Stony Brook University

Structural Insights into Mitochondrial Gene Expression and Disease

A Dissertation Presented

by

Kip Guja

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

May 2015

Stony Brook University

The Graduate School

Kip Guja

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Miguel Garcia-Diaz, Ph.D. – Dissertation Advisor
Associate Professor, Department of Pharmacological Sciences

A. Wali Karzai, Ph.D. – Chairperson of Defense
Professor, Department of Biochemistry and Cell Biology

Daniel F. Bogenhagen, M.D.
Professor, Department of Pharmacological Sciences

Steven E. Glynn, Ph.D.
Assistant Professor, Department of Biochemistry and Cell Biology

Marc Allaire, Ph.D.
Scientist, Berkeley Center for Structural Biology, Lawrence Berkeley National Laboratory

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Structural Insights into Mitochondrial Gene Expression and Disease

by

Kip Guja

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

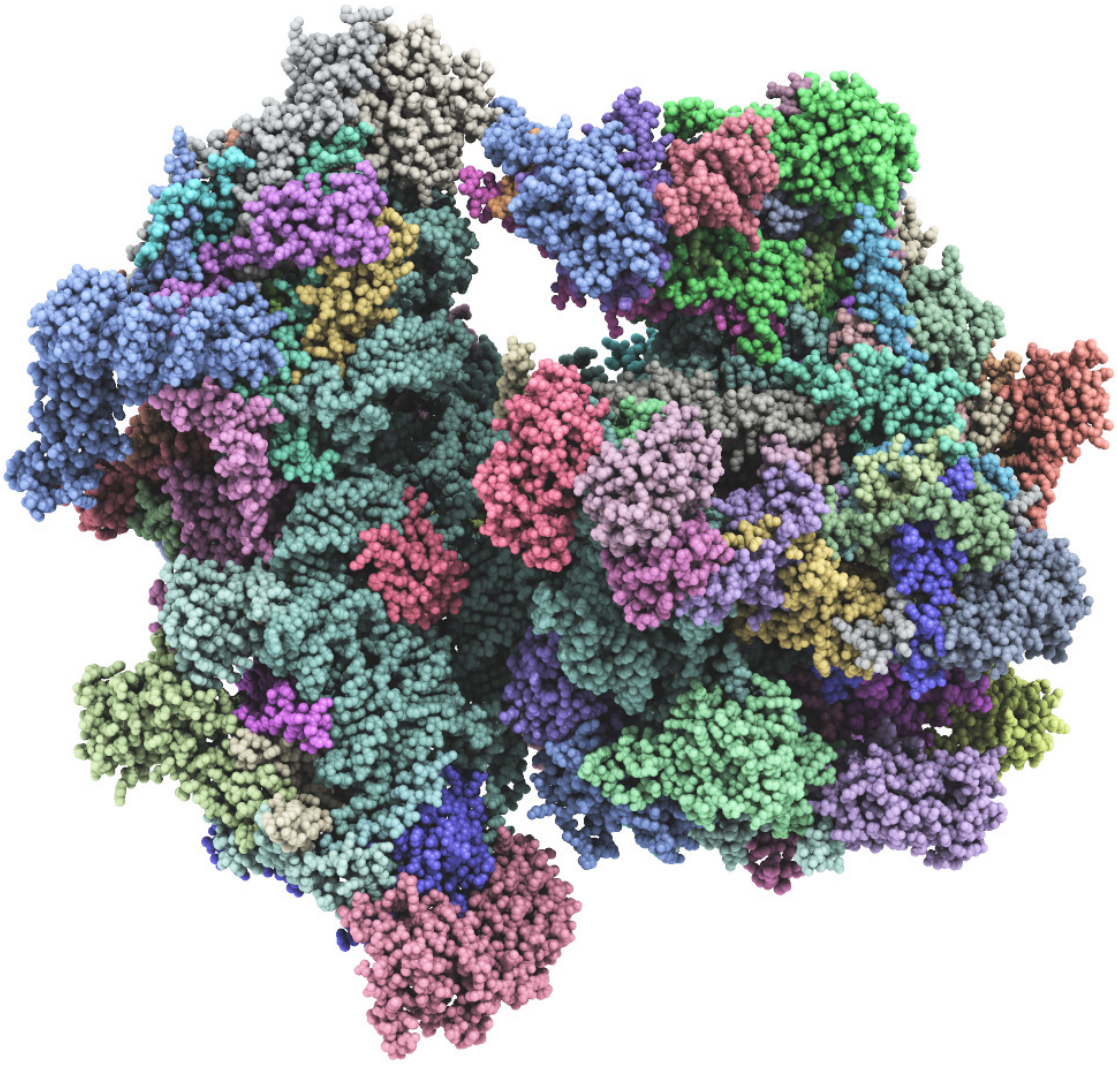
2015

Mitochondria are dual membrane-bound organelles found in the cytoplasm of all eukaryotic cells. Human mitochondria contain a 16.5 kb double-stranded DNA genome that encodes thirteen essential protein subunits of the oxidative phosphorylation system, as well as the two ribosomal RNAs and 22 transfer RNAs needed to translate the thirteen messenger RNAs. Oxidative phosphorylation produces the majority of energy needed for cells to sustain life, and this process is strictly dependent on expression of the mitochondrial genome. Defects in mitochondrial gene expression are linked to many human diseases, age-related pathologies, and the aging process itself. Obtaining a more detailed understanding of the mechanisms of mitochondrial gene expression and their regulation by nuclear encoded proteins is a critical first step towards improving our ability to treat mitochondrial diseases and mitochondrial related pathologies. I have utilized structural biology techniques (mainly X-ray crystallography), complemented with biochemical and genetic experiments to gain novel insights into several key points of regulation, including ribosomal RNA modification, ribosome biogenesis, and termination of transcription. I have obtained novel X-ray crystallographic structures of an essential mitochondrial ribosomal RNA methyltransferase TFB1M, which has isoforms that are linked to maternally inherited deafness and increased risk for developing type II diabetes mellitus. I have demonstrated the methyltransferase activity of TFB1M using an in vivo complementation assay, and obtained a structure of TFB1M bound to the cofactor S-adenosylmethionine, which reveals the structural basis for its activity. I have also obtained novel crystal structures of MTERF family proteins that regulate ribosome assembly and transcription termination in mitochondria, and also modulate pathogenesis of several diseases arising from mutations in the mitochondrial DNA. In sum, my results provide novel structural insights into several members of the mitochondrial gene expression machinery, suggest close links between regulation of transcription and RNA modification in mitochondria, and improve our understanding of the pathogenesis of mitochondrial disease.

Dedication Page

For my parents; *amare et sapere vix deo conceditur.*

Frontispiece



Structure of the mammalian mitochondrial ribosome

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADP – adenosine diphosphate

ATP – adenosine triphosphate

bp – base pair

CsCl – cesium chloride

ETC – electron transport chain

$FADH_2$ – flavin adenine dinucleotide

HS – heavy-strand

HSP – heavy-strand promoter

kb – kilobase

LS – light-strand

LSP – light-strand promoter

MBP – maltose binding protein

mRNA – messenger ribonucleic acid

mtDNA – mitochondrial deoxyribonucleic acid

MTERF – mitochondrial transcription termination factor

NADH – nicotinamide adenine dinucleotide

NCBI – National Center for Biotechnology Information

OXPHOS – oxidative phosphorylation system

POLRMT – mitochondrial RNA polymerase

RNA – ribonucleic acid

rRNA – ribosomal ribonucleic acid

SEC – size-exclusion chromatography

TFAM – mitochondrial transcription factor A

TFB1M – mitochondrial transcription factor B1

TFB2M – mitochondrial transcription factor B2

tRNA – transfer ribonucleic acid

Acknowledgments

Foremost, I am grateful to my advisor Miguel Garcia-Diaz, who has been a better advisor than I could have ever hoped for. I am particularly grateful for his encouragement of my interest in both the experimental and technical aspects of X-ray crystallography and structural biology in general. I feel great gratitude for having the opportunity to train in his laboratory, for the level of independence that he gave me during my training, for his inspiring qualities as a scientist, and for the many enjoyable discussions we had on our shared interests. I thank my colleagues in the laboratory, James Byrnes, Elena Yakubovskaya, Edison Mejia, Woo Suk Choi and Matt Burak, for their support and friendship, and for the opportunity to work with and learn from them. I am also grateful to my committee members, Wali Karzai, Marc Allaire, Dan Bogenhagen, and Steve Glynn for their contribution to my scientific career and for keeping me inspired. I am especially thankful for Wali's guidance and advice throughout my time at Stony Brook, for his straightforward and rigorous approach to science, and for our collaborative projects, which have been very rewarding both scientifically and personally. I am grateful to Marc for his encouragement of my interest in X-ray methods, for introducing me to small-angle X-ray scattering, for giving me the opportunity to become an instructor in the SAXS workbench, and for his inspiring knowledge and enthusiasm for synchrotron science. I thank Dan especially for insightful discussions, particularly those related to the mitochondrial ribosome, and for his extensive and invaluable expertise in mitochondrial biology. I am grateful to Annie Heroux, Howard Robinson, Neil Whalen, and Lin Yang at the NSLS for making much of my scientific work possible. I feel great gratitude to both David P. Roye at Columbia University Medical Center and Joel F. Schildbach at Johns Hopkins University, for inspiring my academic career in medicine and research, and for being extraordinary role models that I hope to emulate as a physician and principal investigator. Last, but most certainly not least, I would like to thank my family and friends, without whom the trajectory of my life would be vastly different.

Publications

1. **Guja KE**, Schildbach JF. (2015). Completing the specificity swap: Single-stranded DNA recognition by F and R100 TraI relaxase domains. *Plasmid*. [Epub 2015 Apr 1]. PMID: 25841886. DOI: 10.1016/j.plasmid.2015.03.006
2. Yang M, Lu R, **Guja KE**, Wipperman MF, St. Clair JR, Bonds AC, Garcia-Diaz M, Sampson NS. (2015). Unraveling Cholesterol Catabolism in Mycobacterium tuberculosis: ChsE4-ChsE5 α2β2 Acyl-CoA Dehydrogenase Initiates β-Oxidation of 3-Oxo-cholest-4-en-26-oyl CoA. *ACS Infectious Diseases*. 1(2):110–125. PMID: N/A. DOI: 10.1021/id500033m
3. Yang M, **Guja KE**, Thomas ST, Garcia-Diaz M, Sampson NS. (2014). A Distinct MaoC-like Enoyl-CoA Hydratase Architecture Mediates Cholesterol Catabolism in Mycobacterium tuberculosis. *ACS Chemical Biology*. 9(11):2632–2645. PMID: 25203216. DOI: 10.1021/cb500232h.
4. Venkataraman K, **Guja KE**, Garcia-Diaz M, Karzai AW. (2014). Nonstop mRNA Decay: a Special Attribute of Trans-Translation Mediated Ribosome Rescue. *Frontiers in Microbiolog*y. 5(93). PMID: 24653719. DOI: 10.3389/fmicb.2014.00093.
5. Yakubovskaya E[1], **Guja KE**[1], Eng ET, Choi WS, Mejia E, Beglov D, Lukin M, Kozakov D, Garcia-Diaz M. (2014). Organization of the human mitochondrial transcription initiation complex. *Nucleic Acids Research*. 42(6):4100-12. PMID: 24413562. DOI: 10.1093/nar/gkt1360. [1]Equal contribution.
6. Jacewicz A, Trzemecka A, **Guja KE**, Plochocka D, Yakubovskaya E, Bebenek A, Garcia-Diaz M. (2013). A remote palm domain residue of RB69 DNA polymerase is critical for enzyme activity and influences the conformation of the active site. *PLoS One*. 8(10):e76700. PMID: 24116139. DOI: 10.1371/journal.pone.0076700.
7. **Guja KE**, Venkataraman K, Yakubovskaya E, Shi H, Mejia E, Hambardjieva E, Karzai AW, Garcia-Diaz M. (2013). Structural basis for S-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1. *Nucleic Acids Research.* 41(16):7947-59. PMID: 23804760. DOI: 10.1093/nar/gkt547.
8. Yakubovskaya E[1], **Guja KE**[1], Mejia E[1] , Castano S, Hambardjieva E, Choi WS, Garcia-Diaz M. (2012). Structure of the Essential MTERF4:NSUN4 Protein Complex Reveals

How an MTERF Protein Collaborates to Facilitate rRNA Modification. *Structure*. 20(11):1940-7. PMID: 23022348. DOI: 10.1016/j.str.2012.08.027. [1]Equal contribution.

9. **Guja KE**, Garcia-Diaz M. (2012). Hitting the brakes: Termination of mitochondrial transcription. *Biochimica et Biophysica Acta*. 1819(9-10):939-947. PMID: 22137970. DOI: 10.1016/j.bbagrm.2011.11.004.
10. Anderson BJ, Larkin C, **Guja K**, Schildbach JF. (2008). Using fluorophore-labeled oligonucleotides to measure affinities of protein-DNA interactions. *Methods in Enzymology*. 450:253-272. PMID: 19152864. DOI: 10.1016/S0076-6879(08)03412-5.
11. Hekman K, **Guja K**, Larkin C, Schildbach JF. (2008). An intrastrand three-DNA-base interaction is a key specificity determinant of F transfer initiation and of F TraI relaxase DNA recognition and cleavage. *Nucleic Acids Research*. 36(14):4565-4572. PMID: 18611948. DOI: 10.1093/nar/gkn422.

Chapter 1: Introduction to mitochondrial gene expression

1.1 The significance of mitochondrial gene expression and mitochondrial disease

Mitochondria are double-membrane bound organelles that are found in human and other eukaryotic cell types (Figure 1-1). According to the endosymbiotic theory, they originated from an α-proteobacterium that entered a symbiotic relationship with an archaebacterium more than one billion years ago [1, 2]. Accordingly, present-day eukaryotes contain both nuclear DNA (nDNA) and mitochondrial DNA (mtDNA), although the size and organization of the mtDNA has become significantly smaller and more compact since the endosymbiotic event. Not surprisingly, many of the mitochondrial proteins found in humans have bacterial origins, but mitochondria have evolved unique mechanisms to regulate expression of their genome [3]. The mitochondrial genome encodes proteins that are essential for cellular energy production via oxidative phosphorylation, and therefore its expression is absolutely required to sustain life. In higher eukaryotes, mitochondria also play a myriad of roles in diverse cellular functions including apoptosis, metabolism, thermogenesis, immunity, cell signaling and steroid synthesis [4-7].

As a consequence of their essential role in cellular energy production, mitochondrial gene expression is regulated at multiple levels. These regulatory processes include replication and genome maintenance, transcription initiation and termination, RNA modification, as well as ribosome biogenesis and translation [8]. Defects in mitochondrial gene expression and its regulation, as well as mutations in the mtDNA itself are major causes of human disease and have also been implicated in age-related diseases and the aging process itself [9]. More than one hundred mitochondrial diseases or syndromes have been characterized, with an incidence of at least 1 in 5,000 [3]. Human pathogenic mtDNA mutations were first described more than 25 years ago and more than 300 mutations in mitochondrial DNA (mtDNA) have been discovered to date [10, 11]. Many of these mutations affect tRNA genes and thereby impair mitochondrial translation [12]. Mitochondrial translation and ribosome biogenesis can also be impaired by mutations in the mitochondrial ribosomal RNA genes; two examples are A1555G and C1494T in

the 12S rRNA, which predispose patients to maternally inherited aminoglycoside-induced deafness [13-15]. These example illustrate that as a consequence of their common ancestor, both the prokaryotic and mitochondrial ribosomes can bind commonly used antibiotics, and thus some of the side effects of these antibiotics arise from impaired mitochondrial translation [16]. Furthermore, many of the proteins that regulate mitochondrial transcription are nuclear encoded and thus nuclear gene mutations can also have a causative role in human mitochondrial disease [17].

Examples of genetic mitochondrial diseases and diseases that involve mitochondrial pathology include mitochondrial encephalopathy lactic acidosis and stroke-like syndrome (MELAS), myoclonic epilepsy with ragged red fibers (MERRF), Leigh syndrome, Leber's hereditary optic neuropathy (LHON), cardiomyopathy, deafness, muscular dysfunction, anemia, diabetes, Parkinson's disease and Alzheimer's disease (Figure 1-2) [3, 18-22]. Some of these diseases arise from mtDNA mutations that can affect more than one aspect of gene expression – for example, MELAS is associated with mutations in a tRNA-Leu gene that result in a functional tRNA defect as well as impaired termination of mitochondrial transcription. Mitochondrial dysfunction is also heavily implicated in aging and mtDNA mutations can undergo clonal expansion, resulting in a mosaic pattern of respiratory chain deficiency in different aging mammalian tissues [9, 23-25]. Mitochondria are heteroplasmic, and there are tissue-specific differences in mitochondrial number, function, protein composition and morphology; as result of this, mitochondrial defects and diseases have organ-specific symptoms and clinical presentations (Figure 1-2) [26]. Typically, organs and tissues with the highest energetic requirements such as brain, heart, and skeletal muscles are the most severely affected.

Current treatments for mitochondrial disorders are palliative rather than curative, and are mostly limited to the administration of vitamins, cofactors, and scavengers of reactive oxygen species [18, 27-30]. Curative treatments would require a detailed mechanistic understanding of pathogenesis and would likely need to be genetic in nature. The development of "genome-editing" and other gene therapy techniques for mitochondrial disease is currently underway and includes expression of mitochondrial-localized restriction enzymes, TALENs, and zinc-finger nucleases [31-34]. Transmission of mammalian mtDNA is strictly maternal, and therefore one method of preventing mitochondrial disease is to prevent transmission of defective mtDNA altogether or to replace defective mtDNA with mtDNA from a healthy donor during in vitro

fertilization. This technique, known as mitochondrial DNA replacement therapy, was recently approved in the United Kingdom, but is not yet approved by the FDA for use in the United States [35-39]. Other methods of genome transfer have also recently shown promise in murine models as potential alternative methods for preventing and circumventing maternal transmission of mitochondrial diseases [40].

1.2 The mitochondrial genome

The human mtDNA is a circular double-stranded DNA molecule of 16.5 kb, and encodes 13 mRNAs, 2 rRNAs and 22 tRNAs (Figure 1-3). The double-strands of the genome are denoted as the heavy strand and light strand, so named because the heavy-strand is G-rich while the light-strand is C-rich and can thus be separated by mass on an alkaline cesium chloride gradient [8]. The tRNAs are dispersed throughout the genome and in all but two cases separate the rRNAs and mRNAs. The rRNAs and tRNAs are components of the mitochondrial ribosome, or mitoribosome and the translation machinery, while the mRNAs are translated on mitoribosomes to generate thirteen essential protein subunits of the oxidative phosphorylation (OXPHOS) complexes. The mtDNA encodes components of Complexes I, III, IV and V, while Complex II is composed entirely of nuclear-encoded subunits. Thus, the mtDNA is entirely dedicated to biogenesis of the electron transport chain (ETC). As the ETC contains a combination of both nuclear- and mitochondrial-encoded subunits, proper coordination of the mitochondrial and nuclear genomes is essential for proper assembly of the ETC. Similar to its prokaryotic ancestor, the mtDNA is packaged into nucleoids by various proteins and unlike the nuclear DNA, is not contained within a membrane [41, 42].

The mammalian mitochondrial genome is extremely compact, as evidenced by the lack of introns or untranslated mRNA regions. Most of the mtDNA-encoded mRNAs require polyadenylation to add or complete the stop codon. The mitochondrial genome utilizes a modified genetic code, where AUA and AUU code for methionine instead of isoleucine, AGA and AGG encodes stop codons instead of arginine, and UGA codes for tryptophan rather than serving as a stop codon [43, 44]. Moreover, mitochondrial-specific tRNA modifications contribute to the usage of only 22 tRNAs rather than the full set of 32 that are required for recognition of the genetic code utilized by the nuclear DNA and cytosolic ribosomes.

Over the course of evolution, most of the ancestral α-proteobacterial genes have been exported to the nucleus, where they evolved along with novel proteins that regulate mitochondrial gene expression and package the mtDNA [45]. These nuclear-encoded mitochondrial proteins are translated by cytosolic ribosomes, imported into the mitochondria and subsequently sorted to the appropriate mitochondrial location [46]. Most of the proteins destined for the mitochondrial matrix are synthesized as precursors containing a positively charged N-terminal mitochondrial localization sequence (MLS) that is cleaved upon entry into the matrix. The process of mitochondrial protein import is dependent on the electrochemical membrane potential, and thus defective mitochondria are selected against by impaired import of the proteins needed for regulation of gene expression [47, 48].

1.3 Mitochondrial DNA transcription

Regulation of mitochondrial gene expression begins with initiation of transcription, which is carried out by a complex of POLRMT, TFAM and TFB2M at one of three promoters: the light-strand promoter (LSP), heavy-strand promoter 1 (HSP1 or simply HSP), or heavy-strand promoter 2 (HSP2) [49, 50]. LSP transcripts are either transcribed as a nearly full-genome length polycistronic RNA encoding mRNAs and tRNAs, or terminated earlier in order to generate primers for mtDNA replication [51]. HSP1 transcripts initiate upstream of the phenylalanine tRNA gene and may be terminated by MTERF1 downstream of the 16S rRNA gene. MTERF1 is proposed to bind to HSP1 and tRNA-Leu (UUR), where termination of transcription prevents the formation of antisense transcripts and may help to regulate rRNA levels. This notion is consistent with the higher expression level of mitochondrial rRNAs relative to mRNAs, and indeed preliminary studies indicate that rRNAs are 50-fold more abundant than mRNA transcripts [52]. Transcripts that initiate from HSP2 begin upstream of the 12S rRNA gene and generate a nearly full-genome length polycistronic RNA. It is important to note that recently, a knockout of the MTERF1 gene in a mouse model has yielded results that conflict with this model, as the loss of MTERF1 did not affect rRNA gene expression, while the amount of antisense RNAs increased [52]. This result suggests that there may only be one transcription unit for the expression of heavy-strand encoded genes that is regulated by a single HSP [52]. In this model, the main function of MTERF1 would be to prevent transcription initiated at LSP from proceeding over the rRNA gene region and thereby producing antisense transcripts. It has been

suggested that production of antisense transcripts would be disadvantageous, as they could prevent proper folding of the 12S and 16S rRNAs [52]. Moreover, antisense transcription has been shown to interfere with LSP activity, possibly by causing a collision or steric hindrance of the transcription machinery that prevents reinitiation [52]. Further studies will be required to clarify the precise role and significance of the second heavy-strand promoter site, HSP2, for transcription initiation.. Regardless of where they initiate, the polycistronic RNAs transcribed from the heavy- and light-strands need to be processed and modified by specific mitochondrial enzymes in order to produce the mature mRNAs, tRNAs and rRNAs required for gene expression.

1.4 Processing of polycistronic RNA transcripts

Mitochondrial mRNAs are essentially leaderless – the 5' ends of most mRNAs begin either with a start codon or very close to a start codon, and the ends are immediately adjacent to tRNA genes [53]. Mitochondrial RNA transcripts are synthesized from both the heavy and light strands as polycistronic precursors. Both the 12S and 16S rRNAs and all but two mRNAs are separated by tRNAs, which must be cleaved out. RNase P and RNase Z are thought to be the endonucleases that recognize the secondary structure of tRNAs within the nascent polycistronic mRNAs and carry out this cleavage at the 5' and 3' ends of the tRNA transcripts. This RNA processing mechanism is also known as the "tRNA punctuation" model and was first proposed more than three decades ago [54-56]. More recently, Bogenhagen and colleagues have reported co-localization of RNase P and RNase Z to mitochondrial nucleoids, which further supports their role in processing nascent tRNAs [57].

RNase P was initially discovered in bacteria as a protein-RNA complex, where the RNA component was thought to be responsible for the observed catalytic activity [58]. The human nuclear form of RNase P is also a protein-RNA complex, which consists of at least ten protein subunits that process precursor tRNAs [59]. However, in human mitochondria RNase P was identified in an RNA-free complex consisting of three proteins, MRPP1, MRPP2 and MRPP3, and its activity was reconstituted in vitro [60]. In addition to its role in RNA processing, MRPP1 was observed to carry out methylations at the m^1G_9 and m^1A_9 positions on tRNAs [61]. Pathogenic mutations that affect tRNA processing and modifications have been found in RNase Z and in tRNA genes themselves. These RNase Z mutations are associated with hypertrophic

cardiomyopathy, and have been shown to result in defective RNA processing [62]. Pathogenic mtDNA mutations in tRNA genes themselves can affect 3' end processing, and impair translation, underscoring the importance of RNA processing for mitochondrial gene expression [63-66].

1.5 mRNA modifications

Mitochondrial mRNAs only undergo one known type of modification – polyadenylation, which is thought to modulate stability of the mRNAs, although the precise mechanism by which this occurs is unclear. The mammalian mitochondrial genome is extremely compact, and contains almost no intronic sequences or untranslated regions [53, 67, 68]. In fact, the majority of mitochondrial mRNAs do not encode complete stop codons and require polyadenylation to complete the termination codon by adding an additional A [68-72]. Polyadenylation is carried out by hmtPAP, which shares homology with the cytosolic PAPs, and adds an average of 45 adenine nucleotides [68, 72, 73]. Although almost all of the mitochondrial mRNAs are polyadenylated, the precise length of that polyadenylation varies between cell types and even between different mitochondria in the same cell [72, 73]. Polyadenylation can either increase or decrease the stability of mRNAs, making it difficult to ascertain whether or not polyadenation has any direct effect on mitochondrial translation [70, 73]. Unlike cytosolic mRNAs, mitochondrial mRNAs lack a 5' 7-methylguanosine cap, and they do utilize an initiator N-formylmethionine, as prokaryotes do. However, mitochondria encode only one methionine tRNA, which is unmodified and used during the elongation phase of translation. The initiator methionine tRNA is formylated by the mitochondrial methionyl-tRNA transformylase, MTFMT [74].

1.6 tRNA modifications

The tRNAs found in prokaryotes and in the eukaryotic cytosol are known as Type 0 tRNAs and have a canonical "cloverleaf" secondary structure (Figure 1-4). This secondary structure is formed by the following elements (in the 5' to 3' direction): an acceptor stem, D-stem, D-loop, anticodon stem, anticodon loop, extra loop, T-stem, T-loop, and CCA terminus. In contrast, mitochondrial tRNAs have three forms of non-canonical tRNAs: Type I, which have a

shortened D-loop and shortened extra loop, Type II, which have variable length and sequence in their D-loop and T-loop, and Type III, which lacks the D-loop (Figure 1-4) [75]. These non-canonical tRNAs lack some of the stabilizing interactions found in the canonical Type 0 tRNAs, and moreover, mitochondrial tRNAs have a higher adenine and uracil content than their cytosolic counterparts. Accordingly, mitochondrial tRNAs have a lower melting temperature and are inherently less stable, which renders them more sensitive to defects in processing and modification. Not surprisingly, there are more than 100 pathogenic mutations of tRNA genes in mitochondria, the majority of which impair recognition by the tRNA modifying enzymes [76].

Numerous species of modified nucleosides have been identified on various positions of mitochondrial tRNAs, three of which are specific to mitochondria and are thus noteworthy: 5-formylcytidine (f^5C), 5-taurinomethyluridine (τm^5U) and 5-taurinomethyl-2-thiouridine (τm^5s^2U) [75, 77]. The remaining modifications include base methylation and pseudouridylation, which are also found on cytosolic and prokaryotic tRNAs. The wobble position nucleosides of mitochondrial tRNAs are: 1) f^5C, 2) queuosine (Q), 3) taurine-containing uridines and 4) unmodified guanosine or uridine. Modifications at these positions are particularly important; for example f^5C at the wobble position is required to recognize the AUA and AUG codons. The existence of modified nucleosides at the wobble position in mitochondria allows for specific decoding of more than 50 different sense codons while using only 22 tRNAs. In contrast, more than 30 tRNAs are required for decoding in the cytosol [75]. The vital importance of tRNA modifications in mitochondria is further evidenced by the fact that defects in taurine modification have been implicated in mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes syndrome (MELAS), and at least one causative mutation in a tRNA-Leu gene prevents recognition by the tRNA-taurine modification system [78-84].

1.7 rRNA modifications and ribosome biogenesis

Mitochondrial transcription produces 12S and 16S rRNAs that are subsequently modified and incorporated into the small and large ribosomal subunits (SSU and LSU), respectively. The mature SSU, or 28S subunit consists of ~30 mitochondrial ribosomal proteins (MRPs) as well as the 12S rRNA, and contains the decoding center where codons in mRNAs are matched with their cognate anticodons on tRNAs. The mature LSU, 39S subunit is made up of ~50 MRPs and the 16S rRNA, and contains the peptidyltransferase center, where a peptide bond is formed between

the incoming aminoacyl-tRNA in the A-site and the elongating peptide chain in the P site. Assembly of the SSU, LSU, and ultimately the 55S monosome is mediated by many accessory factors, including rRNA modifying enzymes, rRNA chaperones, protein chaperones, and GTPases. In prokaryotes, the ribosome assembly process begins with co-transcriptional rRNA processing, modification, and folding, and at least some of these aspects are likely to be conserved in mitochondria [85]. Moreover, roughly half of the mitochondrial ribosomal proteins share some degree of homology with their prokaryotic counterparts (Figure 1-5). However, in comparison to prokaryotic and eukaryotic cytosolic rRNAs, the mitochondrial rRNAs have fewer modifications, although the most common modification is methylation in all three cases. The first three ribose methylation sites were discovered more than three decades ago in hamster cells [86]. At present, five base methylations are known to occur in the 12S rRNA, while the 16S rRNA is known to contain three 2'-O-ribose methylation sites and one pseudouridylation site [73, 86]. These rRNA modifications are generally thought to be crucial for ribosome assembly and/or stability, and may also play a role in RNA folding, stabilization of secondary structure elements, or molecular recognition by RNA-binding proteins [87, 88].

1.7.1 Modifications of the 12S rRNA and biogenesis of the SSU

To date, five base methylations have been identified on the 12S rRNA component of the SSU (Table 1-1). These modifications are likely to be conserved in all mammals, and all but one of them has been attributed to a specific methyltransferase. The five modifications are: two adenine dimethylations ($m^6{}_2A^{937}$ and $m^6{}_2A^{938}$), one uracil methylation (m^5U^{425}), and two cytosine methylations (m^4C^{840} and m^5C^{841}) [86].

The two aforementioned adenine dimethylations occur in a stem-loop structure at the 3' end of the 12S rRNA in the SSU, and are highly conserved in nearly all domains of life [89]. In Escherichia coli, the enzyme responsible for these modifications is KsgA, so named because deletion of this gene leads to resistance to the aminoglycoside kasugamycin, as well as impaired growth of the bacteria [90, 91]. More than 10 years ago, Shadel and colleagues used primer extension assays to demonstrate that these dimethylations in mitochondria ($m^6{}_2A_{936}$ and $m^6{}_2A^{937}$) are carried out by mitochondrial transcription factor B1 (TFB1M). Accordingly, a murine knockout of the TFBIM gene was shown to result in loss of the $m^6{}_2A^{937}$ and $m^6{}_2A^{938}$ modifications and ultimately caused embryonic lethality at just over one week [92]. In the

absence of these modifications, levels of intact SSU are drastically reduced and mitochondrial translation is severely impaired [92]. A heart-specific knockout in mice has been generated for TFB1M and results in cardiomyopathy, decreased 12S rRNA stability and a similar impaired assembly of the SSU [92]. An apparent increase in levels of the LSU was also noted, and perhaps represents an attempt at compensation for decreased numbers of fully assembled monosomes.

The importance of TFB1M's methyltransferase activity is also underscored by the effect of overexpressing a methyltransferase-deficient TFB1M carrying the point mutation G65A in HeLa cells, where a dramatic decrease in mitochondrial biogenesis was observed [93]. The G65A mutation that renders TFB1M catalytically inactive was discovered through systematic mutation of highly conserved residues followed by probing for methylation using primer extension assays, and indeed the subsequently reported crystal structure of TFB1M revealed that a G65A mutation would severely disrupt the SAM-binding pocket [94].

Interestingly, TFB1M is a paralogue of TFB2M, which is a component of the mitochondrial transcription initiation machinery. This has led to some speculation that TFB2M may possess a methyltransferase activity that overlaps with that of TFB1M, however cells overexpressing an analogous G105A point mutation in the catalytic site did not display any obvious defects [93]. This fact is consistent with the absence of methyltransferase activity in mtTFB, the yeast homologue of TFB2M, and supports the notion that TFB2M plays a methyltransferase-independent role in initiation of transcription [8, 93, 95].

It is important to note that the activity of TFB1M is modulated by mtDNA mutations. Patients who are susceptible to aminoglycoside-induced deafness (as well as non-aminoglycoside-induced, or non-syndromic deafness) carry an A1555G mutation in the 12S rRNA gene, which has been reported to result in TFB1M-mediated hypermethylation of the 12S rRNA. Preliminary studies in a mouse model have shown that hypermethylation of the 12S rRNA leads to subsequent ROS-dependent AMPK signaling to E2F1, resulting in apoptosis of hair cells in the inner ear and impaired function of the cochlea [96]. It is important to note that the A1555G mutation is located in the A-site of the assembled ribosome, where the correct tRNA anticodon and mRNA codon are matched during protein synthesis, such that mRNA misreading may also contribute to pathogenesis [97]. Moreover, the recently reported structure of the intact 55S mitoribosome solved at 3.5 Å using cryoelectron microscopy reveals that A1555G would

likely enhance aminoglycoside binding to the decoding center (Figure 1-7) [98, 99]. Further studies are needed to more fully elucidate the pathogenic significance of methylation by TFB1M and any contribution it may have to A1555G-associated deafness.

The m^5C^{841} modification in the 12S rRNA is catalyzed by NSUN4 as recently demonstrated by sequencing of cDNA generated by from bisulfite-treated mouse mitochondrial rRNA [100]. NSUN4 belongs to the same family of m^5C-methyltransferases as the bacterial rRNA m^5C-methyltransferases RsmB, RsmF, and YccW [101]. However, unlike its bacterial homologs, NSUN4 lacks some of the embellishments to the methyltransferase core domain that typically encode RNA binding specificity, and instead forms a stable complex with the RNA-binding protein MTERF4, which interestingly is a member of the mitochondrial transcription termination factor (MTERF) family, so named because the first family identified, MTERF1, mediates termination of mitochondrial transcription [101-104]. The interaction of NSUN4 with MTERF4 ultimately targets the NSUN4:MTERF4 complex to the LSU [101, 104, 105]. Deletion of either MTERF4 or NSUN4 results in embryonic lethality in mice and appears to affect a later stage of ribosome biogenesis, as both the SSU and LSU accumulate but fully assembled monosomes do not [100, 101].

The m_4C^{840} modification in mice is also conserved in E. coli, where it has been proposed to regulate translation by modulating the conformation of the ribosomal P site [106]. However, in the case of E. coli, the analogous position in the 23S rRNA is modified twice, by two distinct methyltransferases, RsmI and RsmHIn [106].

In a mouse model, knocking out NSUN4 abolishes the m^5C^{842} (m^5C^{841} in humans) modification of 12S rRNA, whereas the m^4C^{840} modification is unaffected [100]. Surprisingly, knocking out MTERF4 does not affect the m^5C^{842} modification in mice, despite the fact that human NSUN4 and MTERF4 form a tight heterodimeric complex [104, 105]. This observation has led to a proposed model in which a fraction of free NSUN4 catalyzes the m^5C^{842} / m^5C^{841} modification of the 12S rRNA, while the NSUN4:MTERF4 complex plays an essential role in the final assembly of the monosome via MTERF4 binding to the LSU and acting as a bridge to the SSU, which is modified by NSUN4 [100]. Yakubovskaya and colleagues have shown that the NSUN4:MTERF4 can methylate naked 16S rRNA in vitro, and although the methylation activity they observe is likely unspecific due to the use of naked RNA, it does require MTERF4 and thus supports the notion that MTERF4 binds the LSU [104]. It has been suggested that the m^5C^{842} /

m^5C^{841} modification of the 12S rRNA acts as a quality control mark for ensuring that only fully mature SSUs can interact with the LSU and form a complete monosome [100]. Finally, it is important to note that the adenine dimethylations ($m^6{}_2A^{937}$and $m^6{}_2A^{938}$) of 12S rRNA are established independently of NSUN4, which suggests that they may occur at an earlier stage of the biogenesis of the SSU, or at the very least, they do not depend on the modification catalyzed by NSUN4 [100].

1.7.2 Modifications of the 16S rRNA and biogenesis of the LSU

The initial experiments in hamster cells revealed three ribose methylations of the 16S rRNA, which are, according to the human numbering, Gm^{1145}, Um^{1369}, and Gm^{1370} (Table 1-1) [86]. Recent studies reported by Bogenhagen and colleagues demonstrate that MRM1, MRM2 and MRM3 (RNMTL1) are the enzymes responsible for these modifications in human mitochondria [107, 108]. In yeast mitochondria, methylation of the 21S rRNA by MRM1 has been shown to be crucial for LSU stability, and mitochondrial function is impaired in its absence [109]. Similarly, a deletion of MRM2 in yeast abolishes the Um^{2791} modification, and results in decreased LSU production [110]. However, the Gm^{1370} modification and its corresponding methyltransferase, MRM3 or RNMTL1, appear to be novel and are apparently restricted to the mitochondria of higher eukaryotes [107, 108]. However, the analogous modification (Gm^{2918}) is conserved in the yeast cytosolic rRNA [73, 87].

The sole pseudouridylation site on 16S rRNA, Ψ^{1397}, was not observed in the initial studies of the large mitochondrial rRNA in hamster cells, and was not discovered until nearly 20 years later, using the more sensitive method of N-cyclohexyl-N-(2-morpholinoethyl)-carbodiimid-metho-p-toluolsulfonate (CMC) modification followed by a reverse-transcriptase primer extension assay [111, 112]. This pseudouridylation is conserved in yeast mitochondria and is catalyzed by the enzyme Pus5, which interestingly has a range of human homologs, although it is unclear which of these perform the modification in human mitochondria [113].

The MTERF3 protein, like MTERF4, is a member of the mitochondrial transcription termination factor family [103]. MTERF3 is predicted to bind nucleic acids and shares a high degree of structural similarity with both MTERF1 and MTERF4, including the region responsible for MTERF4's interaction with NSUN4 [103, 104, 114]. MTERF3 has been proposed to act as a negative regulator of mitochondrial DNA transcription initiation and is

essential for embryonic development in mice [115]. Interestingly, it was recently shown that knockout and downregulation of MTERF3 in Drosophila not only activates mitochondrial transcription but also leads to impaired assembly of the LSU [116]. Similarly, biogenesis of the LSU and mitochondrial translation are also severely impaired in MTERF3 knockout mice [116]. RNA immunoprecipitation experiments have shown that both MTERF3 specifically interacts with 16S rRNA, however, it is currently not known whether MTERF3 interacts with an rRNA modifying enzyme, and its precise function in the biogenesis of the LSU remains unclear [116].

1.7.3 rRNA modification and ribosome biogenesis are closely linked to transcription

The mitoribosome biogenesis pathway involves several distinct steps, many of which are closely related (both temporally and spatially) to transcription initiation and termination. The steps of mitoribosome biogenesis include: (1) transcription of the 12S and 16S rRNAs by POLRMT, TFB2M and TFAM; (2) expression of the nuclear-encoded mitoribosomal proteins that assemble with the 12S and 16S rRNAs to generate the LSU and SSU, respectively; (3) modification of the 12S and 16S rRNAs by methylation and pseudouridylation of specific residues; (4) assembly of LSU and SSU of the mitoribosome; (5) assembly of the mature SSU and LSU to functional mitoribosomes.

There are a few notable examples of overlap in these processes, where homologues, paralogues, and/or members of the same protein family mediate different steps. For example, the paralogues TFB1M and TFB2M share more than 30% sequence identity and play different roles in rRNA modification and initiation of transcription, respectively. MTERF3 appears to play a role in regulating both transcription and ribosome biogenesis. The canonical MTERF family member, MTERF1, seems to be involved predominantly in termination of transcription, while MTERF4 appears to be involved exclusively in ribosome biogenesis. In sum, the differing roles of homologous proteins in both mitochondrial transcription and ribosome biogenesis strongly suggest that these processes are temporally and spatially related.

Given that mitochondria evolved from an α-proteobacterial ancestor, the process of ribosome biogenesis in mitochondria is expected to be somewhat similar to prokaryotic ribosome biogenesis. In prokaryotes, polycistronic precursor RNA is transcribed in the nucleoid, cleaved by RNases and folded by helicases, while ribosomal proteins assemble onto the rRNAs [117]. In

mitochondria, transcription also occurs in nucleoids, and at least some ribosomal proteins likely assemble with the rRNAs during the transcription process [57].

In both eukaryotic cytosolic and prokaryotic ribosome biogenesis, the initial rRNA processing and some of the rRNA modifications are co-transcriptional [118, 119]. Not surprisingly, the mitochondrial rRNA methyltransferases RNMTL1, MRM1, and MRM2 were recently localized to foci containing newly transcribed mitochondrial RNA [108]. This observation supports a co-transcriptional aspect of at least the early steps of rRNA modification. Later steps of rRNA modification involve at least partially assembled ribosomal subunits that are specifically recognized by rRNA-modification enzymes, and thus are less likely to be co-transcriptional [92, 100].

In this thesis, I describe novel structural and functional characterizations of TFB2M, a member of the mitochondrial transcription factor B family, and two members of the MTERF family, MTERF1 and MTERF4. This work highlights the close relationship between transcription and ribosome biogenesis, and illustrates how members of the same protein family can mediate different steps in the regulation of mitochondrial gene expression. Furthermore, because the gene expression process plays a critically important role in mitochondrial pathology, this work should ultimately provide insight into mitochondrial disease by improving our basic understanding of the gene expression process and its regulation.

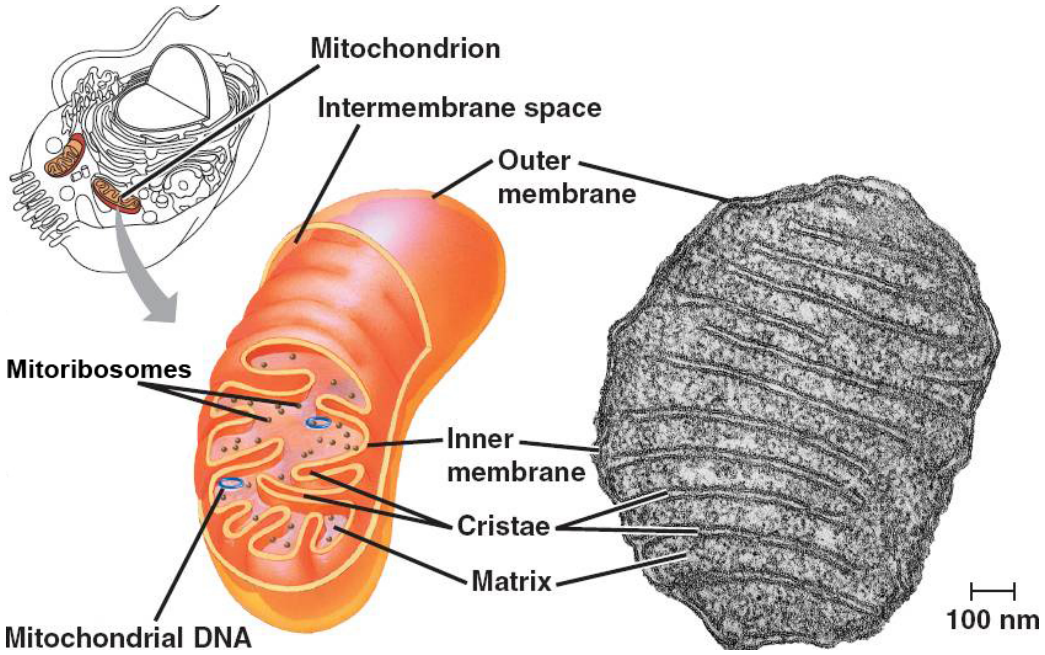


Figure 1-1. The structure of mitochondria
A mitochondrion is a double-membrane bound organelle. The space between the two membranes is referred to as the intermembrane space. The inner membrane has numerous invaginations known as cristae. The cristae serve to increase the exposed surface area between the intermembrane space and the matrix. The matrix contains the mitochondrial DNA and ribosomes (mitoribosomes). On the right, an electron micrograph of a mitochondrion is shown. Figure adapted from: Campbell, N.A.; Reece, J.B.; Biology, 6th ed. San Francisco; Benjamin Cummings, 2002.

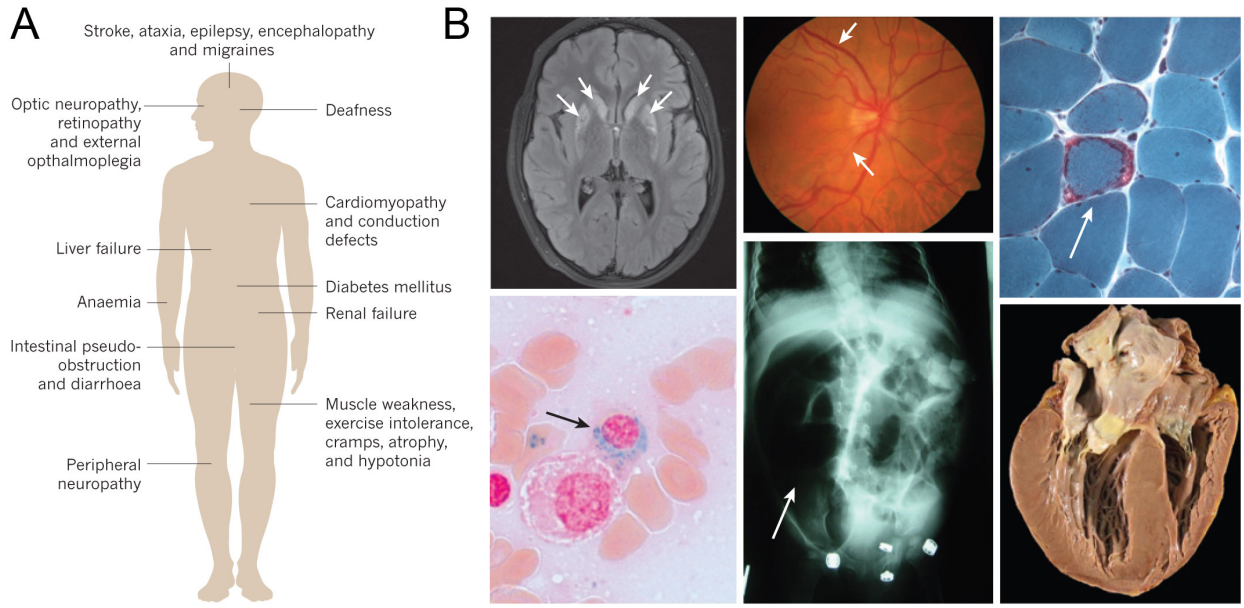


Figure 1-2. Phenotypic spectrum of mitochondrial disorders
A. Common clinical manifestations of mitochondrial disorders. **B.** Clinical images depicting pathology from patients with a variety of mitochondrial disorders. Clockwise from top left, 3-Brain MRI demonstrating Leigh syndrome lesions, which appear as bilateral enhancing lesions in the basal ganglia; retinal image of Leber's hereditary optic neuropathy, demonstrating swollen nerve fibers, as well as engorged and obscured blood vessels (arrows); ragged red fiber (arrow) seen on a modified Gomori-trichrome-stained skeletal-muscle section; anterior four-chamber cross-section of a heart that shows signs of hypertrophic cardiomyopathy, including cardiomegaly and asymmetrical septal hypertrophy; plain abdominal radiograph, showing severe bowel distention (arrow) in the setting of chronic intestinal pseudo-obstruction; and bone-marrow aspirate sample that has been stained for iron demonstrates a ringed sideroblast (arrow), or halo of iron-laden mitochondria around the nucleus of an erythrocyte precursor, which is found in patients with myopathy, lactic acidosis and sideroblastic anemia syndrome. Figure adapted from [3].

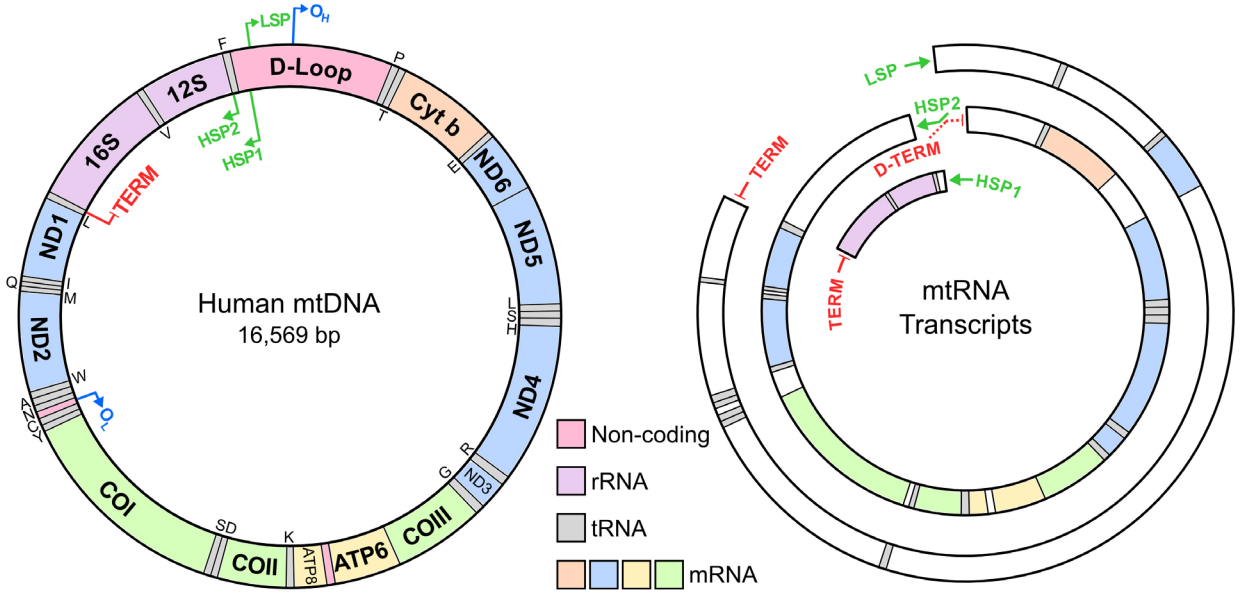


Figure 1-3. Human mtDNA: genes, transcripts, and *cis*-acting elements
Transcription of the L-strand initiates from a single site, LSP, while transcription of the H-strand is initiated from two sites, HSP1 and HSP2. The MTERF1 protein binds to a site in the $tRNA^{Leu}$ gene (TERM) and promotes termination of transcription initiated from LSP and HSP1. The putative termination site for HSP2 transcription (D-TERM) is also shown, although the termination factor acting at this site (if any) has not been identified. The tRNA genes encoded on each of the two strands are indicated with the standard one-letter symbols for amino acids. Abbreviations: COI, cytochrome c oxidase subunit I; COII, cytochrome c oxidase subunit II; COIII, cytochrome c oxidase subunit III; Cytb, cytochrome b; LSP, light-strand promoter; HSP, heavy-strand promoter; ND1-6, NADH dehydrogenase subunits 1-6; O_H, origin of H-strand DNA replication; O_L, origin of L-strand DNA replication.

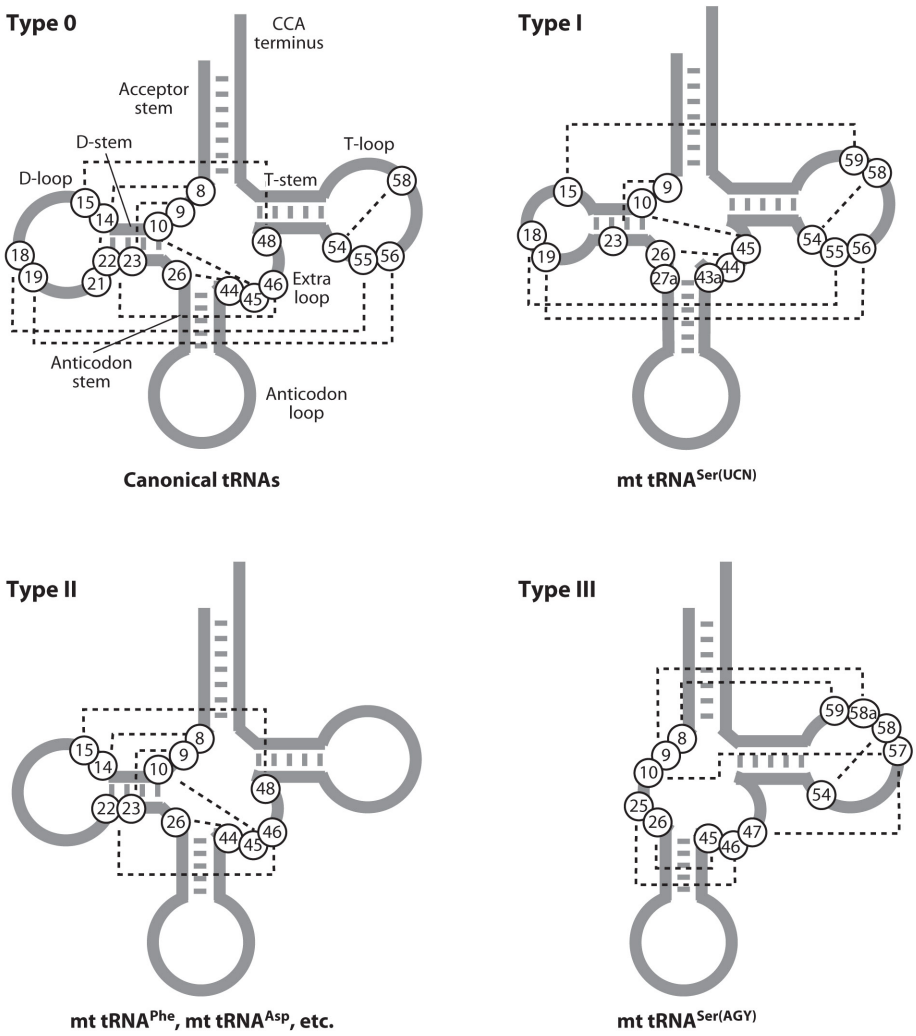


Figure 1-4. Secondary structure diagram of human mitochondrial tRNAs
Canonical ("cloverleaf") tRNA is represented as Type 0. Three types of mt-tRNAs are shown: types I, II, and III. Circled numbers represent the nucleotide positions according to the tRNAdb numbering system [120]. Tertiary interactions between nucleobases are indicated by dotted lines. Figure adapted from [75].

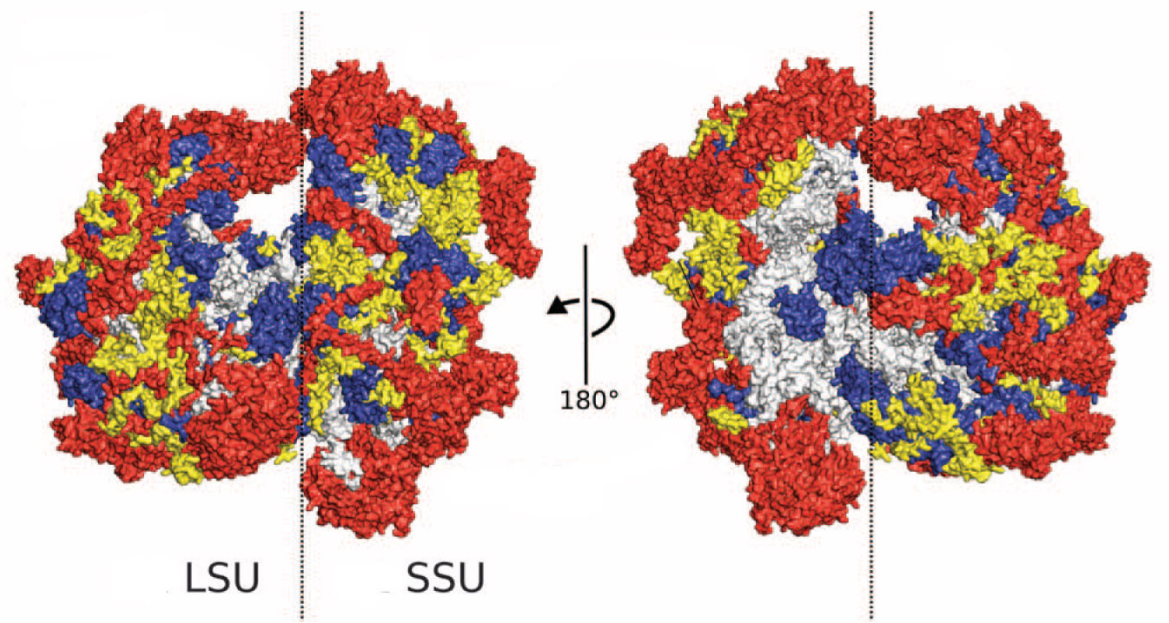


Figure 1-5. Overview of the 55S human mitoribosome
A. Proteins conserved with bacteria (blue), extensions of homologous proteins (yellow), and mitochondria-specific proteins (red). rRNA is shown in gray. Figure adapted from [98].

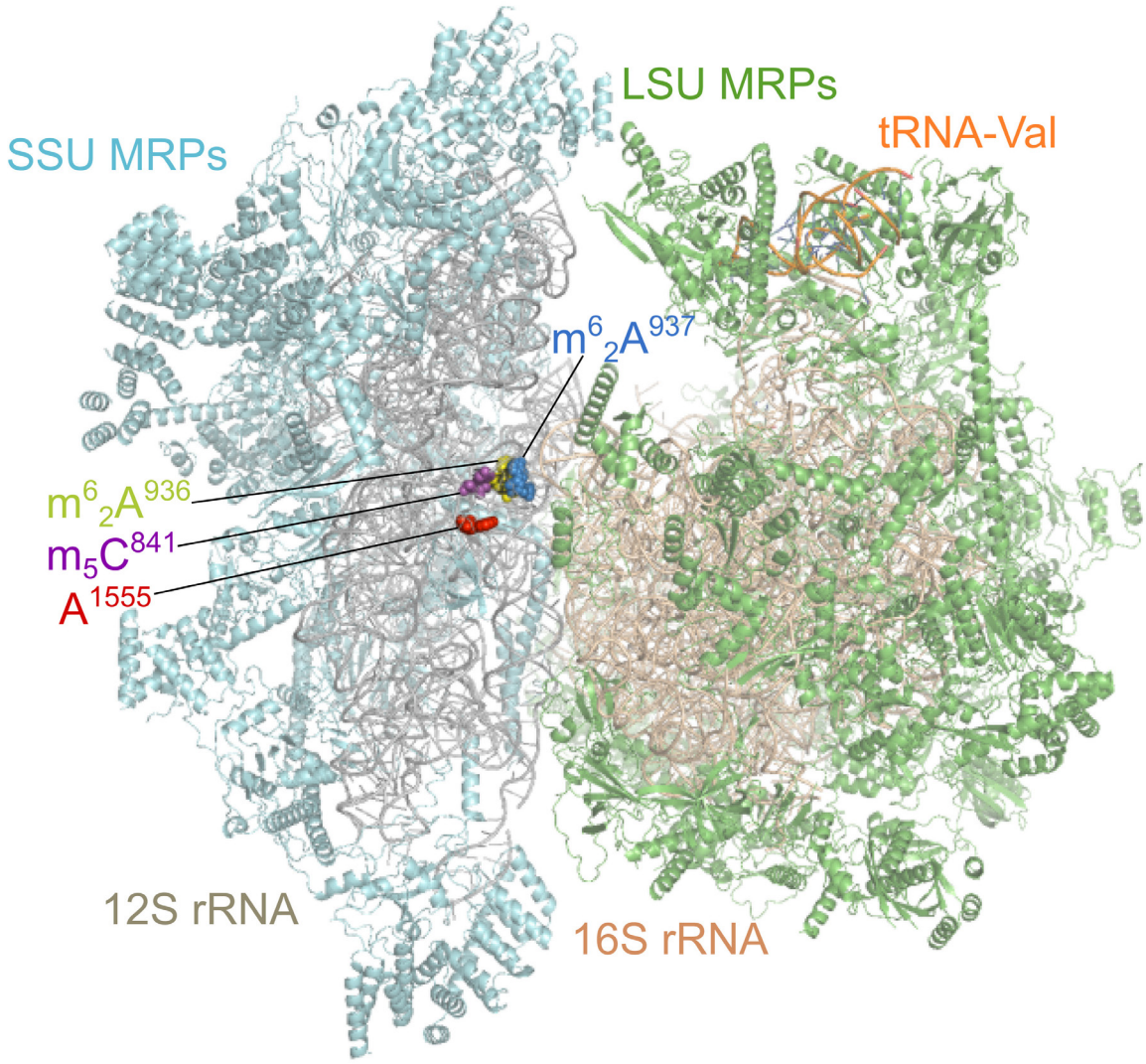


Figure 1-6. Location of the NSUN4 and TFB1M modification sites, and A1555G on the mitoribosome

The residues methylated by NSUN4 and TFB1M are in intimate contact on the 55S mitoribosome, and all three are in close proximity to the mRNA channel [98]. The residues methylated by TFB1M are also in proximity to, but not in contact with, the residue at position 1555 on the 12S rRNA. A mutation at this position, A1555G, has been associated with maternally inherited deafness and hypermethylation activity of TFB1M in a mouse model [96]. The 12S and 16S rRNAs are depicted as grey and tan cartoons, respectively. The residue methylated by NSUN4 is depicted as purples spheres, while the residues methylated by TFB1M are depicted as blue and yellow spheres. Proteins of the SSU and LSU are depicted as cyan and green cartoons, respectively. A tRNA-Val that is a component of the LSU is depicted as an orange cartoon.

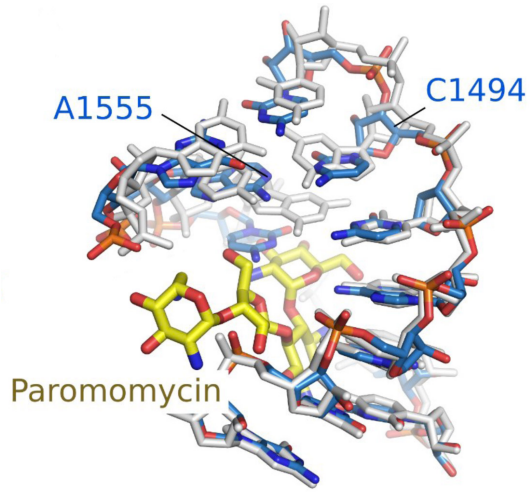


Figure 1-7. The aminoglycoside-binding site in the human mitoribosome
Paromomycin is a member of the aminoglycoside class of antibiotics. It binds adjacent to the A site of the decoding center of the bacterial rRNA (grey) and induces codon-anticodon misreading and prevents translocation of mRNA:tRNA to the P site. The binding site in the human mt-rRNA (blue) is more flexible due to the loss of a G:C base pair between the 1555 and 1494 positions. An A1555G mutation, which would reintroduce the G:C base pair and result in a less flexible, more prokaryotic-like conformation has been implicated in aminoglycoside hypersensitivity. Figure adapted from [98].

Table 1-1. Positions of rRNA modifications relative to the mtDNA and rRNA sequences
The numbering of nucleotide positions in 12S and 16S rRNA are relative to the first nucleotide at the 5' end of the mature transcript. The nucleotide positions of the human and mouse mtDNA sequences are numbered according to NCBI NC_012920 and NC_005089, respectively.

	Human			Mouse		
Modification	**12S rRNA**	**16S rRNA**	**mtDNA**	**12S rRNA**	**16S rRNA**	**mtDNA**
m^4C	839		1,486	840		909
m^5C	841		1,488	842		911
$m^6{}_2$	936, 937		1,583 1584	937, 938		1,006 1,007
Gm		1,145 1,370	2,815 3,040		1,160 1,389	2,253 2,482
Um		1,369	3,039		1,388	2,481
pseudouridine		1,397	3,067		1,416	2,509

Chapter 2: Materials and methods

2.1 Human and murine TFB1M, and *E. coli* KsgA

2.1.1 Protein cloning, expression, and purification

Human (residues 20-396; UniProt Q9H5Q4) and mouse TFB1M (residues 1–346; UniProt Q8WVM0) were cloned into pTEV-HMBP3, allowing expression of a fusion with His-tagged maltose binding protein (MBP) cleavable by TEV protease. TFB proteins were overexpressed in Arctic Xpress *E. coli* (DE3) cells (Stratagene) at 16° C for 20 hours, and purified using ProBond Resin (Invitrogen), followed by overnight TEV protease cleavage, Heparin and Superdex 200 size-exclusion chromatography. *E. coli* KsgA (residues 1-273; UniProt P06992) was purified as previously described [121]. A selenomethionine (SeMet; Sigma-Aldrich) protein derivative of TFB1M (SeMet-TFB1M) was produced by adding SeMet and the other 19 amino acids to minimal medium, and purified as for the native protein. Proteins were concentrated using a 10,000 MWCO Amicon Ultra-15 device. Concentrated proteins were stored in 20 mM HEPES (pH 8.0), 200 mM KCl, 5% glycerol, and 1 mM DTT.

2.1.2 Crystallization and structure determination of TFB1M

Orthorhombic (*I*222) crystals with one molecule per asymmetric unit were obtained by hanging-drop vapor diffusion (4 µl at 26 mg/ml) at 4 °C for SeMet-TFB1M (resolution 2.2 Å), HighRes-TFB1M (resolution 1.8 Å), and TFB1M:SAM (resolution 2.1 Å) with 2.9 M sodium acetate, pH 7.0 as reservoir solution. TFB1M:SAM crystals were subsequently soaked in mother liquor containing 10 mM *S*-adenosyl-L-methionine for 10 hours. Diffraction data for all crystals was collected on beamline X29 at the National Synchrotron Light Source (BNL, Upton, New York). All datasets were processed using XDS [122] and SCALA [123] as implemented in autoPROC [124]. The SeMet-TFB1M structure was determined with multiwavelength anomalous dispersion phasing [125] using intensities measured at the Se-absorption peak, inflection, and high-energy remote wavelengths (Table 3-1). All six selenium sites were located and initial phases were determined using SOLVE [126], followed by density modification using

RESOLVE [127]. Manual model building in COOT [128] using the 2.2 Å experimental map generated a nearly complete model. Refinement was carried out in PHENIX [129], and the resulting model was subsequently refined against a second higher resolution (1.8 Å) data set (HighRes-TFB1M). Data collection and refinement statistics are summarized in Table 3-1. R_{work} converged to 18.4% and R_{free} to 22.4%. The final model includes residues 10–328 of TFB1M, and no Ramachandran outliers, as assessed by MOLPROBITY [130]. The TFB1M:SAM structure was solved with molecular replacement, using the HighRes-TFB1M structure as a model. Manual model building and adjustment was carried out in COOT [128], followed by refinement in PHENIX [129]. R_{work} and R_{free} converged to 20.5% and 23.4%, respectively. The final model for TFB1M:SAM includes residues 10-327 of TFB1M, one bound SAM molecule, and no Ramachandran outliers as assessed by MOLPROBITY [130].

Anisotropic diffraction of both TFB1M and TFB1M:SAM crystals was addressed by carrying out anisotropic scaling and ellipsoidal truncation using the Diffraction Anisotropy Server [131]. Briefly, data residing outside an ellipse centered at the reciprocal lattice origin and having vertices at 1/1.8, 1/2.7, and 1/1.8 $Å^{-1}$ for TFB1M or 1/2.1, 1/2.7, 1/2.1 $Å^{-1}$ for TFB1M:SAM along a*, b* and c*, respectively, were removed. This treatment resulted in improved refinement statistics and electron density maps for both models.

2.1.3 Kasugamycin sensitivity assay

The *ksgA* gene of the *E. coli* MG1655 strain was disrupted by P_1 transduction of the *ksgA*::Kan cassette from the Keio single gene deletion library [132]. The full-length *ksgA* gene was amplified from *E. coli* MG1655 genomic DNA and cloned into the pBAD24 expression vector. The KsgA Y116A variant was generated by PCR based site-directed mutagenesis and confirmed by sequencing. Plasmids encoding the human mitochondrial TFB1M (hsTFB1M) or the mouse mitochondrial TFB1M (mmTFB1M) and its variants were expressed in MG1655 *ksgA*::Kan strain in the presence of the pRARE plasmid. Antibiotics Ampicillin (100 μg/ml), Kanamycin (50 μg/ml) and Chloramphenicol (30 μg/ml) were added to the growth medium.

E. coli ΔksgA cells with or without plasmid expressing KsgA, mmTFB1M, hsTFB1M, or their indicated variants were grown overnight at 37°C in Luria Bertani (LB) broth with appropriate antibiotics. The OD_{600} of overnight cultures was adjusted to 3.5 and diluted 1:400 in LB broth containing 0.1% arabinose and appropriate antibiotics. Freshly prepared kasugamycin

(Santa Cruz Biotechnology) was added to the broth to a final concentration of 400 μg/ml. The cells were allowed to grow at 37° C and OD_{600} of the cultures was measured after 5 hours.

All experiments were performed in triplicate and analyzed for statistical significance using a two-tailed t-test in Prism (GraphPad Software).

2.2 Human MTERF4 and NSUN4

2.2.1 Protein cloning, expression, and purification

Wild-type MTERF4 (residues 66–332) and NSUN4 (residues 26 –384) were cloned into pTEV-HMBP3, allowing expression of a fusion with his-tagged maltose binding protein (MBP) that is cleavable with TEV protease. Proteins were overexpressed in Arctic Xpress (DE3) *E. coli* cells (Stratagene) at 16°C for 20 hr. MTERF4 and NSUN4 were purified using ProBond Resin (Invitrogen), followed by overnight TEV protease cleavage and Heparin chromatography. Subsequently, ProBond Resin was used to remove TEV. The MTERF4-NSUN4 complex was created by mixing the two proteins in equimolar concentration after the Heparin column. All proteins were subjected to gel filtration on a Superdex 200 16/60 column (GE Healthcare) in a buffer containing 20 mM Hepes, pH 8.0, 2.5% glycerol, 0.5 M KCl, 1mM DTT and 1 mM EDTA. The gel filtration profile shows that MTERF4 and NSUN4 form a complex of 70-80kDa.

SeMet-substituted protein was expressed in the Arctic Xpress (DE3) *E. coli*, which is not auxotrophic for methionine. Methionine biosynthesis was inhibited by growth conditions as described previously (Van Duyne et al., 1993). The protein was subsequently purified as described above. Proteins were concentrated using a 10,000 MWCO Amicon Ultra-15 device up to 20mg/ml. Concentrated proteins were stored in 20 mM HEPES (pH 8.0), 300 mM KCl, 2.5% glycerol, and 1mM DTT.

2.2.2 Crystallization and structure determination

MTERF4:NSUN4 crystals were grown by hanging drop vapor diffusion at room temperature in 4 μL drops of a 1:1 ratio of protein and crystallization solution (0.1-0.2M Mg Formate and 50 mM Bis-Tris, pH 5.0-6.0) in the presence of SAM. Crystals were cryoprotected with 35% ethylene glycol and cryo-cooled in liquid nitrogen. Diffraction data were collected on beamlines X25 and X29 of the National Synchrotron Light Source (Upton, NY) at the peak,

inflection point and high-energy remote wavelengths of the K edge of selenium. Datasets were processed using XDS (Kabsch, 2010) and Scala as implemented in the autoPROC pipeline (Vonrhein et al., 2011). All seven selenium sites in NSUN4 were identified with SHELX (Sheldrick, 2008) and phases were determined via Multiwavelegth Anomalous Dispersion (Hendrickson, 1997). Phases were calculated to 2.0 Å using SHARP (Vonrhein et al., 2007) and automated modeling building with ARP-wARP (Langer et al., 2008) produced a starting model consisting of 75% of NSUN4 and 30% of MTERF4. Manual model building was carried out in COOT (Emsley and Cowtan, 2004) using the experimental map and the complete final model was refined with PHENIX (Terwilliger, 2002). Model quality was assessed using MOLPROBITY (Davis et al., 2007). Data collection and refinement statistics are shown in Table 4-1.

2.2.3 Binding measurements

Isothermal titration calorimetry (ITC) experiments were performed with a VP-ITC calorimeter (Microcal). NSUN4 (20-25 μM) was titrated with 10 μl injections of 200–250 μM MTERF4. Samples were prepared by dialyzing all interacting components against a buffer containing 20 mM HEPES, pH 8.0, 200 mM KCl, 2.5% glycerol, 1 mM EDTA. All runs were made at a constant stirring speed of 310 rpm, and all experiments were performed at 25 °C. Data were analyzed using ORIGIN (OriginLab).

2.2.4 In vitro methylation assay

A 1.5kb fragment of the human 16S mitochondrial DNA was cloned through PCR of HeLa mitochondrial DNA in the pBlueScriptII SK+ vector between the XhoI and EcoRI sites. 16S ribosomal RNA was transcribed using T7 RNA polymerase using DNA linearized with EcoRI as a template. The corresponding antisense RNA was transcribed using T3 RNA polymerase using DNA linearized with XhoI. The resulting products were purified using the RNeasy Mini Kit (Qiagen). *In vitro* methylation assays were carried out according to a published procedure (Kuratani et al., 2010) with some modifications. ^{3}H-SAM (MP Biomedical) was used as the methyl donor in the reaction. The methylation reaction was carried out in 100μl and contained 25mM Hepes pH 8.0, 50mM KCl, 0.1mM EDTA, 5mM $MgCl_2$, 1mM DTT, 2.5%

glycerol, 5 pmol RNA and 70 pmol SAM (78Ci/mmol). The reaction was initiated by adding the enzyme (0.5-5 pmol). A time course with different enzyme concentrations was carried out to identify linear conditions. Final experiments were carried out using 2 pmol of enzyme (NSUN4, MTERF4 or MTERF4:NSUN4 holoenzyme). Reactions with no enzyme were performed as a control. After 30 min, samples were transferred to 5ml of 10% trichloroacetic acid (TCA). The acid-insoluble fraction was then transferred to filter paper (Whatman), washed 3 times with 5ml of ice-cold 10% TCA and with 100% ethanol. The radioactivity on each filter disk was measured on a scintillation counter (Beckman Coulter). The amount of methyl groups incorporated into the RNA was calculated from the measured decays per minute (DPM). Experiments were independently repeated at least three times.

2.3 Human MTERF1

2.3.1 Protein cloning, expression and purification

Wild-type human MTERF1 (Uniprot Q99551; residues 57–399), TFAM (UniProt Q00059; residues 43–246), TFB2M (UniProt Q9H5Q4; residues 20–396) and POLRMT (UniProt O00411; residues 42–1230) were expressed as fusions with histidine-tagged maltose binding protein and purified as described [133].

All MTERF1 mutants were constructed by QuikChange mutagenesis (Stratagene). His-tagged TFAM was cloned in pET-22. Proteins were overexpressed in Arctic Xpress E.coli (DE3) cells (Stratagene) at 16°C for 20 hr. WT and mutant MTERF1 proteins and TFB2M were purified using ProBond Resin (Invitrogen), followed by overnight TEV protease cleavage, Heparin and Mono S chromatography. POLRMT was purified by Heparin chromatography, overnight TEV protease cleavage, second Heparin and Mono S chromatography. TFAM was purified using ProBond Resin, Heparin and Mono S chromatography. Proteins were concentrated using a 10,000 MWCO Amicon Ultra-15 device. Concentrated proteins were stored in 20 mM HEPES (pH 8.0), 200 mM KCl, 5% glycerol, and 1mM DTT.

Human TFAM, TFB2M, and POLRMT were expressed as fusions with his-tagged maltose binding protein and purified as described [103]. TFAM and TFB2M were purified using ProBond Resin (Invitrogen), followed by overnight TEV protease cleavage, Heparin, Mono S, and size exclusion chromatography. POLRMT was purified by Heparin chromatography,

overnight TEV protease cleavage, followed by a second Heparin, Mono S, and size exclusion chromatography. Proteins were concentrated using a 10,000 MWCO Amicon Ultra-15 device. Concentrated proteins were stored in 20 mM HEPES (pH 8.0), 200 mM KCl, 5% glycerol, and 1 mM DTT.

2.3.2 Transcription termination assays

Termination assays were carried out as previously described, with some modifications [103, 134]. Nucleotides 491-790 of the human mitochondrial DNA sequence (containing the HSP) were cloned between the NcoI and HindIII sites in a pET-22 vector (Novagen). 100bp downstream of the HSP the MTERF1 DNA binding sequence was inserted in reverse orientation with respect to the HSP (5′-TAAGATGGCAGAGCCCGGTAAT-3′). Plasmid DNA substrate was linearized by digestion with HindIII at 37°C for 3 hours. *In vitro* runoff transcription termination assays were carried out in 20 µL total volume and included 20 mM HEPES (pH 8.0), 80 mM KCl, 10 mM $MgCl_2$, 0.4 nM DNA template, 5 mM DTT, 1 mM EDTA and 5% glycerol. rNTPS were used at the following concentrations: 0.4mM ATP, 0.15mM CTP,GTP and 0.1mM UTP, 1 µL of [α-^{32}P]UTP (3000 µCi/mmole) was used. MTERF1 was diluted to final concentrations in a low salt dilution buffer (PDB buffer). Aliquots of a mix containing labeled UTP, rNTPs, digested DNA and transcription buffer were mixed with 2 µL of MTERF1 at 0.1 µM, 0.2 µM, 0.4 µM and incubated at room temperature for 30 minutes. A control experiment did not include MTERF1 but used 2 µL of PDB buffer. The reaction was initiated by adding 3 µL of a mix for a final concentration of TFAM (0.375 µM), TFB2M (0.15 µM) and POLRMT (0.15 µM) and incubated at 32°C for 30 minutes. The reaction was stopped by adding 100 µL of a stop solution containing 1% SDS, 20 mM EDTA, 200 mM sodium acetate, 0.12 mg/mL glycogen followed by ethanol precipitation. The precipitated RNA was run on a 5% Urea-PAGE gel in a loading buffer containing urea, BB, XyleneFF and run at 200 V for 65 minutes. Gels were dried and exposed to a phosphorimager screen (Amersham Biosciences). Scanning was done using a Typhoon FLA9000 (GE Healthcare) followed by analysis in ImageQuant TL Toolbox Version 7.0 (GE Healthcare). Statistical analysis of termination activity was carried out with Prism (GraphPad Software) based on the results of at least three independent experiments.

2.3.3 Crystallization and structure determination

Oligonucleotides were synthesized, annealed and combined with the protein in a 2:1 ratio. Crystals grew at 4 C in a solution derived from Index screen (Hampton Research) condition 88: 200 mM ammonium acetate tribasic (pH 5.0), 18% (w/v) polyethylene glycol 3350, 5.5 mM HEPES, 10 mM MgCl2. X-ray diffraction data were collected on beamline X25 of the National Synchrotron Light Source at Brookhaven National Laboratory (Upton, NY) at 100 K using a wavelength of 1.1 Å. Data were processed using XDS and Aimless as implemented in the autoPROC pipeline [122, 124, 135]. Phases were obtained by molecular replacement using the structure of wild-type human MTERF1 bound to the termination sequence (PDB 3MVA) as the search model in Phaser [103, 136]. Model building was carried out in Coot, followed by refinement in Phenix [137, 138]. The geometric quality of the refined models was assessed with MolProbity and the structure validation tools in the Phenix suite [139]. Data collection and refinement statistics are show in Table 5-1.

Chapter 3: Structural basis for *S*-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1

This chapter has been published as **Guja KE**, Venkataraman K, Yakubovskaya E, Shi H, Mejia E, Hambardjieva E, Karzai AW, Garcia-Diaz M. (2013). Structural basis for S-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1. *Nucleic Acids Research.* 41(16):7947-59. PMID: 23804760

Author contributions: KEG and MGD conceived of the study; VK and AWK designed and carried out the complementation assay; KEG, EY, HS, EM, and EH carried out cloning, expression and purification; KEG carried out crystallization, data collection, experimental phasing, refinement and analysis of the crystal structure; KEG drafted the initial manuscript and prepared all the figures; MGD, VK, and AWK helped revise the manuscript.

3.1 Abstract

Mitochondrial transcription factor B (TFB) proteins are homologous to KsgA/Dim1 rRNA methyltransferases. The mammalian TFB protein, TFB1M, is an essential protein necessary for mitochondrial gene expression. TFB1M mediates an rRNA modification in the small ribosomal subunit and thus plays a role analogous to KsgA/Dim1 proteins. This modification has been linked to mitochondrial dysfunctions leading to maternally inherited deafness, aminoglycoside sensitivity, and diabetes. Here we present the first structural characterization of the mammalian mitochondrial TFB1 factor. We have solved two X-ray crystallographic structures of TFB1M with (2.1 Å) and without (2.0 Å) its cofactor S-adenosyl-L-methionine (SAM). These structures reveal that TFB1M shares a conserved methyltransferase core with other KsgA/Dim1 methyltransferases and shed light on the structural basis of SAM binding and methyltransferase activity. Together with mutagenesis studies, these data suggest a model for substrate binding and provide insight into the mechanism of methyl transfer, clarifying the role of this factor in an essential process for mitochondrial function.

3.2 Introduction

Mitochondria are cellular organelles responsible for the bulk of eukaryotic cellular energy production via oxidative phosphorylation. This process is strictly dependent on the coordinated expression of nearly one hundred proteins that constitute the respiratory chain, which are encoded in either the nuclear or mitochondrial genome [140]. Deficiencies in mitochondrial gene expression and energy production have been implicated in a variety of genetic disorders, age-related chronic diseases and the aging process itself [141]. Such deficiencies may be caused by mutations in mtDNA-encoded proteins, tRNAs, or rRNAs as well as by mutations in nuclear-encoded respiratory chain components or regulatory factors that control mitochondrial transcription and translation [142-145]. Given the role of mitochondrial dysfunction in human disease pathology, unraveling the complexities of mitochondrial gene expression that allow for coordinated cellular energy production is of significant interest.

A critical requirement for mitochondrial gene expression is proper ribosome biogenesis. This process, which is not well understood, involves several posttranscriptional RNA modifications that are thought to be crucial for ribosome assembly. Recent evidence has highlighted the importance of these modifications in mitochondria [92, 146]. Although several modification sites are known in both the 12S and 16S rRNAs [147], the enzymology responsible for the modifications remains unknown, with the exception of a conserved methylation in the small ribosomal subunit [148].

KsgA and Dim1 methyltransferases dimethylate two adjacent adenine residues in a stem-loop close to the 3′ end of the small-subunit rRNA in prokaryotes and eukaryotes, respectively [149, 150]. This modification is highly conserved in nature and is one of the few that is present in all but a few organisms [151-153]. Accordingly, both the stem-loop structure and methylation events are conserved in the human 12S mitochondrial rRNA. Seidel-Rogol and colleagues elegantly demonstrated that TFB1M (transcription factor B1, mitochondrial), a member of the KsgA/Dim1 family of methyltransferases, catalyzes the analogous rRNA modification in mitochondria [148]. Methylation by KsgA is related to bacterial aminoglycoside sensitivity and TFB1M activity could similarly be related to the effect of aminoglycoside antibiotics on the mitochondrial ribosome and aminoglycoside-induced hearing loss [154]. Consistently, aminoglycoside sensitivity in humans is often maternally inherited [155, 156]. Furthermore, TFB1M modulates the effects of a pathogenic mtDNA mutation linked to deafness [141, 157].

Hypermethylation of the 12S rRNA has been shown to underlie these effects, which were recapitulated in a mouse model [96]. Other TFB1M polymorphisms are associated with reduced insulin secretion and increased risk of type II diabetes mellitus [158], further highlighting the importance of this modification for normal mitochondrial function.

Interestingly, metazoan cells contain two TFB (transcription factor B) factors, TFB1M and TFB2M (transcription factor B2, mitochondrial). TFB2M is an essential mitochondrial initiation factor [159, 160], but despite potentially having methyltransferase activity [95] does not appear to be responsible for dimethylation of the stem-loop in the 12S mitochondrial rRNA [92]. Although the extent to which TFB1M participates in transcription is controversial, it has been reported to interact with both the mitochondrial RNA polymerase, POLRMT [159], and mitochondrial transcription factor A (TFAM) [161]. The relevance of these interactions for the transcription process remains unknown [159, 162]. However, it has recently been reported that TFB1M can associate with POLRMT in a transcription-independent manner, and that this interaction can modulate the activity of TFB1M on the mitochondrial 12S rRNA [163]. Hence, TFB1M might provide a regulatory link between transcriptional regulation and ribosome biogenesis [93].

Here we report the first x-ray crystallographic structures of the mammalian TFB1M factor. Our results demonstrate the conservation of the KsgA-like fold in TFB1M, characterize the structural basis for SAM binding, and provide structural insight into the mechanism of rRNA methylation, helping to clarify how this protein mediates a critical process for mitochondrial gene expression.

3.3 Results and discussion

3.3.1 Human and mouse TFB1M are functionally conserved

In order to investigate how the putative methyltransferase fold of TFB proteins contributes to the mitochondrial gene expression process we decided to crystallize full-length TFB1M. Despite developing a robust purification strategy for human TFB1M (hsTFB1M), all crystallization attempts were unsuccessful. We then considered characterizing TFB1M from an alternative mammalian species. Mouse TFB1M (mmTFB1M) is essential for mitochondrial function and mice deficient in TFB1M display a phenotype consistent with its role as a 12S

rRNA methyltransferase [92]. Moreover, mmTFB1M shows a high degree of sequence conservation (85% identity with the human protein; Figure 3-7). We therefore decided to determine if the mouse protein also displays rRNA methyltransferase activity and if this activity is comparable to that of human TFB1M.

In *E. coli,* loss of KsgA activity confers resistance to the translation inhibitor antibiotic kasugamycin [164]. Consistent with the evolutionary conservation of this modification, hsTFB1M was shown to be able to functionally complement an *E. coli ksgA*$^{-}$ mutation [148]. We therefore employed the same complementation strategy to study mmTFB1M. As reported previously, expression of human TFB1M restored sensitivity to kasugamycin in a *ksgA*$^{-}$ strain (Figure 3-1A). Likewise, expression of wild-type mmTFB1M resulted in significantly increased sensitivity to kasugamycin compared with a control expressing wild-type KsgA, suggesting that mouse and human TFB1M have comparable methyltransferase activities. Interestingly, mmTFB1M displayed stronger complementation activity than the human protein. Although this might be due to differential stability of the protein in bacteria, it could also be a consequence of sequence differences between the human and mouse small rRNA hairpins (Figure 3-1B).

3.3.2 Overall structure of ligand-free TFB1M

Since we confirmed that mmTFB1M is a suitable model for studying the human protein, we decided to determine its crystal structure. We were able to obtain crystals of mmTFB1M that diffracted to high resolution. In order to facilitate phasing of the structure, we expressed and purified selenomethionine-substituted mmTFB1M protein from *E. coli* and obtained crystals that diffracted to a resolution of 2.0 Å (Table 3-1). We solved the structure using multiwavelength anomalous dispersion phasing methods [125]. The resulting electron density map was of sufficient quality to permit building of a structural model for almost all of the protein (residues 10-328; residues 1-9 and 329-346 are not resolved in the electron density). This likely includes a portion of the mitochondrial targeting peptide, as the first thirteen residues of mmTFB1M are predicted to be cleaved upon import into the mitochondria [165]. Our X-ray crystallographic structure (Figure 3-2A) represents the first crystal structure of a mammalian TFB protein, and definitively confirms that TFB1M shares a common fold with rRNA adenine methyltransferases. The fold displays a typical two-domain architecture (Figure 3-2B). The larger N-terminal domain (residues 32-236; blue in Figure 3-2A) that contains the putative active site forms a canonical

Rossman-like methyltransferase fold with a central seven-stranded beta-sheet flanked by three α-helices on each side (Figure 3-2C). Two additional N-terminal α-helices (α2 and α3) and a loop region that is inserted between strands β6 and β7 define the active-site region (Figure 3-2D). This fold is characteristic of SAM-dependent methyltransferases [166]. The smaller C-terminal domain spans residues 237-328 (yellow in Figure 3-2A) and consists of five α-helices (α8 to α12). TFB1M contains an additional N-terminal extension beyond the predicted methyltransferase fold. The N-terminal extension (residues 14-31; pink in Figure 3-2A) includes a protruding N-terminal α-helix (residues 18-31).

Comparison of the TFB1M structure with other methyltransferases reveals that indeed the N-terminal methyltransferase core domain is relatively well conserved: RNA and DNA methyltransferases as well as small molecule methyltransferases all contain a central beta sheet surrounded by two pairs of three α-helices (Figure 3-2E). However, different methyltransferases contain differential additions to this conserved fold. These additions are located in very different positions with respect to the central core. These embellishments are likely related to the substrate specificity of the enzymes (Figure 3-2F) [166]. For instance, TFB1M contains a C-terminal addition that is also present in ErmC′, another adenine N6-specific RNA methyltransferase. Similarly, the catechol methyltransferase (COMT) contains the smallest of these additions, consistent with the small size of its substrate: the catechol molecule presumably does not require additional contacts outside the binding pocket in the core methyltransferase domain [167].

In addition to its C-terminal lobe, TFB1M also appears to contain an additional N-terminal α-helix that is not part of the conserved methyltransferase core or the predicted mitochondrial localization sequence. Given the peripheral position of this helix with respect to the overall fold, it is tempting to speculate that this tail might be involved in the establishment of putative protein-protein or protein-RNA interactions to help localize TFB1M to the appropriate site in the small ribosomal subunit.

3.3.3 A conserved pocket for SAM binding

Unlike KsgA, which does not bind SAM in the absence of the RNA substrate [132, 168, 169], TFB1M has been reported to directly associate with the SAM cofactor [161]. In an effort to identify the structural features important for recognition and binding of the putative cofactor SAM and substrate adenine, we attempted co-crystallization and apoenzyme crystal soaks with

S-adenosyl-L-methionine (SAM), *S*-adenosyl-L-homocysteine (SAH), 5′-methylthioadenosine (MTA; the major hydrolysis product of SAM), and adenosine monophosphate (AMP). The crystal soaks with SAM were successful and we obtained co-crystals that diffracted to a resolution of 2.1 Å. We solved the structure of the TFB1M:SAM complex using molecular replacement with the apo-TFB1M structure as a search model. The final electron density allowed building a structural model for most of the protein (residues 10-328) and the unambiguous placement of the bound SAM molecule (Figure 3-3). Data collection and refinement statistics are summarized in Table 3-1.

SAM is bound in a conserved acidic binding pocket in the N-terminal methyltransferase domain (Figure 3-3 and Figure 3-8). The structure allows us to identify several active site residues that stabilize the bound SAM molecule by a variety of hydrogen bonding and van der Waals interactions (Figure 3-3B). The N6 of the adenosine moiety makes a hydrogen bond to Asp 111, and the adenine base is also bracketed by van der Waals interactions with two adjacent residues, Lys 86 and Val 112. In addition, Glu85 establishes a bidentate hydrogen bonding interaction with both hydroxyl groups of the ribose. Two hydrogen-bonding interactions exist between Glu 61, Asn 141 and the amino group of the methionine moiety. Finally, the side chain carbonyl oxygen (OE1) of Gln 35 is in position to interact with the positively charged sulfur atom of the cofactor. These interacting residues are among the most highly conserved in the KsgA family (Figure 3-7) [170].

The fold of the SAM binding pocket is extremely well conserved among different methyltransferases (Figure 3-3C). Interestingly, this conservation is more marked in the regions of the binding pocket that contact the ribose and the adenine ring of the cofactor, while a greater degree of structural divergence is observed in the distal portion of the pocket, surrounding the methionine moiety (Figure 3-3C). Since the methyl group acceptor of the substrate will need to be located in proximity to the methionine moiety, this asymmetry in the structural conservation of the pocket is likely related to the structural variability of the different substrates bound by different methyltransferases. Furthermore, the side chain interactions to the ribose and base of the nucleotide appear to be well conserved, in particular a highly conserved glutamate that interacts with the 2′ and 3′ ribose oxygens (Glu 85 for TFB1M) and a conserved interaction with the N6 atom of the adenine ring (Asp 111 for TFB1M).

3.3.4 Conformational change upon ligand binding

In order to determine whether SAM binding resulted in a conformational change we decided to compare the SAM-bound and apo-TFB1M structures. The SAM-bound TFB1M structure is essentially identical to the ligand-free structure (RMSD of 0.26 Å for 318 C-α atoms; Figure 3-3A), implying that no large-scale conformational change takes place upon SAM binding. However, closer inspection of the SAM binding pocket reveals that while SAM binding does not induce any major conformational change, slight structural perturbations are observed in the SAM-bound complex. In the apo structure, an acetate molecule is observed in the binding pocket, interacting with residues Gln35 and Glu85 (Figure 3-9). Upon SAM binding, however, the loop containing Gln35 (which connects α-helices 1 and 2) undergoes a shift in its position (see inset in Figure 3-3A), effectively opening the SAM binding pocket in order to accommodate the SAM molecule. This shift affects the position of Gln35, which in the apo structure is in a position that would preclude SAM from binding (Figure 3-3D). This type of subtle rearrangement of loop regions is typically observed in other methyltransferases [166].

3.3.5 TFB1M is homologous to KsgA/Dim1 proteins, but diverges from sc-mtTFB

In order to identify the closest structural homologs to TFB1M we analyzed the TFB1M fold using DALI [171]. The analysis revealed structural homology to KsgA/Dim1 proteins and to the *S. cerevisiae* mtTFB (sc-mtTFB). Perhaps surprisingly, TFB1M exhibits the highest similarity to KsgA methyltransferases, followed by Dim1 proteins, and not with mtTFB. This is consistent with the degree of sequence identity between those proteins (25.6% with KsgA, 22.0% with Dim1 and 13.3% with mtTFB; Figure 3-7) and generally in agreement with previous phylogenetic analyses [95]. Comparison of the TFB1M fold with *E. coli* KsgA (RMSD of 1.26 Å for 192 C-α atoms; Figure 3-4A) stresses the conservation of the methyltransferase N-terminal domain: the 7-stranded β-sheet is fully conserved, as are some of the adjacent α-helices, although several divergent insertions are present in TFB1M. Nevertheless, the conservation is much lower in the C-terminal domain, perhaps emphasizing the differences between the methylation substrates in the *E. coli* and mammalian mitochondrial ribosome. A similar overlay can be produced with human Dim1 (RMSD of 1.72 Å for 211 C-α atoms; Figure 3-4B), again highlighting the conservation of the N-terminal domain, although in this case much more

substantial differences exist in the C-terminal lobe, which is perhaps consistent with the larger differences that likely exist between the substrates of Dim1 and TFB1M. The subtle structural differences that exist between prokaryotic KsgA and human Dim1 might be related to the presence of additional non-methyltransferase functions in the eukaryotic enzyme [172]. Interestingly, sc-mtTFB, although still maintaining a similar fold (RMSD of 2.39 Å for 212 C-α atoms; Figure 3-4C) is the most divergent protein of known structure in the KsgA/Dim1/TFB family. This divergence is not limited to the C-terminal lobe, but is also quite substantial in the N-terminal domain. It is tempting to speculate that these differences may be related to the transcriptional functions of the yeast factor [172]. Other methyltransferases that exhibited structural similarity to TFB1M were enzymes from the Erm family of rRNA methyltransferases [173, 174] and the tRNA methyltransferase Trm14 [175, 176].

Interestingly, inspection of the *E. coli* and *T.thermophilus* KsgA structures [177] suggests that an N-terminal α-helix, similar to that observed in TFB1M, might be present in these enzymes. This α-helix adopts a different conformation in the *T. thermophilus* structure and is not observed in the *E.coli* enzyme, but this might be due to disorder in that region of the crystal structure. It is important to note that, when resolved in the electron density, this α-helix participates in crystal lattice contacts and thus its conformation is likely affected by crystal packing. Hence, a similar extension might in fact be present in other enzymes from this family.

Inspection of the electrostatic surface potential of TFB1M (Figure 3-5A) clearly reveals that the surface of the protein contains a large basic groove and a negatively charged binding pocket (inset in Figure 3-5A) where SAM binds. The large basic groove spans the entire protein fold and is likely responsible for mediating the interaction with the RNA substrate. A similar (although less markedly basic) groove is present in KsgA (Figure 3-5B) and yeast mtTFB (Figure 3-5C), suggesting that the nucleic acid binding mode might be conserved throughout the family, and stressing that the entire fold, including the C-terminal domain, appears to be involved in nucleic acid binding. Notable in the electrostatic surface potential map is the large region of positive charge in the cleft between the N-terminal and C-terminal domains. This region is present in TFB1M, KsgA and sc-mtTFB. Residues on the N-terminal side of this cleft were shown to be important for RNA binding by a similar methyltransferase, ErmC [178]. However, it is interesting to note that some Erm family members, such as the *Mycobacterium tuberculosis* ermMT [179], lack the entire C-terminal domain.

Importantly, the clear acidic SAM binding pocket that can be observed in both TFB1M and KsgA is conspicuously absent in sc-mtTFB (see insets in Figure 3-5). Comparison of the TFB1M and sc-mtTFB structures reveals noticeable differences around the SAM binding pocket. Of the 7 residues in TFB1M that interact directly with SAM (Gln 35, Glu 61, Glu 85, Lys86, Asp 111, Val112 and Asn 141), all appear to be conserved in KsgA (Figure 3-7), human Dim1, and human TFB1M (all proteins with demonstrated methyltransferase activity), while only four are conserved in sc-mtTFB. Moreover, the putative SAM binding pocket in sc-mtTFB contains several bulky side chains that are not present in the known methyltransferases and that are likely to preclude SAM binding. This is in agreement with the fact that the yeast mitochondrial ribosome does not conserve the stem-loop modification catalyzed by KsgA/Dim1/TFB proteins [153] and the observation that sc-mtTFB is unable to complement a KsgA deficiency in *E. coli* [95]. Hence, the divergence of the sc-mtTFB structure is in part likely to be a consequence of the loss of its methyltransferase activity and the adaptation to facilitating transcription initiation and association with the RNA polymerase [180].

3.3.6 Model for substrate binding and catalysis

The major substrate of KsgA/Dim1/TFB proteins is a conserved hairpin close to the 3′ end of the small subunit rRNA (Figure 3-1B). The structure of the *E. coli* ribosome [181] reveals that the two nucleotides that are substrates for these proteins are located in a region of the RNA with marked single-stranded character (Figure 3-6A). Since no catalytically relevant complexes of KsgA/Dim1/TFB proteins have been crystallized in complex with SAM and their nucleic acid substrate, we reasoned that comparisons with nucleic acid methyltransferases from other families might shed some light on the mechanism of RNA binding. The MTaqI methyltransferase is responsible for N6-adenine methylation within a double-stranded DNA TCGA sequence [182]. Comparison of TFB1M with the substrate-bound complex structure of MTaqI [183] reveals that while the C-terminal domains of both proteins are quite divergent, consistent with the large differences in their substrates, the N-terminal catalytic domains of both enzymes are quite similar, with an overall RMSD of 3.9 Å for 136 C-α atoms. When the active sites are superimposed (Figure 3-6B), the SAM molecule bound to TFB1M and the cofactor analog bound in the MTaqI ternary complex have nearly identical orientations. Most interestingly, the substrate adenine fits into a putative substrate-binding pocket on TFB1M, and displays an orientation that

seems to be consistent with catalysis. It is important to note that MTaqI and KsgA-related methyltransferases accommodate one substrate base at a time, although the processivity (or lack thereof) differs between enzymes [184]. Much like Tyr 108 in MTaqI is engaged in a base-stacking interaction with the substrate adenine, Phe 144 of TFB1M appears capable of forming a similar interaction with the putative substrate (Figure 3-6C). Significantly, this aromatic residue is conserved in KsgA/Dim1, ErmC′ [178], and other DNA methyltransferases [185].

We therefore hypothesized that TFB1M utilizes a binding mode that involves base flipping in order to gain access to the substrate adenines. Nucleic acid methyltransferases like TFB1M and MTaqI are often faced with the problem of substrate access – if the target base is involved in secondary/tertiary structure (as is quite common in the case of RNA) then the target base must be everted for catalysis and methylation to occur. Moreover, the electrostatic surface of TFB1M is also consistent with this binding mode, showing a positively charged groove that could interact with the RNA backbone (Figure 3-6D).

In order to further investigate the validity of this model, we generated TFB1M mutants and studied their ability to complement a KsgA deficiency. We first decided to eliminate the benzene ring of Phe 144. This ring would presumably stabilize the everted nucleotide in the active site by forming a π-stacking interaction with the adenine base. Consistently, a F144A mutant was significantly impaired in its ability to rescue the phenotype of a *ksgA*$^{-}$ strain. This result was consistent with the effect of an analogous substitution in KsgA (Y116A), indicating that this conserved aromatic residue plays an important catalytic role (Figure 3-6E). Furthermore, we decided to test the importance of Arg 291, a residue in the C-terminal domain that appears to be involved in forming the positively charged cleft in TFB1M. The analogous KsgA residue (Arg248) has been shown to be important for substrate binding in KsgA [186]. Significantly, substitution of this residue also led to an impaired ability to complement the KsgA deficiency, indicating that the C-terminal domain is likely involved in RNA binding. This is consistent with both our model and with EM docking studies of KsgA binding to the bacterial ribosome [186].

3.4 Conclusions

TFB proteins belong to a large superfamily of methyltransferases that mediate a highly conserved rRNA modification near the 3′ end of the small ribosomal subunit. In prokaryotes this modification is catalyzed by KsgA proteins, while in eukaryotes Dim1 proteins catalyze the modification in the cytoplasmic ribosome and the mitochondrial ribosome of plants [150, 151, 187]. The TFB branch is present in fungal and animal mitochondria [95]. Its function appears to be essential for mitochondrial function in mammals [92, 93] and has been implicated in the pathogenesis of different human disorders [141]. In fungi and animals, this gene family has evolved an additional function related to the initiation of transcription. The two mammalian TFB proteins thus appear to unevenly distribute both functions, with TFB1M acting primarily as a KsgA-like methyltransferase and TFB2M as a transcription initiation factor. A similar functional distribution might take place in invertebrates [188-190]. However, sc-mtTFB appears to have completely lost its methyltransferase activity and remain solely as a transcription factor, as the rRNA modification is not conserved in the yeast mitochondrial ribosome [95]. Accordingly, sc-mtTFB is unable to complement a KsgA deficiency in E. coli [95]. Consistently, sc-mtTFB appears to be the most structurally divergent member of the family, while mammalian TFB1M is structurally and functionally more closely related to bacterial KsgA. On the other hand, there are reports that both mammalian TFB factors, TFB1M and TFB2M, act as methyltransferases – they both have the capacity to bind *S*-adenosyl-L-methionine (SAM), the methyl donor for the methyltransferase reaction, and expression of either protein can complement bacteria lacking KsgA [95, 148]. However, the extent to which TFB2M functions as a methyltransferase *in vivo*, if at all, is currently unclear.

As shown here, KsgA/Dim1/TFB proteins likely take advantage of a nucleic acid binding mechanism that involves base-flipping. While our results do not yet confirm such a mechanism, the fact that the methyltransferase activities of TFB1M and KsgA depend on a conserved aromatic residue proposed to stack with the flipped out base supports this notion. Moreover, base flipping has been demonstrated for several DNA methyltransferases, such as MTaqI [191], HhaI [192], and EcoRI [193], and has been proposed as a general mechanism for nucleic acid methyltransferases [194]. It would therefore not be surprising if the KsgA/Dim1/TFB family of rRNA methyltransferases utilized a similar mechanism.

3.5 Accession codes

The atomic coordinates for ligand-free TFB1M and TFB1M:SAM have been deposited in the Protein Data Bank with accession codes 4GC5 and 4GC9, respectively.

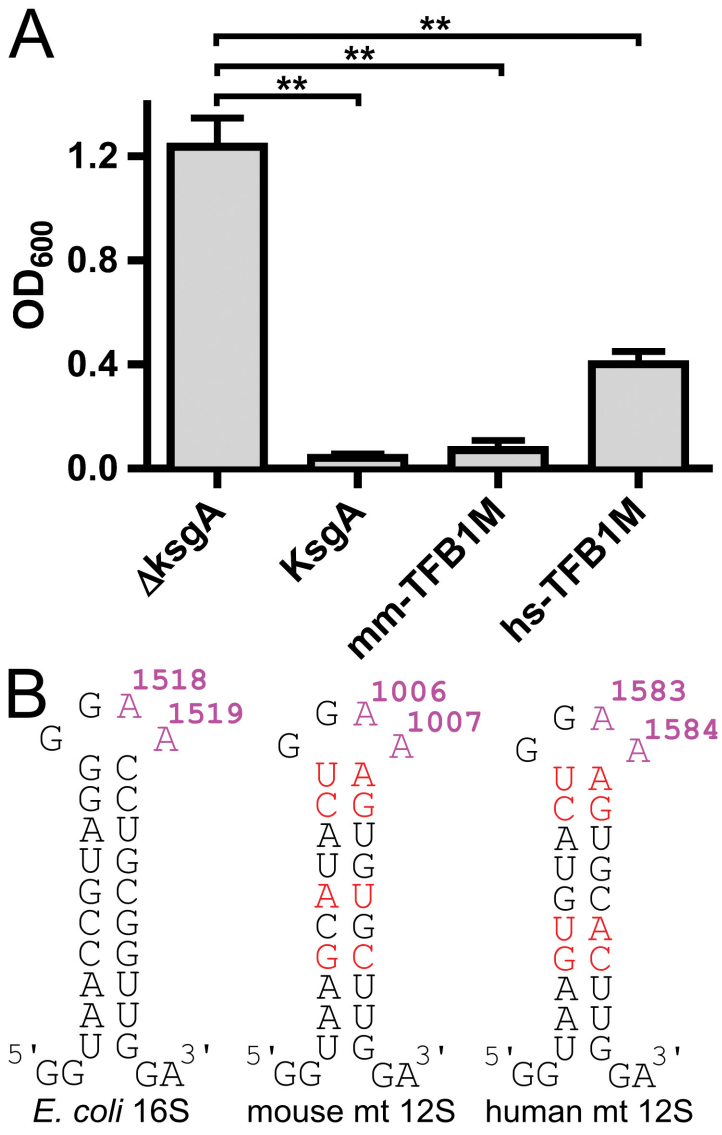


Figure 3-1. Functional complementation of *E. coli* KsgA rRNA methyltransferase activity by human and mouse TFB1M

A. The results of a kasugamycin-sensitivity assay are shown. Plotted is the optical density (OD_{600}) after 5 hours of culture growth in the presence of kasugamycin. Error bars represent the standard deviation (s.d.) of three replicates. ** indicates a P value of ≤ 0.005 (see Methods) **B**. Schematic representation of the conserved stem-loops in the 16S *E. coli* rRNA as well as the

human and mouse 12S mitochondrial rRNA. The two dimethylated adenines are highlighted in magenta. Non-conserved residues are shown in red.

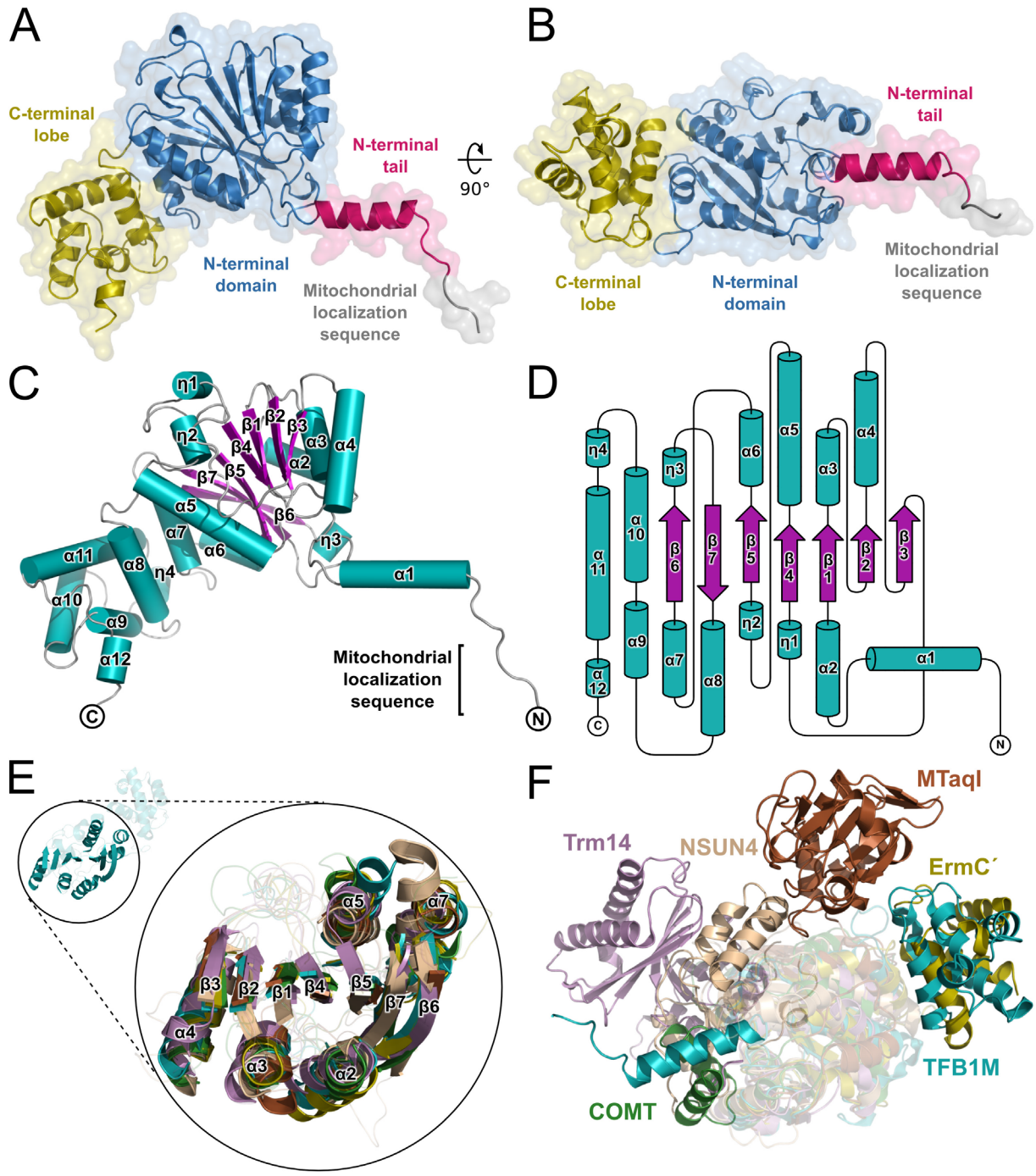


Figure 3-2. Overall architechure of TFB1M

A. TFB1M adopts a methyltransferase fold with a two-domain architecture. A conserved N-terminal domain (blue) constitutes a Rossman-like methyltransferase fold and contains the active site. A C-terminal lobe (yellow) is believed to mediate RNA substrate specificity. A short N-terminal extension (magenta) is resolved in the crystal structure. The structure contains a portion of the mitochondrial localization sequence that is predicted to be cleaved upon import (grey). The molecular surface of the protein is rendered transparent. A 90° rotation is shown in **B**. **C.** Schematic representation of mmTFB1M. β-strands are shown as magenta arrows, α-helices are shown as cyan ribbons, and coiled regions are shown as gray lines. The N- and C-termini are marked. **D**. Topology diagram of the protein fold using the same color coding as in **C**. Secondary structure elements were identified using DSSP [195]. Coiled regions are indicated by solid black lines. **E.** Structural conservation of the methyltransferase domain. The inset shows the conservation of the central seven-stranded beta-sheet and the three flanking α-helices on each side in several classes of methyltransferases. Shown are another adenine N6-specific rRNA methyltransferase (ErmC′; yellow; PDB 1QAO; [196], a large rRNA subunit methyltransferase (NSUN4, beige; PDB 4FZV; [104], a tRNA methyltransferase (Trm14; pink; PDB 3TM4; [197], a DNA methyltransferase (MTaqI; brown; PBD 2ADM; [198] and a small-molecule methyltransferase (COMT, green; PDB 1VID [167]. The secondary structure elements corresponding to TFB1M are labeled. **F.** Divergence of the C-terminal lobe in different methyltransferases. The figure shows an overlay of the enzymes rendered in **E** maintaining the same color scheme. The C-terminal domain in the ErmC′ rRNA methyltransferase is similar to that in TFB1M. Other methyltransferases exhibit structurally distinct domains that are oriented differently with respect to the methyltransferase catalytic domain (see text).

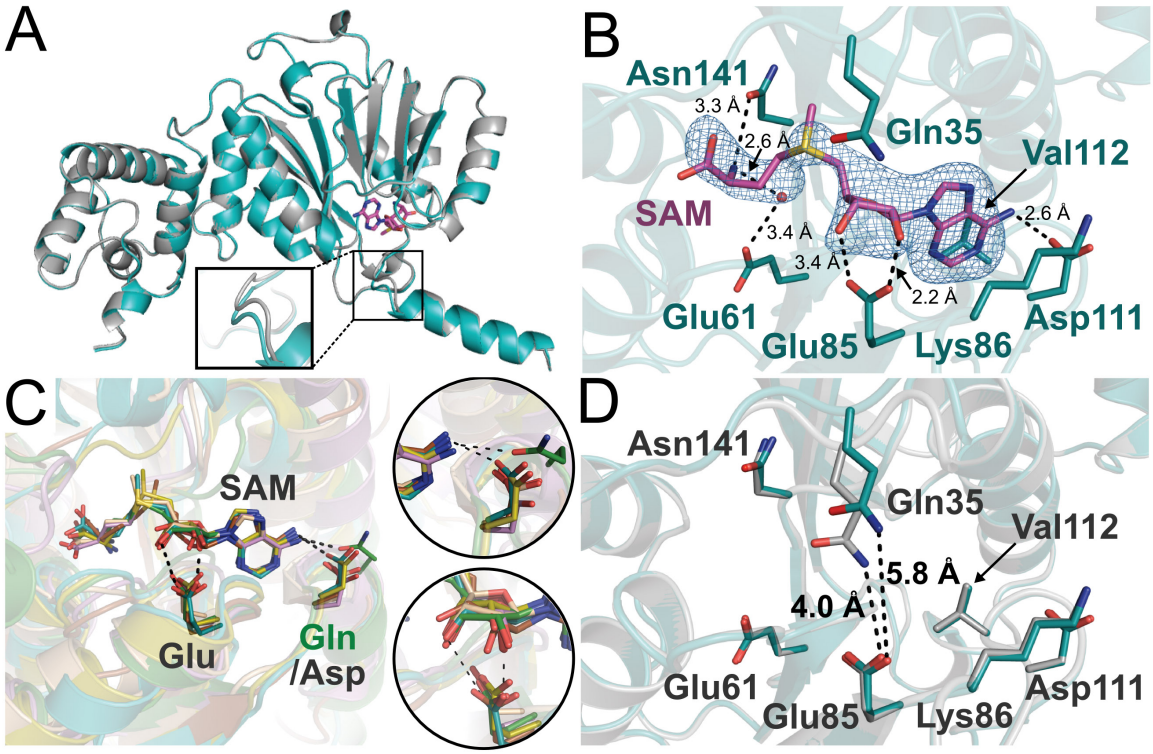


Figure 3-3. The SAM binding pocket in TFB1M

A. Overlay between the ligand-free (gray) and SAM-bound TFB1M (cyan) structures. Both structures are essentially identical, with an rmsd of 0.26 Å for 326 C-α atoms. The inset highlights a subtle shift in the loop connecting α-helices 1 and 2 (see text). **B**. The SAM-binding pocket in TFB1M. SAM is bound in a negatively charged binding pocket in the N-terminal domain. Several residues establish contacts with the cofactor (see text). A simulated-annealing F_o-F_c omit electron density map is shown (blue), contoured at 3σ. A water molecule (red sphere) bridges the interaction between SAM and Glu61. **C**. Overlay of the SAM binding pocket in the TFB1M (cyan), ErmC' (yellow), NSUN4 (beige), MTaqI (brown), Trm14 (pink) and COMT (green) methyltransferases (see Figure 2). The insets highlight two conserved interactions with the SAM cofactor. A conserved hydrogen bond with the N6 atom of the adenine base is observed in all methyltransferases (upper inset), although the protein residue involved in the interaction is not always the same and is not strictly conserved (Gln in COMT and Asp in all the others). An absolutely conserved glutamate residue forms at least one hydrogen bond to the ribose of the SAM cofactor in every structure examined (lower inset). **D.** Overlay between the ligand-free and SAM-bound TFB1M highlighting differences in the SAM binding pocket. The loop between α-helices 1 and 2 undergoes a subtle repositioning upon SAM binding, resulting in a substantial shift of the Gln35 side chain.

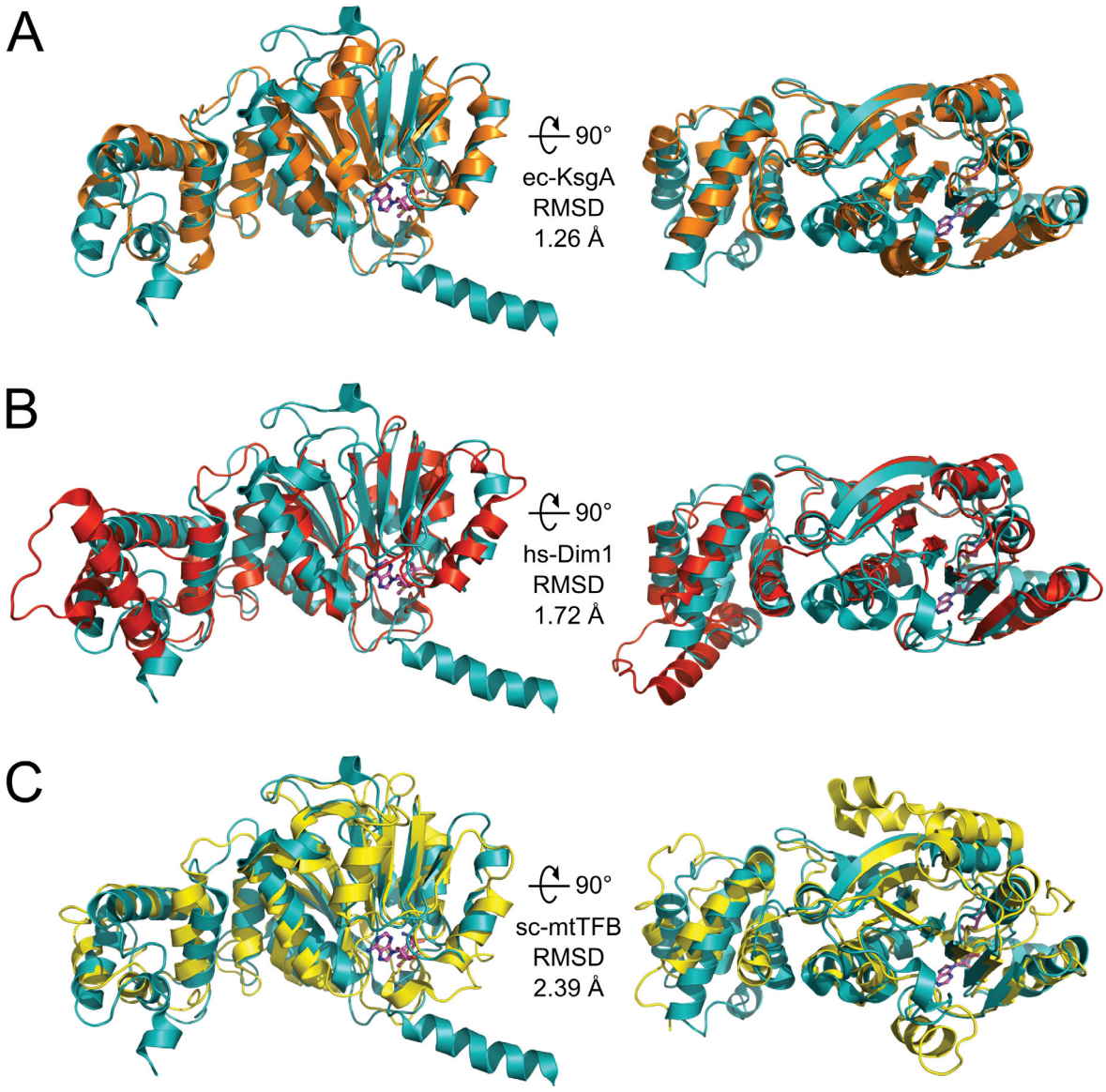


Figure 3-4. TFB1M is structurally similar to KsgA and Dim1 methyltransferases
A. Overlay between TFB1M (cyan) and *E. coli* KsgA (orange; PDB ID 1QYR) [170]. **B**. Overlay between TFB1M and human Dim1 (red; PDB ID 1ZQ9; A. Dong, H. Wu, H. Zeng, P. Loppnau, M. Sundstrom, C. Arrowsmith, A. Edwards, A. Bochkarev and A. Plotnikov, unpublished data). **C**. Overlay between TFB1M and sc-mtTFB (yellow; PDB ID 1I4W) [199]. A 90 degree rotation is shown on the right. The SAM molecule from the TB1M structure is shown in all three panels for reference.

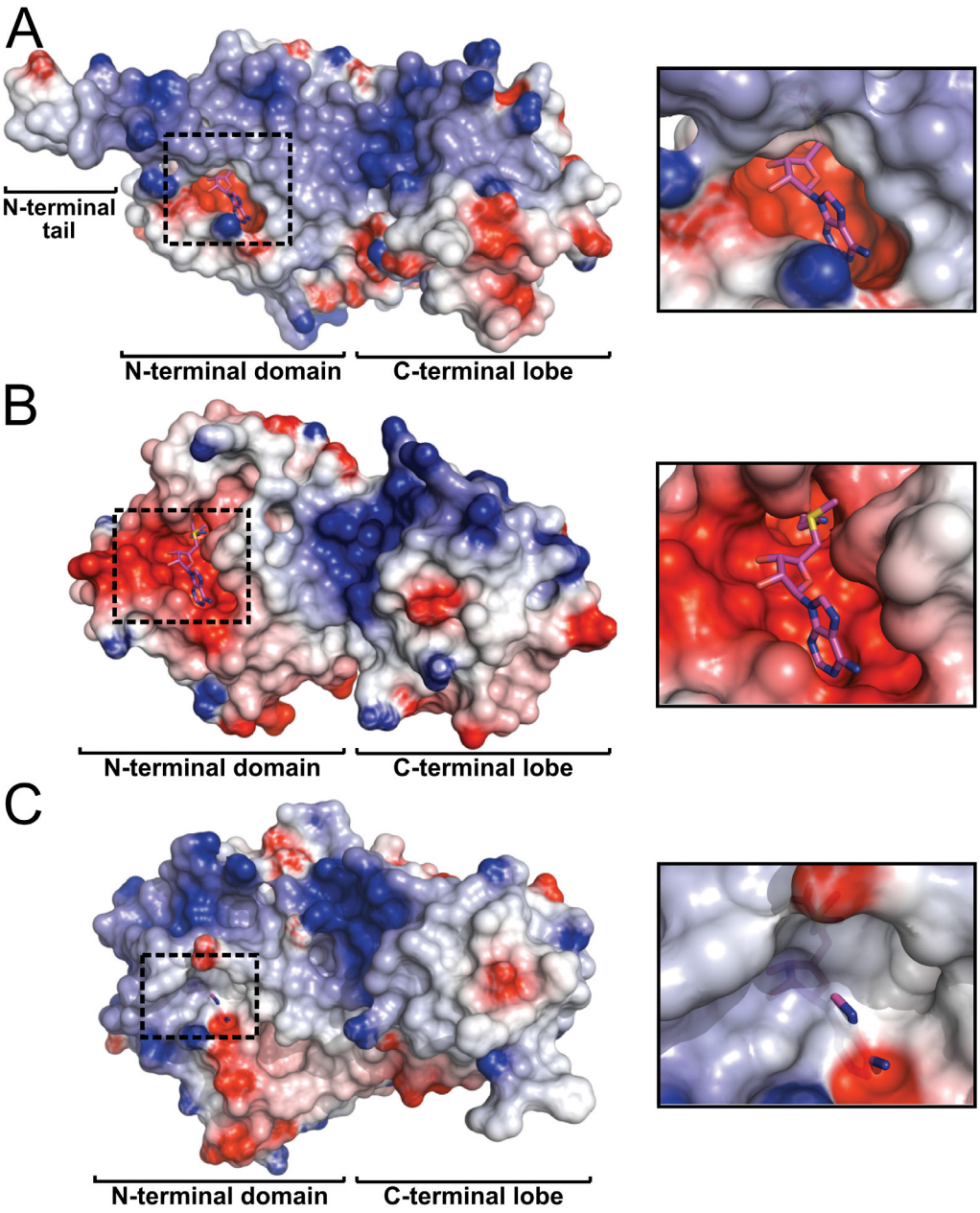


Figure 3-5. Electrostatic surface potential maps of TFB1M, KsgA and mtTFB
Electrostatic surface potential maps of (**A**) TFB1M in complex with SAM, shown as magenta sticks (**B**) *E. coli* KsgA [PDB ID 1QYR] and (**C**) *S. cerevisiae* mtTFB [PDB ID 1I4W]. The TFB1M:SAM structure was overlaid with both KsgA and mtTFB and the SAM molecule bound in the TFB1M structure is shown in B and C for reference (magenta). The insets on the right of the figure highlight the observed SAM binding pocket in TFB1M and the putative SAM binding pockets in KsgA and mtTFB. The position of the N-terminal and C-terminal domains (and the N-terminal extension for TFB1M) is indicated. The electrostatic surface potential maps were generated with Delphi [168] and are colored from -7 kTe^{-1} (blue) to +7 kTe^{-1} (red).

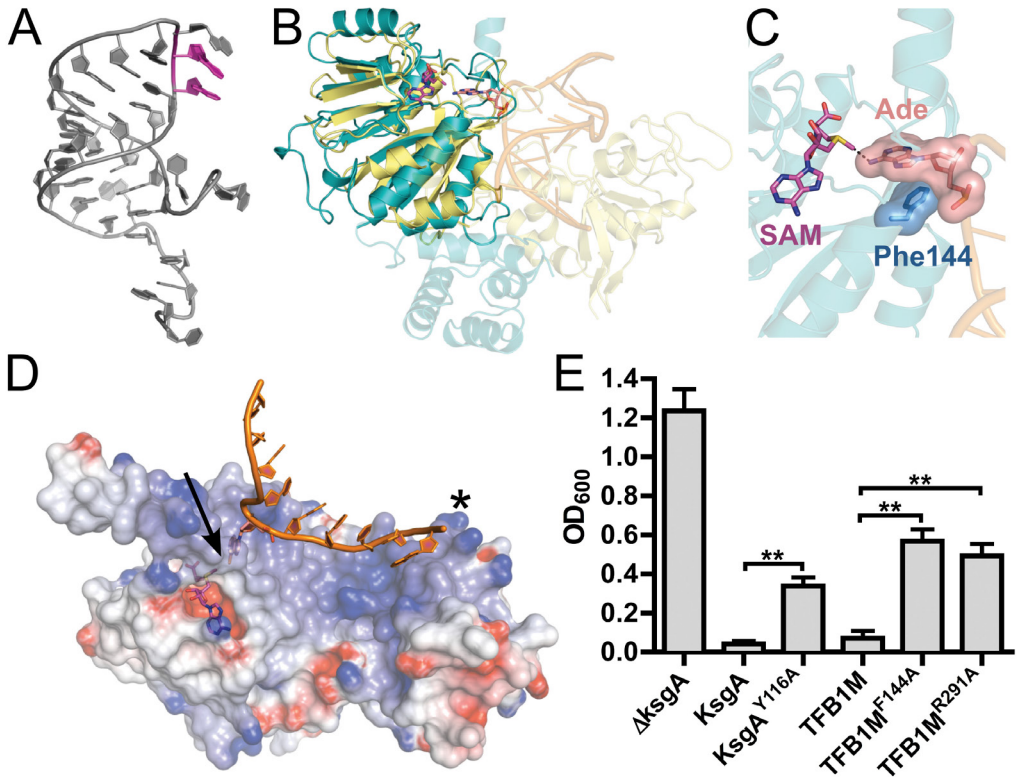


Figure 3-6. A model for RNA binding by TFB1M

A. Structure of the conserved stem-loop in the *E. coli* 16S rRNA that is methylated by KsgA. The two substrate adenine bases are shown in magenta. **B**. Overview of the TFB1M:SAM structure superposed with the structure of MTaqI in complex with substrate DNA and a cofactor analog [PDB entry 1G38]. TFB1M is shown in teal, with the bound SAM molecule and substrate adenine shown as magenta and pink sticks, respectively. MTaqI is shown in yellow, with the bound DNA shown in orange. The conserved methyltransferase domain of each protein is shown as a solid cartoon while the remainder of each protein and the DNA is transparent. **C**. The putative catalytic center of TFB1M. The overlay of TFB1M and MTaqI suggests a putative substrate binding mechanism that involves base flipping, with the substrate adenine residue (pink) forming a π-stacking interaction with residue Phe144 (blue) of TFB1M. The methyl leaving group of SAM and N6 of the substrate adenine are separated by a distance of 2.4 Å (dashed line). **D**. Model for RNA binding by TFB1M. The electrostatic surface potential of TFB1M is rendered transparent, and a single DNA strand containing the substrate adenine from the MTaqI structure is shown in orange. The black arrow indicates where the methyl group of SAM and the N6 of the adenine come into close proximity. The black asterisk denotes the location of Arg291 near the backbone of the nucleic acid. **E**. Phe144 and Arg291 are important for catalysis, suggesting that they play a role in substrate binding. The results of a kasugamycin-sensitivity assay, with optical density (OD_{600}) after 5 hours of culture growth in the presence of kasugamycin is plotted in gray. The experiments were carried out using mouse TFB1M. Error bars represent the s.d. of three replicates. ** indicates a P value of ≤ 0.005

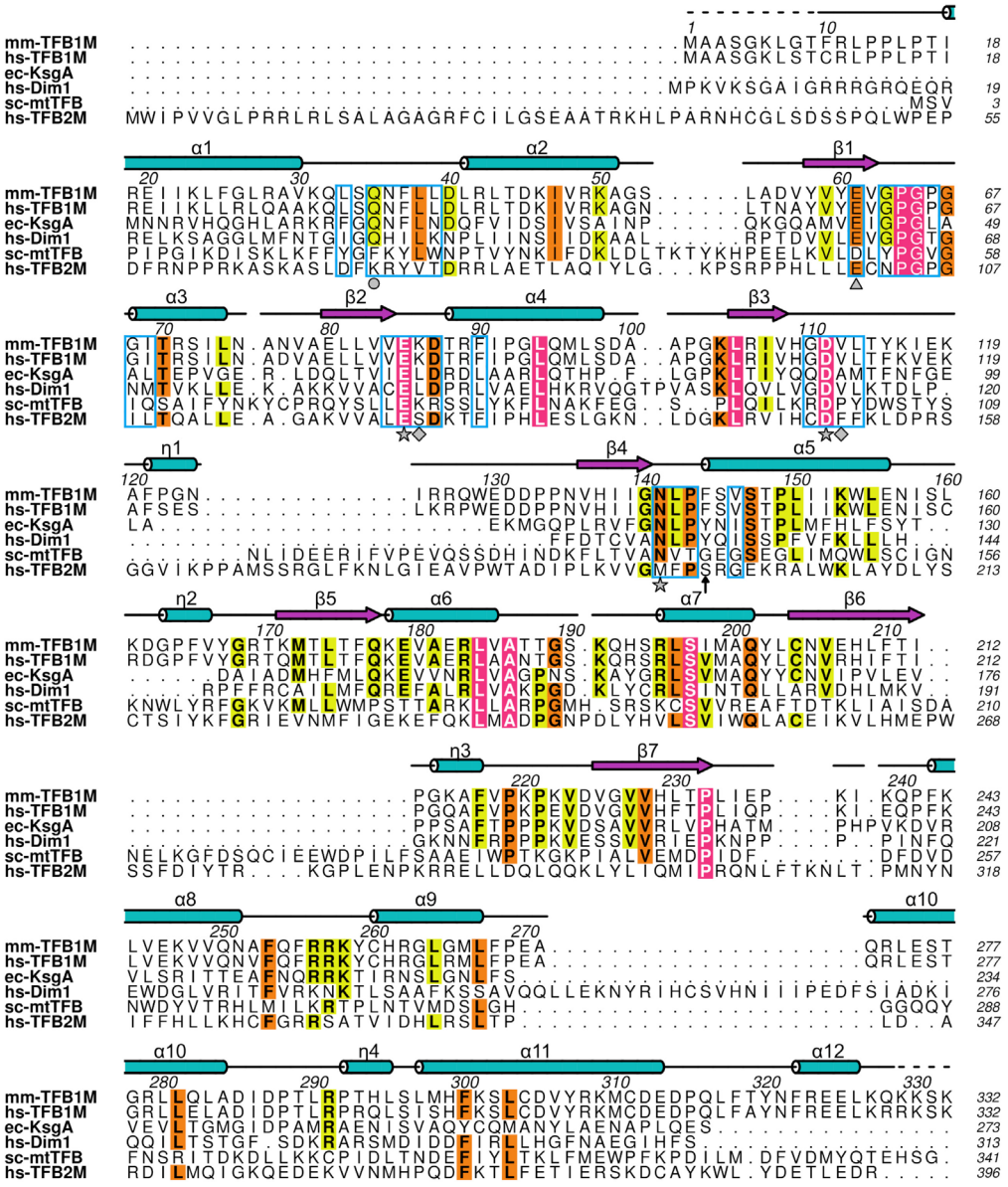


Figure 3-7. **Structure-based amino acid sequence alignment of KsgA/Dim1/TFB family proteins**

Secondary structure elements corresponding to the TFB1M structure are shown on top of the alignment. Residues are color-coded based on the degree of conservation. Fully conserved residues are shown in white with a magenta background. Residues identical in five of the six proteins are shown with an orange background and residues identical in four proteins are shown in yellow. The blue boxes mark residues that constitute the SAM binding pocket. Residues marked with a star at the bottom of the alignment indicate residues that hydrogen bond with the SAM molecule via side chains. The diamonds indicate residues that establish Van der Waals interactions with the SAM molecule. The circle indicates the position of Glu35 and the arrow that of Phe144, a residue postulated to be involved in base-flipping (see main text and Figure 3-6). Mouse TFB1M shares 25.6%, 22.0%, 17.6% and 13.3% amino acid identities with *E.coli* KsgA, human Dim1, human TFB2M and yeast mtTFB, respectively.

.

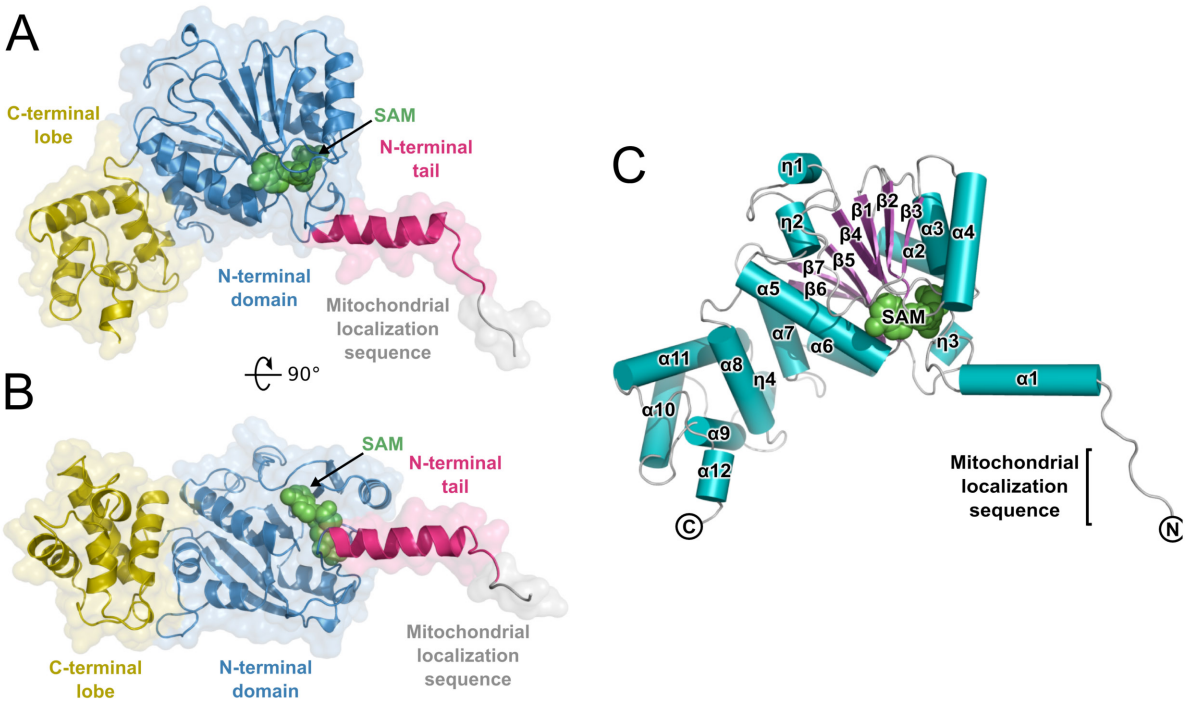


Figure 3-8. TFB1M in complex with *S*-adenosylmethionine
A. Overview of the TFB1M fold in the SAM-bound structure (analogous to the view in Figure 3-2 but showing the position of the bound SAM). The N-terminal domain is blue, the C-terminal lobe is yellow, the N-terminal extension is magenta and the SAM molecule is green. The structure contains a portion of the mitochondrial localization sequence that is predicted to be cleaved upon import to the mitochondria (grey). The molecular surface of the protein is rendered transparent. **B.** A 90° rotation of the view shown in panel A. **C.** Schematic representation of mmTFB1M showing the position of the SAM molecule. The color-coding is analogous to Figure 3-2C.

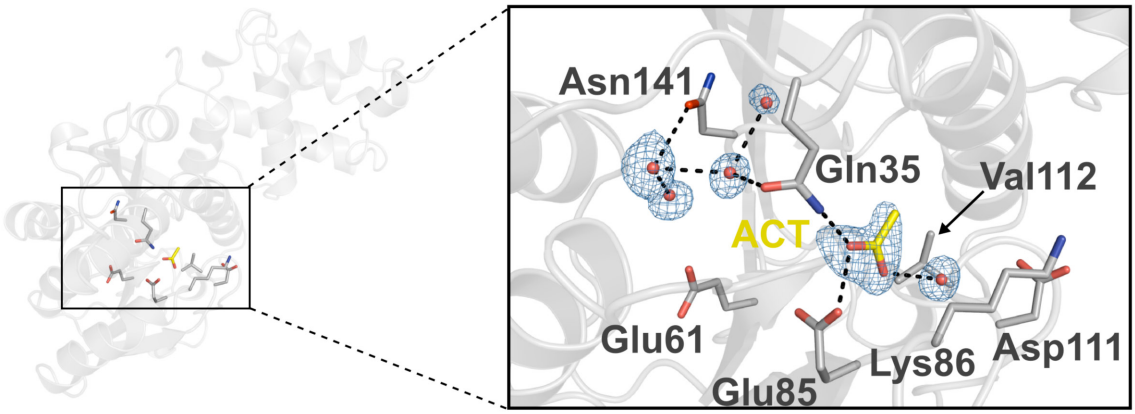


Figure 3-9. The SAM-binding pocket in the ligand-free TFB1M structure
A simulated-annealing F_o-F_c omit electron density map (blue) contoured at 3σ shows the absence of electron density corresponding to a SAM molecule. Instead, the binding pocket is occupied by several solvent molecules and an acetate molecule that hydrogen-bonds to Gln35 and Glu85.

Table 3-1. Data collection and refinement statistics for TFB1M and TFB1M:SAM

Crystal	SeMet-TFB1M[a]			HighRes-TFB1M[a]	TFB1M:SAM[a]
Data collection					
Space group		*I*222		*I*222	*I*222
Cell dimensions					
a, *b*, *c* (Å)		47.6, 99.7, 211		47.6, 101.3, 211.6	47.7, 101.7, 211.7
α, β, γ (°)		90, 90, 90		90, 90, 90	90 90 90
	Peak	*Inflection*	*Remote*		
Resolution (Å)	39.8−2.2 (2.24−2.20)[b]	39.8−2.2 (2.24−2.20)[b]	39.8−2.2 (2.24−2.20)[b]	33.42−2.02 (2.03−2.02)[b]	33.47−2.08 (2.088−2.081)[b]
R_{sym}	0.106 (0.676)	0.096 (0.721)	0.107 (0.796)	0.096 (0.832)	0.046 (0.606)
$I / \sigma I$	29.39 (2.67)	36.23 (2.54)	36.38 (3.17)	12.12 (2.4)	22.7 (4.1)
Completeness (%)	95.8 (81.6)	95.2 (80.4)	97.8 (85.3)	99.5 (99.7)	97.8 (100)
Redundancy	13.5 (11.4)	13.4 (11.2)	13.8 (12.2)	7.0 (7.4)	7.2 (7.4)
Refinement					
Resolution (Å)				33.42−1.80	33.47−2.10
No. reflections				33,225	23,221
Completeness (%)[c]				69.2 (17.2)	76.0 (29.8)
R_{work} / R_{free}				0.1911 / 0.2280	0.2052 / 0.2336
No. atoms					
Total				2,787	2,708
Protein				2,546	2,554
SAM				-	27
Water				233	150
B-factors					
Protein				51.7	35.3
SAM				-	53.1
Water				55.0	35.5
R.m.s deviations					
Bond lengths (Å)				0.008	0.005
Bond angles (°)				1.22	0.69
PDB ID				4GC5	4GC9

[a] One crystal was used for each dataset
[b] Values in parentheses are for the highest-resolution shell
[c] Completeness after elliptical truncation

Chapter 4: Structure of the essential MTERF4:NSUN4 protein complex reveals how an MTERF protein collaborates to facilitate rRNA modification

This chapter has been published as Yakubovskaya E[1], **Guja KE[1]**, Mejia E[1], Castano S, Hambardjieva E, Choi WS, Garcia-Diaz M. (2012). Structure of the Essential MTERF4:NSUN4 Protein Complex Reveals How an MTERF Protein Collaborates to Facilitate rRNA Modification. *Structure*. 20(11):1940-7. PMID: 23022348. **[1]Equal contribution.**

Author contributions: EY and MGD conceived of the study; EY, EM, SC, WSC, and EH carried out cloning, expression and purification; KEG and EY designed the methyltransferase assay; EY carried out the methyltransferase assay and SEC experiments; KEG and EY designed the crystallization experiments; EY and EM carried out the crystallization experiments, KEG carried out data collection, experimental phasing, refinement and analysis of the crystal structure; KEG, MGD, and WSC prepared figures; KEG, EY, and MGD drafted and revised the manuscript. Mark Lukin is gratefully acknowledged for oligonucleotide synthesis.

4.1 Abstract

MTERF4 is the first mitochondrial transcription termination factor (MTERF) family member shown to bind RNA, and plays an essential role as a regulator of ribosomal biogenesis in mammalian mitochondria. It forms a complex with the rRNA methyltransferase NSUN4 and recruits it to the large ribosomal subunit. In this paper, we characterize the interaction between both proteins, demonstrate that MTERF4 strongly stimulates the specificity of NSUN4 during *in vitro* methylation experiments and present the 2.0 Å resolution crystal structure of the MTERF4:NSUN4 protein complex, containing a truncated MTERF4, that lacks 48 residues of the C-terminal acidic tail, bound to *S*-adenosyl-L-methionine, revealing the nature of the interaction between both proteins and the structural conservation of the most divergent of the human MTERF family members. Moreover, the structure suggests a model for RNA binding by

the MTERF4:NSUN4 complex, providing insight into the mechanism by which an MTERF family member facilitates rRNA methylation.

4.2 Introduction

The MTERF (Mitochondrial TERmination Factor) family of proteins is conserved in the plastids and mitochondria of metazoans [200, 201]. Proteins of the MTERF family, named after the human mitochondrial transcription termination factor MTERF1, are believed to be nucleic acid binding proteins and play various roles regulating organellar gene expression [202]. Most species contain multiple MTERF family members (four in humans, more than thirty in some plant genomes) and, in the cases studied, their function appears to be non-redundant [203]. In mammalian cells, MTERF proteins have been clearly demonstrated to regulate mitochondrial gene expression [101, 115, 201, 204]. The founding member of the family, MTERF1, mediates termination of transcription in a site-specific manner [102]. It recognizes a sequence in the mitochondrial tRNA-Leu(UUR) gene, which is downstream and adjacent to the 16S rRNA gene and promotes bidirectional transcriptional termination [103, 134]. MTERF2 and MTERF3 have been proposed to associate with mitochondrial DNA in a non sequence-specific manner and regulate mtDNA transcription [201, 204, 205].

Recently, crystal structures of MTERF1 and MTERF3 have revealed that these proteins share a unique superhelical fold [103, 114, 206]. Both proteins have a modular architecture and are composed of a series of MTERF repeats that are organized in a left-handed helical fold that conforms a large central groove capable of mediating the association with nucleic acids, suggesting that MTERF proteins belong to a larger class of tandem helical repeat nucleic acid binding proteins [207]. MTERF1 is thus far the only MTERF protein that has been crystallized in complex with nucleic acid. The crystal structure of MTERF1 in complex with the termination sequence highlights the mechanism of sequence-specific double-stranded DNA recognition by MTERF1 and reveals that binding of MTERF1 perturbs the DNA structure by bending and unwinding the duplex and everting three nucleotides [103]. This perturbation results in a higher kinetic stability of the sequence-specific complex, and is thought to be crucial for efficient termination of transcription [103, 202]. The remarkable structural similarity between MTERF1 and MTERF3 suggests that this fold, and perhaps some aspects of the nucleic acid binding mechanism, might be conserved in other MTERF family proteins.

Until recently, tandem α-helical repeat motifs, like those found in MTERF1 and MTERF3, were believed to be primarily associated with protein scaffolds or RNA recognition enzymes [207]. Thus, MTERF proteins are an example of an analogous structural solution that is dedicated to DNA binding. However, the extensive similarities with RNA binding tandem α-helical repeat proteins from the PUF family, suggested the possibility that some MTERF proteins may be capable of binding RNA as well [103, 207-210]. Indeed, in contrast with other MTERF family members, MTERF4, the fourth human MTERF protein, was recently shown to bind RNA and be an essential regulator of translation in mammalian mitochondria by mediating a critical rRNA modification in the large mitochondrial ribosomal subunit [101]. Modifications of the mitochondrial rRNAs appear to be essential for ribosome biogenesis and normal mitochondrial function. However, with the exception of the adenine dimethylation catalyzed by the TFB1M methyltransferase, mitochondrial rRNA modifications are at present poorly understood [92, 93, 96, 148, 161]. Importantly, the role of MTERF4 in this process depends on its ability to form a complex with the rRNA methyltransferase NSUN4. Unlike TFB1M, NSUN4 appears to be devoid of any ability to bind RNA in a sequence-specific manner, and therefore needs to be targeted to its methylation site [101]. While the MTERF4/NSUN4 interaction with RNA has not yet been characterized and might involve interactions also with the ribosomal proteins, these observations led to a model whereby MTERF4 recruits NSUN4 to the large ribosomal subunit and mediates the interaction with the ribosomal RNA, explaining why the loss of MTERF4 leads to severe defects in ribosomal assembly and translation.

The ability of MTERF4 to bind RNA highlights the versatility of the MTERF fold and the wide spectrum of substrate specificity that it can confer, thus making MTERF4 a particularly interesting subject for more detailed structural and functional studies. In addition, the fact that a defined binding partner has been identified allows for the study of the mechanisms that facilitate this interaction. Protein-protein interactions have been postulated to be essential for the functions of MTERF proteins [202, 211]. However, how these interactions might be mediated by the MTERF fold, which appears to be fully engaged in nucleic acid binding, has thus far been a matter of debate.

Here, we characterize the interaction between MTERF4 and NSUN4, demonstrate that the association between these proteins leads to a strong stimulation of specificity and efficiency of *in vitro* mitochondrial 16S rRNA methylation, and describe the high resolution x-ray

crystallographic structure of a complex between both proteins in the presence of the methyl donor *S*-adenosyl-L-methionine (SAM). The structure, which lacks 48 residues of the C-terminal acidic tail of MTERF4 (that were purposely truncated in our construct), reveals the mechanism of the interaction between the two proteins, the mechanism of SAM binding and the conservation of the MTERF fold in the most divergent of the human MTERF family members. It suggests a model that explains how MTERF4 facilitates rRNA methylation by NSUN4.

4.3 Results and discussion

4.3.1 MTERF4 and NSUN4 form a tight complex

The proposed function of MTERF4 in mitochondrial ribosome biogenesis is dependent on the formation of a ternary complex with RNA and the methyltransferase NSUN4. It was previously demonstrated that MTERF4 and NSUN4 form a heterodimer *in vitro* [101]. In order to better assess the likelihood that these proteins will form a complex *in vivo*, we decided to experimentally measure the stability of this complex through isothermal titration calorimetry (ITC). We expressed both human proteins in *E. coli* without their N-terminal targeting peptides (see Methods in Chapter 2) and developed a successful purification strategy for full-length NSUN4. However, we were unable to purify full-length MTERF4 in sufficient amounts due to poor protein stability. Inspection of the MTERF4 amino acid sequence reveals a 60 residue long C-terminal acidic tail of which the last 50 residues are predicted to be highly disordered (Figure 4-1A). Systematic truncations in this region allowed us to generate a construct missing the C-terminal 48 amino acids that resulted in the production of high yields of stable protein. We initially hypothesized that the truncated region might be involved in mediating the interaction with NSUN4. However, although it is possible that this tail might affect the interaction with NSUN4, removal of this C-terminal region did not prevent complex formation, suggesting that it might be dispensable for the interaction.

Subsequent ITC studies demonstrated that MTERF4 and NSUN4 form a stable complex with 1:1 stoichiometry (Figure 4-1B), and that this interaction does not depend on the presence of the SAM cofactor (Figure 4-8). The association of MTERF4 and NSUN4 is significantly exothermic (ΔH = -11.04 kcal/mol), whereas the entropy of association (ΔS = -1.05 cal/mol K) is very small. Low entropy of association usually indicates a smaller contribution of hydrophobic

interactions to protein-protein complex formation, suggesting that the majority of the contacts between the two proteins are largely established between polar amino acids [212]. Moreover, the affinity of MTERF4 and NSUN4 binding appears to be extremely high (K_D = 13.3 nM). Such a small dissociation constant allows us to conclude that MTERF4 and NSUN4 form an obligate heterodimer in mitochondria, and that it is reasonable to consider both proteins as subunits of a stable holoenzyme as suggested previously [101].

4.3.2 MTERF4 enhances the activity and specificity of NSUN4 *in vitro*

We then decided to address the functional importance of the interaction between MTERF4 and NSUN4. Since NSUN4 has been postulated to methylate the mitochondrial 16S rRNA, we analyzed the ability of both proteins to directly methylate *in vitro* transcribed 16S rRNA (see Methods in Chapter 2) [101]. Incubation of the purified 16S rRNA with MTERF4 and NSUN4 resulted in incorporation of ^{3}H-methyl groups into the 16S rRNA that was significantly above background (Figure 4-2). This activity appeared to be mostly specific, since using an RNA antisense to the 16S rRNA as a substrate resulted in considerably lower activity, although still above that of a control with no protein, suggesting a certain degree of non-specific methylation. Interestingly, NSUN4 alone was also able to promote methylation of the 16S rRNA, but the activity observed was distinctly lower and, importantly, equivalent with both substrates, indicating that this activity was devoid of specificity. We therefore conclude that MTERF4 stimulates the activity of NSUN4 *in vitro* and is essential to target it to its specific methylation site, which is consistent with the proposed model of site-specific association of MTERF4 with the mitochondrial 16S rRNA. This constitutes the first formal demonstration that NSUN4 is endowed with an intrinsic RNA methyltransferase activity, and that it is capable of methylating the naked 16S rRNA. However, it is important to note that the endogenous substrate is unlikely to be naked rRNA and thus our observed in vitro activity may differ from the in vivo activity of NSUN4 or the MTERF4:NSUN4 complex. Moreover, the ratio of the two proteins in vivo may not be 1:1, such that a free fraction of either NSUN4 or MTERF4 may exist and play additional role(s) in ribosome biogenesis.

4.3.3 Crystal structure of the MTERF4:NSUN4 complex

MTERF4 is the most divergent human member of the MTERF family [201]. Therefore, we initially decided to determine its structure to assess whether it indeed folds like other characterized MTERF proteins. However, all attempts to crystallize MTERF4 alone were unsuccessful. Since MTERF4 and NSUN4 interact very tightly, we then decided to attempt cocrystallization of both proteins, as such a structure could yield important functional information regarding the role of the MTERF4:NSUN4 complex. We were able to obtain native crystals of a MTERF4:NSUN4 heterodimer that, after optimization, diffracted to 2.3 Å. Attempts to solve the structure by molecular replacement using the available MTERF structures and/or methyltransferases thought to be related to NSUN4 were unsuccessful. We then prepared selenomethionione (SeMet)-derived MTERF4 protein and obtained crystals that diffracted to 2.9 Å and allowed us to perform initial phasing and build a preliminary model. Preparation of SeMet-derived NSUN4 resulted in crystals that diffracted to high resolution (2.0 Å; Table 4-1), contained one molecule of the complex in the asymmetric unit, and allowed us to build a complete model (Figure 4-3). The electron density was of sufficient quality to build a full-length model for NSUN4 (residues 38-384), with the exception of a small disordered loop (residues 111-115), and unambiguously place the bound SAM molecule. However, the lack of extensive lattice interactions in the MTERF4 region, combined with the fact that the MTERF fold likely exhibits some flexibility in the absence of its nucleic acid substrate resulted in significant disorder in the C-terminal region of MTERF4 (distal to the interaction site) [103, 206]. Nevertheless, we were able to build side chains for most of the protein (residues 122-328) as well as the backbone for one additional MTERF motif. The overall structure of the complex reveals that both proteins are arranged end to end in an extended conformation and thus the interaction between them is restricted to a relatively small region. Significantly, this structure represents the first structural characterization of an MTERF protein participating in a protein-protein interaction.

4.3.4 MTERF4 adopts a canonical MTERF fold

MTERF1 and MTERF3 have been shown to adopt very similar modular superhelical folds composed of repeating MTERF motifs, despite the fact that MTERF3 was crystallized in

the absence of its nucleic acid substrate. This fold has been postulated to be a common feature of MTERF proteins and the basis for their ability to bind nucleic acids. Our current structure reveals that the MTERF fold is clearly conserved in MTERF4 and is not perturbed by the interaction with NSUN4 (Figure 4-3). Overlays with the MTERF3 (Figure 4-4A) and MTERF1 (Figure 4-4B) structures reveal extensive structural similarity between MTERF4 and the other MTERF proteins. However, MTERF4 displays much higher structural homology to MTERF3 (rmsd of 2.37 Å for 165 C-α atoms) than MTERF1 (rmsd of 4.33 Å for 165 C-α atoms; Figure 4-4), although this might simply be a consequence of the fact that both MTERF3 and MTERF4 were crystallized in the absence of nucleic acid. Six MTERF motifs can be recognized in MTERF4 (Figure 4-4C and Figure 4-9), and these motifs are nearly identical to those found in MTERF1 or MTERF3 (Figure 4-4D). However, the C-terminal MTERF motif in MTERF4 (motif 6 in Figure 4-4C) is distorted, as the second α-helix in the motif is replaced by a large loop, as seen in MTERF3 (Figure 4-4A). Moreover, two additional α-helices are present C-terminal to the last MTERF motif. These two helices are present in both MTERF4 and the C-terminus of MTERF3. Interestingly, these two additional helices are involved in the interaction between MTERF4 and NSUN4 (see below).

4.3.5 NSUN4 is structurally related to m5C methyltransferases and binds SAM

NSUN4 exhibits a classical Rossman-like fold that is characteristic of SAM-dependent methyltransferases, containing a core seven-stranded β-sheet (Figure 4-10) [166]. Structural homology searches with the DALI server indicated that, as suggested previously on the basis of sequence homology, NSUN4 shares extensive structural homology with m5C methyltransferases [101, 171]. The lowest rmsd (2.3 Å for 235 C-α atoms) was obtained with the tRNA methyltransferase Trm4 from *M. jannaschi* (PDB 3A4T; Figure 4-5A), although similar overlays can be obtained with other available m5C methyltransferase structures (PDB IDs 2FRX, 3M6V, 2YXL, 1IXK, etc). Lower structural homology was found to other classes of methyltransferases. Despite the fact that the structural overlap is extensive between Trm4 and NSUN4, NSUN4 contains a number of extensions that are not conserved in other m5C methyltransferases. Interestingly, the majority of these insertions are directly involved in the interaction with MTERF4 (see below).

The active site of NSUN4 displayed clear density for the bound SAM molecule (Figure 4-5B). As could be expected, SAM is located in a position analogous to the one observed in the Trm4 structure. NSUN4 has a relatively open SAM binding site that is similar to that of other methyltransferases. The SAM molecule is bound in a negatively charged pocket and makes extensive contacts with protein side chains (Figure 4-5B), with the *S*-methyl group oriented towards a large groove in the surface of the protein. The bound SAM is stabilized by a variety of hydrogen-bonding and van der Waals interactions. The N6 amino group of the adenosine is stabilized by a hydrogen bond with Asp237, while the adenine ring is sandwiched between two leucine residues. The ribose is bound through a bidentate interaction between the O2′ and O3′ groups and Asp204 and an additional hydrogen bond between the O3′ group and Arg209. The methionine moiety is stabilized by interactions with the main chain as well as a hydrogen bond with Asp255.

4.3.6 A conserved interaction motif adjacent to the MTERF fold

All the regions involved in the interaction between MTERF4 and NSUN4 are well ordered, clearly revealing the molecular details of the interaction between both proteins. This is the first structure depicting a protein-protein interaction involving an MTERF family member, and it is particularly interesting that the interaction involves only regions of MTERF4 that are C-terminal to the conserved MTERF motifs. As a consequence, most of the fold remains available to participate in RNA binding. This observation is consistent with the idea that the MTERF fold itself is dedicated to nucleic acid binding. MTERF4 interacts with NSUN4 utilizing the last distorted MTERF motif in the C-terminus of the protein (6 in Figure 4-4C) as well as the two α-helices C-terminal to this motif (Figure 4-6A). In NSUN4, the interaction involves an α-helix (residues 58-69; Figure 4-6B and Figure 4-6D) and two small non-conserved loops (residues 132-142 and 364-375; Figure 4-6B and Figure 4-6C). It is interesting that all of the regions of NSUN4 that are involved in the interactions are not conserved in Trm4. The aforementioned α-helix is part of an N-terminal extension (although this extension is present in some m5C methyltransferases other than Trm4), while the two loops are present in Trm4 but are longer in NSUN4.

In accordance with our thermodynamic data, most of the contacts established between MTERF4 and NSUN4 are electrostatic, although they surround a patch of hydrophobic residues that forms the center of the interaction surface (Figure 4-6C). The contact surface between both proteins has an area of 732.5 $Å^2$, and involves 73 atoms in 18 residues.

Residues in the C-terminal α-helix of MTERF4 are responsible for establishing most of the contacts with NSUN4 and are thus likely crucial for the interaction (Figure 4-6C). Importantly, as mentioned above, this helix, and the region surrounding it, seems to be conserved in MTERF3 (arrow in Figure 4-4A). This suggests that this region might be a conserved motif to mediate protein-protein interactions in MTERF3 and MTERF4. Moreover, a similar C-terminal extension is likely present in MTERF2 and thus could similarly constitute a potential interaction site [103]. Interestingly, the only human MTERF protein in which this motif is definitely absent is the transcriptional terminator MTERF1, perhaps highlighting the fact that this protein is able to mediate transcriptional termination in the absence of a protein partner.

4.3.7 A model for RNA binding

MTERF1 binds double-stranded DNA through its central groove. However, the vast majority of the contacts are established with only one of the two DNA strands, suggesting that a similar mechanism could be employed by MTERF4 to associate with a single strand of RNA. Calculation of an electrostatic surface potential map for MTERF4 revealed a large basic patch in the expected interaction region based on the MTERF1 structure. Accordingly, an overlay of MTERF1 and MTERF4 suggests a binding mode that is consistent with the electrostatic potential (Figure 4-7A).

Significantly, a large prominent groove can be clearly seen that is formed first by the MTERF4:NSUN4 interaction region and then by NSUN4 alone and that leads into the SAM binding pocket. This readily suggests a potential path for the RNA (yellow dashed line in Figure 4-7) and a model for how MTERF4 can direct binding of the RNA substrate by NSUN4. This model is consistent with the orientation of the SAM cofactor in the groove. Moreover, the model suggests that, in addition to the extensive structural conservation between MTERF proteins, the mechanism of nucleic acid binding might also be conserved, at least in its general features. Given the strongly structured nature of rRNAs, it is possible that the RNA molecule bound by MTERF4 might exhibit some double-stranded character, increasing the similarities with the

MTERF1 binding mechanism. The binding groove in MTERF4 is sufficiently wide to accommodate a second RNA strand (Figure 4-7A).

It is interesting to note that the long acidic MTERF4 C-terminal tail that is absent in our structure would be expected to be located near the putative RNA binding groove (asterisk in Figure 4-7B). It is tempting to speculate that a function of this tail might be to increase the binding specificity by competing for the same binding surface with the RNA, as has been previously shown for other proteins [213]. Alternatively, it is also possible that this acidic tail serves to mediate additional protein-protein interactions that might help coordinate the MTERF4:NSUN4 complex with the rest of the mitochondrial rRNA modification machinery.

4.4 Conclusions

In this paper, we have described the structure of a MTERF4:NSUN4 complex that is essential to direct the activity of the NSUN4 methyltransferase to its methylation site in the mitochondrial 16S rRNA. This novel structure is the first example of a protein-protein interaction that involves an MTERF protein. MTERF family members have been previously postulated to interact with other proteins to help control their regulatory activities or explain the strong polarity observed in transcription termination by MTERF1 [202, 211, 214]. Our structure reveals that the interaction does not involve the MTERF fold, which is likely exclusively devoted to nucleic acid binding. Rather, the interaction mostly involves a region C-terminal to the MTERF fold that appears to be conserved in other MTERF proteins, suggesting that they could utilize an analogous binding mode to mediate protein-protein interactions. Moreover, the structure suggests a model for how the MTERF4:NSUN4 complex utilizes the MTERF fold to associate with RNA and direct methylation by NSUN4. This highlights how an MTERF protein can cooperate with a separate enzymatic activity by endowing it with substrate and/or sequence specificity. In this respect, it is important to note that NSUN4 appears devoid of any intrinsic sequence specificity, therefore suggesting that this specificity is conferred by the MTERF fold. Future experiments are needed to fully understand how these proteins, despite sharing a very similar fold, can nevertheless have very different substrate specificities, including affinity for a different nucleic acid.

4.5 Accession numbers

Coordinates and structure factor files have been deposited in the RCSB Protein Data Bank with accession code 4FZV.

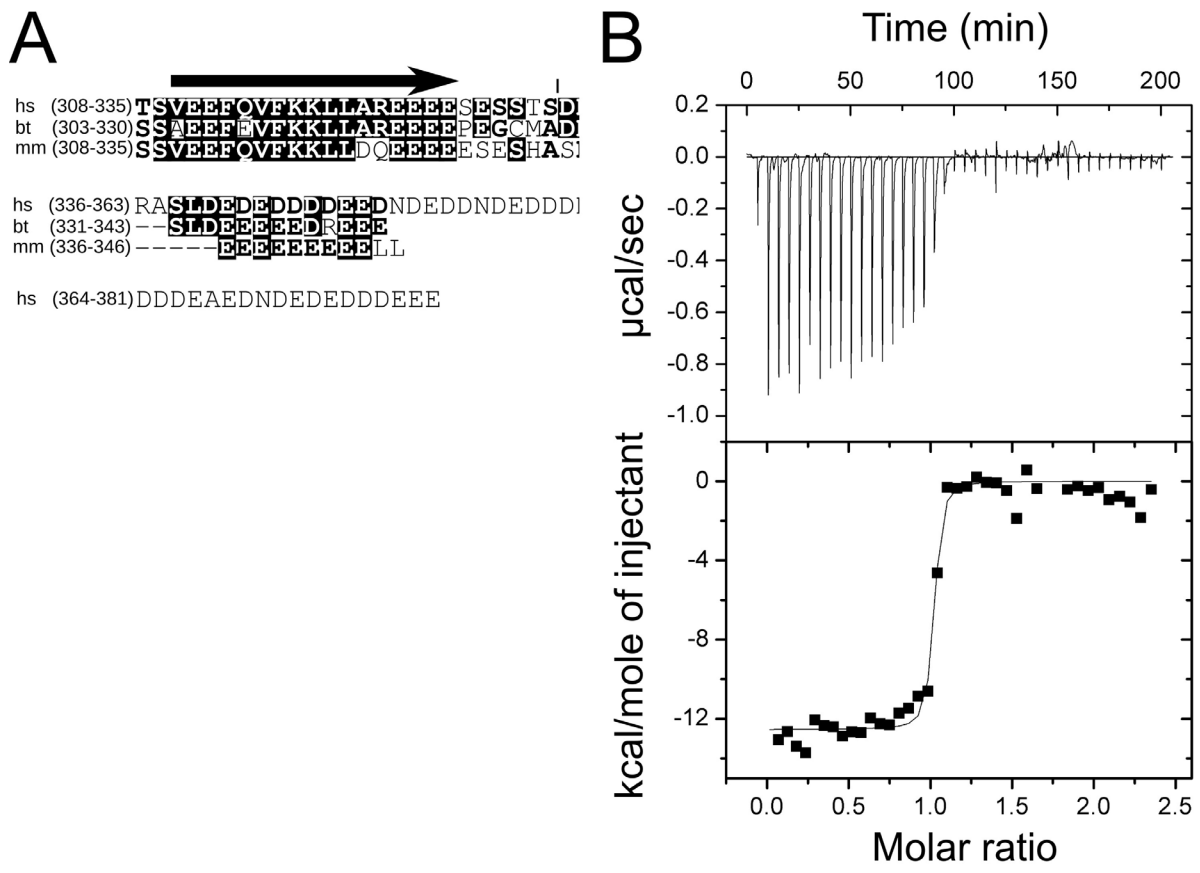


Figure 4-1. Interaction between MTERF4 and NSUN4

A. Alignment of the C-terminal regions of human (hs), mouse (mm) and bovine (bt) MTERF4. Despite differences in length, the acidic character of the C-terminus is conserved. The acidic region begins at the C-terminal end of a conserved α-helix and spans ca 30 residues, although 30 additional residues are present in the human protein (45 of the final 59 residues are acidic). The arrow represents a predicted α-helix. The vertical line represents the truncation site in the construct reported in this manuscript. **B.** Isothermal titration calorimetry of MTERF4 binding to NSUN4 at 25 °C. Integrated heat measurements from 3µl injections of 220 µM MTERF4 into the calorimeter cell containing NSUN4 at an initial concentration of 20 µM (top panel). A standard one-site model was used for curve fitting (bottom panel). The interaction between NSUN4 and MTERF4 is characterized as an exothermic process with ΔH of -11.04 (±0.22) kcal/mol, a K_D of 13.3 nM (±0.4), ΔS of -1.05 cal/mol K and a stoichiometry of N=1.08 (±0.02). See also Figure S1.

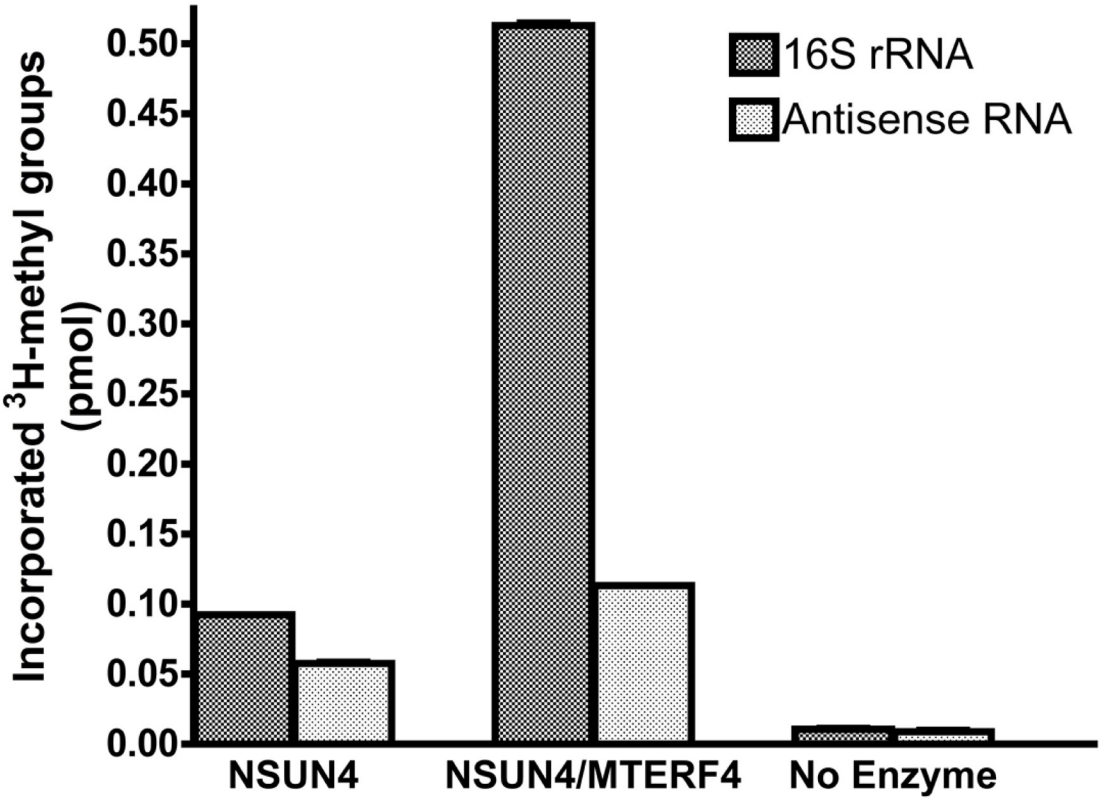


Figure 4-2. MTERF4 stimulates the *in vitro* activity and specificity of NSUN4 on the mitochondrial 16S rRNA

In vitro transcribed 16S mitochondrial rRNA and the corresponding antisense RNA were used for an *in vitro* ^{3}H-methyl incorporation assay in order to test the activity of purified NSUN4 and the MTERF4:NSUN4 complex. The number of moles of methyl groups incorporated into the rRNA substrate was calculated from the measured disintegrations per minute (DPM). Vertical bars correspond to the mean ± SD of three independent experiments. Note that the error bars on the graph are small and may be difficult to see. The same reaction performed in the absence of enzyme was used as a control.

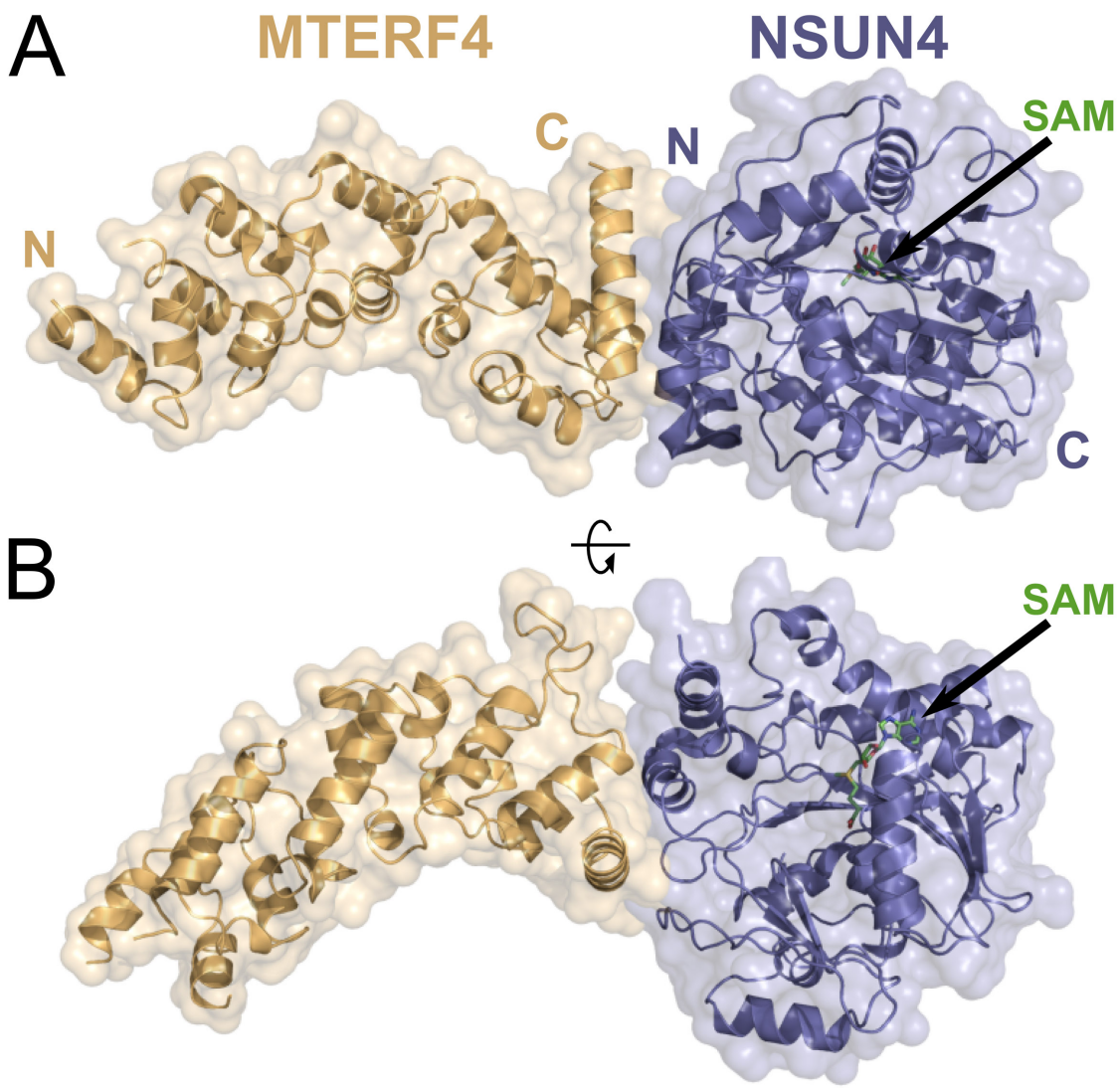


Figure 4-3. Crystal structure of the MTERF4:NSUN4 complex
A. Global view of the MTERF4:NSUN4 complex. MTERF4 is rendered in orange. NSUN4 is blue. The molecular surface is rendered transparent. N, N-terminal end; C, C-terminal end. The black arrow points to the position of the SAM residue. **B.** A 90° rotation of the complex around the long axis.

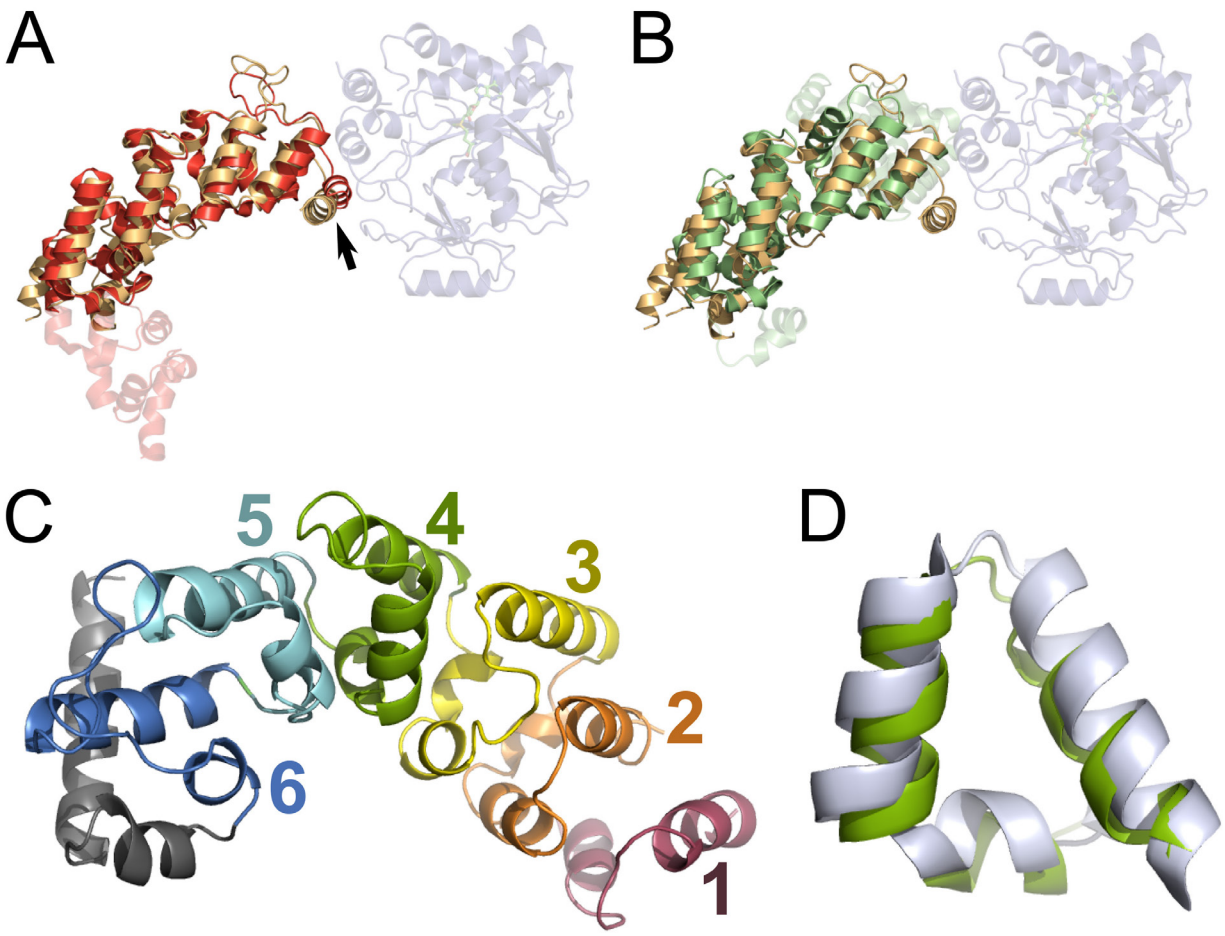


Figure 4-4. MTERF4 adopts the canonical MTERF fold

A. Overlay of MTERF4 (orange) and MTERF3 (red). The rmsd was 2.37 Å for 165 C-α atoms. The arrow points to the α-helix involved in heterodimer formation. **B.** Overlay of MTERF4 (orange) and MTERF1 (green). The rmsd was 4.33 Å for 165 C-α atoms. **C.** Five canonical MTERF motifs (1-5) could be resolved in MTERF4. Only the backbone could be assigned for motif 1 (see text). The final C-terminal MTERF motif (6) is distorted and followed by two α-helices (gray). See also Figure S2. **D.** Structural conservation of the MTERF motif. Overlay of MTERF motifs from MTERF4 (green) and MTERF1 (light gray).

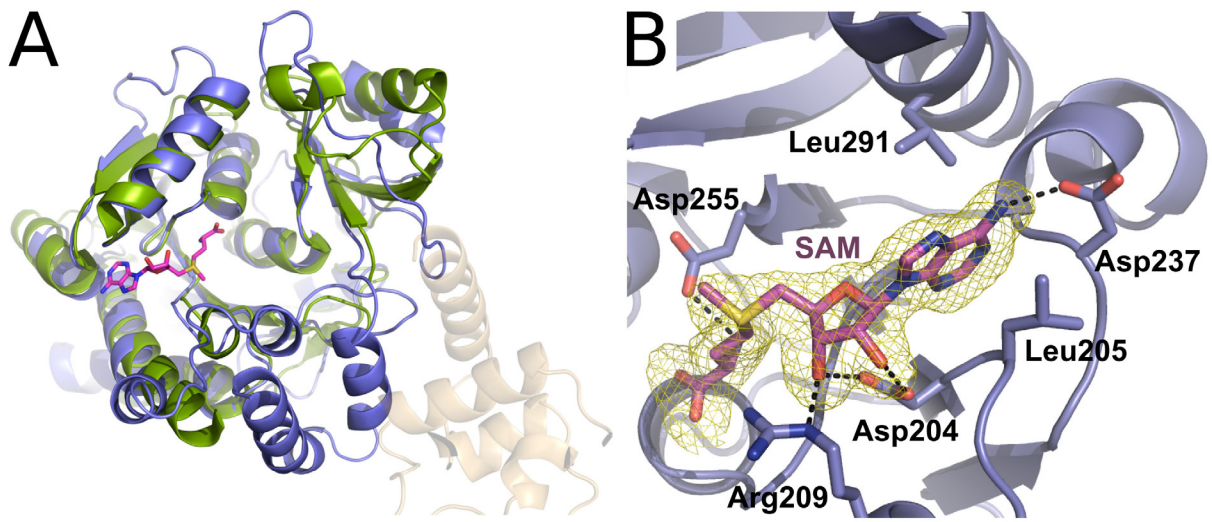


Figure 4-5. NSUN4 is related to m5C methyltransferases
A. Overlay of NSUN4 (blue) with the Trm4 m5C methyltransferase from *M. jannaschii* (green; PDB 3A4T). The rmsd was 0.78 Å for 116 C-α atoms. The SAM molecule corresponding to the MTERF4:NSUN4 structure is shown (magenta). See also Figure S3. **B.** View of the SAM binding pocket in NSUN4. SAM (magenta) is bound in a positively charged binding pocket and establishes several interactions with NSUN4 side chains. A simulated annealing F_o-F_c omit electron density map is shown (yellow) contoured at 4σ. Hydrogen bonds are shown as dashed lines.

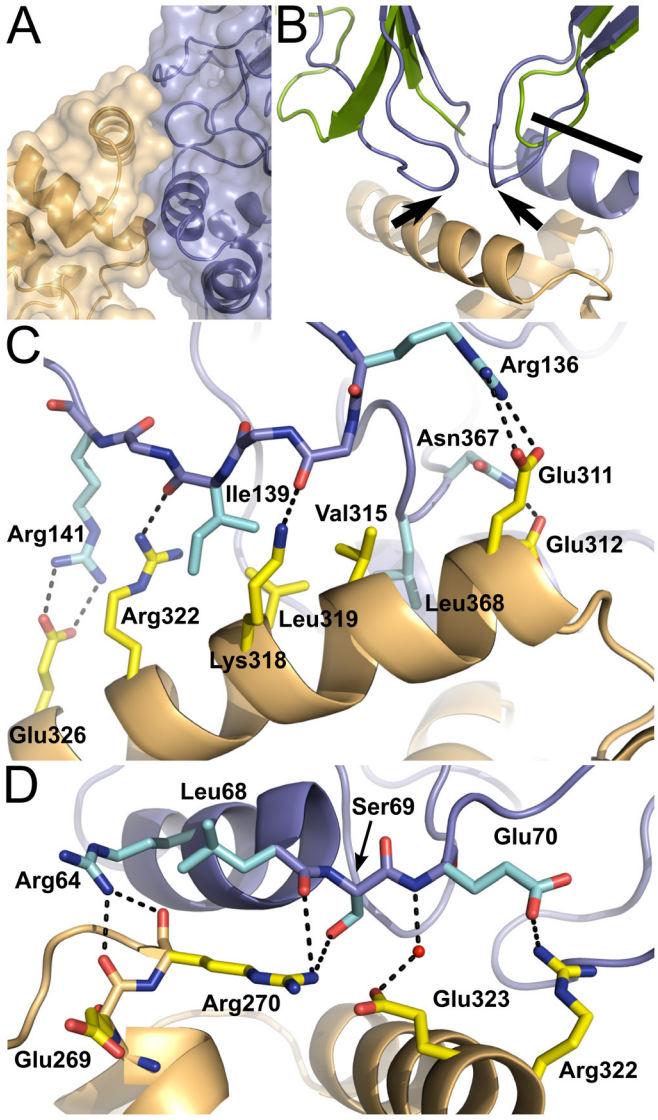


Figure 4-6. The interaction with NSUN4 involves the region C-terminal of the MTERF fold
A. Region of interaction between MTERF4 and NSUN4. The C-terminal distorted MTERF motif (6 in Figure 4C) of MTERF4 is involved in the interaction. Contacts are mainly established with one α-helix and two loops in NSUN4. The molecular surface is rendered transparent. **B.** The two loops in NSUN4 involved in the interaction with MTERF4 are not conserved in other m5C methyltransferases. Overlay of NSUN4 with the methyltransferase Trm4 from *M. jannaschii* (green), showing the two non-conserved loops (black arrows) and the N-terminal helix involved in the interaction (black bar). **C.** Interactions between MTERF4 and NSUN4. Most interactions are established between a C-terminal α-helix in MTERF4 (orange) and one of the non-conserved loops in NSUN4 (blue; see text). NSUN4 side chains are cyan. MTERF4 side chains are yellow. Hydrogen bonds are shown as dashed lines. **D.** Additional region of interaction between MTERF4 and NSUN4. Contacts also involve an α-helix in NSUN4 (in addition to the two loops in NSUN4 shown in panel C). NSUN4 side chains are cyan. MTERF4 side chains are yellow. Hydrogen bonds are shown as dashed lines.

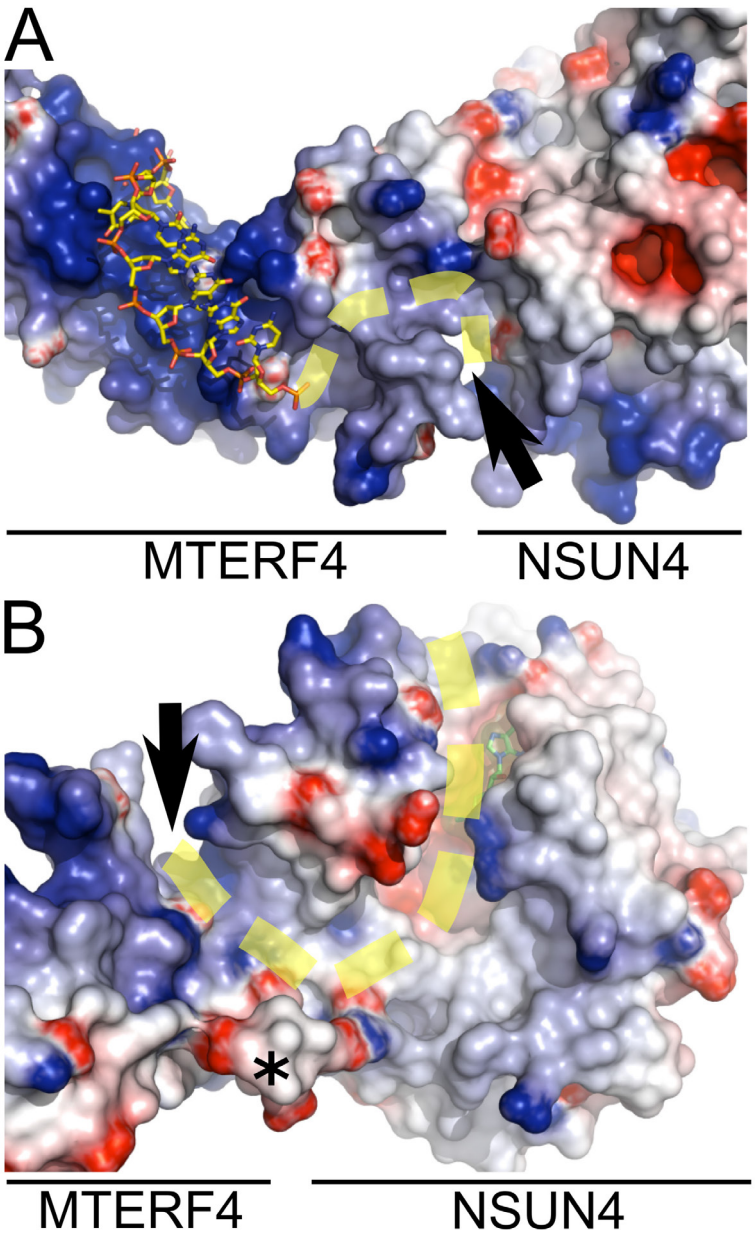


Figure 4-7. A proposed model for RNA binding by the MTERF4:NSUN4 complex
A. MTERF4 has a basic surface in the central groove postulated to mediate nucleic acid binding. Part of one of the DNA chains in the MTERF1 structure (3MVA) is shown (yellow). The RNA could bind in a similar mode and track through a groove (black arrows) present in the complex that leads to the SAM binding site. The proposed path of the RNA is shown as a dashed yellow line. **B.** A different orientation is shown, highlighting the groove in NSUN4. SAM is shown in green. The last residue modeled in the C-terminus of MTERF is shown with an asterisk. The electrostatic surface potential map was generated with Delphi (Honig and Nicholls, 1995) and is colored from -7 kTe^{-1} (blue) to +7 kTe^{-1} (red).

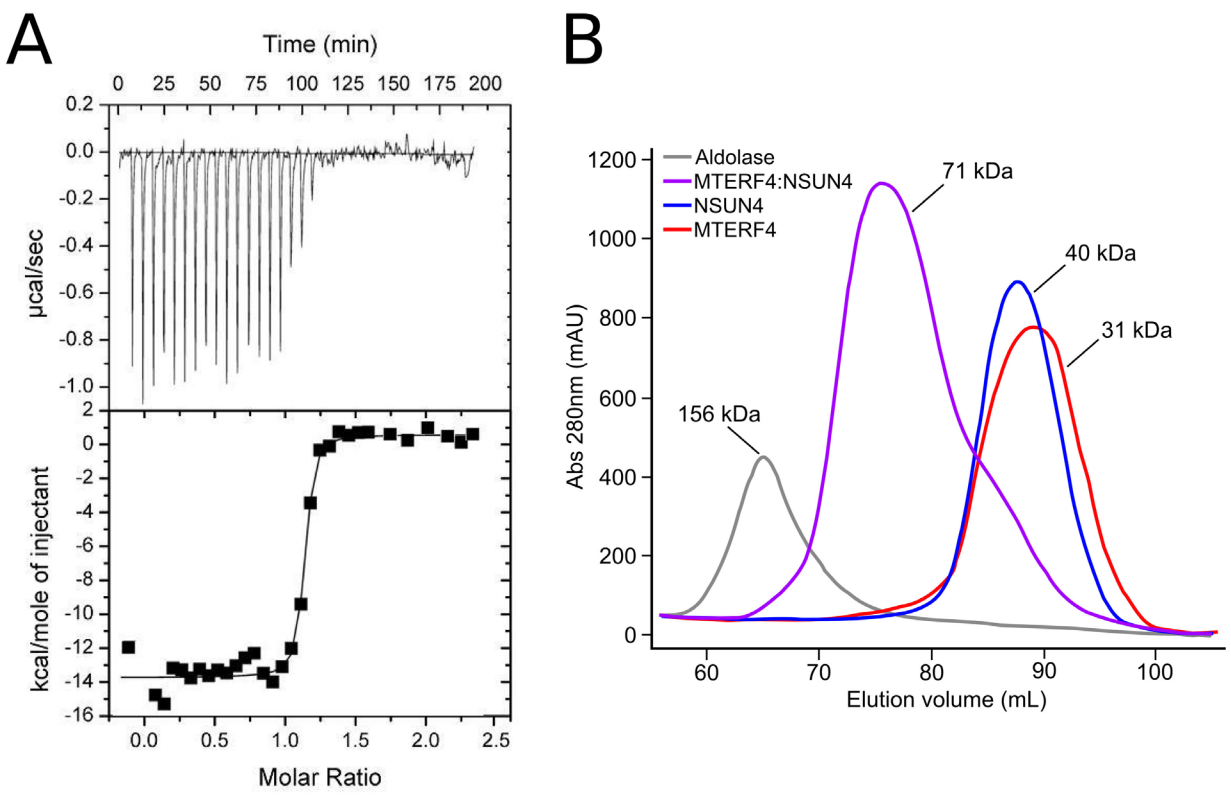


Figure 4-8. MTERF4 and NSUN4 form a tight complex with 1:1 stoichiometry

A. Isothermal titration calorimetry of MTERF4 binding to NSUN4 in the presence of SAM at 25 °C. Integrated heat measurements from 5 µl injections of 250 µM MTERF4 into the calorimeter cell containing NSUN4 and SAM at initial concentrations of 25 µM and 50 µM, respectively (top panel). A standard one-site model was used for curve fitting (bottom panel). The interaction between NSUN4 and MTERF4 is characterized as an exothermic process with ΔH of −15.71 kcal/mol, a K_D of 18.70 nM , ΔS of −0.80 cal/mol K and a stoichiometry of N = 1.1

B. Stoichiometry of the MTERF4:NSUN4 complex determined by size exclusion chromatography. Elution profiles of MTERF4 (red), NSUN4 (blue), the MTERF4:NSUN4 complex (purple) and aldolase (gray; control). Labels indicate the theoretical molecular mass predicted by amino acid sequence. Chromatography was performed on a HiLoad Superdex 200 16/60 (GE Healthcare).

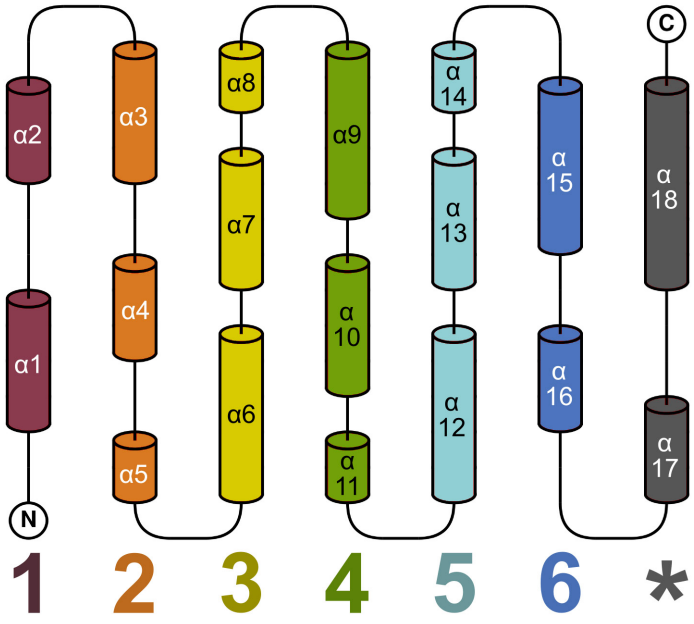


Figure 4-9. The overall fold of MTERF4
Topology diagram of the protein fold, colored and numbered by MTERF motif (as in Figure 4C). Five canonical MTERF motifs (1-5) could be resolved in MTERF4. The final C-terminal MTERF motif (6) is distorted and followed by two α-helices (gray; asterisk). Helices were identified using DSSP. Coiled regions are indicated by solid black lines.

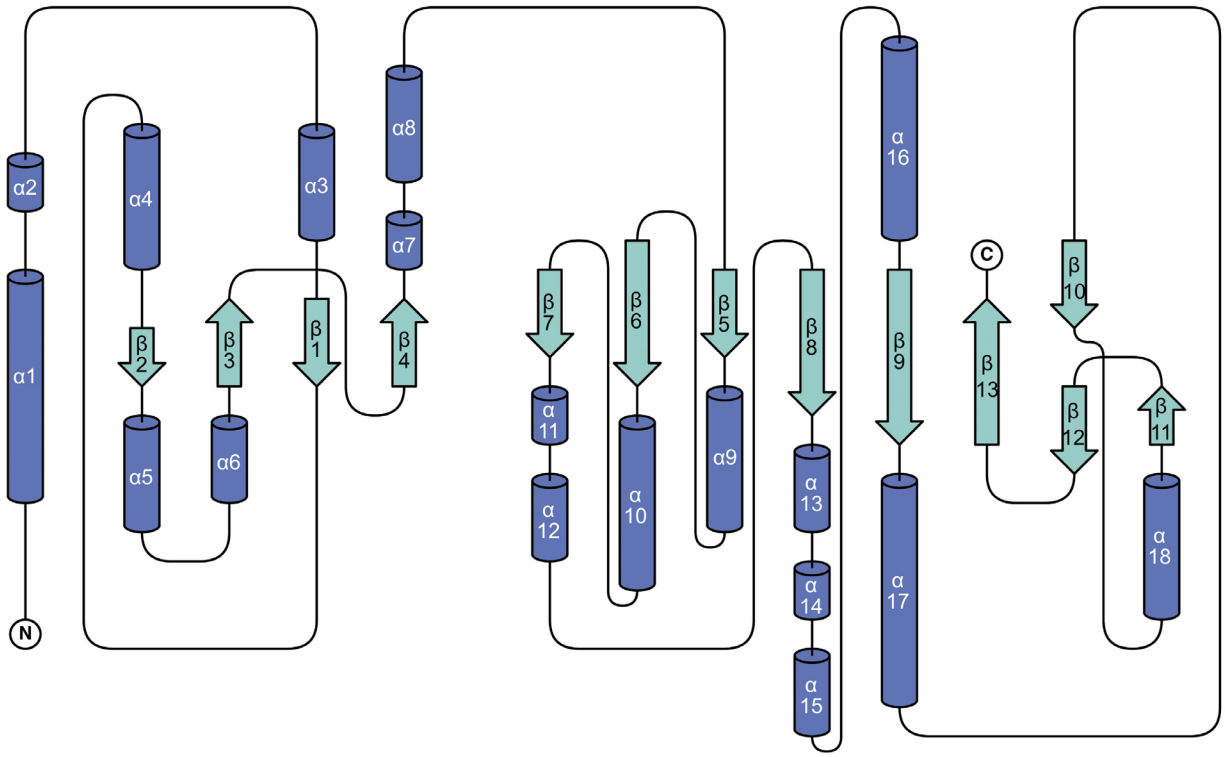


Figure 4-10. The overall fold of NSUN4
Topology diagram of the protein fold. Strands (green) and helices (blue) were identified using DSSP. Coiled regions are indicated by solid black lines.

Table 4-1. Data collection and refinement statistics for the MTERF4:NSUN4 complex

Crystal	Native-MTERF4:SeMet-NSUN4		
Data collection			
Space group	*C 1 2 1*		
Cell dimensions			
a, *b*, *c* (Å)	298.68, 53.25, 53.04		
α, β, γ (°)	90.00, 98.76, 90.00		
Wavelength (Å)	*Peak*	*Inflection*	*Remote*
	0.9790	0.9795	0.9500
Resolution (Å)	46.82−2.00 (2.06−2.00)	49.21−2.09 (2.10−2.09)	46.84−2.18 (2.19−2.18)
R_{merge}	0.058 (0.623)	0.041 (0.548)	0.049 (0.546)
I / σI	23.7 (2.4)	26.3 (3.0)	25.6 (3.0)
Completeness (%)	99.9 (99.6)	99.8 (99.8)	99.9 (99.3)
Redundancy	6.6 (6.1)	6.7 (6.8)	6.7 (6.8)
Refinement			
Resolution (Å)	46.82−2.00		
Unique reflections	55,979		
R_{work} / R_{free}	0.1688 / 0.1960		
No. atoms			
Total	4633		
Protein	4195		
Water	301		
SAM	26		
B-factors			
Protein	54.8		
SAM	20.0		
Water	48.4		
RMS deviations			
Bond lengths (Å)	0.010		
Bond angles (°)	1.29		
Ramachandran			
Favored (%)	95.0		
Outliers (%)	0		
PDB ID	4FZV		

* Values in parentheses are for highest-resolution shell.

Chapter 5: Sequence-specific hydrogen bonds modulate base-flipping and termination of transcription by MTERF1

This chapter contains a description of results that will be incorporated in a manuscript currently in preparation: **Guja KE**, Byrnes J, Mejia E, Kalkenings L, Yakubovskaya E, Garcia-Diaz M. 2015. "Sequence-specific hydrogen bonds modulate base-flipping and termination activity by MTERF1."

Author contributions: Cloning of variant proteins was carried out by KEG and EM; expression and purification was done by KEG, EM, LK, and EY; crystallization was done by KEG, EM, and LK; X-ray data collection, structure solution and refinement was done by KEG; termination assays and oligonucleotide synthesis were done by JB; design of the experiments was done by KEG, JB, and MGD.

Note: this chapter contains a broader and lengthier introduction in consideration of the change in focus from rRNA modification to transcription. Some aspects of mitochondrial transcription are reviewed and new experimental results are also described.

5.1 Abstract

Defects in mitochondrial gene expression are associated with aging and disease. Mitochondrial transcription termination factor (MTERF) proteins have been implicated in modulating transcription, replication and ribosome biogenesis. Previously, we reported the first crystal structure of a member of this family, the human mitochondrial transcriptional terminator MTERF1, bound to double-stranded DNA encoding the termination sequence [103]. The structure indicates that upon sequence recognition MTERF1 unwinds the DNA molecule, promoting eversion of three nucleotides. Base flipping is critical for stable binding and transcriptional termination. Here we report four crystal structures of MTERF1 variants, which reveal that sequence-specific hydrogen bonds play an essential role in this unique DNA binding mechanism, while in vitro transcription termination assays with these variants demonstrate that sequence specific hydrogen bonds also play a critical role in the termination activity of

MTERF1. Our results provide further insight into the mechanism of transcription termination in mitochondria, and help improve our understanding of pathogenic mtDNA mutations that affect termination.

5.2 Introduction

Genetic or age-related defects in mitochondrial gene expression can reduce or eliminate mitochondrial function and cause multiple human pathologies, including neurodegenerative and cardiovascular disease, diabetes and cancer. Mitochondrial transcription is a key process for gene expression, and accordingly some of these pathogenic alterations have been associated with proteins involved in transcription. However, our functional understanding of the different components of the transcription machinery and its regulation in mitochondria is relatively poor. Characterizing the molecular mechanisms underlying mitochondrial transcription if of significant interest, as it will allow us to understand how small genetic or environmental alterations in this process can result in gene expression deficiencies and the development of mitochondrial pathology.

Mitochondrial transcription is preferentially terminated within the $tRNA^{Leu}$ gene immediately downstream of the 16S rRNA gene [215, 216]. This site-specific termination event is mediated by a DNA-binding protein, mitochondrial transcription termination factor 1 (MTERF1), first identified by Attardi and colleagues [102], and later shown to be sufficient to mediate transcriptional termination in vitro [134]. Subsequent in vitro work has demonstrated that termination by MTERF1 is bidirectional [134, 217], and shows even higher efficiency when POLRMT proceeds in the direction of L-strand transcription [134].

In addition to binding the termination site in $tRNA^{Leu}$, MTERF1 may also bind an additional site in the HSP1 promoter region and stimulate transcription [211]. In a manner that bears some similarities with prokaryotic transcription, it has been proposed that simultaneous binding of MTERF1 to the HSP1 site and the canonical $tRNA^{Leu}$ termination site causes a looping-out of the rDNA. This loop would allow recycling of POLRMT and other transcription components from the $tRNA^{Leu}$ site to the HSP1 initiation site for efficiently starting another transcription cycle after termination occurs. This ‘‘ribomotor’’ model [218] is one way that the greater abundance of rRNAs compared to the downstream H-strand mRNAs can be explained [219], although differential stability of rRNA and mRNA species is thought to be a significant

factor [219]. Moreover, this model would also explain the selective termination of transcription originating at HSP1 and suggests a mechanism for regulating the balance between termination and read-through. In addition, some evidence also suggests that MTERF1 may be involved in transcription initiation from the HSP2 [211]. The mechanistic basis for this remains unknown as does the precise protein components required for regulation at HSP2, which is the least studied of the three known human mtDNA promoters. Recently, Jacobs and colleagues showed that MTERF1 binds several sites in mtDNA in vivo and that altered MTERF1 protein levels affect mtDNA replication pausing at these sites [220], leading to a model in which MTERF1 may mediate transcription and replication passage on the same mtDNA molecule. It is important to note that the existence of such alternative MTERF1 binding sites is controversial, and subsequent studies have not been able to reproduce binding of MTERF1 to the HSP [115]. Furthermore, in vitro experiments indicate that the specificity of MTERF1 seems to be exquisite for the $tRNA^{Leu}$ site [103], implying that the affinity for other sites (if they exist) would be lower. This hypothesis is consistent with the striking degree of protection of the MTERF1 $tRNA^{Leu}$ site observed in in vivo methylation [221] and DNaseI footprinting assays [222], as well as with ChIP studies that have predominately found MTERF1 bound to the $tRNA^{Leu}$ site.

5.2.1 Structure of MTERF1

The X-ray crystallographic structure of human MTERF1 reveals that the 342-residue protein (mature form) has an all-α-helical structure that binds as a monomer to a 22-nucleotide termination sequence in the $tRNA^{Leu}$ gene [103, 206]. The MTERF1 protein structure is modular, being configured around a motif of two α-helices and a 3_{10} helix repeat known as the MTERF repeat or MTERF motif. This MTERF repeat configuration was previously predicted by Roberti et al [223] and is similar in structure to other all-α-helical domains such as the HEAT domain [224], suggested to play a role in duplex DNA binding in DNA PK-cs [225], and the PUM/PUF domain, which is involved in RNA binding by PUMILIO proteins [226]. The 8 MTERF repeats found in MTERF1 constitute a helical fold that allows the protein to bind the major groove of duplex DNA containing the termination sequence. The extensive surface interactions observed between MTERF1 and the substrate DNA in the crystal structure indicate that the protein fold is likely dedicated to duplex DNA binding.

The MTERF1 crystal structure demonstrates that upon binding its target sequence, the protein alleviates the DNA duplex twist and promotes duplex melting and eversion of three nucleotide bases, leading to a novel and unique DNA binding mode (Figure 5-1). Although the mechanism by which it promotes this eversion, or base-flipping is not yet clear, MTERF1 stabilizes these three nucleotides in an extra-helical conformation through stacking interactions with three amino acid residues (Phe243, Tyr288 and Arg162; see Figure 5-1 inset). We have demonstrated, using an MTERF1 triple mutant, RFY, that these three side chains are essential to maintaining the conformation observed in the crystal structure and although the MTERF1 triple mutant can still bind to the termination sequence, affinity for this termination site is significantly reduced. Moreover, the ability of the triple mutant MTERF1 to promote transcriptional termination in vitro is dramatically reduced. Although it is clear that the pi-stacking interactions provided by Phe243, Tyr288, and Arg162 are essential for both base-flipping and in vitro termination of transcription by MTERF1, it is not clear if any other mechanisms contribute to the stabilization of base-flipping and termination activity.

5.3 Results and discussion

The wild-type MTERF1 crystal structure suggests that sequence-specific hydrogen bonds, in addition to pi-stacking interactions, may play a role in stabilizing the everted nucleotides. As shown in Figure 5-2A, the everted nucleotides Thy3243, Ade3243, and Cyt3244 form hydrogen bonds with Asn199, Ser285, and Glu280, respectively. To assess the role of these amino acid residues in base-flipping and termination by MTERF1, we constructed individual variants with alanine substitutions at these positions, as well as a fourth variant that combining all three substitutions. In an attempt to visualize and thus better understand the role of these residues and their cognate hydrogen bonds in the base flipping process, we obtained crystal structures of N199A, E280A, S285A, and N199A/E280A/S285A (NES) variant MTERF1 proteins in complex with the termination sequence (data collection and refinement statistics shown in Table 5-1).

5.3.1 Crystal structures of the MTERF1 variants reveal perturbations of base-flipping

The crystal structure of the MTERF1-N199A variant in complex with a 22mer encoding the termination sequence in the tRNA-Leu gene demonstrates that base-flipping is nearly eliminated (Figure 5-2B). Thy3243, Cyt3244, and Ade3243 all show dramatic shifts back into the DNA helix, and Thy3243 returns to an intrahelical position, as it is no longer stabilized in the everted position by Asn199. Cyt3244 and Ade3243 are also significantly shifted back into the helix, however the geometry of all three bases remains somewhat distorted, displaying clear tilts or rotations relative to canonical DNA stacking. Moreover, the dramatic distortion of the DNA backbone imposed by MTERF1 binding remains, and perhaps contributes to the perturbed geometry of Cyt3244 and Ade3243. In comparison to the RFY variant, Cyt3244 is shifted farther back into the helix, and although Thy3242 is also intrahelical, it is clearly shifted upstream towards the 3242 position (Figure 5-3B).

Base-flipping is somewhat less affected in the E280A variant, and Ade3243 remains in an everted position that is nearly identical to the wild-type MTERF1 structure, where it is stabilized by a hydrogen bond with Ser285 (Figure 5-2C). Thy3243 on the other hand, shifts dramatically back into the helix, as it did in the N199A variant – despite the presence Asn199 in this case. The effect on Cyt3244 falls somewhere in the middle, as it adopts alternative conformations – one in the wild-type position, and interestingly, one that is even more everted than wild-type.

The S285A substitution, as one might expect, allows Ade3243 to adopt a more intra-helical position, as it is no longer stabilized outside the helix by a hydrogen bond with Ser285 (Figure 5-2D). Interestingly, the S285A variant displays an even more severe perturbation of Cyt3244 than E280A, which directly abolishes its hydrogen-bonding partner. Cyt3244 is displaced farther outside the helix, while Thy3243 returns to an intra-helical position.

The MTERF1-NES variant, which combines all three substitutions, surprisingly does not display the most severe base-flipping defect (Figure 5-2D). Both Ade3243 and Thy3243 shift back into the helix, but Cyt3244 remains in the wild-type position. All four variants display some level of base-flipping defect, although the effect of substitutions are not additive as one might expect, and moreover, the structures all demonstrate some perturbation of base-flipping. Moreover, all of the variant structures display some degree of shift relative to the RFY variant as

well (Figure 5-3). One might expect that both the RFY and NES variants would be completely deficient in base-flipping and thus have similar conformations, yet that is not the case.

5.3.2 Impaired in vitro termination activity of the MTERF1 variants

To analyze the functional importance of the hydrogen bonds formed between Asn199, Glu280, Ser285 and their cognate DNA base partners, we examined the ability of the different variant MTERF1 proteins to promote transcriptional termination in a reconstituted in vitro system. We utilized the assay described in our previous work and generated a substrate for run-off transcription from the HSP where the termination sequence – the same 22-mer sequence used for crystallization – has been inserted 100 bp downstream from the promoter (see Methods described in Chapter 2)[103]. The transcription initiation machinery (consisting of purified TFAM, TFB2M, and POLRMT) generates a unique run-off transcription product on this substrate, while the addition of MTERF1 results in the appearance of a specific, shorter termination product (Figure 5-4). Our results are essentially equivalent to previously published data [103, 134], confirming that our purified MTERF1 is active and, at least in vitro, MTERF1 displays robust termination activity in the LSP direction.

The most severe termination defects were evident in the NES variant, and the S285A variant. The individual N199A and E280A variants might be expected to show more moderate results, given that there are two flipped bases on the heavy strand, but only one on the light strand (which is stabilized by Ser285). Surprisingly, the N199A variant retains nearly wild-type termination activity. This unexpected finding underscores the complex nature of the base-flipping process and emphasizes the non-additive effects of substitutions on base-flipping.

5.3.3 Additional factors that contribute to DNA binding and termination by MTERF1

With the exception of those described above, the majority of the interactions observed between MTERF1 and DNA involve the phosphate backbone of the substrate DNA sequence, and are therefore mostly electrostatic and non-specific. Sequence specificity appears to be determined in large part by a small number of key interactions between six arginine residues and guanine bases in the termination sequence. This type of major groove interaction is frequently

seen in sequence-specific DNA binding proteins and is thought to be critical for MTERF function – we have shown that eliminating even a single one of these interactions can drastically affect both DNA binding and transcription termination [103]. Interestingly, this mechanism of sequence recognition implies that while the interaction between MTERF1 and its binding sequence involves contacts with 20 base pairs, only six of the 40 bases appear to be initially actively recognized by the protein. The extensive total number of protein-DNA interactions suggest that the entirety of the MTERF1 fold is involved in binding the termination sequence, but it is not immediately apparent how a single MTERF1 molecule could simultaneously bind both the HSP initiation and termination sites in the transcriptional loop model [211]. One possibility is that an additional molecule of MTERF1 or other factor(s) may mediate the association of these two sites and facilitate loop formation.

5.3.4 Model for termination by MTERF1

The MTERF1 crystal structure suggests a binding mechanism that involves establishment of site-specific interactions for sequence recognition followed by melting and unwinding of the DNA duplex. This unwinding would presumably destabilize base-pairing of the central nucleotides in the recognition sequence and facilitate subsequent base flipping, thereby stabilizing MTERF1 on the substrate DNA. We have demonstrated that stable MTERF1 binding and termination in vitro is dependent on base flipping, and the ability of MTERF1 to promote termination appears to be at least partially dependent on the strength of the cumulative interactions between MTERF1 and the termination sequence. This would suggest a model of transcription in which MTERF1 acts as a "roadblock," preventing or interfering with transcriptional elongation. This model is consistent with the observation that MTERF1 terminates transcription bi-directionally and can arrest elongation by heterologous polymerases ([217] and our own unpublished results). However, in vitro termination by MTERF1 displays a distinct polarity, with MTERF1 being more efficient when terminating transcription originating from the light strand promoter than the heavy strand promoter [103, 134]. This polarity may be a result of the majority of protein-DNA interactions being established with the light strand [103] (ie, the strand transcribed from the LSP promoter) as well as the higher affinity of MTERF1 for this strand, as observed by Nam et al [227]. This suggests a mechanism in which MTERF1

might transiently bind to single-stranded DNA, although the measured affinity of MTERF1 for single-stranded DNA is extremely low ([227] and our own unpublished observations). Nevertheless, except for the asymmetrical distribution of interactions, no obvious structural feature provides an explanation for the observed polarity of termination. One possibility is that the observed orientation dependence of termination activity is due to the unique conformation of MTERF1 on DNA, although interactions between MTERF1 and mitochondrial POLRMT (or additional elongation factors) may also influence the polarity of termination events. Furthermore, it is not yet known whether other proteins can modulate termination polarity in vivo.

5.3.5 Significance of termination at $tRNA^{Leu}$

The strong polarity observed in termination assays in vitro combined with the lack of in vivo evidence supporting a role in HSP termination suggest that MTERF1-mediated HSP termination might only be a secondary role and that the main function of MTERF1 is termination of LSP transcription at the $tRNA^{Leu}$ site. This hypothesis is supported by the fact that the light-strand does not encode any additional genes beyond $tRNA^{Leu}$. Recent observations in vivo have shown that manipulation of MTERF1 expression levels is strongly correlated with alterations in the relative amounts of antisense transcripts on both sides of the $tRNA^{Leu}$ termination site [228]. These findings are consistent with the notion that termination at this site may be important for preventing the accumulation of antisense transcripts that would otherwise interfere with the assembly of the rRNAs into ribosomes. This idea would imply that MTERF1 function is, as originally thought, important for ribosome biogenesis, albeit by a different mechanism than first proposed.

5.3.6 Termination of HSP transcription at the distal site

Any discussion of mitochondrial termination is incomplete when only considering termination at the $tRNA^{Leu}$ site. Evidently some degree of read-through must occur at the $tRNA^{Leu}$ site so that HSP transcription can progress beyond this site and generate most of the mitochondrial mRNAs. How such read-through is regulated, and whether it is specific to transcription originating from HSP2 is in itself an interesting question for which no answer yet

exists. In any case, read-through HSP transcription is thought to ultimately terminate at a distal site within the D-loop. The mouse HSP distal termination site was originally identified as a 22 bp region within the D-loop containing a conserved A/T rich sequence motif. Initial in vitro studies suggested that termination at this site is unidirectional and requires sequence-specific DNA binding proteins [229], but the mechanisms controlling distal termination in human mitochondria are far from being well understood. Since MTERF1 predominantly binds to the $tRNA^{Leu}$ site, it is possible that distal termination may not depend on MTERF1, therefore raising the question of how termination is achieved at that site. Recent work by Wanrooij et al suggests that the mitochondrial polymerase, POLRMT, utilizes a mechanism similar to bacterial rho-independent termination to generate the RNA primers necessary for mitochondrial replication [230]. In T7 bacteriophages transcription termination occurs as a consequence of the polymerase interacting with spontaneously generated double-stranded RNA hairpins and falling off DNA due to the weak affinity of T7 RNA polymerase for double stranded structures [231]. In the case of mitochondrial POLRMT, Wanrooij et al have shown that stable G-quadruplex structures can form in the nascent RNA, suggesting that these structures then mediate transcription termination in a way reminiscent of RNA hairpins in rho-independent termination [230]. It is therefore possible that a similar mechanism is responsible for termination of transcription at the HSP distal site. Nevertheless, it cannot be excluded that MTERF1 might influence in termination at that site. As mentioned earlier, alternative MTERF1 binding sites have been described in the proximity of the distal site [220], and these sites will need to be further characterized in order to elucidate how termination is achieved and what role, if any, MTERF1 plays in the process. Finally, it is also possible that proteins other than MTERF1 might contribute to distal HSP termination. Preliminary experiments have identified several proteins that appear to bind at the mouse HSP distal termination region, including the Leucine-rich pentatricopeptide-repeat containing protein (LRPPRC) [232]. This protein is essential for the expression of mtDNA encoded respiratory chain subunits, and LRPPC deficient cells were reported to have a reduction in both oxygen consumption and expression of mRNA and tRNA [232, 233]. Recent in vivo work in an RNAi system further supports the concept that LRPPRC is involved in regulating expression of mitochondrial mRNAs [234]. Moreover, several homologues of MTERF1 have been identified in recent years, and like LRPPRC, they have also been implicated in the regulation of mitochondrial transcription and gene expression.

Overall, these new findings on MTERF family members underscore the importance of these proteins in regulating mitochondrial gene expression. The results strongly suggest that these proteins have evolved to bind nucleic acids, although with remarkable differences in their substrate specificity, being capable of specific and/or unspecific binding to ssDNA, dsDNA and RNA in the case of MTERF4. Furthermore, the results of the aforementioned in vivo experiments highlight their potential to serve multiple functional roles.

5.4 Termination defects and implications for disease

Pathogenic mtDNA mutations were first reported in human patients more than two decades ago [10, 11, 235], and these important discoveries provided a genetic basis for the classification of mitochondrial disease. To date, a significant number of pathogenic mtDNA mutations have been identified and many of them have associated clinical phenotypes that have been characterized (reviewed in refs [12, 236, 237]). These mutations in the mtDNA often result in altered gene expression and therefore contribute to respiratory defects [12, 236, 238] and other mitochondrial pathologies that have been implicated in aging [9, 239, 240] as well as human disease. Most mtDNA mutations are thought to result in either defects in mitochondrial DNA maintenance or defects in mitochondrial translation. Such pathogenic defects in mitochondrial translation are frequently associated with mutations in tRNA genes, and interestingly, the $tRNA^{Leu}$ gene contains the highest number of identified mutations of any mitochondrial tRNA gene. Not surprisingly, several of these mutations occur in the MTERF1 binding sequence (Table 5-2). One such mutation, A3243G, is associated with mitochondrial myopathy, encephalomyopathy, lactic acidosis, and stroke-like episodes (MELAS), a syndrome which is characterized by lactic acidosis, episodic vomiting, seizures and recurrent cerebral insults resembling strokes that lead to hemiparesis, hemianopsia or cortical blindness [241]. The A3243G mutation, which has been shown to affect transcription in vitro [242], is the most common mtDNA mutation with a prevalence of more than 200 per 100,000 in a recent Australian study [243], and accounts for the majority (80%) of MELAS cases [244, 245]. Despite the observed in vitro termination defect, in vivo studies have shown that the A3243G mutation does not affect the balance between mitochondrial HSP transcripts upstream and downstream of the termination sequence [246]. Moreover, it has been shown that A3243G leads to defects in tRNA function, suggesting that these are the main cause of MELAS pathology

rather than transcriptional alterations [80]. However, it is important to note that the transcript ratios measured in these in vivo studies only provide a measure of HSP termination, and therefore it is possible that defects in LSP termination might exist in A3243G carriers and contribute to MELAS pathogenesis. Furthermore, it has also been suggested that termination may be regulated in a tissue-specific manner in vivo [247], such that termination defects occur predominately in particular cell types that have not yet been studied in vivo.

After the MTERF1 structure was solved, revealing the basis for MTERF1 sequence specificity, transcription termination assays performed on seven other mutations within the termination sequence also revealed effects on transcriptional termination (Table 5-2) [103]. Of these, two showed an effect that was significantly stronger than that observed with the A3243G mutation [103]. These mutations include the G3249A mutation that causes a variant of Kearns-Sayre syndrome [248] as well as the G3242A mutation, associated with an uncharacterized mitochondrial disorder [249]. Although tRNA mutations are generally thought to predominately affect the function of the mature tRNAs [250], these in vitro observations raise the possibility that the pathogenic effects of these mutations are also related to their effect on termination by MTERF1. Furthermore, given that MTERF1 appears to be needed for initiation at HSP2 and recycling of transcription at HSP1, it is possible that MTERF1 might play a role in promoter regulation, which could also be impacted by A3243G and other mtDNA mutations.

In considering the effects of these mutations it is important to keep in mind that, in addition to potential effects in transcription, MTERF1 appears to be able to modulate replication and therefore affect mitochondrial DNA maintenance [251]. Moreover, patients harboring pathogenic mtDNA mutations often are heteroplasmic, i.e. have a mixture of wild-type and mutated mtDNA molecules [21]. Furthermore, the appearance and severity of clinical symptoms seems to correlate with the mutation load [252]. Interestingly, it has been reported that long-term culture of human cells with the A3243G mtDNA mutation results in an increase in the relative levels of A3243G mutant mtDNA [253, 254], suggesting the mutation confers a replicative advantage. With this in mind, it is tempting to speculate that in some cases, the primary deficiency caused by mutations in the $tRNA^{Leu}$ MTERF1 binding site may be a defect in tRNA function, but the additional effect of these mutations on MTERF1 binding and/or function results in more rapid replication of the mutated genomes, thus facilitating an increase in the

mutation load. This, in turn, would contribute to the appearance of clinical symptoms and/or an increase in their severity.

Given the unusual and important roles of MTERF family members in transcription termination and initiation, as well as replication, more in-depth study of this interesting class of proteins is clearly warranted. A more detailed molecular understanding of how these proteins interact with the known transcription machinery, how they interact with each other, and how they are utilized differentially to control mtDNA expression in vivo will also provide insight into the pathogenesis of MELAS and other mitochondrial diseases

Table 5-1. Data collection and refinement statistics for MTERF1 variants

Crystal	MTERF1-N199A	MTERF1-E280A	MTERF1-S285A	MTERF1-N199A/E280A/S285A
Data collection				
Space group	*C 2 2 2₁*	*C 2 2 2₁*	*C 2 2 2₁*	*C 2 2 2₁*
Cell dimensions				
a, *b*, *c* (Å)	88.50 89.10 160.25	88.96 90.74 160.11	88.75 90.76 159.66	88.57 89.55 161.31
α, β, γ (˚)	90 90 90	90, 90, 90	90 90 90	90 90 90
Resolution (Å)	40.06 − 2.26 (2.269 − 2.261)	59.05 − 2.69 (2.703 − 2.695)	58.97 − 2.24 (2.244 − 2.237)	40.64 − 2.80 (2.893 − 2.804)
R_{merge}	0.045 (0.757)	0.037 (0.570)	0.034 (0.671)	0.033 (0.424)
R_{meas}	0.050 (0.832)	0.043 (0.656)	0.038 (0.764)	0.038 (0.526)
R_{pim}	0.021 (0.342)	0.020 (0.319)	0.18 (0.360)	0.018 (0.307)
$CC_{1/2}$	0.999 (0.809)	0.999 (0.841)	0.998 (0.834)	1.0 (0.780)
I / σI	20 (2.1)	19.5 (2.3)	22.3 (2.2)	31 (2.2)
Completeness (%)	99.4 (100)	98.9 (99.6)	99.0 (98.8)	97.6 (98.9)
Multiplicity	5.9 (5.9)	4.2 (4.1)	4.3 (4.1)	3.6 (2.7)
Wilson B (Å^2)	54.7	72.1	57.0	70.9
Refinement				
Resolution (Å)	40.06 − 2.26	59.05 − 2.69	58.97 − 2.24	40.64 − 2.80
No. reflections	28,151	18,049	30,710	15,829
R_{work} / R_{free}	0.1797 / 0.2202	0.2003 / 0.2606	0.2138 / 0.2472	0.2283 / 0.2624
No. atoms				
Total	3,661	3,560	3,578	3,540
Protein/DNA	3,590	3,521	3,525	3,508
Solvent	71	39	53	32
B-factors				
Protein/DNA	65.5	72.4	64.5	74.9
Solvent	59.5	68.1	63.7	65.3
R.M.S deviations				
Bond lengths (Å)	0.007	0.015	0.007	0.015
Bond angles (˚)	1.160	1.035	1.023	1.077
Ramachandran Favored/outliers (%)	98 / 0	96 / 0	98 / 0	95 / 0

Table 5-2 Pathogenic mutations in the MTERF1 binding site at tRNA-Leu

mtDNA mutation	Termination defect	Functional tRNA defect	Clinical disease
A3236G	+ [103]	–	Sporadic bilateral optic neuropathy [255]
G3242A	++ [103]	–	Uncharacterized mitochondrial myopathy [249]
A3243G	+++ [103, 242]	+ [80, 82, 256]	MELAS (>80% of cases) [244, 245]
A3243T	++ [103]	+ [257]	Mitochondrial encephalopathy [258]
G3244A	++ [103]	+ [80]	MELAS associated, not fully characterized [249]
G3249A	+++ [103]	–	Variant of Kearns-Sayre syndrome [248]
T3250C	+ [103]	+ [63]	Mitochondrial myopathy [259]
A3251G	+ [103]	–	Mitochondrial myopathy [260]
A3252G	– [103]	–	MELAS [261]

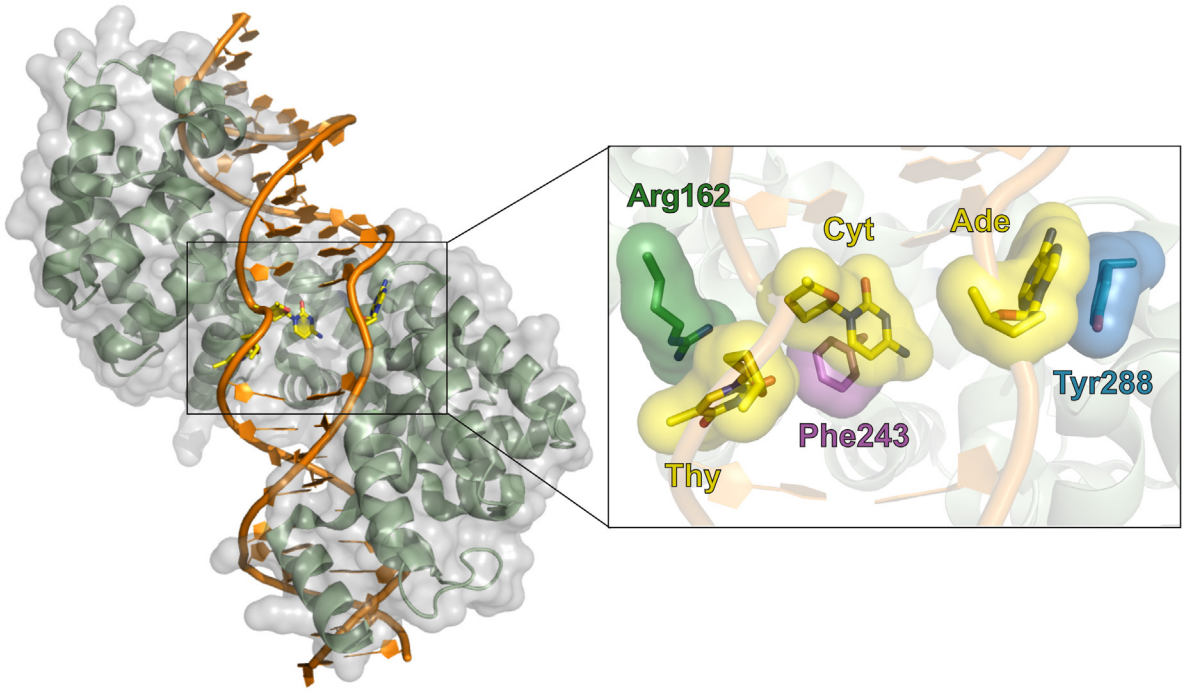


Figure 5-1. Crystal structure of the human MTERF1:mtDNA complex
MTERF1 utilizes a unique DNA binding mode that results in duplex unwinding and eversion of three nucleotide bases. MTERF1 is shown in green with its molecular surface shown in transparent grey; the DNA termination site is shown in orange; everted bases are shown as yellow sticks. *Inset:* π-stacking interactions with the MTERF1 residues (dark green, purple, and blue) stabilize the three corresponding DNA bases (yellow) in an extra-helical position. These interactions are essential for termination activity.

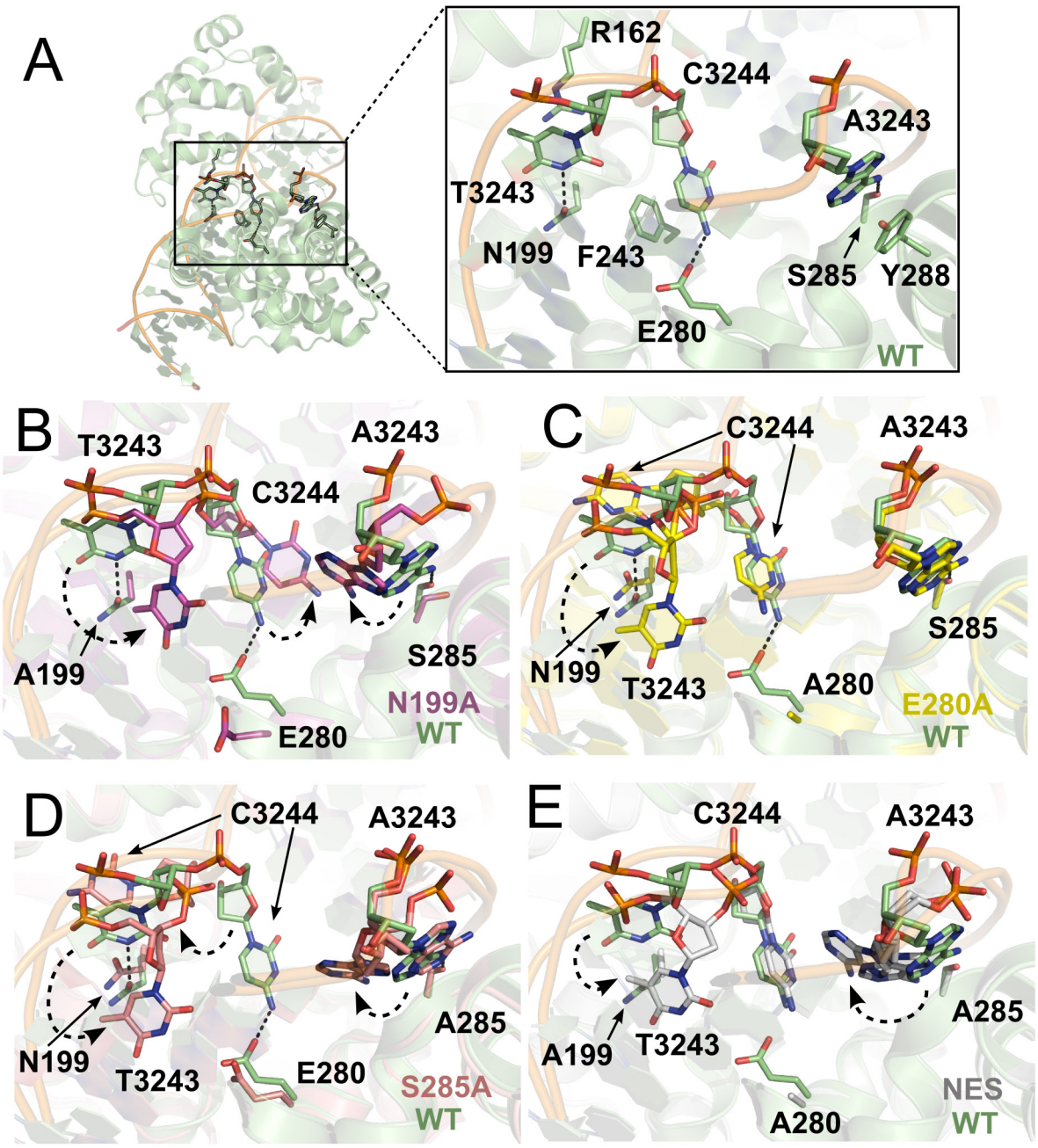


Figure 5-2. Structural comparison of wild-type and variant MTERF1 bound to the termination sequence

X-ray crystallographic structures of wild-type (panel A) and variant MTERF1 proteins (panels B-E) are shown in complex with an oligonucleotide encoding the termination sequence in the tRNA-Leu gene. Protein and DNA are depicted as transparent cartoons, with relevant amino acid sidechains and DNA bases shown as solid sticks. Dashed lines indicate hydrogen bonds, and dashed arrows highlight the relative change in position of DNA bases.

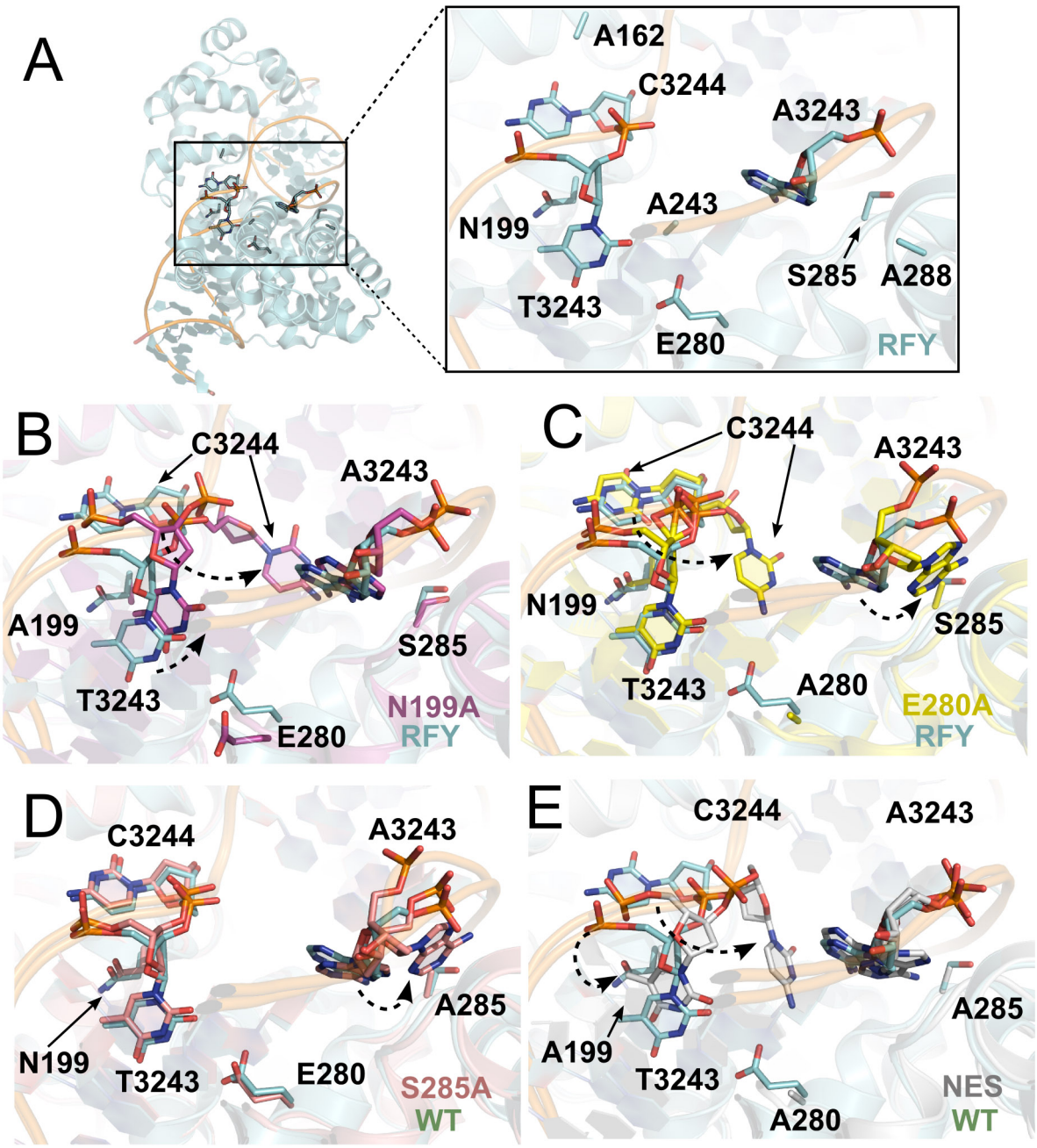


Figure 5-3. Structural comparison of MTERF1 variants that lack pi-stacking interactions or sequence-specific hydrogen bonds that are required for base-flipping

X-ray crystallographic structures of MTERF1 variants are shown in complex with an oligonucleotide encoding the termination sequence in the tRNA-Leu gene. Protein and DNA are depicted as transparent cartoons, with relevant amino acid sidechains and DNA bases shown as solid sticks. Dashed lines indicate hydrogen bonds, and dashed arrows highlight the relative change in position of DNA bases. The structure of the MTERF1-RFY variant that lacks pi-stacking interactions is shown in panel A. The N199A, E280A, S285A, and NES variants are superposed with the RFY variants in panels B, C, D, and E, respectively.

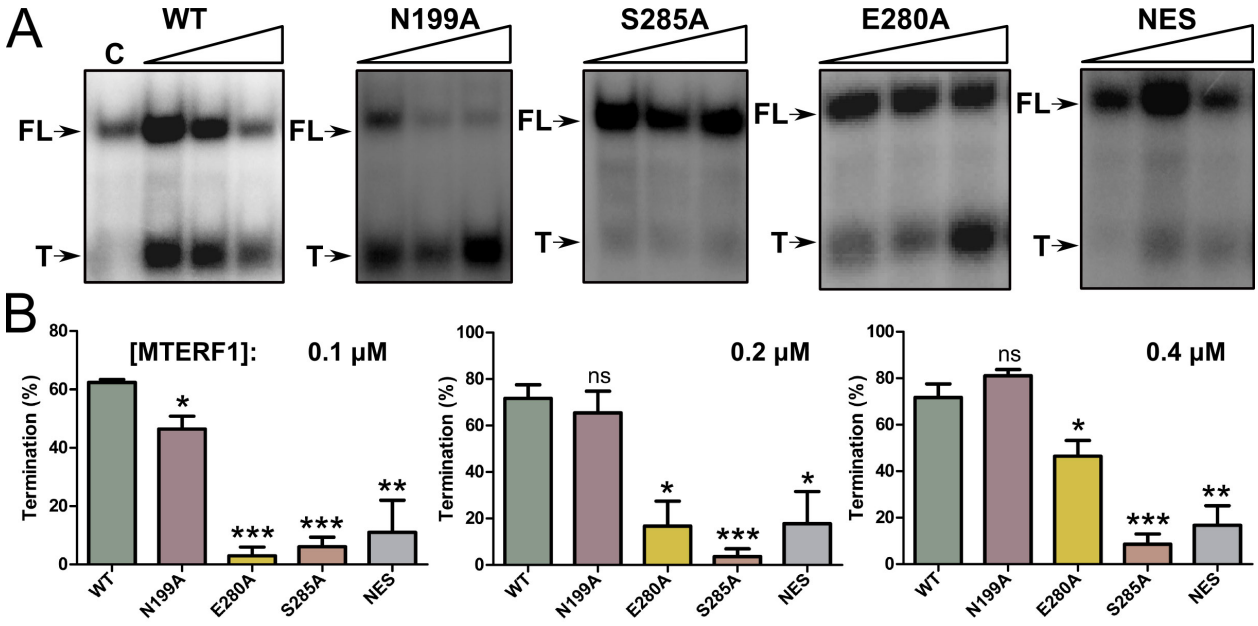


Figure 5-4. In vitro transcription termination activity of wild-type and variant MTERF1
(A) Wild-type and variant MTERF1 proteins were assayed for their ability to terminate transcription in vitro (see Methods described in Chapter 2). The results show clear termination for wild-type MTERF1 but only residual termination for S285A, E280A, the triple (NES) mutant which contains all three N199A, E280A, and S285A substitutions. Full-length, run-off transcription is denoted by "FL", the letter "T" denotes transcription termination product, and the letter "C" denotes a control without MTERF1 protein. (B) Quantification of the termination activity. The bar graph shows the percent termination observed in in vitro termination experiments. Values correspond to the mean ± SD of at least three independent experiments.

References

1. Dyall, S.D., M.T. Brown, and P.J. Johnson, *Ancient invasions: from endosymbionts to organelles.* Science, 2004. **304**(5668): p. 253-7.
2. Gray, M.W., G. Burger, and B.F. Lang, *Mitochondrial evolution.* Science, 1999. **283**(5407): p. 1476-81.
3. Vafai, S.B. and V.K. Mootha, *Mitochondrial disorders as windows into an ancient organelle.* Nature, 2012. **491**(7424): p. 374-83.
4. Finkel, T., *Signal transduction by mitochondrial oxidants.* J Biol Chem, 2012. **287**(7): p. 4434-40.
5. Finkel, T., *Signal transduction by reactive oxygen species.* J Cell Biol, 2011. **194**(1): p. 7-15.
6. Finkel, T., *Telomeres and mitochondrial function.* Circ Res, 2011. **108**(8): p. 903-4.
7. Miller, W.L., *Steroid hormone synthesis in mitochondria.* Mol Cell Endocrinol, 2013. **379**(1-2): p. 62-73.
8. Falkenberg, M., N.-G. Larsson, and C.M. Gustafsson, *DNA Replication and Transcription in Mammalian Mitochondria.* Annual review of biochemistry, 2007. **76**: p. 679-699.
9. Larsson, N.-G., *Somatic Mitochondrial DNA Mutations in Mammalian Aging.* Annual review of biochemistry, 2010. **79**(1): p. 683-706.
10. Holt, I.J., A.E. Harding, and J.A. Morgan-Hughes, *Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies.* Nature, 1988. **331**(6158): p. 717-9.
11. Wallace, D.C., et al., *Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy.* Science, 1988. **242**(4884): p. 1427-30.
12. Larsson, N.G. and D.A. Clayton, *Molecular genetic aspects of human mitochondrial disorders.* Annual review of genetics, 1995. **29**: p. 151-78.
13. Prezant, T.R., et al., *Mitochondrial ribosomal RNA mutation associated with both antibiotic-induced and non-syndromic deafness.* Nat Genet, 1993. **4**(3): p. 289-94.
14. Vandebona, H., et al., *Prevalence of mitochondrial 1555A-->G mutation in adults of European descent.* N Engl J Med, 2009. **360**(6): p. 642-4.
15. Zhao, H., et al., *Maternally inherited aminoglycoside-induced and nonsyndromic deafness is associated with the novel C1494T mutation in the mitochondrial 12S rRNA gene in a large Chinese family.* Am J Hum Genet, 2004. **74**(1): p. 139-52.
16. Boczonadi, V. and R. Horvath, *Mitochondria: impaired mitochondrial translation in human disease.* Int J Biochem Cell Biol, 2014. **48**: p. 77-84.
17. Rotig, A., *Human diseases with impaired mitochondrial protein synthesis.* Biochim Biophys Acta, 2011. **1807**(9): p. 1198-205.
18. DiMauro, S. and E.A. Schon, *Mitochondrial respiratory-chain diseases.* N Engl J Med, 2003. **348**(26): p. 2656-68.
19. Stumpf, J.D. and W.C. Copeland, *Mitochondrial DNA replication and disease: insights from DNA polymerase gamma mutations.* Cell Mol Life Sci, 2011. **68**(2): p. 219-33.
20. Taylor, R.W. and D.M. Turnbull, *Mitochondrial DNA mutations in human disease.* Nat Rev Genet, 2005. **6**(5): p. 389-402.
21. Wallace, D.C., *Mitochondrial diseases in man and mouse.* Science, 1999. **283**(5407): p. 1482-8.
22. Wallace, D.C., *Mitochondria and cancer.* Nat Rev Cancer, 2012. **12**(10): p. 685-98.

23. Bogenhagen, D.F., *Does mtDNA nucleoid organization impact aging?* Exp Gerontol, 2010. **45**(7-8): p. 473-7.
24. Bratic, A. and N.G. Larsson, *The role of mitochondria in aging.* J Clin Invest, 2013. **123**(3): p. 951-7.
25. Trifunovic, A., et al., *Premature ageing in mice expressing defective mitochondrial DNA polymerase.* Nature, 2004. **429**(6990): p. 417-23.
26. Fernandez-Vizarra, E., et al., *Tissue-specific differences in mitochondrial activity and biogenesis.* Mitochondrion, 2011. **11**(1): p. 207-13.
27. Gardner, J.L., et al., *Experimental strategies towards treating mitochondrial DNA disorders.* Biosci Rep, 2007. **27**(1-3): p. 139-50.
28. Koene, S. and J. Smeitink, *Mitochondrial medicine: entering the era of treatment.* J Intern Med, 2009. **265**(2): p. 193-209.
29. Smeitink, J.A., et al., *Distinct clinical phenotypes associated with a mutation in the mitochondrial translation elongation factor EFTs.* Am J Hum Genet, 2006. **79**(5): p. 869-77.
30. Smeitink, J.A., et al., *Mitochondrial medicine: a metabolic perspective on the pathology of oxidative phosphorylation disorders.* Cell Metab, 2006. **3**(1): p. 9-13.
31. Reddy, P., et al., *Selective elimination of mitochondrial mutations in the germline by genome editing.* Cell, 2015. **161**(3): p. 459-69.
32. Srivastava, S. and C.T. Moraes, *Manipulating mitochondrial DNA heteroplasmy by a mitochondrially targeted restriction endonuclease.* Hum Mol Genet, 2001. **10**(26): p. 3093-9.
33. Minczuk, M., *Engineered zinc finger proteins for manipulation of the human mitochondrial genome.* Methods Mol Biol, 2010. **649**: p. 257-70.
34. Minczuk, M., et al., *Construction and testing of engineered zinc-finger proteins for sequence-specific modification of mtDNA.* Nat Protoc, 2010. **5**(2): p. 342-56.
35. Amato, P., et al., *Three-parent in vitro fertilization: gene replacement for the prevention of inherited mitochondrial diseases.* Fertil Steril, 2014. **101**(1): p. 31-5.
36. Craven, L., et al., *Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease.* Nature, 2010. **465**(7294): p. 82-5.
37. Tachibana, M., et al., *Human embryonic stem cells derived by somatic cell nuclear transfer.* Cell, 2013. **153**(6): p. 1228-38.
38. Tachibana, M., et al., *Mitochondrial gene replacement in primate offspring and embryonic stem cells.* Nature, 2009. **461**(7262): p. 367-72.
39. Vogel, G., *Stem cells. Therapeutic cloning reaches milestone.* Science, 2014. **344**(6183): p. 462-3.
40. Wang, T., et al., *Polar body genome transfer for preventing the transmission of inherited mitochondrial diseases.* Cell, 2014. **157**(7): p. 1591-604.
41. Bogenhagen, D.F., *Mitochondrial DNA nucleoid structure.* Biochim Biophys Acta, 2012. **1819**(9-10): p. 914-20.
42. Bogenhagen, D.F., D. Rousseau, and S. Burke, *The layered structure of human mitochondrial DNA nucleoids.* The Journal of biological chemistry, 2008. **283**: p. 3665-75.
43. Barrell, B.G., et al., *Different pattern of codon recognition by mammalian mitochondrial tRNAs.* Proc Natl Acad Sci U S A, 1980. **77**(6): p. 3164-6.
44. Barrell, B.G., A.T. Bankier, and J. Drouin, *A different genetic code in human mitochondria.* Nature, 1979. **282**(5735): p. 189-94.
45. Kucej, M. and R.A. Butow, *Evolutionary tinkering with mitochondrial nucleoids.* Trends Cell Biol, 2007. **17**(12): p. 586-92.

46. Schmidt, O., N. Pfanner, and C. Meisinger, *Mitochondrial protein import: from proteomics to functional mechanisms.* Nat Rev Mol Cell Biol, 2010. **11**(9): p. 655-67.
47. Stewart, J.B., et al., *Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease.* Nat Rev Genet, 2008. **9**(9): p. 657-62.
48. Stewart, J.B., et al., *Strong purifying selection in transmission of mammalian mitochondrial DNA.* PLoS Biol, 2008. **6**(1): p. e10.
49. Clayton, D.A., *Mitochondrial DNA replication: what we know.* IUBMB Life, 2003. **55**(4-5): p. 213-7.
50. Yakubovskaya, E., et al., *Organization of the human mitochondrial transcription initiation complex.* Nucleic Acids Res, 2014. **42**(6): p. 4100-12.
51. Pham, X.H., et al., *Conserved sequence box II directs transcription termination and primer formation in mitochondria.* J Biol Chem, 2006. **281**(34): p. 24647-52.
52. Terzioglu, M., et al., *MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation.* Cell Metab, 2013. **17**(4): p. 618-26.
53. Montoya, J., D. Ojala, and G. Attardi, *Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs.* Nature, 1981. **290**(5806): p. 465-70.
54. Battey, J. and D.A. Clayton, *The transcription map of human mitochondrial DNA implicates transfer RNA excision as a major processing event.* J Biol Chem, 1980. **255**(23): p. 11599-606.
55. Ojala, D., et al., *The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA.* Cell, 1980. **22**(2 Pt 2): p. 393-403.
56. Ojala, D., J. Montoya, and G. Attardi, *tRNA punctuation model of RNA processing in human mitochondria.* Nature, 1981. **290**(5806): p. 470-4.
57. Bogenhagen, D.F., D.W. Martin, and A. Koller, *Initial steps in RNA processing and ribosome assembly occur at mitochondrial DNA nucleoids.* Cell Metab, 2014. **19**(4): p. 618-29.
58. Guerrier-Takada, C., et al., *The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme.* Cell, 1983. **35**(3 Pt 2): p. 849-57.
59. Jarrous, N., *Human ribonuclease P: subunits, function, and intranuclear localization.* RNA, 2002. **8**(1): p. 1-7.
60. Holzmann, J., et al., *RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme.* Cell, 2008. **135**(3): p. 462-74.
61. Vilardo, E., et al., *A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase--extensive moonlighting in mitochondrial tRNA biogenesis.* Nucleic Acids Res, 2012. **40**(22): p. 11583-93.
62. Haack, T.B., et al., *ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy.* Am J Hum Genet, 2013. **93**(2): p. 211-23.
63. Levinger, L., M. Morl, and C. Florentz, *Mitochondrial tRNA 3' end metabolism and human disease.* Nucleic Acids Research, 2004. **32**(18): p. 5430-41.
64. Levinger, L., et al., *A pathogenesis-associated mutation in human mitochondrial tRNALeu(UUR) leads to reduced 3'-end processing and CCA addition.* J Mol Biol, 2004. **337**(3): p. 535-44.
65. Levinger, L. and D. Serjanov, *Pathogenesis-related mutations in the T-loops of human mitochondrial tRNAs affect 3' end processing and tRNA structure.* RNA Biol, 2012. **9**(3): p. 283-91.

66. Yan, H., N. Zareen, and L. Levinger, *Naturally occurring mutations in human mitochondrial pre-tRNASer(UCN) can affect the transfer ribonuclease Z cleavage site, processing kinetics, and substrate secondary structure.* J Biol Chem, 2006. **281**(7): p. 3926-35.
67. Temperley, R., et al., *Hungry codons promote frameshifting in human mitochondrial ribosomes.* Science, 2010. **327**(5963): p. 301.
68. Temperley, R.J., et al., *Human mitochondrial mRNAs--like members of all families, similar but different.* Biochim Biophys Acta, 2010. **1797**(6-7): p. 1081-5.
69. Chang, J.H. and L. Tong, *Mitochondrial poly(A) polymerase and polyadenylation.* Biochim Biophys Acta, 2012. **1819**(9-10): p. 992-7.
70. Gagliardi, D., et al., *Messenger RNA stability in mitochondria: different means to an end.* Trends Genet, 2004. **20**(6): p. 260-7.
71. Rackham, O. and A. Filipovska, *The role of mammalian PPR domain proteins in the regulation of mitochondrial gene expression.* Biochim Biophys Acta, 2012. **1819**(9-10): p. 1008-16.
72. Rackham, O., T.R. Mercer, and A. Filipovska, *The human mitochondrial transcriptome and the RNA-binding proteins that regulate its expression.* Wiley Interdiscip Rev RNA, 2012. **3**(5): p. 675-95.
73. Rorbach, J. and M. Minczuk, *The post-transcriptional life of mammalian mitochondrial RNA.* Biochem J, 2012. **444**(3): p. 357-73.
74. Tucker, E.J., et al., *Mutations in MTFMT underlie a human disorder of formylation causing impaired mitochondrial translation.* Cell Metab, 2011. **14**(3): p. 428-34.
75. Suzuki, T., A. Nagao, and T. Suzuki, *Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases.* Annu Rev Genet, 2011. **45**: p. 299-329.
76. Yarham, J.W., et al., *Mitochondrial tRNA mutations and disease.* Wiley Interdiscip Rev RNA, 2010. **1**(2): p. 304-24.
77. Suzuki, T. and T. Suzuki, *A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs.* Nucleic Acids Res, 2014. **42**(11): p. 7346-57.
78. Kirino, Y., et al., *Specific correlation between the wobble modification deficiency in mutant tRNAs and the clinical features of a human mitochondrial disease.* Proc Natl Acad Sci U S A, 2005. **102**(20): p. 7127-32.
79. Kirino, Y. and T. Suzuki, *Human mitochondrial diseases associated with tRNA wobble modification deficiency.* RNA Biol, 2005. **2**(2): p. 41-4.
80. Kirino, Y., et al., *Codon-specific translational defect caused by a wobble modification deficiency in mutant tRNA from a human mitochondrial disease.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(42): p. 15070-5.
81. Yasukawa, T., et al., *A pathogenic point mutation reduces stability of mitochondrial mutant tRNA(Ile).* Nucleic Acids Res, 2000. **28**(19): p. 3779-84.
82. Yasukawa, T., et al., *Defect in modification at the anticodon wobble nucleotide of mitochondrial tRNA(Lys) with the MERRF encephalomyopathy pathogenic mutation.* FEBS letters, 2000. **467**(2-3): p. 175-8.
83. Yasukawa, T., et al., *Modification defect at anticodon wobble nucleotide of mitochondrial tRNAs(Leu)(UUR) with pathogenic mutations of mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes.* J Biol Chem, 2000. **275**(6): p. 4251-7.
84. Suzuki, T., A. Nagao, and T. Suzuki, *Human mitochondrial diseases caused by lack of taurine modification in mitochondrial tRNAs.* Wiley Interdiscip Rev RNA, 2011. **2**(3): p. 376-86.

85. Woodson, S.A., *RNA folding pathways and the self-assembly of ribosomes.* Acc Chem Res, 2011. **44**(12): p. 1312-9.
86. Baer, R.J. and D.T. Dubin, *Methylated regions of hamster mitochondrial ribosomal RNA: structural and functional correlates.* Nucleic Acids Res, 1981. **9**(2): p. 323-37.
87. Decatur, W.A. and M.J. Fournier, *rRNA modifications and ribosome function.* Trends Biochem Sci, 2002. **27**(7): p. 344-51.
88. Motorin, Y. and M. Helm, *RNA nucleotide methylation.* Wiley Interdiscip Rev RNA, 2011. **2**(5): p. 611-31.
89. Shutt, T.E. and M.W. Gray, *Homologs of mitochondrial transcription factor B, sparsely distributed within the eukaryotic radiation, are likely derived from the dimethyladenosine methyltransferase of the mitochondrial endosymbiont.* Mol Biol Evol, 2006. **23**(6): p. 1169-79.
90. Helser, T.L., J.E. Davies, and J.E. Dahlberg, *Change in methylation of 16S ribosomal RNA associated with mutation to kasugamycin resistance in Escherichia coli.* Nat New Biol, 1971. **233**(35): p. 12-4.
91. Poldermans, B., H. Bakker, and P.H. Van Knippenberg, *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, 1980. **8**(1): p. 143-51.
92. Metodiev, M.D., et al., *Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome.* Cell Metab, 2009. **9**(4): p. 386-97.
93. Cotney, J., S.E. McKay, and G.S. Shadel, *Elucidation of separate, but collaborative functions of the rRNA methyltransferase-related human mitochondrial transcription factors B1 and B2 in mitochondrial biogenesis reveals new insight into maternally inherited deafness.* Hum Mol Genet, 2009. **18**(14): p. 2670-82.
94. Guja, K.E., et al., *Structural basis for S-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1.* Nucleic Acids Res, 2013. **41**(16): p. 7947-59.
95. Cotney, J. and G.S. Shadel, *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, 2006. **63**(5): p. 707-17.
96. Raimundo, N., et al., *Mitochondrial stress engages E2F1 apoptotic signaling to cause deafness.* Cell, 2012. **148**(4): p. 716-26.
97. O'Sullivan, M., et al., *Mitochondrial m.1584A 12S m62A rRNA methylation in families with m.1555A>G associated hearing loss.* Hum Mol Genet, 2015. **24**(4): p. 1036-44.
98. Amunts, A., et al., *Ribosome. The structure of the human mitochondrial ribosome.* Science, 2015. **348**(6230): p. 95-8.
99. Greber, B.J., et al., *Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome.* Science, 2015. **348**(6232): p. 303-8.
100. Metodiev, M.D., et al., *NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly.* PLoS Genet, 2014. **10**(2): p. e1004110.
101. Cámara, Y., et al., *MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome.* Cell metabolism, 2011. **13**(5): p. 527-539.

102. Kruse, B., N. Narasimhan, and G. Attardi, *Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination.* Cell, 1989. **58**(2): p. 391-7.
103. Yakubovskaya, E., et al., *Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription.* Cell, 2010. **141**(6): p. 982-993.
104. Yakubovskaya, E., et al., *Structure of the essential MTERF4:NSUN4 protein complex reveals how an MTERF protein collaborates to facilitate rRNA modification.* Structure, 2012. **20**(11): p. 1940-7.
105. Spahr, H., et al., *Structure of the human MTERF4-NSUN4 protein complex that regulates mitochondrial ribosome biogenesis.* Proc Natl Acad Sci U S A, 2012. **109**(38): p. 15253-8.
106. Kimura, S. and T. Suzuki, *Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the Escherichia coli 16S rRNA.* Nucleic Acids Res, 2010. **38**(4): p. 1341-52.
107. Lee, K.W. and D.F. Bogenhagen, *Assignment of 2'-O-methyltransferases to modification sites on the mammalian mitochondrial large subunit 16 S ribosomal RNA (rRNA).* J Biol Chem, 2014. **289**(36): p. 24936-42.
108. Lee, K.W., et al., *Mitochondrial ribosomal RNA (rRNA) methyltransferase family members are positioned to modify nascent rRNA in foci near the mitochondrial DNA nucleoid.* J Biol Chem, 2013. **288**(43): p. 31386-99.
109. Sirum-Connolly, K. and T.L. Mason, *Functional requirement of a site-specific ribose methylation in ribosomal RNA.* Science, 1993. **262**(5141): p. 1886-9.
110. Pintard, L., et al., *MRM2 encodes a novel yeast mitochondrial 21S rRNA methyltransferase.* EMBO J, 2002. **21**(5): p. 1139-47.
111. Dubin, D.T. and R.H. Taylor, *Modification of mitochondrial ribosomal RNA from hamster cells: the presence of GmG and late-methylated UmGmU in the large subunit (17S) RNA.* J Mol Biol, 1978. **121**(4): p. 523-40.
112. Ofengand, J. and A. Bakin, *Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts.* J Mol Biol, 1997. **266**(2): p. 246-68.
113. Ansmant, I., et al., *Identification of the Saccharomyces cerevisiae RNA:pseudouridine synthase responsible for formation of psi(2819) in 21S mitochondrial ribosomal RNA.* Nucleic Acids Res, 2000. **28**(9): p. 1941-6.
114. Spåhr, H., et al., *Structure of mitochondrial transcription termination factor 3 reveals a novel nucleic acid-binding domain.* Biochemical and biophysical research communications, 2010. **397**(3): p. 386-390.
115. Park, C.B., et al., *MTERF3 is a negative regulator of mammalian mtDNA transcription.* Cell, 2007. **130**(2): p. 273-285.
116. Wredenberg, A., et al., *MTERF3 regulates mitochondrial ribosome biogenesis in invertebrates and mammals.* PLoS Genet, 2013. **9**(1): p. e1003178.
117. Shajani, Z., M.T. Sykes, and J.R. Williamson, *Assembly of bacterial ribosomes.* Annu Rev Biochem, 2011. **80**: p. 501-26.
118. Henras, A.K., et al., *The post-transcriptional steps of eukaryotic ribosome biogenesis.* Cell Mol Life Sci, 2008. **65**(15): p. 2334-59.
119. Kaczanowska, M. and M. Ryden-Aulin, *Ribosome biogenesis and the translation process in Escherichia coli.* Microbiol Mol Biol Rev, 2007. **71**(3): p. 477-94.

120. Juhling, F., et al., *tRNAdb 2009: compilation of tRNA sequences and tRNA genes.* Nucleic Acids Res, 2009. **37**(Database issue): p. D159-62.
121. O'Farrell, H.C., et al., *Crystallization and preliminary X-ray diffraction analysis of KsgA, a universally conserved RNA adenine dimethyltransferase in Escherichia coli.* Acta crystallographica. Section D, Biological crystallography, 2003. **59**(Pt 8): p. 1490-2.
122. Kabsch, W., *Xds.* Acta crystallographica. Section D, Biological crystallography, 2010. **66**(Pt 2): p. 125-32.
123. Evans, P., *Scaling and assessment of data quality.* Acta Crystallographica Section D, 2006. **62**(1): p. 72-82.
124. Vonrhein, C., et al., *Data processing and analysis with the autoPROC toolbox.* Acta crystallographica. Section D, Biological crystallography, 2011. **67**(Pt 4): p. 293-302.
125. Hendrickson, W.M. and C. Ogata, *Phase determination from multiwavelength anomalous diffraction measurements.* Meth Enzymol, 1997(276): p. 494-523.
126. Terwilliger, T.C. and J. Berendzen, *Automated MAD and MIR structure solution.* Acta crystallographica. Section D, Biological crystallography, 1999. **55**(Pt 4): p. 849-61.
127. Terwilliger, T.C., *Maximum-likelihood density modification.* Acta crystallographica. Section D, Biological crystallography, 2000. **56**(Pt 8): p. 965-72.
128. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta crystallographica. Section D, Biological crystallography, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
129. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta crystallographica. Section D, Biological crystallography, 2010. **66**(Pt 2): p. 213-21.
130. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids.* Nucleic acids research, 2007. **35**(Web Server issue): p. W375-83.
131. Strong, M., et al., *Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(21): p. 8060-5.
132. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.* Molecular systems biology, 2006. **2**: p. 2006 0008.
133. Yakubovskaya, E., et al., *Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription.* Cell, 2010. **141**(6): p. 982-93.
134. Asin-Cayuela, J., et al., *The human mitochondrial transcription termination factor (mTERF) is fully active in vitro in the non-phosphorylated form.* Journal of Biological Chemistry, 2005. **280**(27): p. 25499-25505.
135. Evans, P.R. and G.N. Murshudov, *How good are my data and what is the resolution?* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 7): p. 1204-14.
136. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
137. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
138. Zwart, P.H., et al., *Automated structure solution with the PHENIX suite.* Methods Mol Biol, 2008. **426**: p. 419-35.
139. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.
140. Ryan, M.T. and N.J. Hoogenraad, *Mitochondrial-nuclear communications.* Annual review of biochemistry, 2007. **76**: p. 701-22.

141. Shutt, T.E. and G.S. Shadel, *A compendium of human mitochondrial gene expression machinery with links to disease.* Environ Mol Mutagen, 2010. **51**(5): p. 360-79.
142. Zhu, X., et al., *Pathogenic mutations of nuclear genes associated with mitochondrial disorders.* Acta Biochim Biophys Sin (Shanghai), 2009. **41**(3): p. 179-87.
143. Ylikallio, E. and A. Suomalainen, *Mechanisms of mitochondrial diseases.* Ann Med, 2012. **44**(1): p. 41-59.
144. Tuppen, H.A., et al., *Mitochondrial DNA mutations and human disease.* Biochim Biophys Acta, 2010. **1797**(2): p. 113-28.
145. Copeland, W.C., *Inherited mitochondrial diseases of DNA replication.* Annu Rev Med, 2008. **59**: p. 131-46.
146. Camara, Y., et al., *MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome.* Cell Metab, 2011. **13**(5): p. 527-39.
147. Liu, J., et al., *Crystal structure of tRNA (m1G37) methyltransferase from Aquifex aeolicus at 2.6 A resolution: a novel methyltransferase fold.* Proteins, 2003. **53**(2): p. 326-8.
148. Seidel-Rogol, B.L., V. McCulloch, and G.S. Shadel, *Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop.* Nat Genet, 2003. **33**(1): p. 23-4.
149. O'Farrell, H.C., et al., *Sequence and structural evolution of the KsgA/Dim1 methyltransferase family.* BMC Res Notes, 2008. **1**: p. 108.
150. Lim, K., et al., *Crystal structure of YecO from Haemophilus influenzae (HI0319) reveals a methyltransferase fold and a bound S-adenosylhomocysteine.* Proteins, 2001. **45**(4): p. 397-407.
151. Michel, G., et al., *The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot.* Structure, 2002. **10**(10): p. 1303-15.
152. Bujnicki, J.M., *In silico analysis of the tRNA:m1A58 methyltransferase family: homology-based fold prediction and identification of new members from Eubacteria and Archaea.* FEBS letters, 2001. **507**(2): p. 123-7.
153. Dlakic, M., *Chromatin silencing protein and pachytene checkpoint regulator Dot1p has a methyltransferase fold.* Trends in biochemical sciences, 2001. **26**(7): p. 405-7.
154. Guan, M.X., *Mitochondrial 12S rRNA mutations associated with aminoglycoside ototoxicity.* Mitochondrion, 2011. **11**(2): p. 237-45.
155. Hu, D.N., et al., *Genetic aspects of antibiotic induced deafness: mitochondrial inheritance.* J Med Genet, 1991. **28**(2): p. 79-83.
156. Fischel-Ghodsian, N., *Genetic factors in aminoglycoside toxicity.* Pharmacogenomics, 2005. **6**(1): p. 27-36.
157. Bykhovskaya, Y., et al., *Human mitochondrial transcription factor B1 as a modifier gene for hearing loss associated with the mitochondrial A1555G mutation.* Mol Genet Metab, 2004. **82**(1): p. 27-32.
158. Koeck, T., et al., *A common variant in TFB1M is associated with reduced insulin secretion and increased future risk of type 2 diabetes.* Cell Metab, 2011. **13**(1): p. 80-91.
159. Falkenberg, M., et al., *Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA.* Nat Genet, 2002. **31**(3): p. 289-94.
160. Shutt, T.E., et al., *Core human mitochondrial transcription apparatus is a regulated two-component system in vitro.* Proc Natl Acad Sci U S A, 2010. **107**(27): p. 12133-8.

161. McCulloch, V., B.L. Seidel-Rogol, and G.S. Shadel, *A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine.* Mol Cell Biol, 2002. **22**(4): p. 1116-25.
162. Shutt, T.E., et al., *Core human mitochondrial transcription apparatus is a regulated two-component system in vitro.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(27): p. 12133-8.
163. Surovtseva, Y.V. and G.S. Shadel, *Transcription-independent role for human mitochondrial RNA polymerase in mitochondrial ribosome biogenesis.* Nucleic acids research, 2013. **41**(4): p. 2479-88.
164. van Buul, C.P. and P.H. van Knippenberg, *Nucleotide sequence of the ksgA gene of Escherichia coli: comparison of methyltransferases effecting dimethylation of adenosine in ribosomal RNA.* Gene, 1985. **38**(1-3): p. 65-72.
165. Claros, M.G. and P. Vincens, *Computational method to predict mitochondrially imported proteins and their targeting sequences.* European journal of biochemistry / FEBS, 1996. **241**(3): p. 779-86.
166. Martin, J.L. and F.M. McMillan, *SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold.* Current opinion in structural biology, 2002. **12**(6): p. 783-93.
167. Vidgren, J., L.A. Svensson, and A. Liljas, *Crystal structure of catechol O-methyltransferase.* Nature, 1994. **368**(6469): p. 354-8.
168. Li, L., et al., *DelPhi: a comprehensive suite for DelPhi software and associated resources.* BMC biophysics, 2012. **5**(1): p. 9.
169. Fellous, T.G., et al., *Locating the stem cell niche and tracing hepatocyte lineages in human liver.* Hepatology, 2009. **49**(5): p. 1655-63.
170. O'Farrell, H.C., J.N. Scarsdale, and J.P. Rife, *Crystal structure of KsgA, a universally conserved rRNA adenine dimethyltransferase in Escherichia coli.* J Mol Biol, 2004. **339**(2): p. 337-53.
171. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.
172. Pulicherla, N., et al., *Structural and functional divergence within the Dim1/KsgA family of rRNA methyltransferases.* J Mol Biol, 2009. **391**(5): p. 884-93.
173. Bussiere, D.E., et al., *Crystal structure of ErmC', an rRNA methyltransferase which mediates antibiotic resistance in bacteria.* Biochemistry, 1998. **37**(20): p. 7103-12.
174. Schluckebier, G., et al., *The 2.2 A structure of the rRNA methyltransferase ErmC' and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism.* J Mol Biol, 1999. **289**(2): p. 277-91.
175. Menezes, S., et al., *Formation of m2G6 in Methanocaldococcus jannaschii tRNA catalyzed by the novel methyltransferase Trm14.* Nucleic Acids Res, 2011. **39**(17): p. 7641-55.
176. Schneidman-Duhovny, D., M. Hammel, and A. Sali, *FoXS: a web server for rapid computation and fitting of SAXS profiles.* Nucleic acids research, 2010. **38**(Web Server issue): p. W540-4.
177. Demirci, H., et al., *Structural rearrangements in the active site of the Thermus thermophilus 16S rRNA methyltransferase KsgA in a binary complex with 5'-methylthioadenosine.* Journal of molecular biology, 2009. **388**(2): p. 271-82.

178. Maravic, G., et al., *Alanine-scanning mutagenesis of the predicted rRNA-binding domain of ErmC' redefines the substrate-binding site and suggests a model for protein-RNA interactions.* Nucleic Acids Res, 2003. **31**(16): p. 4941-9.
179. Buriankova, K., et al., *Molecular basis of intrinsic macrolide resistance in the Mycobacterium tuberculosis complex.* Antimicrob Agents Chemother, 2004. **48**(1): p. 143-50.
180. Schinkel, A.H., et al., *Specificity factor of yeast mitochondrial RNA polymerase. Purification and interaction with core RNA polymerase.* The Journal of biological chemistry, 1987. **262**(26): p. 12785-91.
181. Dunkle, J.A., et al., *Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(40): p. 17152-7.
182. McClelland, M., *Purification and characterization of two new modification methylases: MClaI from Caryophanon latum L and MTaqI from Thermus aquaticus YTI.* Nucleic Acids Res, 1981. **9**(24): p. 6795-804.
183. Goedecke, K., et al., *Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog.* Nat Struct Biol, 2001. **8**(2): p. 121-5.
184. O'Farrell, H.C., et al., *Control of substrate specificity by a single active site residue of the KsgA methyltransferase.* Biochemistry, 2012. **51**(1): p. 466-74.
185. Pues, H., et al., *Functional roles of the conserved aromatic amino acid residues at position 108 (motif IV) and position 196 (motif VIII) in base flipping and catalysis by the N6-adenine DNA methyltransferase from Thermus aquaticus.* Biochemistry, 1999. **38**(5): p. 1426-34.
186. Boehringer, D., et al., *Structural insights into methyltransferase KsgA function in 30S ribosomal subunit biogenesis.* J Biol Chem, 2012. **287**(13): p. 10453-9.
187. Richter, U., et al., *A mitochondrial rRNA dimethyladenosine methyltransferase in Arabidopsis.* Plant J, 2010. **61**(4): p. 558-69.
188. Adan, C., et al., *Mitochondrial transcription factor B2 is essential for metabolic function in Drosophila melanogaster development.* J Biol Chem, 2008. **283**(18): p. 12333-42.
189. Matsushima, Y., R. Garesse, and L.S. Kaguni, *Drosophila mitochondrial transcription factor B2 regulates mitochondrial DNA copy number and transcription in schneider cells.* J Biol Chem, 2004. **279**(26): p. 26900-5.
190. Matsushima, Y., et al., *Drosophila mitochondrial transcription factor B1 modulates mitochondrial translation but not transcription or DNA copy number in Schneider cells.* J Biol Chem, 2005. **280**(17): p. 16815-20.
191. Goedecke, K., et al., *Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog.* Nature structural biology, 2001. **8**(2): p. 121-5.
192. Klimasauskas, S., et al., *HhaI methyltransferase flips its target base out of the DNA helix.* Cell, 1994. **76**(2): p. 357-69.
193. Allan, B.W., et al., *Direct real time observation of base flipping by the EcoRI DNA methyltransferase.* The Journal of biological chemistry, 1998. **273**(4): p. 2368-73.
194. Cheng, X. and R.J. Roberts, *AdoMet-dependent methylation, DNA methyltransferases and base flipping.* Nucleic acids research, 2001. **29**(18): p. 3784-95.
195. Kabsch, W. and C. Sander, *Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.* Biopolymers, 1983. **22**(12): p. 2577-637.

196. Schluckebier, G., et al., *The 2.2 A structure of the rRNA methyltransferase ErmC' and its complexes with cofactor and cofactor analogs: implications for the reaction mechanism.* Journal of molecular biology, 1999. **289**(2): p. 277-91.
197. Fislage, M., et al., *Crystal structures of the tRNA:m2G6 methyltransferase Trm14/TrmN from two domains of life.* Nucleic acids research, 2012. **40**(11): p. 5149-61.
198. Schluckebier, G., et al., *Differential binding of S-adenosylmethionine S-adenosylhomocysteine and Sinefungin to the adenine-specific DNA methyltransferase M.TaqI.* Journal of molecular biology, 1997. **265**(1): p. 56-67.
199. Schubot, F.D., et al., *Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription.* Protein science : a publication of the Protein Society, 2001. **10**(10): p. 1980-8.
200. Linder, T., et al., *A family of putative transcription termination factors shared amongst metazoans and plants.* Current genetics, 2005. **48**(4): p. 265-269.
201. Roberti, M., et al., *The MTERF family proteins: mitochondrial transcription regulators and beyond.* Biochimica et Biophysica Acta, 2009. **1787**(5): p. 303-311.
202. Guja, K.E. and M. Garcia-Diaz, *Hitting the brakes: termination of mitochondrial transcription.* Biochim Biophys Acta, 2012. **1819**(9-10): p. 939-47.
203. Babiychuk, E., et al., *Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family.* Proc Natl Acad Sci U S A, 2011. **108**(16): p. 6674-9.
204. Wenz, T., et al., *mTERF2 regulates oxidative phosphorylation by modulating mtDNA transcription.* Cell metabolism, 2009. **9**(6): p. 499-511.
205. Park, C.B., et al., *MTERF3 is a negative regulator of mammalian mtDNA transcription.* Cell, 2007. **130**(2): p. 273-85.
206. Jiménez-Menéndez, N., et al., *Human mitochondrial mTERF wraps around DNA through a left-handed superhelical tandem repeat.* Nature structural & molecular biology, 2010. **17**(7): p. 891-893.
207. Rubinson, E.H. and B.F. Eichman, *Nucleic acid recognition by tandem helical repeats.* Curr Opin Struct Biol, 2012. **22**(1): p. 101-9.
208. Edwards, T.A., et al., *Structure of Pumilio reveals similarity between RNA and peptide binding motifs.* Cell, 2001. **105**(2): p. 281-9.
209. Wang, X., et al., *Modular recognition of RNA by a human pumilio-homology domain.* Cell, 2002. **110**(4): p. 501-12.
210. Wang, X., P.D. Zamore, and T.M. Hall, *Crystal structure of a Pumilio homology domain.* Mol Cell, 2001. **7**(4): p. 855-65.
211. Martin, M., et al., *Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis.* Cell, 2005. **123**(7): p. 1227-1240.
212. Pierce, M.M., C.S. Raman, and B.T. Nall, *Isothermal titration calorimetry of protein-protein interactions.* Methods, 1999. **19**(2): p. 213-21.
213. Marintcheva, B., et al., *Acidic C-terminal tail of the ssDNA-binding protein of bacteriophage T7 and ssDNA compete for the same binding surface.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 1855-60.
214. Byrnes, J. and M. Garcia-Diaz, *Mitochondrial transcription: how does it end?* Transcription, 2011. **2**(1): p. 32-6.

215. Montoya, J., G.L. Gaines, and G. Attardi, *The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units.* Cell, 1983. **34**(1): p. 151-159.
216. Martínez-Azorín, F., *The mitochondrial ribomotor hypothesis.* IUBMB life, 2005. **57**(1): p. 27-30.
217. Shang, J. and D.A. Clayton, *Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro.* The Journal of biological chemistry, 1994. **269**(46): p. 29112-20.
218. Martinez-Azorin, F., *The mitochondrial ribomotor hypothesis.* IUBMB life, 2005. **57**(1): p. 27-30.
219. Gelfand, R. and G. Attardi, *Synthesis and turnover of mitochondrial ribonucleic acid in HeLa cells: the mature ribosomal and messenger ribonucleic acid species are metabolically unstable.* Molecular and cellular biology, 1981. **1**(6): p. 497-511.
220. Hyvärinen, A.K., et al., *The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA.* Nucleic Acids Res, 2007. **35**(19): p. 6458-6474.
221. Rebelo, A.P., S.L. Williams, and C.T. Moraes, *In vivo methylation of mtDNA reveals the dynamics of protein-mtDNA interactions.* Nucleic Acids Res, 2009. **37**(20): p. 6701-6715.
222. Mercer, T.R., et al., *The human mitochondrial transcriptome.* Cell, 2011. **146**(4): p. 645-58.
223. Roberti, M., et al., *MTERF3, the most conserved member of the mTERF-family, is a modular factor involved in mitochondrial protein synthesis.* Biochimica et Biophysica Acta, 2006. **1757**(9-10): p. 1199-1206.
224. Groves, M.R. and D. Barford, *Topological characteristics of helical repeat proteins.* Current opinion in structural biology, 1999. **9**(3): p. 383-9.
225. Sibanda, B.L., D.Y. Chirgadze, and T.L. Blundell, *Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats.* Nature, 2010. **463**(7277): p. 118-21.
226. Lu, G., S.J. Dolgner, and T.M. Hall, *Understanding and engineering RNA sequence specificity of PUF proteins.* Current opinion in structural biology, 2009. **19**(1): p. 110-5.
227. Nam, S.-C. and C. Kang, *DNA light-strand preferential recognition of human mitochondria transcription termination factor mTERF.* Journal of biochemistry and molecular biology, 2005. **38**(6): p. 690-694.
228. Hyvärinen, A.K., et al., *Effects on mitochondrial transcription of manipulating mTERF protein levels in cultured human HEK293 cells.* BMC molecular biology, 2010. **11**: p. 72.
229. Camasamudram, V., J.-K. Fang, and N.G. Avadhani, *Transcription termination at the mouse mitochondrial H-strand promoter distal site requires an A/T rich sequence motif and sequence specific DNA binding proteins.* Eur J Biochem, 2003. **270**(6): p. 1128-1140.
230. Wanrooij, P.H., et al., *G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(37): p. 16072-16077.
231. Kochetkov, S.N., E.E. Rusakova, and V.L. Tunitskaya, *Recent studies of T7 RNA polymerase mechanism.* FEBS letters, 1998. **440**(3): p. 264-7.
232. Sondheimer, N., et al., *Leucine-rich pentatricopeptide-repeat containing protein regulates mitochondrial transcription.* Biochemistry, 2010. **49**(35): p. 7467-7473.
233. Sasarman, F., et al., *LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria.* Molecular biology of the cell, 2010. **21**(8): p. 1315-23.

234. Gohil, V.M., et al., *Mitochondrial and nuclear genomic responses to loss of LRPPRC expression.* The Journal of biological chemistry, 2010. **285**(18): p. 13742-7.
235. Zeviani, M., et al., *Deletions of mitochondrial DNA in Kearns-Sayre syndrome.* Neurology, 1988. **38**(9): p. 1339-46.
236. Munnich, A. and P. Rustin, *Clinical spectrum and diagnosis of mitochondrial disorders.* American journal of medical genetics, 2001. **106**(1): p. 4-17.
237. Zeviani, M., *Mitochondrial disorders.* Supplements to Clinical neurophysiology, 2004. **57**: p. 304-12.
238. Zeviani, M. and S. Di Donato, *Mitochondrial disorders.* Brain : a journal of neurology, 2004. **127**(Pt 10): p. 2153-72.
239. Krishnan, K.J., et al., *The ageing mitochondrial genome.* Nucleic Acids Research, 2007. **35**(22): p. 7399-405.
240. Krishnan, K.J., et al., *Mitochondrial DNA mutations and aging.* Annals of the New York Academy of Sciences, 2007. **1100**: p. 227-40.
241. Pavlakis, S.G., et al., *Mitochondrial myopathy, encephalopathy, lactic acidosis, and strokelike episodes: a distinctive clinical syndrome.* Annals of neurology, 1984. **16**(4): p. 481-8.
242. Hess, J.F., et al., *Impairment of mitochondrial transcription termination by a point mutation associated with the MELAS subgroup of mitochondrial encephalomyopathies.* Nature, 1991. **351**(6323): p. 236-9.
243. Manwaring, N., et al., *Population prevalence of the MELAS A3243G mutation.* Mitochondrion, 2007. **7**(3): p. 230-3.
244. Goto, Y., I. Nonaka, and S. Horai, *A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies.* Nature, 1990. **348**(6302): p. 651-3.
245. Kobayashi, Y., et al., *A point mutation in the mitochondrial tRNA(Leu)(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes).* Biochemical and biophysical research communications, 1990. **173**(3): p. 816-22.
246. Chomyn, A., et al., *MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4221-4225.
247. Shutt, T.E. and G.S. Shadel, *A compendium of human mitochondrial gene expression machinery with links to disease.* Environmental and molecular mutagenesis, 2010. **51**(5): p. 360-379.
248. Seneca, S., et al., *A new mitochondrial point mutation in the transfer RNA(Leu) gene in a patient with a clinical phenotype resembling Kearns-Sayre syndrome.* Archives of neurology, 2001. **58**(7): p. 1113-8.
249. Mimaki, M., et al., *Different effects of novel mtDNA G3242A and G3244A base changes adjacent to a common A3243G mutation in patients with mitochondrial disorders.* Mitochondrion, 2009. **9**(2): p. 115-22.
250. Sasarman, F., H. Antonicka, and E.A. Shoubridge, *The A3243G tRNALeu(UUR) MELAS mutation causes amino acid misincorporation and a combined respiratory chain assembly defect partially suppressed by overexpression of EFTu and EFG2.* Human molecular genetics, 2008. **17**(23): p. 3697-707.

251. Hyvarinen, A.K., et al., *The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA.* Nucleic Acids Research, 2007. **35**(19): p. 6458-74.
252. Brown, M.D. and M.T. Lott, *Emery and Rimoin's Principles and Practice of Medical Genetics*, ed. D. Rimoin, et al. Vol. 1. 1996, London: Churchill Livingstone.
253. Yoneda, M., et al., *Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy.* Proc Natl Acad Sci U S A, 1992. **89**(23): p. 11164-11168.
254. Shoubridge, E.A., *Segregation of mitochondrial DNAs carrying a pathogenic point mutation (tRNA(leu3243)) in cybrid cells.* Biochemical and biophysical research communications, 1995. **213**(1): p. 189-95.
255. Bosley, T.M., et al., *Sporadic bilateral optic neuropathy in children: the role of mitochondrial abnormalities.* Investigative ophthalmology & visual science, 2008. **49**(12): p. 5250-6.
256. Chomyn, A., et al., *The mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episode syndrome-associated human mitochondrial tRNALeu(UUR) mutation causes aminoacylation deficiency and concomitant reduced association of mRNA with ribosomes.* The Journal of biological chemistry, 2000. **275**(25): p. 19198-209.
257. Sohm, B., et al., *Towards understanding human mitochondrial leucine aminoacylation identity.* Journal of Molecular Biology, 2003. **328**(5): p. 995-1010.
258. Alston, C.L., et al., *The pathogenic m.3243A>T mitochondrial DNA mutation is associated with a variable neurological phenotype.* Neuromuscular disorders : NMD, 2010. **20**(6): p. 403-6.
259. Goto, Y., et al., *A novel point mutation in the mitochondrial tRNA(Leu)(UUR) gene in a family with mitochondrial myopathy.* Annals of neurology, 1992. **31**(6): p. 672-5.
260. Houshmand, M., et al., *Fatal mitochondrial myopathy, lactic acidosis, and complex I deficiency associated with a heteroplasmic A --> G mutation at position 3251 in the mitochondrial tRNALeu(UUR) gne.* Human genetics, 1996. **97**(3): p. 269-73.
261. Goto, Y., *Clinical features of MELAS and mitochondrial DNA mutations.* Muscle & nerve, 1995. **3**: p. S107-12.