN KDFYIMANLT EKKH-EGLLM QWMNCVGNRN
EDRVLTP--- -SMQKRDHIH NEFLIAANLT NKKG-EQLYV QYLQCIANQN
EKKIMQP--- -EVQTRDHIH DSFIVMGNLT DKRG-EQLYM QYLQCIANKN
KDKLIQP--- -QIKSFDTPH DELLILANWT GNKE-ESVLA QWIKCCGHRN
KDKIIEP--- -SIKPRDEIH DELLIIGNLT PQKFGESLLA QWIMCSVYQN
---------- ---------- -QPLRVFGNL PYNISTPLMF HLFS------
PRIQTEE--- ---EQKLSPH RELLFFAHLP HGYAGLLFVS QILDFLSARD

SIYKFGRIEV NMFIGEKEFQ KLMADPGN-P DLYHVLSVIW QLACEIKVLH
SIYKFGRIEV NMFIGEKEFQ KLMADPGN-P DLYHVLSVIW QLACEIKVLH
SVYSYGRVQL NMFITEREYE KLVASPET-P HLYQVLSVLW QVACKIKLLH
SVYKYGRAEL NLFISEKECR KLTANPQT-P ALYQSLSVLG QTACGIKVLC
SIYRYGRVEL NMFVSEKEFR KLIATPKR-P DLYQVMAVLW QVACDVKFLH
SIYRYGRVEL NMFISEKEFR KLIATPKR-P DLYQVLGVLW QVACEIKFLH
SIFRYGRVEL IMFISQKEYT KLVTRPRD-Y KNYQAFSALA QMAFDIELLH
SIYRYGRIEL NVFISEKQYT KLVSQPGD-M RNYQALSALF QASCDIQLLH
SIYRYGRIEL NLFMSEKEYL KLSSQPGD-M MNYRASSVLW QMACDIELLH
SIYRYGRVEL NIFISEKEYK VLTAKPGE-L RIYQALSVLW QAGCDIQLLH
SLFNLGRCEM ILAMPPPIYI HLTCNNEIGY LIYRSTSVLF QILFEHKFIA
ELFSLGRYEM FLVMSPLLFS HIASTKDAGY KLYRGGTIVF QLYFEHEFLG
GPFTYGRTQM TLTFQQEVAE RLTASTKN-- KQRSRLSIMS QYLCNVKNCF
GPFTYGRTQM TLTFQKEVAE RLTASTSS-- KQRSRLSIMA QYLCNVKNCF
GPFVYGRTQM TLTFQKEVAE RLAANTGS-- KQRSRLSVMA QYLCNVRHIF
GPFVYGRTQM TLTFQKEVAE RLAANTGS-- KQRSRLSVMA QYLCNVRHIF
GPFVYGRTRM TLTFQKEVAE RLTANTGS-- KQRSRLSIMA QYLCNVQHIL
GPFAYGRTRM MLTFQKEVAE RLTATTGS-- KQRSRLSIMA QYLCDVQHIL
GPFVYGRTKM TLTFQKEVAE RLVATTGS-- KQHSRLSIMA QYLCNVEHLF
GPFVYGRTKM TLTFQKEVAE RLVATTGS-- KQRSRLSIMA QYLCNVEHLF
GPFVYGRTQM TLTFQKEVAE RLTAGTGN-- KQRSRLSIMA QYLCKVDNSF
GPFIYGRTQM TLTFQKEVAE RLTANTGG-- KQRSRLSIMS QHLCTVDNCF
GIFMFGRTRL TLTFQKEVAE RLTASTGS-- RQRSRLSVMA QYLTTVKSCF
GPFAYGRTRL TLTFQKEVAE RLTASTGS-- RQRSRLSIMA QYLCTIHSCF
GAFRRIDTCM TLTFQQEVAE RICAPVGG-- EQRCRLSVMS QVWTEPVMKF
GAWSYGRASL TLTFQKEVAE RIVAPILS-- DQRCRLSVMN QIWSTPELRF
DAWTFGKTRM TLTFQKEVAE RLIAQPLD-- VQRCRLSVMA QAWTHPVLHF
GVWQYGRVPL TLTFQLEVAK RLCSPIAC-- DTRSRISIMS QYVAEPKMVF
GVWEYGRVPL TLTFQLEVAK RLCSPIAC-- DTRSRISIMA QYVAEPKLVF
WLYRFGKVKM LLWMPSTTAR KLLARPGM-- HSRSKCSVVR EAFTDTKLIA
WLYRFGKVKM LLWMPSTTAK KLLARPST-- HARSKCSVVR EAFTDTKLIA
WLQRFGRVKM LVWVPEATAH KVLARPGS-- LIRAKCSVVT EAFTDTKLVA
WLFRFGRSPM LIWMPTPTAS KLLADSGD-- HSRHKCSLVR EAFTDTKLVA
```

```
WMQRFGLVKM LVWIPQQTAR KLFAPFSN-- KDRNRLTLLS ELATNTKLVA
WMQRFGLVRM LFWVPQTTAI KLLSPCGF-- KSRSRCSVIT EAVTDTRLIA
WLMKYGKVRM VIFAPSVSAM KFLGEPGF-- KKRRRTGLKR DLYTDSRLIG
WLQKYGRIKM VCFVPDSTAQ KFLSGPRF-- PKRNKSAIKR ELFTDSKLIG
--YTDAIADM HFMLQKEVVN RLVAGPNS-- KAYGRLSVMA QYYCNVIPVL
WLGIFGRVRV LLWLPCSPTV TLLGSRGF-- SKRSKTSVFR EAFTDSRVLA

MEPWSSFDIY TRLYLIQMIP FTKNLTPMNY NIFFHLLKHC FGRRSATVID
MEPWSSFDIY TRLYLIQMTP FTKNLTPMNY NIFFHLLKHC FGRRSATVID
VEPWSSFGVY GQMCLIQLTP FTENLTSINY DVFFLMLKQC FMKRNFKLMD
TEPSSLFDTY AILCFVQLTP FTGTLTPFNY DVFFHMLRQC FMKRNAKLID
MEPWSSFSVH TELYLVRMTP FTENLSPLNY DIFFHLVKHC FGKRNAPIIR
MEPWSSFSVH AELYLVRMTP FTENLSPLNY DMFFHLVKHC FGKRNAPIIH
EEPLSSFLTT TNLCLVRITP FSSHLTPLNG STLVLMVKQC LAKRKGKLIQ
MEPWSSFLTP LRLCLVRITP FTDSLTRENG NTFIVMVKQC LGKRKAKLVD
KESWDSFVMS SRLCLVRLRP FSAGLTSSNA STLLMMVKQC LAKRKVKLID
MEPWSSFLTN MKLCLVRLTP FTGGLKPANS ATFIFMVKQC LTKPRSKLID
KVPREDFLPQ QMLYLVKFTP HELCQSQ-DL PALWFFIKQN YVSRRNRIIP
KVPSKHFLPW CTWYLVRIVP FDHLLPD-NL SLFASFVTQH YVSRRNRIIP
TIPGRAFIPK PKVGVVHLTP FVQPKIEQPF KLVEKVVRCI FQFRRKYCHH
TIPGRAFVPK PEVGVVHFTP FIQPKIEQPF KVVEKVVRSV FQFRRKYCHH
TIPGQAFVPK PEVGVVHFTP LIQPKIEQPF KLVEKVVQNV FQFRRKYCHR
TIPGRAFVPK PEVGVVHFTP LIQPKIEQPF KLVEKVVQNV FQFRRKYCHR
TIPGQAFVPK PEVGVVHFTP LVRPRIEQPF KLVEKVVQNT FQFRRKYCYR
TIPGQAFVPK PESGVVHFTP LTRPRIKQPF KLVEKVVQNA FQFRRKYCHR
TIPGKAFVPK PKVGVVHLTP LIEPKIKQPF KLVEKVVQNA FQFRRKYCHR
TIPGKAFVPK PEVGVVHLMP LVQPKIKQPF KLVEKVVQNV FQFRRKYCHR
IIPGRAFVPK PEVGVVHFTP LVQPQINQPF KLVEKVVQST FQFRRKYCHH
VIPGQAFVPK PDVAVVHFTP LVQPKIQQPF ELVEKVVQSV FQFRRKYCYR
TIPGQAFVPK PNVGVVHFTP LAQPQIQQPF KLVEKIVKNA FQFRRKHCRR
TIPGRAFVPK PEVGVVHFTP LVKAQIQQPF KLVEKVVRNV FQFRRKHCHK
TIPGKAFVPK PQVGVVKLIP LKRPKTQLPF HLVERVVRHI FSMRQKYCRR
MISGRAFVPK PEVGVVTIVP LQTPLTQVHF DTVEKVVRHI FSMRQKYCRR
IIPGTAFIPK PKVGLVTFVP LTIPRTKHEF SIFEKVTRHI FSFRQKYGIR
QISGSCFVPR PQVGVVRFVP RKTPLVNTSF EVLEKVCRQV FHYRQKYVTK
QISGSCFVPR PQVGVVRFVP RKTPLVKTSF EVLEKLCRQV FHYRQKYVEK
ISDANELKGF DSIALVEMDP ---IDFDFDV DNWDYVTRHL MILKRTPLNT
ISDANELKGF DSIALVEMDP ---IDFDFDV DNWDYVTRHL MILKRTPLNT
TSDSSTLQKF SSISLLEVNP ---IDHDIDL DNWDYVTKHL LILKSTPLHT
LSQEEDMKSF NSIALVEFTP ---KKHDIDL DNWEFVTKHL MVLYKTPLID
TSE-NSVKKF LPLSLVEINP ---RDHSIDL DHWDFVTQKL MILKSKPVEE
TTP-DNLASF GPYALLEVLP ---LNHNMKL EYWDYCMQRL LVCKSTPLED
VVNSEK---Y DAYSVIEISP ---KYTASQI ENIEHFLS-A IYLTNKKLVD
LVESSGDHSF DPLAVVEIDP ---NLLEIDF EMMEYILQIL MYKSTGKLSE
EVPPSAFTPP PKSAVVRLVP ATMPHPVKDV RVLSRITTEA FNQRRKTIRN
ASESTLQKLK ENLTLVEMCS -KPQDKQLSI PVFESIVRIL LTCKATSLSK

HLRSLTPLDA RDILMQIGKQ EDEKVVNMHP QDFKTLFETI ERSKDCAYKW
HLRSLTPLDA RDILMQIGKQ EDEKVVNMHP QDFKTLFETI ERSKDCAYKW
HLQLLSPVNA AAILKQIGKD GDMKITDLYP QDFKQLFEAI EYFR-ESCRW
HLPSLSPIDA VHILKQIKKK KDVRVVDMYP KDFLRLFETI ECSKDDTCKW
HLRSLSTVDP INILRQIRKN PGDTAARMYP HDFKKLFETI EQSEDSVFKW
HLRSLSTVDP INILRQIRKR PGDTAAKMYP HDFKRLFETI ERSEDSVFKW
```

```
QINSWSPGMG SELISNLGFL DDTLTGDVYP DEYKRLFELM EQSGSFAESW
RLNSWDPGNG HKLLRQLALP EDIQTGNVSP EQYKQLFEIM ECSEEFNKSW
RLNLWSPDSG SKLLAEMGMQ EDILTGHVHP EEYLQLFQMM DKSQEFTQSW
RLNSWSLDNA DKLLKALEIP EYVETGNVYP EDYKRLFCAL QNSSVFTESW
NLEKWVPGCG PRLIINPKSS ESTQFGDLLP SQILTLFSQF RQWPEYGESS
SLEHWIPHCG ARLILNSNYT RKTEFGELTP AQVLTLFNEF INWSEFHQSP
GVSILFPEEI RIQLTEQMLR LALRPTELTM THFKKLCNVY REMCDQNPHL
GVSILFPEES RLKCTEQMLR LALRPTELTM THFKKLCNVY REMCDQNPNL
GLRMLFPEAQ RLESTGRLLE LALRPRQLSI SHFKSLCDVY RKMCDEDPQL
GLRMLFPEAQ RLESTGRLLE LALRPCQPSI SHFKSLCDVY RKMCDEDPQL
GLGMLFPQIH RVESTGKLLE LALRPSQLSV SHFKSLCDVY RKMCDEDPHL
GLGMLFPEAR RLESTGKLLE LALRPTQLTV SHFKSLCDVY RKMCDEDPHL
GLGMLFPEAQ RLESTGRLLQ LALRPTHLSL MHFKSLCDVY RKMCDEDPQL
GLGMLFPEAQ RLESTGRLLQ LALRPTHLSL MHFKSLCDVY RKMCDEDPQL
GIGILFPQAE RSKNTEKMLM LALRPSQLSI LHFKNLCDVY RKMCDENPNL
GVEILFPERE RLKKTEQLMM AALRPFQLTM AHFRNLCNTY RKMCDEDPSL
GLGMLFPESQ RQELTEELLC SALRPTEISI SEFRALADAY SKLCNENQGL
GIEKLFPEAC RPEMTQEVMQ RALRPTELTI PQIRALADAY AHLCTLEPDL
GYGTLLPPED REEVAEKLFQ RALRPFELTV EQCLRLAEVY SEHLVTRPEV
GVANLYPPAV REELTEQTFK RAARSFQLSV AECLRIVEAY DGLVRERPEI
GVETLFPLEY RTELAQMMYK LSTRPVELTI ENID------ --------NL
GLKTLYPEEL EDELSDDLLK KCTTSIRLGI EQFADLAEGY NEQCIRYPGL
GLKTLYPPEM ENEMADELLK KCTTSIKLGM EQFADLAEGY NEQCIRHPGL
VMDSLGHGGQ QYFNSRITDK --KCPIDLTN DEFIYLTKLF MEWPFKPDIL
VMDSLGHGGQ QYFNSRITDK --KCPIDLTN DEFIYLTKLF MEWPFKPDIL
AIDSLGHGGK QYFSEKVEDK --KCPKDLTN KEFVYLTSIF NNWPFKPDIY
AIDCLGHGAR DYFSANIDKD HIKCPQDFTN NDFVYLTNLF HLWPFKPDVY
MIEILGHGAR DWFISRLDPV --KKPYELTY LEIDEIAKVF ALWPFKPSLL
ILEVLGHGAS DFLKCRIDPE --KKPMQLTN QEFTKIASLY ALWPFKPS-I
ILPTLAPG-A MYMAKDLPEE --KSSYEFTC EDIIKLSEAY ENWPFKPS-V
ALAQLAPGAE IDLAPKLSKE --KCPRDLLA DEIWEVFHVF DKWAFKPA-L
SLGNLFSVEV LTGMGIDP-- --MRAENISV AQYCQMANYL AENAPLQES-
SIYYLGPGAE TLLPSFTQCG INMPVGLLSA ADFLTISKII QKYPFKHHLH

LYDETLEDR
LYDETLEDR
LLDDCMEEM
LYDEFMEDA
IYDYCPEDM
IYDYCSDDS
LYQETLETT
IFDETLEDL
LYEEILEN-
FHVD-----
FLASLENAL
FMQAVDS--
FSYNYREEL
FAYNFREEL
FAYNFREEL
FAYNFREEL
FAYNFREEL
FAYNFREEL
FTYNFREEL
```

```
FAYNFREEL
FAYNFREEL
FAYNYREEL
FDYDFREEL
QSYEFREEL
AAYDYRAPK
AAYDYRAPK
ISYHYLIQK
FLYDYTNKL
FLYDYTNKQ
MDFVDMYQT
MDFVDMYQT
MDFIDVFQE
MDFLDIYQE
VDFYDENED
YDFYDPSDD
ADLYDFDIT
TDTIGIVQE
---------
LGTIIEDS-
```

Appendix F: References

Achanta, G., R. Sasaki, L. Feng, J. S. Carew, W. Lu, H. Pelicano, M. J. Keating, and P. Huang. 2005. Novel role of p53 in maintaining mitochondrial genetic stability through interaction with DNA Pol gamma. *Embo J* 24 (19):3482-92.

Alam, T. I., T. Kanki, T. Muta, K. Ukaji, Y. Abe, H. Nakayama, K. Takio, N. Hamasaki, and D. Kang. 2003. Human mitochondrial DNA is packaged with TFAM. *Nucleic Acids Res* 31 (6):1640-5.

Alberts, Bruce. 2002. *Molecular biology of the cell*. 4th ed. New York: Garland Science.

Albring, M., J. D. Griffith, and G. Attardi. 1977. Association of a protein structure of probable membrane derivation with HeLa cell mitochondrial DNA near its origin of replication. *Proc Natl Acad Sci USA* 74 (4):1348-52.

Aloni, Y., and G. Attardi. 1971. Expression of the mitochondrial genome in HeLa cells. II. Evidence for complete transcription of mitochondrial DNA. *J Mol Biol* 55 (2):251-67.

Aloni, Y., and G. Attardi. 1972. Expression of the mitochondrial genome in HeLa cells. XI. Isolation and characterization of transcription complexes of mitochondrial DNA. *J Mol Biol* 70 (2):363-73.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreier, A. J. Smith, R. Staden, and I. G. Young. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290 (5806):457-65.

Andersson, S. G., and C. G. Kurland. 1999. Origins of mitochondria and hydrogenosomes. *Curr Opin Microbiol* 2 (5):535-41.

Antoshechkin, I., and D. F. Bogenhagen. 1995. Distinct roles for two purified factors in transcription of Xenopus mitochondrial DNA. *Mol Cell Biol* 15 (12):7032-42.

Antoshechkin, I., D. F. Bogenhagen, and I. A. Mastrangelo. 1997. The HMG-box mitochondrial transcription factor xl-mtTFA binds DNA as a tetramer to activate bidirectional transcription. *Embo J* 16 (11):3198-206.

Attardi, B., and G. Attardi. 1967. A membrane-associated RNA of cytoplasmic origin in HeLa cells. *Proc Natl Acad Sci USA* 58 (3):1051-8.

Attardi, G., and B. Attardi. 1968. Mitochondrial origin of membrane-associated heterogeneous RNA in HeLa cells. *Proc Natl Acad Sci USA* 61 (1):261-8.

Baer, R., and D. T. Dubin. 1980. The 3'-terminal sequence of the small subunit ribosomal RNA from hamster mitochondria. *Nucleic Acids Res* 8 (21):4927-41.

Baross, J. A., and S. E. Hoffman. 1985. Submarine hydrothermal vents and associated gradient environments as sites for the origin and …. *Origins of Life and Evolution of Biospheres (Formerly ….*

Belin, A. C., B. F. Björk, M. Westerlund, D. Galter, O. Sydow, C. Lind, K. Pernold, L. Rosvall, A. Håkansson, B. Winblad, H. Nissbrandt, C. Graff, and L. Olson. 2007. Association study of two genetic variants in mitochondrial transcription factor A (TFAM) in Alzheimer's and Parkinson's disease. *Neurosci Lett* 420 (3):257-62.

Bendtsen, J. D., H. Nielsen, G. von Heijne, and S. Brunak. 2004. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340 (4):783-95.

Bilinski, T. 1991. Oxygen toxicity and microbial evolution. *Biosystems*.

Blankenship, R. E. 1992. Origin and early evolution of photosynthesis. *Photosyn Res* 33:91-111.

Bogenhagen, D. F. 1996. Interaction of mtTFB and mtRNA polymerase at core promoters for transcription of Xenopus laevis mtDNA. *J Biol Chem* 271 (20):12036-41.

Boldogh, I. R., D. W. Nowakowski, H. C. Yang, H. Chung, S. Karmon, P. Royes, and L. A. Pon. 2003. A protein complex containing Mdm10p, Mdm12p, and Mmm1p links mitochondrial membranes and DNA to the cytoskeleton-based segregation machinery. *Mol Biol Cell* 14 (11):4618-27.

Bonawitz, N. D., M. Chatenay-Lapointe, Y. Pan, and G. S. Shadel. 2007. Reduced TOR signaling extends chronological life span via increased respiration and upregulation of mitochondrial gene expression. *Cell Metab* 5 (4):265-77.

Bonawitz, N. D., D. A. Clayton, and G. S. Shadel. 2006. Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol Cell* 24 (6):813-25.

Bonawitz, N. D., M. S. Rodeheffer, and G. S. Shadel. 2006. Defective mitochondrial gene expression results in reactive oxygen species-mediated inhibition of respiration and reduction of yeast life span. *Mol Cell Biol* 26 (13):4818-29.

Bonawitz, N. D., and G. S. Shadel. 2007. Rethinking the mitochondrial theory of aging: the role of mitochondrial gene expression in lifespan determination. *Cell Cycle* 6 (13):1574-8.

Boore, J. L. 1999. Animal mitochondrial genomes. *Nucleic Acids Res* 27 (8):1767-80.

Brown, T. A., C. Cecconi, A. N. Tkachuk, C. Bustamante, and D. A. Clayton. 2005. Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes Dev* 19 (20):2466-76.

Bryan, A. C., M. S. Rodeheffer, C. M. Wearn, and G. S. Shadel. 2002. Sls1p is a membrane-bound regulator of transcription-coupled processes involved in Saccharomyces cerevisiae mitochondrial gene expression. *Genetics* 160 (1):75-82.

Butow, R. A., and N. G. Avadhani. 2004. Mitochondrial signaling: the retrograde response. *Mol Cell* 14 (1):1-15.

Bykhovskaya, Y., E. Mengesha, D. Wang, H. Yang, X. Estivill, M. Shohat, and N. Fischel-Ghodsian. 2004. Human mitochondrial transcription factor B1 as a modifier gene for hearing loss associated with the mitochondrial A1555G mutation. *Mol Genet Metab* 82 (1):27-32.

Bykhovskaya, Y., E. Mengesha, D. Wang, H. Yang, X. Estivill, M. Shohat, and N. Fischel-Ghodsian. 2004. Phenotype of non-syndromic deafness associated with the mitochondrial A1555G mutation is modulated by mitochondrial RNA modifying enzymes MTO1 and GTPBP3. *Mol Genet Metab* 83 (3):199-206.

Carnahan, R. H., A. Feoktistova, L. Ren, S. Niessen, J. R. Yates, and K. L. Gould. 2005. Dim1p is required for efficient splicing and export of mRNA encoding lid1p, a component of the fission yeast anaphase-promoting complex. *Eukaryot Cell* 4 (3):577-87.

Carroll, J., I. M. Fearnley, R. J. Shannon, J. Hirst, and J. E. Walker. 2003. Analysis of the subunit composition of complex I from bovine heart mitochondria. *Mol Cell Proteomics* 2 (2):117-26.

Catlett, N. L., and L. S. Weisman. 2000. Divide and multiply: organelle partitioning in yeast. *Curr Opin Cell Biol* 12 (4):509-16.

Cavalier-Smith, T. 2006. Cell evolution and Earth history: stasis and revolution. *Philos Trans R Soc Lond, B, Biol Sci* 361 (1470):969-1006.

Cavalier-Smith, T. 2006. Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. *Proc Biol Sci* 273 (1596):1943-52.

Chan, D. C. 2006. Mitochondrial fusion and fission in mammals. *Annu Rev Cell Dev Biol* 22:79-99.

Chen, D., Z. Yu, Z. Zhu, and C. D. Lopez. 2006. The p53 pathway promotes efficient mitochondrial DNA base excision repair in colorectal cancer cells. *Cancer Res* 66 (7):3485-94.

Chen, Y. S., A. Torroni, L. Excoffier, A. S. Santachiara-Benerecetti, and D. C. Wallace. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am J Hum Genet* 57 (1):133-49.

Choi, Y. S., S. Kim, and Y. K. Pak. 2001. Mitochondrial transcription factor A (mtTFA) and diabetes. *Diabetes Res Clin Pract* 54 Suppl 2:S3-9.

Chomyn, A. 1996. In vivo labeling and analysis of human mitochondrial translation products. In *Methods in Enzymology*, edited by G. M. Attardi and A. Chomyn. San Diego, CA: Academic Press.

Christianson, T. W., and D. A. Clayton. 1986. In vitro transcription of human mitochondrial DNA: accurate termination requires a region of DNA sequence that can function bidirectionally. *Proc Natl Acad Sci USA* 83 (17):6277-81.

Cliften, P. F., S. H. Jang, and J. A. Jaehning. 2000. Identifying a core RNA polymerase surface critical for interactions with a sigma-like specificity factor. *Mol Cell Biol* 20 (18):7013-23.

Cliften, P. F., J. Y. Park, B. P. Davis, S. H. Jang, and J. A. Jaehning. 1997. Identification of three regions essential for interaction between a sigma-like factor and core RNA polymerase. *Genes Dev* 11 (21):2897-909.

Copeland, W. 2007. Inherited Mitochondrial Diseases of DNA Replication. *Annu Rev Med*.

Correa-Rotter, R, CN Mariash, and ME Rosenberg. 1992. Loading and transfer control for Northern hybridization. *Meth Enzymol* 12 (2):154-8.

Coskun, P. E., E. Ruiz-Pesini, and D. C. Wallace. 2003. Control region mtDNA variants: longevity, climatic adaptation, and a forensic conundrum. *Proc Natl Acad Sci USA* 100 (5):2174-6.

Cotney, J., and G. S. Shadel. 2006. Evidence for an early gene duplication event in the evolution of the mitochondrial transcription factor B family and maintenance of rRNA methyltransferase activity in human mtTFB1 and mtTFB2. *J Mol Evol* 63 (5):707-17.

Cotney, J., Z. Wang, and G. S. Shadel. 2007. Relative abundance of the human mitochondrial transcription system and distinct roles for h-mtTFB1 and h-

mtTFB2 in mitochondrial biogenesis and gene expression. *Nucleic Acids Res* 35 (12):4042-54.

Dairaghi, D. J., G. S. Shadel, and D. A. Clayton. 1995. Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. *J Mol Biol* 249 (1):11-28.

Dairaghi, D. J., G. S. Shadel, and D. A. Clayton. 1995. Human mitochondrial transcription factor A and promoter spacing integrity are required for transcription initiation. *Biochim Biophys Acta* 1271 (1):127-34.

Dairaghi, D. J., G. S. Shadel, and D. A. Clayton. 1995. The alternatively spliced form of human mitochondrial transcription factor A displays altered DNA binding properties. In *Progress in Cell Research - Symposium on "Thirty Years of Progress in Mitochondrial Bioenergetics and Molecular Biology"*, edited by F. Palmieri, S. Papa, C. Saccone and M. N. Gadaleta.

de Souza-Pinto, N. C., C. C. Harris, and V. A. Bohr. 2004. p53 functions in the incorporation step in DNA base excision repair in mouse liver mitochondria. *Oncogene* 23 (39):6559-68.

Decatur, W. A., and M. J. Fournier. 2002. rRNA modifications and ribosome function. *Trends Biochem Sci* 27 (7):344-51.

Desai, P. M., and J. P. Rife. 2006. The adenosine dimethyltransferase KsgA recognizes a specific conformational state of the 30S ribosomal subunit. *Arch Biochem Biophys* 449 (1-2):57-63.

Diffley, J. F., and B. Stillman. 1991. A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc Natl Acad Sci USA* 88 (17):7864-8.

Doudican, N. A., B. Song, G. S. Shadel, and P. W. Doetsch. 2005. Oxidative DNA damage causes mitochondrial genomic instability in Saccharomyces cerevisiae. *Mol Cell Biol* 25 (12):5196-204.

Duan, S., P. Hajek, C. Lin, S. K. Shin, G. Attardi, and A. Chomyn. 2003. Mitochondrial outer membrane permeability change and hypersensitivity to digitonin early in staurosporine-induced apoptosis. *J Biol Chem* 278 (2):1346-53.

Dubin, D. T. 1974. Methylated nucleotide content of mitochondrial ribosomal RNA from hamster cells. *J Mol Biol* 84 (2):257-73.

Duhl, DM, DD Gillespie, and F Sulser. 1992. Ethidium bromide fluorescence of 28S ribosomal RNA can be used to normalize samples in northern or dot blots when analyzing small drug-induced changes in specific mRNA. *J Neurosci Methods* 42 (3):211-8.

Eaton, J. S., Z. P. Lin, A. C. Sartorelli, N. D. Bonawitz, and G. S. Shadel. 2007. Ataxia-telangiectasia mutated kinase regulates ribonucleotide reductase and mitochondrial homeostasis. *J Clin Invest* 117 (9):2723-2734.

Ekstrand, M. I., M. Falkenberg, A. Rantanen, C. B. Park, M. Gaspari, K. Hultenby, P. Rustin, C. M. Gustafsson, and N. G. Larsson. 2004. Mitochondrial transcription factor A regulates mtDNA copy number in mammals. *Hum Mol Genet* 13 (9):935-44.

Elder, G. H. 1993. Molecular genetics of disorders of haem biosynthesis. *J Clin Pathol* 46 (11):977-81.

Eykholt, RL, MD Mitchell, and KW Marvin. 2000. Direct imaging of northern blots on an optical scanner using ethidium bromide. *Meth Enzymol* 28 (5):864-6, 868, 870.

Falkenberg, M., M. Gaspari, A. Rantanen, A. Trifunovic, N. G. Larsson, and C. M. Gustafsson. 2002. Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat Genet* 31 (3):289-94.

Fatica, A., and D. Tollervey. 2002. Making ribosomes. *Curr Opin Cell Biol* 14 (3):313-8.

PHYLIP (Phylogeny Inference Package) version 3.65. Department of Genome Sciences, University of Washington, Seattle, WA.

Fischel-Ghodsian, N. 1999. Genetic factors in aminoglycoside toxicity. *Ann N Y Acad Sci* 884:99-109.

Fisher, R. P., and D. A. Clayton. 1985. A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. *J Biol Chem* 260 (20):11330-8.

Fisher, R. P., and D. A. Clayton. 1988. Purification and characterization of human mitochondrial transcription factor 1. *Mol Cell Biol* 8 (8):3496-509.

Fisher, R. P., T. Lisowsky, G. A. Breen, and D. A. Clayton. 1991. A rapid, efficient method for purifying DNA-binding proteins. Denaturation-renaturation chromatography of human and yeast mitochondrial extracts. *J Biol Chem* 266 (14):9153-60.

Fisher, R. P., T. Lisowsky, M. A. Parisi, and D. A. Clayton. 1992. DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J Biol Chem* 267 (5):3358-67.

Franco, M., M. Johansson, and A. Karlsson. 2007. Depletion of mitochondrial DNA by down-regulation of deoxyguanosine kinase expression in non-proliferating HeLa cells. *Exp Cell Res* 313 (12):2687-94.

Garstka, H. L., W. E. Schmitt, J. Schultz, B. Sogl, B. Silakowski, A. Perez-Martos, J. Montoya, and R. J. Wiesner. 2003. Import of mitochondrial transcription factor A (TFAM) into rat liver mitochondria stimulates transcription of mitochondrial DNA. *Nucleic Acids Res* 31 (17):5039-47.

Gaspari, M., M. Falkenberg, N. G. Larsson, and C. M. Gustafsson. 2004. The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. *Embo J* 23 (23):4606-14.

Giles, R. E., H. Blanc, H. M. Cann, and D. C. Wallace. 1980. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA* 77 (11):6715-9.

Gleyzer, N., K. Vercauteren, and R. C. Scarpulla. 2005. Control of mitochondrial transcription specificity factors (TFB1M and TFB2M) by nuclear respiratory factors (NRF-1 and NRF-2) and PGC-1 family coactivators. *Mol Cell Biol* 25 (4):1354-66.

Gray, M. W. 1989. The evolutionary origins of organelles. *Trends Genet* 5 (9):294-9.

Gray, M. W. 1999. Evolution of organellar genomes. *Curr Opin Genet Dev* 9 (6):678-87.

Gray, M. W., G. Burger, and B. F. Lang. 1999. Mitochondrial evolution. *Science* 283 (5407):1476-81.

Gray, M. W., B. F. Lang, R. J. Cedergren, G. B. Golding, C. Lemieux, D. Sankoff, M. Turmel, N. Brossard, E. Delage, T. G. Littlejohn, I. Plante, P. Rioux, D. Saint-Louis, Y. Zhu, and G. Burger. 1998. Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res* 26 (4):865-78.

Greenleaf, A. L., J. L. Kelly, and I. R. Lehman. 1986. Yeast RPO41 gene product is required for transcription and maintenance of the mitochondrial genome. *Proc Natl Acad Sci USA* 83 (10):3391-4.

Guan, M. X., N. Fischel-Ghodsian, and G. Attardi. 2000. A biochemical basis for the inherited susceptibility to aminoglycoside ototoxicity. *Hum Mol Genet* 9 (12):1787-93.

Guan, M. X., N. Fischel-Ghodsian, and G. Attardi. 2001. Nuclear background determines biochemical phenotype in the deafness-associated mitochondrial 12S rRNA mutation. *Hum Mol Genet* 10 (6):573-80.

BioEdit version 7.05. Distributed by author., Carlsbad, CA.

Harada, T., M. Iwamori, Y. Nagai, and Y. Nomura. 1986. Ototoxicity of neomycin and its penetration through the round window membrane into the perilymph. *Ann Otol Rhinol Laryngol* 95 (4 Pt 1):404-8.

Helser, T. L., J. E. Davies, and J. E. Dahlberg. 1971. Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in Escherichia coli. *Nature New Biol* 233 (35):12-4.

Helser, T. L., J. E. Davies, and J. E. Dahlberg. 1972. Mechanism of kasugamycin resistance in Escherichia coli. *Nat New Biol* 235 (53):6-9.

Holm, N. G., and E. M. Andersson. 1995. Abiotic synthesis of organic compounds under the conditions of submarine hydrothermal systems: a perspective. *Planetary and space science* 43 (1-2):153-9.

Hoppins, S., L. Lackner, and J. Nunnari. 2007. The machines that divide and fuse mitochondria. *Annu Rev Biochem* 76:751-80.

Huntington, K. M., T. Yi, Y. Wei, and D. Pei. 2000. Synthesis and antibacterial activity of peptide deformylase inhibitors. *Biochemistry* 39 (15):4543-51.

Hutson, S. M., A. J. Sweatt, and K. F. Lanoue. 2005. Branched-chain amino acid metabolism: implications for establishing safe intakes. *J Nutr* 135 (6 Suppl):1557S-64S.

Igarashi, K., K. Kishida, K. Kashiwagi, I. Tatokoro, T. Kakegawa, and S. Hirose. 1981. Relationship between methylation of adenine near the 3' end of 16-S ribosomal RNA and the activity of 30-S ribosomal subunits. *Eur J Biochem* 113 (3):587-93.

Ikeda, R. A., and C. C. Richardson. 1987. Enzymatic properties of a proteolytically nicked RNA polymerase of bacteriophage T7. *J Biol Chem* 262 (8):3790-9.

Inoue, K., S. Basu, and M. Inouye. 2007. Dissection of 16S rRNA methyltransferase (KsgA) function in Escherichia coli. *J Bacteriol*.

Jang, S. H., and J. A. Jaehning. 1991. The yeast mitochondrial RNA polymerase specificity factor, MTF1, is similar to bacterial sigma factors. *J Biol Chem* 266 (33):22671-7.

Kaguni, L. S. 2004. DNA polymerase gamma, the mitochondrial replicase. *Annu Rev Biochem* 73:293-320.

Kang, D., S. H. Kim, and N. Hamasaki. 2007. Mitochondrial transcription factor A (TFAM): roles in maintenance of mtDNA and cellular functions. *Mitochondrion* 7 (1-2):39-44.

Kanki, T., H. Nakayama, N. Sasaki, K. Takio, T. I. Alam, N. Hamasaki, and D. Kang. 2004. Mitochondrial nucleoid and transcription factor A. *Ann N Y Acad Sci* 1011:61-8.

Kanki, T., K. Ohgaki, M. Gaspari, C. M. Gustafsson, A. Fukuoh, N. Sasaki, N. Hamasaki, and D. Kang. 2004. Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA. *Mol Cell Biol* 24 (22):9823-34.

Kasting, J. F. 1993. Earth's early atmosphere. *Science* 259 (5097):920-6.

Kelly, J. L., and I. R. Lehman. 1986. Yeast mitochondrial RNA polymerase. Purification and properties of the catalytic subunit. *J Biol Chem* 261 (22):10340-7.

Klootwijk, J., I. Klein, and L. A. Grivell. 1975. Minimal post-transcriptional modification of yeast mitochondrial ribosomal RNA. *J Mol Biol* 97 (3):337-50.

Kooijman, S. A., and T. A. Troost. 2007. Quantitative steps in the evolution of metabolic organization as specified by the Dynamic Energy Budget theory. *Biological reviews of the Cambridge Philosophical Society* 82 (1):113-42.

Kravchenko, I. u. E., and P. M. Chumakov. 2005. [Alternative transcripts from POLRMT responsible for synthesis of nuclear RNA polymerase IV]. *Mol Biol (Mosk)* 39 (1):67-71.

Kravchenko, I. u. E., I. B. Rogozin, E. V. Koonin, and P. M. Chumakov. 2005. Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. *Nature* 436 (7051):735-9.

Kruse, B., N. Narasimhan, and G. Attardi. 1989. Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. *Cell* 58 (2):391-7.

Lafontaine, D., J. Delcour, A. L. Glasser, J. Desgrès, and J. Vandenhaute. 1994. The DIM1 gene responsible for the conserved m6(2)Am6(2)A dimethylation in the 3'-terminal loop of 18 S rRNA is essential in yeast. *J Mol Biol* 241 (3):492-7.

Lafontaine, D. L., J. Vandenhaute, and D. Tollervey. 1995. The 18S rRNA dimethylase Dim1p is required for pre-ribosomal RNA processing in yeast. *Genes Dev* 9 (20):2470-81.

Lafontaine, DL, T Preiss, and D Tollervey. 1998. Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? *Mol Cell Biol* 18 (4):2360-70.

Lang, B. F., G. Burger, C. J. O'Kelly, R. J. Cedergren, G. B. Golding, C. Lemieux, D. Sankoff, M. Turmel, and M. W. Gray. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387 (6632):493-7.

Larsson, N. G., J. D. Garman, A. Oldfors, G. S. Barsh, and D. A. Clayton. 1996. A single mouse gene encodes the mitochondrial transcription factor A and a testis-specific nuclear HMG-box protein. *Nat Genet* 13 (3):296-302.

Larsson, N. G., A. Oldfors, J. D. Garman, G. S. Barsh, and D. A. Clayton. 1997. Down-regulation of mitochondrial transcription factor A during spermatogenesis in humans. *Hum Mol Genet* 6 (2):185-91.

Larsson, N. G., A. Oldfors, E. Holme, and D. A. Clayton. 1994. Low levels of mitochondrial transcription factor A in mitochondrial DNA depletion. *Biochem Biophys Res Commun* 200 (3):1374-81.

Larsson, N. G., J. Wang, H. Wilhelmsson, A. Oldfors, P. Rustin, M. Lewandoski, G. S. Barsh, and D. A. Clayton. 1998. Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet* 18 (3):231-6.

Lee, S., S. Kim, X. Sun, J. H. Lee, and H. Cho. 2007. Cell cycle-dependent mitochondrial biogenesis and dynamics in mammalian cells. *Biochem Biophys Res Commun* 357 (1):111-7.

Legros, F., F. Malka, P. Frachon, A. Lombès, and M. Rojo. 2004. Organization and dynamics of human mitochondrial DNA. *J Cell Sci* 117 (Pt 13):2653-62.

Lesage, G., J. Shapiro, C. A. Specht, A. M. Sdicu, P. Ménard, S. Hussein, A. H. Tong, C. Boone, and H. Bussey. 2005. An interactional network of genes involved in chitin synthesis in Saccharomyces cerevisiae. *BMC Genet* 6 (1):8.

Li, F., Y. Wang, K. I. Zeller, J. J. Potter, D. R. Wonsey, K. A. O'Donnell, J. W. Kim, J. T. Yustein, L. A. Lee, and C. V. Dang. 2005. Myc stimulates nuclearly encoded mitochondrial genes and mitochondrial biogenesis. *Mol Cell Biol* 25 (14):6225-34.

Lightowlers, R. N., P. F. Chinnery, D. M. Turnbull, and N. Howell. 1997. Mammalian mitochondrial genetics: heredity, heteroplasmy and disease. *Trends Genet* 13 (11):450-5.

Lill, R., K. Diekert, A. Kaut, H. Lange, W. Pelzer, C. Prohl, and G. Kispal. 1999. The essential role of mitochondria in the biogenesis of cellular iron-sulfur proteins. *Biol Chem* 380 (10):1157-66.

Lisowsky, T., and G. Michaelis. 1989. Mutations in the genes for mitochondrial RNA polymerase and a second mitochondrial transcription factor of Saccharomyces cerevisiae. *Mol Gen Genet* 219 (1-2):125-8.

Little, R., N. P. Fiil, and P. P. Dennis. 1981. Transcriptional and post-transcriptional control of ribosomal protein and ribonucleic acid polymerase genes. *J Bacteriol* 147 (1):25-35.

Liu, Z., and R. A. Butow. 2006. Mitochondrial retrograde signaling. *Annu Rev Genet* 40:159-85.

Lloyd, D. 2006. Hydrogen sulfide: clandestine microbial messenger? *Trends in Microbiology*.

Mangus, D. A., S. H. Jang, and J. A. Jaehning. 1994. Release of the yeast mitochondrial RNA polymerase specificity factor from transcription complexes. *J Biol Chem* 269 (42):26568-74.

Margulis, L. 1971. Origin of Eukaryotic Cells.

Margulis, L., and M. J. Chapman. 1998. Endosymbioses: cyclical and permanent in evolution. *Trends Microbiol* 6 (9):342-5; discussion 345-6.

Margulis, L., M. J. Chapman, R. Guerrero, and J. Hall. 2006. The last eukaryotic common ancestor (LECA): acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic Eon. *Proc Natl Acad Sci USA* 103 (35):13080-5.

Margulis, L., and M. McMenamin. 1990. Marriage of convenience: the motility of the modern cell may reflect an ancient symbiotic union. *The Sciences* 30 (5):30-7.

Margulis, Lynn. 1970. *Origin of eukaryotic cells; evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth*. New Haven: Yale University Press.

Martin, M., J. Cho, A. J. Cesare, J. D. Griffith, and G. Attardi. 2005. Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. *Cell* 123 (7):1227-40.

Martin, W., and M. Müller. 1998. The hydrogen hypothesis for the first eukaryote. *Nature* 392 (6671):37-41.

Masters, B. S., L. L. Stohl, and D. A. Clayton. 1987. Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell* 51 (1):89-99.

Matsunaga, M., and J. A. Jaehning. 2004. Intrinsic promoter recognition by a "core" RNA polymerase. *J Biol Chem* 279 (43):44239-42.

Matsunaga, T., H. Kumanomido, M. Shiroma, A. Ohtsuka, K. Asamura, and S. Usami. 2004. Deafness due to A1555G mitochondrial mutation without use of aminoglycoside. *Laryngoscope* 114 (6):1085-91.

Matsushima, Y., C. Adan, R. Garesse, and L. S. Kaguni. 2005. Drosophila mitochondrial transcription factor B1 modulates mitochondrial translation but not transcription or DNA copy number in Schneider cells. *J Biol Chem* 280 (17):16815-20.

Matsushima, Y., R. Garesse, and L. S. Kaguni. 2004. Drosophila mitochondrial transcription factor B2 regulates mitochondrial DNA copy number and transcription in schneider cells. *J Biol Chem* 279 (26):26900-5.

McCulloch, V, and GS Shadel. 2003. Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity. *Mol Cell Biol* 23 (16):5816-24.

McCulloch, V., B. L. Seidel-Rogol, and G. S. Shadel. 2002. A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. *Mol Cell Biol* 22 (4):1116-25.

McCulloch, V., and G. S. Shadel. 2003. Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity. *Mol Cell Biol* 23 (16):5816-24.

Mears, J. A., J. J. Cannone, S. M. Stagg, R. R. Gutell, R. K. Agrawal, and S. C. Harvey. 2002. Modeling a minimal ribosome based on comparative sequence analysis. *J Mol Biol* 321 (2):215-34.

Meeusen, S., and J. Nunnari. 2003. Evidence for a two membrane-spanning autonomous mitochondrial DNA replisome. *J Cell Biol* 163 (3):503-10.

Mereschkowsky, C. 1905. Uber Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol. centralb.* 25:593-604.

Michel, H., J. Behr, A. Harrenga, and A. Kannt. 1998. Cytochrome c oxidase: structure and spectroscopy. *Annual review of biophysics and biomolecular structure* 27:329-56.

Micura, R., W. Pils, C. Höbartner, K. Grubmayr, M. O. Ebert, and B. Jaun. 2001. Methylation of the nucleobases in RNA oligonucleotides mediates duplex-hairpin conversion. *Nucleic Acids Res* 29 (19):3997-4005.

Miller, M. L., T. J. Antes, F. Qian, and D. L. Miller. 2006. Identification of a putative mitochondrial RNA polymerase from Physarum polycephalum: characterization, expression, purification, and transcription in vitro. *Curr Genet* 49 (4):259-71.

Moreira, D., and P. Lopez-Garcia. 1998. Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis *J Mol Evol* 47 (5):517-30.

Morrish, F., C. Giedt, and D. Hockenbery. 2003. c-MYC apoptotic function is mediated by NRF-1 target genes. *Genes Dev* 17 (2):240-55.

Nam, S. C., and C. Kang. 2005. DNA light-strand preferential recognition of human mitochondria transcription termination factor mTERF. *J Biochem Mol Biol* 38 (6):690-4.

Nielsen, H., and A. Krogh. 1998. Prediction of signal peptides and signal anchors by a hidden Markov model. Paper read at Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology (ISMB 6), at Manlo Park, CA.

Nierhaus, K. H. 1991. The assembly of prokaryotic ribosomes. *Biochimie* 73 (6):739-55.

Norrbom, J., C. J. Sundberg, H. Ameln, W. E. Kraus, E. Jansson, and T. Gustafsson. 2004. PGC-1alpha mRNA expression is influenced by metabolic perturbation in exercising human skeletal muscle. *J Appl Physiol* 96 (1):189-94.

O'Brien, T. W. 2003. Properties of human mitochondrial ribosomes. *IUBMB Life* 55 (9):505-13.

O'Farrell, H. C., N. Pulicherla, P. M. Desai, and J. P. Rife. 2006. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. *Rna* 12 (5):725-33.

O'Farrell, H. C., J. N. Scarsdale, and J. P. Rife. 2004. Crystal structure of KsgA, a universally conserved rRNA adenine dimethyltransferase in Escherichia coli. *J Mol Biol* 339 (2):337-53.

Ohgaki, K., T. Kanki, A. Fukuoh, H. Kurisaki, Y. Aoki, M. Ikeuchi, S. H. Kim, N. Hamasaki, and D. Kang. 2007. The C-terminal tail of mitochondrial transcription factor a markedly strengthens its general binding to DNA. *J Biochem* 141 (2):201-11.

Organelle genomes database. NCBI, 11-2-2007 [cited 11-6-2007. Available from http://www.ncbi.nlm.nih.gov/genomes/static/euk_o.html.

Osteryoung, K. W., and J. Nunnari. 2003. The division of endosymbiotic organelles. *Science* 302 (5651):1698-704.

Osyczka, A., C. C. Moser, and P. L. Dutton. 2005. Fixing the Q cycle. *Trends Biochem Sci* 30 (4):176-82.

Page, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* 12 (4):357-8.

Parisi, M. A., and D. A. Clayton. 1991. Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science* 252 (5008):965-9.

Parisi, M. A., B. Xu, and D. A. Clayton. 1993. A human mitochondrial transcriptional activator can functionally replace a yeast mitochondrial HMG-box protein both in vivo and in vitro. *Mol Cell Biol* 13 (3):1951-61.

Pica-Mattoccia, L., and G. Attardi. 1971. Expression of the mitochondrial genome in HeLa cells. V. Transcription of mitochondrial DNA in relationship to the cell cycle. *J Mol Biol* 57 (3):615-21.

Pica-Mattoccia, L., and G. Attardi. 1972. Expression of the mitochondrial genome in HeLa cells. IX. Replication of mitochondrial DNA in relationship to cell cycle in HeLa cells. *J Mol Biol* 64 (2):465-84.

Pohjoismäki, J. L., S. Wanrooij, A. K. Hyvärinen, S. Goffart, I. J. Holt, J. N. Spelbrink, and H. T. Jacobs. 2006. Alterations to the expression level of mitochondrial transcription factor A, TFAM, modify the mode of mitochondrial DNA replication in cultured human cells. *Nucleic Acids Res* 34 (20):5815-28.

Poldermans, B., H. D. Bakker, and P. H. van Knippenberg. 1980. Studies on the function of two adjacent N6,N6-dimethyladenosines near the 3' end of 16S ribosomal RNA of Escherichia coli. IV. The effect of the methylgroups on ribosomal subunit interaction. *Nucleic Acids Res* 8 (1):143-51.

Poldermans, B., C. P. Van Buul, and P. H. Van Knippenberg. 1979. Studies on the function of two adjacent N6,N6-dimethyladenosines near the 3' end of 16 S ribosomal RNA of Escherichia coli. II. The effect of the absence of the methyl groups on initiation of protein biosynthesis. *J Biol Chem* 254 (18):9090-3.

Prieto-Martín, A., J. Montoya, and F. Martínez-Azorín. 2004. New DNA-binding activity of rat mitochondrial transcription termination factor (mTERF). *J Biochem* 136 (6):825-30.

Psarra, A. M., S. Solakidi, and C. E. Sekeris. 2006. The mitochondrion as a primary site of action of steroid and thyroid hormones: presence and action of steroid and thyroid hormone receptors in mitochondria of animal cells. *Mol Cell Endocrinol* 246 (1-2):21-33.

Rabinowitz, M., J. Sinclair, L. DeSalle, R. Haselkorn, and H. H. Swift. 1965. Isolation of deoxyribonucleic acid from mitochondria of chick embryo heart and liver. *Proc Natl Acad Sci USA* 53 (5):1126-33.

Rantanen, A., M. Gaspari, M. Falkenberg, C. M. Gustafsson, and N. G. Larsson. 2003. Characterization of the mouse genes for mitochondrial transcription factors B1 and B2. *Mamm Genome* 14 (1):1-6.

Rodeheffer, M. S., B. E. Boone, A. C. Bryan, and G. S. Shadel. 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 (11):8616-22.

Rodeheffer, M. S., and G. S. Shadel. 2003. Multiple interactions involving the amino-terminal domain of yeast mtRNA polymerase determine the efficiency of mitochondrial protein synthesis. *J Biol Chem* 278 (20):18695-701.

Ruiz-Pesini, E., M. T. Lott, V. Procaccio, J. C. Poole, M. C. Brandon, D. Mishmar, C. Yi, J. Kreuziger, P. Baldi, and D. C. Wallace. 2007. An enhanced MITOMAP with a

global mtDNA mutational phylogeny. *Nucleic Acids Res* 35 (Database issue):D823-8.

Scarpulla, R. C. 2006. Nuclear control of respiratory gene expression in mammalian cells. *J Cell Biochem* 97 (4):673-83.

Schinkel, A. H., M. J. Koerkamp, E. P. Touw, and H. F. Tabak. 1987. Specificity factor of yeast mitochondrial RNA polymerase. Purification and interaction with core RNA polymerase. *J Biol Chem* 262 (26):12785-91.

Schluenzen, F., C. Takemoto, D. N. Wilson, T. Kaminishi, J. M. Harms, K. Hanawa-Suetsugu, W. Szaflarski, M. Kawazoe, M. Shirouzu, M. Shirouzo, K. H. Nierhaus, S. Yokoyama, and P. Fucini. 2006. The antibiotic kasugamycin mimics mRNA nucleotides to destabilize tRNA binding and inhibit canonical translation initiation. *Nat Struct Mol Biol* 13 (10):871-8.

Schmelzle, T., and M. N. Hall. 2000. TOR, a central controller of cell growth. *Cell* 103 (2):253-62.

Schubot, F. D., C. J. Chen, J. P. Rose, T. A. Dailey, H. A. Dailey, and B. C. Wang. 2001. Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. *Protein Sci* 10 (10):1980-8.

Schubot, FD, CJ Chen, JP Rose, TA Dailey, HA Dailey, and BC Wang. 2001. Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. *Protein Sci* 10 (10):1980-8.

Schulz, H. 1991. Beta-oxidation of fatty acids. *Biochim Biophys Acta* 1081 (2):109-20.

Schulz, T. J., K. Zarse, A. Voigt, N. Urban, M. Birringer, and M. Ristow. 2007. Glucose Restriction Extends Caenorhabditis elegans Life Span by Inducing Mitochondrial Respiration and Increasing Oxidative Stress. *Cell Metab* 6 (4):280-93.

Schuwirth, B. S., J. M. Day, C. W. Hau, G. R. Janssen, A. E. Dahlberg, J. H. Cate, and A. Vila-Sanjurjo. 2006. Structural analysis of kasugamycin inhibition of translation. *Nat Struct Mol Biol* 13 (10):879-86.

Schwerzmann, K., and P. L. Pedersen. 1986. Regulation of the mitochondrial ATP synthase/ATPase complex. *Arch Biochem Biophys* 250 (1):1-18.

Seidel-Rogol, B. L., V. McCulloch, and G. S. Shadel. 2003. Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. *Nat Genet* 33 (1):23-4.

Seidel-Rogol, B. L., and G. S. Shadel. 2002. Modulation of mitochondrial transcription in response to mtDNA depletion and repletion in HeLa cells. *Nucleic Acids Res* 30 (9):1929-34.

Shadel, G. S., and D. A. Clayton. 1995. A Saccharomyces cerevisiae mitochondrial transcription factor, sc-mtTFB, shares features with sigma factors but is functionally distinct. *Mol Cell Biol* 15 (4):2101-8.

Shang, J., and D. A. Clayton. 1994. Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro. *J Biol Chem* 269 (46):29112-20.

Sharma, M. R., E. C. Koc, P. P. Datta, T. M. Booth, L. L. Spremulli, and R. K. Agrawal. 2003. Structure of the mammalian mitochondrial ribosome reveals an expanded functional role for its component proteins. *Cell* 115 (1):97-108.

Shen, E. L., and D. F. Bogenhagen. 2001. Developmentally-regulated packaging of mitochondrial DNA by the HMG-box protein mtTFA during Xenopus oogenesis. *Nucleic Acids Res* 29 (13):2822-8.

Shutt, T. E., and M. W. Gray. 2006. Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet* 22 (2):90-5.

Shutt, T. E., and M. W. Gray. 2006. Homologs of Mitochondrial Transcription Factor B, Sparsely Distributed within the Eukaryotic Radiation, are Likely Derived from the Dimethlyadenosine Methyltransferase of the Mitochondrial Endosymbiont. *Mol Biol Evol* 23 (6):1169-79.

Siciliano, G., M. Mancuso, L. Pasquali, M. L. Manca, A. Tessa, and A. Iudice. 2000. Abnormal levels of human mitochondrial transcription factor A in skeletal muscle in mitochondrial encephalomyopathies. *Neurol Sci* 21 (5 Suppl):S985-7.

Singh, K. K. 2006. Mitochondria damage checkpoint, aging, and cancer. *Ann N Y Acad Sci* 1067:182-90.

Sleep, NH, K Zahnle, and PS Neuhoff. 2001. Initiation of clement surface conditions on the earliest Earth. *Proc Natl Acad Sci USA* 98 (7):3666-72.

Stryer, Lubert. 1995. *Biochemistry*. 4th ed. New York: W.H. Freeman.

Takamatsu, C., S. Umeda, T. Ohsato, T. Ohno, Y. Abe, A. Fukuoh, H. Shinagawa, N. Hamasaki, and D. Kang. 2002. Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein. *EMBO Rep* 3 (5):451-6.

Taylor, S. D., H. Zhang, J. S. Eaton, M. S. Rodeheffer, M. A. Lebedeva, W. O'Rourke T, W. Siede, and G. S. Shadel. 2005. The conserved Mec1/Rad53 nuclear checkpoint

pathway regulates mitochondrial DNA copy number in Saccharomyces cerevisiae. *Mol Biol Cell* 16 (6):3010-8.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25 (24):4876-82.

Thorburn, D. R. 2004. Mitochondrial disorders: prevalence, myths and advances. *J Inherit Metab Dis* 27 (3):349-62.

Tiranti, V., A. Savoia, F. Forti, M. F. D'Apolito, M. Centra, M. Rocchi, and M. Zeviani. 1997. Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. *Hum Mol Genet* 6 (4):615-25.

Tominaga, K., J. Hayashi, Y. Kagawa, and S. Ohta. 1993. Smaller isoform of human mitochondrial transcription factor 1: its wide distribution and production by alternative splicing. *Biochem Biophys Res Commun* 194 (1):544-51.

Tong, A. H., G. Lesage, G. D. Bader, H. Ding, H. Xu, X. Xin, J. Young, G. F. Berriz, R. L. Brost, M. Chang, Y. Chen, X. Cheng, G. Chua, H. Friesen, D. S. Goldberg, J. Haynes, C. Humphries, G. He, S. Hussein, L. Ke, N. Krogan, Z. Li, J. N. Levinson, H. Lu, P. Ménard, C. Munyana, A. B. Parsons, O. Ryan, R. Tonikian, T. Roberts, A. M. Sdicu, J. Shapiro, B. Sheikh, B. Suter, S. L. Wong, L. V. Zhang, H. Zhu, C. G. Burd, S. Munro, C. Sander, J. Rine, J. Greenblatt, M. Peter, A. Bretscher, G. Bell, F. P. Roth, G. W. Brown, B. Andrews, H. Bussey, and C. Boone. 2004. Global mapping of the yeast genetic interaction network. *Science* 303 (5659):808-13.

van Buul, C. P., J. B. Damm, and P. H. van Knippenberg. 1983. Kasugamycin resistant mutants of Bacillus stearothermophilus lacking the enzyme for the methylation of two adjacent adenosines in 16S ribosomal RNA. *Mol Gen Genet* 189 (3):475-8.

van Buul, C. P., M. Hamersma, W. Visser, and P. H. Van Knippenberg. 1984. Partial methylation of two adjacent adenosines in ribosomes from Euglena gracilis chloroplasts suggests evolutionary loss of an intermediate stage in the methyl-transfer reaction. *Nucleic Acids Res* 12 (23):9205-8.

van Buul, C. P., and P. H. van Knippenberg. 1985. Nucleotide sequence of the ksgA gene of Escherichia coli: comparison of methyltransferases effecting dimethylation of adenosine in ribosomal RNA. *Gene* 38 (1-3):65-72.

van Buul, C. P., W. Visser, and P. H. Van Knippenberg. 1984. Increased translational fidelity caused by the antibiotic kasugamycin and ribosomal ambiguity in mutants harbouring the ksgA gene. *FEBS Lett* 177 (1):119-24.

van Gemen, B., J. Twisk, and P. H. Van Knippenberg. 1989. Autogenous regulation of the Escherichia coli ksgA gene at the level of translation. *J Bacteriol* 171 (7):4002-8.

Vellai, T., and G. Vida. 1999. The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells. *Proc Biol Sci* 266 (1428):1571-7.

Vesteg, M., J. Krajcovic, and L. Ebringer. 2006. On the origin of eukaryotic cells and their endomembranes. *Riv Biol* 99 (3):499-519.

Vila-Sanjurjo, A., C. L. Squires, and A. E. Dahlberg. 1999. Isolation of kasugamycin resistant mutants in the 16 S ribosomal RNA of Escherichia coli. *J Mol Biol* 293 (1):1-8.

Wallin, IE. 1922. A Note on the Morphology of Bacteria Symbiotic in the Tissues of Higher Organisms. *J Bacteriol* 7 (5):471-4.

Wallin, IE. 1924. On the nature of mitochondria VII The independent growth of mitochondria in culture media. *Am J Anatomy* 33 (1):147-173.

Wallin, IE. 1926. Bacteria and the origin of species. *Science* 64 (1651):173-175.

Wang, Y., and D. F. Bogenhagen. 2006. Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane. *J Biol Chem* 281 (35):25791-802.

Wang, Y., and G. S. Shadel. 1999. Stability of the mitochondrial genome requires an amino-terminal domain of yeast mitochondrial RNA polymerase. *Proc Natl Acad Sci USA* 96 (14):8046-51.

Wang, Z., J. Cotney, and G. S. Shadel. 2007. Human mitochondrial ribosomal protein MRPL12 interacts directly with mitochondrial RNA polymerase to modulate mitochondrial gene expression. *J Biol Chem* 282 (17):12610-8.

Weissman, L., N. C. de Souza-Pinto, T. Stevnsner, and V. A. Bohr. 2007. DNA repair, mitochondria, and neurodegeneration. *Neuroscience* 145 (4):1318-29.

Wiesner, R. J., G. Zsurka, and W. S. Kunz. 2006. Mitochondrial DNA damage and the aging process: facts and imaginations. *Free Radic Res* 40 (12):1284-94.

Williamson, J. R., and R. H. Cooper. 1980. Regulation of the citric acid cycle in mammalian systems. *FEBS Lett* 117 Suppl:K73-85.

Winkley, C. S., M. J. Keller, and J. A. Jaehning. 1985. A multicomponent mitochondrial RNA polymerase from Saccharomyces cerevisiae. *J Biol Chem* 260 (26):14214-23.

Wong, T. W., and D. A. Clayton. 1985. In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. *Cell* 42 (3):951-8.

Wong, T. W., and D. A. Clayton. 1985. Isolation and characterization of a DNA primase from human mitochondria. *J Biol Chem* 260 (21):11530-5.

Wu, Z., P. Puigserver, U. Andersson, C. Zhang, G. Adelmant, V. Mootha, A. Troy, S. Cinti, B. Lowell, R. C. Scarpulla, and B. M. Spiegelman. 1999. Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* 98 (1):115-24.

Xu, B., and D. A. Clayton. 1992. Assignment of a yeast protein necessary for mitochondrial transcription initiation. *Nucleic Acids Res* 20 (5):1053-9.

Xu, B., and D. A. Clayton. 1996. RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. *Embo J* 15 (12):3135-43.

Yoshida, Y., H. Izumi, T. Ise, H. Uramoto, T. Torigoe, H. Ishiguchi, T. Murakami, M. Tanabe, Y. Nakayama, H. Itoh, H. Kasai, and K. Kohno. 2002. Human mitochondrial transcription factor A binds preferentially to oxidatively damaged DNA. *Biochem Biophys Res Commun* 295 (4):945-51.

Yoshida, Y., H. Izumi, T. Torigoe, H. Ishiguchi, H. Itoh, D. Kang, and K. Kohno. 2003. P53 physically interacts with mitochondrial transcription factor A and differentially regulates binding to damaged DNA. *Cancer Res* 63 (13):3729-34.

Zhang, H., P. Gao, R. Fukuda, G. Kumar, B. Krishnamachary, K. I. Zeller, C. V. Dang, and G. L. Semenza. 2007. HIF-1 inhibits mitochondrial biogenesis and cellular

respiration in VHL-deficient renal cell carcinoma by repression of C-MYC activity. *Cancer Cell* 11 (5):407-20.

Zhang, X. Y., L. M. DeSalle, and S. B. McMahon. 2006. Identification of novel targets of MYC whose transcription requires the essential MbII domain. *Cell Cycle* 5 (3):238-41.

Vita

Justin Lee Cotney was born August 14, 1980 to Ronald and Julie Cotney. He was raised along with his younger brother Matthew in the small town of Daviston, Alabama. He graduated from Benjamin Russell High School in Alexander City, Alabama in May 1998. Justin then attended Birmingham-Southern College in Birmingham, Alabama; majored in biology, and received his Bachelor of Science degree in May 2002. He was then accepted to the Genetics and Molecular Biology Graduate program at Emory University in Atlanta, Georgia. There he joined the laboratory of Dr. Gerald Shadel and subsequently moved with his advisor to Yale University in April 2004. There he focused on research aimed at understanding the biological contributions of two critical components of the human mitochondrial transcription machinery, h-mtTFB1 and h-mtTFB2. He completed his graduate research in the spring of 2008 and was accepted for a post-doctoral position with Dr. Michael Snyder at Yale University.