10-2016

Replication-Transciption Switch in Human Mitochondria

Karen Agaronyan
Rowan University

REPLICATION-TRANSCIPTION SWITCH IN HUMAN MITOCHONDRIA

Karen Agaronyan, M.S.

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

Stratford, New Jersey 08084

October 2016

Table of Contents

Acknowledgements

I would like to thank my scientific advisor, Dr. Dmitry Temiakov, for the opportunity to work in his lab in such a great environment, his continuous support and guidance.

I thank Dr. Anikin for invaluable expertise and comments that helped to shape this work.

I thank the members of my committee: Dr. McAllister, Dr. Moss and Dr. Shcherbik – for their criticism, comments and helpful suggestions during my work on thesis.

I would like to thank Julia Jones and Drs. Ander Parshin, Yaroslav Morozov, Dmitry Markov for their assistance and help during experiments.

Finally, I thank my family and friends for their support.

This work was funded by NIH.

Abstract

Coordinated replication and expression of mitochondrial genome is critical for metabolically active cells during various stages of development. However, it is not known whether replication and transcription can occur simultaneously without interfering with each other and whether mtDNA copy number can be regulated by the transcription machinery. Human mitochondrial RNA polymerase (mtRNAP) is a central enzyme involved in gene expression in mitochondria. It generates genome-size polycistronic transcripts and also makes replication primers at two origins of replication. MtRNAP is distantly related to phage T7 RNAP. While T7 RNAP is optimized to produce large amounts of transcripts to overcompete the bacterial RNAP, mtRNAP must coordinate RNA synthesis with processing and translation. We hypothesized that mtRNAP must be slower than T7 RNAP and measured elongation rates for these RNAPs. We found that mtRNAP is about 20 times slower than T7 RNAP. We also found that mtRNAP is inherently non-processive and cannot synthesize long transcripts. We proposed that low processivity and slow elongation rates of mtRNAP requires assistance of an additional elongation factor. We show that interaction of a recently identified human transcription elongation factor, TEFM, with mtRNAP dramatically increases processivity and elongation rates of the mitochondrial transcription machinery. Importantly, we found that TEFM prevents premature transcription termination and thus generation of replication primers by mtRNAP. Thus, TEFM serves as a component of a molecular switch between replication and transcription, which appear to be mutually exclusive processes in mitochondria. The switch likely allows avoiding the detrimental consequences of

head-on collisions between replication and transcription machineries. Regulation of TEFM may explain how mtRNAP transcription rates and, as consequence, respiration and ATP production, can be increased in mitochondria without the need to replicate mtDNA, which has been observed during different developmental processes.

Introduction

1. Mitochondrial functions.

Mitochondria are ubiquitous organelles present in all nucleated cells. The general structure of the organelle consists of the outer mitochondrial membrane (OMM), the inner mitochondrial membrane (IMM), the soluble matrix and the intermembrane space. The IMM can be divided into two parts: the inner boundary membrane and the cristae membrane. The cristae membrane forms large tubular invaginations that extend to the mitochondrial matrix. (*2*, *3*). Additionally, discovery of the mitochondrial contact site and cristae organizing system at the cristae junctions shed light onto communication between IMM and OMM (*4*, *5*).

The primary function of mitochondria is the generation of energy in the form of ATP by means of oxidative phosphorylation. The electron transport chain responsible for ATP generation is located in the cristae membrane (*2*, *5*). It involves the macromolecular assembly of proteins including 4 main respiratory chain complexes: complex I – NADH dehydrogenase, complex II – succinate dehydrogenase, complex III – cytochrome c oxidoreductase, complex IV – cytochrome c oxidase. The final Complex V is ATP synthase (*6*). Complexes I-IV are essentially proton pumps create the electrochemical gradient of H^+ across the IMM according to chemiosmotic theory proposed by P. Mitchell (*7*). That gradient is then used by ATP synthase to produce the energy equivalent of ATP (*8*). Necessary metabolites and reducing equivalents (NADH, $FADH_2$) fueling the electron transport chain are produced by another biochemical reaction in mitochondria – the

tricarboxylic acid (or Krebs) cycle (*9*). Metabolites produced by the Krebs cycle are utilized in multiple biosynthetic pathways for amino acids, lipids, carbohydrates, nucleotides (*10*). Mitochondria are also involved in heme biosynthesis and formation of iron-sulfur clusters (*11, 12*).

Mitochondria have long been regarded as predominantly bioenergetic and biosynthetic organelles. However, new mitochondrial functions have emerged that changed our view of the role of these organelles in various cellular processes. The regulation of calcium homeostasis is one of the first known alternative functions of mitochondria. Accordingly, mitochondria serve as Ca^{2+} buffers for cytosol and possess several transport systems for calcium (*13, 14*). Mitochondrial Ca^{2+} has also been implicated in a number of processes such as regulation of bioenergetics, protein quality control, and mitophagy (*15, 16*). In general, mitochondrial Ca^{2+} is considered as a second messenger molecule associated with a number of signal transduction pathways (*17, 18*).

The main organelle responsible for Ca^{2+} uptake and storage is the endoplasmic reticulum (ER). Involvement of mitochondria in calcium signaling revealed an important interplay between mitochondria and ER (*19*). The physical interaction of ER and mitochondria has been discovered, and termed mitochondria-associated ER membranes. These organelles contact each other using proteins and lipids located on the OMM and ER membrane, forming juxtaposed tethering structures (*20, 21*). The ER-mitochondria structure identified in yeast is the best characterized to date (*22*). Overall, ER-mitochondria contact sites are important

formations that permit the reciprocal regulation of both organelles, modulating Ca^{2+} and insulin signaling, lipid metabolism, cell death and survival signals (*23*, *24*).

Mitochondria employ numerous means to communicate with the rest of the cell involving a number of different activities including apoptosis, reactive oxygen species (mtROS) generation, and protein quality control (*10*). Each of these pathways uses a particular mitochondrial response in order to accomplish the appropriate physiological outcome. The defining signs of apoptosis are mitochondrial cytochrome c release, OMM permeabilization and activation of caspases (*25*). These events are regulated by pro- (BAX, BAK) and anti-(Bcl-2, Bcl-XL)-apoptotic proteins located on the OMM (*26, 27*).

Mitochondrial reactive oxygen species initially have been considered as pathological consequences of respiratory chain dysfunction (*28*). However, it has been demonstrated that they act as second messengers in the induction and regulation of genes responsible for metabolic control during low oxygen concentration (hypoxia-inducible factors) (*29*). In addition, different pattern recognition receptor signaling cascades such as toll-like, nuclear oligomerization, domain-like and inflammasome pathways are regulated by mtROS, which make them significant contributors to innate immunity (*30, 31*).

The identification of a mitochondrial antiviral- signaling protein on the OMM signifies the role of that compartment in establishing a platform for anti-viral immune responses (*31, 32*).

Another critical function of the OMM is related to localization of A-kinase anchoring proteins, which bind cAMP-dependent protein kinase A forming a signaling hub and phosphorylating substrates on OMM (*10*, *33*).

Finally, overwhelming of mitochondrial function results in the mitochondrial unfolded protein response. The mitochondrial unfolded protein response induces the expression of nuclear-encoded mitochondrial chaperones and proteases that monitor quality control mechanisms by removing faulty proteins and restoring mitochondrial homeostasis (*34*, *35*).

In order to perform the functions discussed above, mitochondria need the necessary protein machineries that are mostly encoded by nucleus (*36*). As a consequence, mitochondria possess intricate protein import pathways. The general entry route to mitochondria is through a translocase of the outer membrane called TOM. After translocation, the imported proteins are distributed between five major protein sorting pathways depending on their sequence and structure. Proteins with a cleavable mitochondrial targeting sequence destined to the matrix or IMM are transported through TIM23/22 translocases. Separate routes are available for imported proteins with cystein-rich motifs, β-barrels or α-helical transmembrane domains (*37*-*39*).

Finally, proper functioning and maintenance of mitochondrial homeostasis is dependent on mitochondrial dynamics which falls into four categories: fusion (the joining of two organelles into one), fission (the division of one organelle to two), transport (the movement of an organelle within a cell) and mitophagy (autophagocytic destruction of an organelle) (*40*). Maintenance of mitochondrial

dynamics and, consequently, homeostasis is critical for faithful reproduction and expression of the mitochondrial genome by mitochondrial replication and transcription machineries.

2. Structure of the Human Mitochondrial Genome.

The uniqueness of mitochondria as organelles is underlined by the presence of a dedicated genome. The human mitochondrial genome is a double-stranded circular molecule of DNA approximately 16 kbps composed of heavy (H) and light (L) strands as a result of differences in their base composition. Human mtDNA encodes 13 mRNA genes of components of respiratory chain complexes, 22 tRNAs and 2 rRNAS genes (12S and 16S) (Fig. 1A) (*41, 42*). The genome lacks introns and contains one large (~1 kbs) non-coding region (NCR), also known as the control region. The control region is composed of several critical sequence elements, which include promoters for L and H strand transcription (LSP and HSP), three conserved sequence boxes (CSBI, CSBII and CSBIII), the displacement loop (D-loop), termination-associated sequences (TAS), and the origin of H strand replication (O_H) (Fig. 1B) (*41, 43*). The origin of L strand replication (O_L) (30 bp region) is located at a distance of about two-thirds of mtDNA relative to O_H (*44*). A characteristic structure of mtDNA, the displacement loop (D-loop), is believed to form due to premature termination of replication evens that start at O_H and concludes some 650 bps downstream, within the termination-associated sequence region (TAS). The terminated nascent H strand, called 7S DNA, remains annealed to the template strand,

forming a triple-stranded DNA segment (Fig. 1B). The D-loop region extends from O_H (the 5′ end of 7S DNA) to the TAS (the 3′ end of 7S DNA). The 5′ end of 7S DNA can be placed at a number of defined sites in human mtDNA, suggesting that it comprises a family of molecules of different sizes (*45-47*). It is generally thought that the synthesis of 7S DNA is primed by the 7S RNA molecule generated from LSP (*48*). 7S DNA may thus represent an intermediate product of prematurely terminated heavy-strand replication (*49*). TAS is located at 3' end of nascent D-loop H strands and thought to participate in termination of replication and transcription of both strands (*50*).

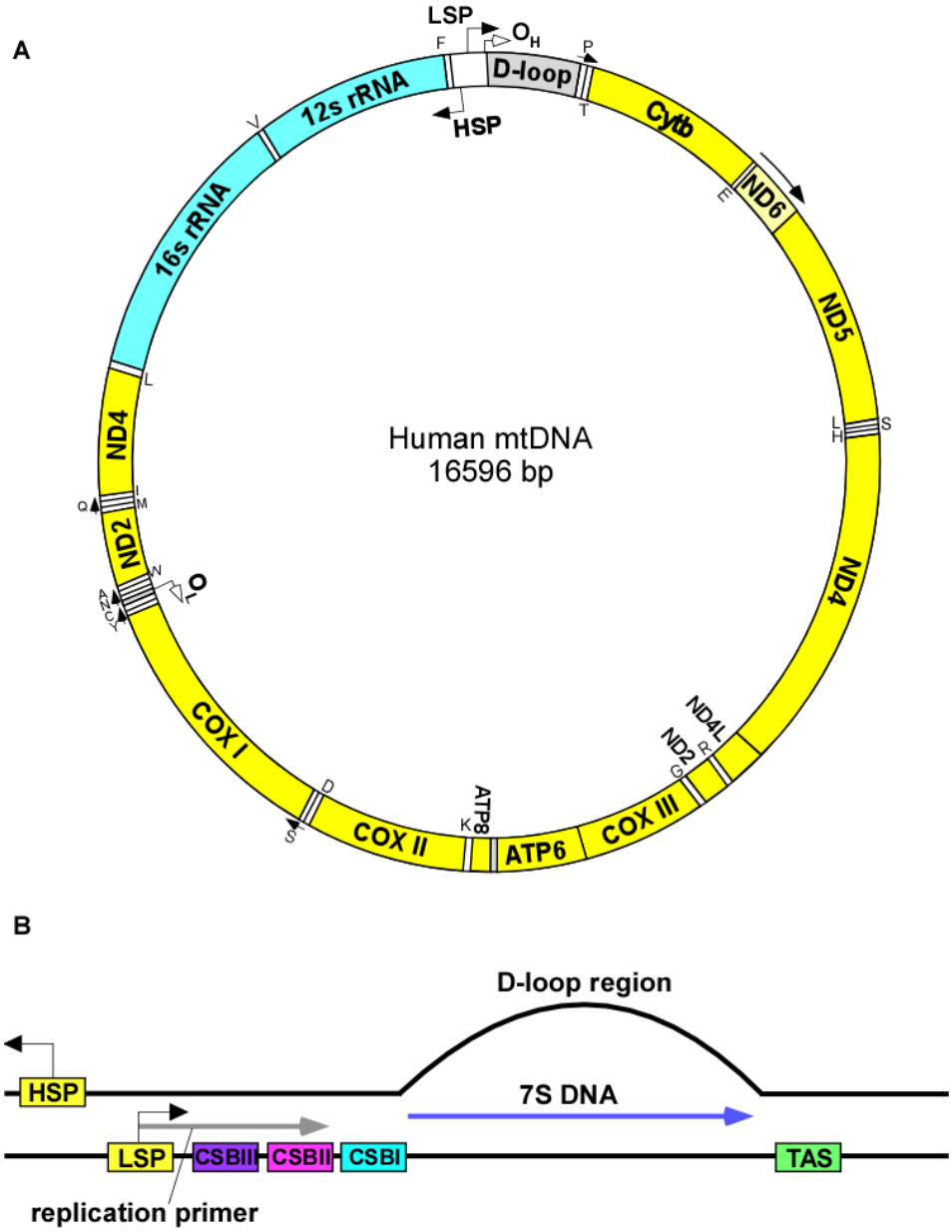

Figure 1. Organization of mitochondrial genome. A Structure of human mtDNA (Falkenberg et al., 2007 with modifications). B. Noncoding region of mtDNA

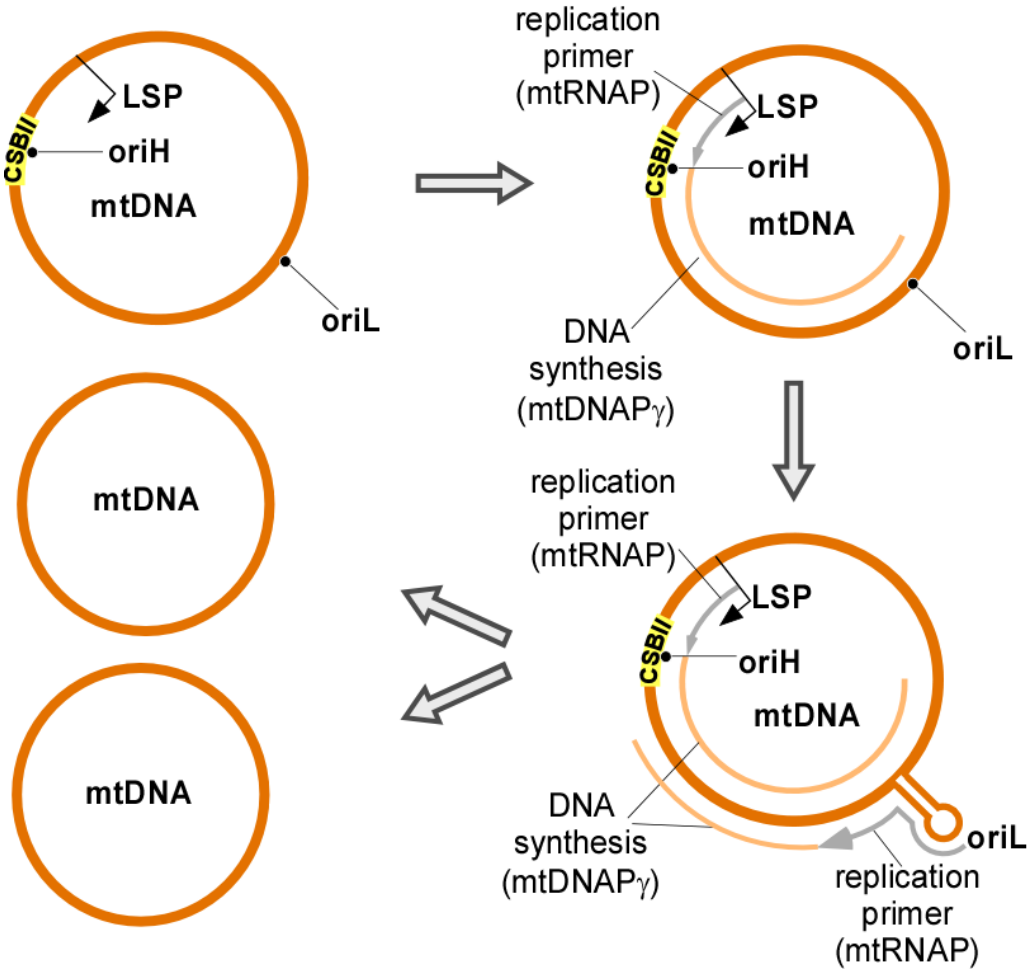

Figure 2. Asymmetric model of mtDNA replication (see the text for details).

MtDNA is packaged in nucleoid structures similarly to bacterial cells. Mitochondrial nucleoids have been visualized by different microscopy methodologies. The diameter of the nucleoid has been shown to be ~100 nM in various cell lines (Brown et al., 2011; Kukat et al., 2011). It is estimated that each nucleoid in human cells contains about one mtDNA molecule (*51*). The major constituent of nucleoid is mitochondrial transcription factor A (TFAM), a histone-like protein belonging to HMG-box family, however other mitochondrial proteins have also been detected (*52*). TFAM is absolutely required for nucleoid stability and

structure, as TFAM knockout in mice results in severe depletion of mtDNA and embryonic lethality (*53*). MtDNA is believed to be completely covered by TFAM as one molecule of TFAM is present for every 20-25 bps of mtDNA (*43, 54, 55*). TFAM alone can fully compact mtDNA in nucleoid structure by cross-strand binding and loop formation. Packaging of the mitochondrial genome is critical for regulation of mtDNA maintenance and expression.(*55*).

3. Regulation of mtDNA copy number.

MtDNA copy number regulation is crucial for cell function. Cells maintain different number of mitochondrial DNA molecules ranging from 100 up to 100,000 copies per cell (*56*). Depending on physiological conditions cells can either increase or decrease mtDNA copy number in order to maintain a homeostatic set point value. The set point becomes an important parameter of cellular metabolism as replication and segregation of mitochondrial genome is involved in efficiency of mitochondrial function. Deviations from the set point value result in a number of abnormalities and pathological conditions affecting cellular bioenergetics. Cell development and differentiation strategies involve preserving a particular mtDNA copy number during the transition to successive developmental stages. Oogenesis and embryogenesis present one of the best studied examples where regulation of mtDNA copy number is directly connected to the productive physiological outcome (*56, 57*).

3.1. MtDNA copy number regulation during development.

MtDNA copy number is strictly regulated during embryogenesis and oogenesis. Precursors of oocytes – primordial germ cells - contain about 200 copies

of mtDNA at basal state (*57, 58*). During oogenesis, these cells accumulate mtDNA at metaphase I. Upon transition to metaphase II the mtDNA copy number increases exponentially, reaching 10^5 copies and providing a critical threshold necessary for fertilization (*56*). . After fertilization, mitochondrial replication ceases and does not occur again until the zygote stage (*57*).

Subsequent stages of embryogenesis are also affected by replication of mtDNA. Thus, a developed embryo at the blastocyst stage undergoes compaction and expansion transitioning to the morula stage. At that moment, key developmental processes regulate formation of two different compartments of the embryo: the outer layer ring that is throphectoderm, and the inner cell mass that eventually transforms into the fetus. Trophectoderm cells increase mtDNA copy number at the blastocyst stage whereas the inner cell mass gradually reduces mtDNA quantities (*57, 59*). Thus, the latter cells restrict mitochondrial replication and establish an mtDNA set point. Establishment of the mtDNA set point is a crucial event for embryonic stem cells as it is the basis for fast proliferation events. Stem cells primarily rely on glycolysis for their metabolic needs as it generates energy faster than oxidative phosphorylation (*56, 60*). Accordingly, a low mtDNA copy number is related to inefficient oxidative phosphorylation and ATP production that ensures maintenance of glycolytic potential and rapid proliferation of stem cells before organogenesis and differentiation. Further, during differentiation events the mtDNA copy number is regulated in a cell-specific manner. In the majority of cases, commitment to a specific differentiation program such as neurogenesis and myogenesis is accompanied by increase in mtDNA (*57*).

3.2. MtDNA copy number regulation during pathogenesis.

MtDNA copy number plays a key role not only in developmental processes but also affects a number of pathologic conditions defining mtDNA as a potential biomarker of various diseases. Mitochondrial diseases that include MELAS, Pearson syndrome and mtDNA depletion syndrome, are associated with reduction of mtDNA copies. These diseases are caused by particular mutations in protein or tRNA genes that may cause reduction of replication potential (*57, 61*). Mitochondrial depletion syndrome is caused by large deletions of about 5 kb that affect origins of replication (*57*). A low mtDNA copy number is important for efficient glycolytic metabolism and, consequently, proliferation. Therefore, diseases of inflammatory origin downregulate mtDNA copy number to promote efficient energy production and cell expansion. Thus, Parkinson disease neurons are associated with reduction of mtDNA (*62*). Additionally, numerous tumor types such as kidney, breast, esophageal cancers, preferentially decrease mtDNA number, promoting the Warburg effect associated with aberrant mitochondrial function (*63, 64*). However, certain cancer types such as lung adenocarcinoma show significant accumulation of mtDNA that may indicate a mitochondrial-dependent mechanism of tumorigenesis (*65*).

3.3. Molecular mechanisms of mtDNA copy number regulation.

Replication of mtDNA occurs irrespective of cell cycle and replication of nuclear genes (*66*). Currently three different models of mtDNA replication are proposed (*43, 67*). The prevailing model, called strand the displacement model or asymmetrical model of replication (Fig. 2), was put forth by David Clayton (*49, 68, 69*). According to this model, mtDNA is continuously replicated on both strands

without formation of Okazaki fragments. Replication is initiated at the LSP by mtRNAP, which functions as a primase (*41, 43*). MtRNAP primes replication of the H strand of mtDNA by generating a ~200 nt primer, synthesis which is terminated at conserved sequence block II (CSB II) (Fig. 1B) (*70*). CSBII is a G-rich sequence conserved in vertebrates that forms a G-quadruplex structure in the RNA, resulting in transcription termination by mtRNAP (*48, 71, 72*). The replication primer at CSBII is then extended by the replisome, which consists of mtDNA polymerase (mtDNAP), TWINKLE helicase and mitochondrial single-strand binding protein (mtSSB), until it reaches two-thirds of mtDNA, passing O_L (*41, 43, 68*). At that point, O_L becomes single-stranded and forms a stem-loop structure. This structure is recognized by mtRNAP that primes the synthesis of the L strand of DNA. After synthesis of a 25 nt primer at O_L, mtRNAP is replaced by mtDNAP (Fig. 2) (*73, 74*). Subsequently, the synthesis of both mtDNA strands proceeds continuously and results in two molecules of mtDNA. The support for this model comes from numerous experiments, which include atomic force microscopy, mapping of free 5' ends of O_H and O_L, and reconstitution of initiation of replication at O_L (*74-77*). Two other models of replication are: ribonucleotide incorporation throughout the lagging strand (RITOLS) and general strand-coupled DNA replication (*78, 79*). The supporting evidence for these models is based entirely on observation of DNA intermediates in two-dimensional gels. However, the presence of such intermediates is not supported by microscopy experiments or *in organelle* labeling (*67*).

Regulation of mtDNA copy number could be achieved by a number of different mechanisms. First, mtDNA replication could be regulated through the

initiation of transcription at LSP as mtRNAP generates the replication primer (*41*). Second, replication may be regulated by termination of the LSP transcript at the CSB2 region to initiate DNA synthesis (*71, 72*). Third, regulation may occur by termination of mtDNA replication at TAS to form the 7S DNA (*50*). Finally, replication can be regulated by controlling the degree of mtDNA compaction in order to produce mtDNA molecules available for initiation of DNA synthesis (*80*).

Generation of the replication primer by mtRNAP could be an important stage of regulation of mtDNA replication as it is absolutely necessary for initiation of mtDNA replication. It has been suggested that levels of mtRNAP differentially regulate transcription initiation events at promoters (*81*). Lower mtRNAP levels favor transcription initiation at LSP in order to maintain replication primer synthesis. Under normal conditions, most of the replication initiation events at O_H are prematurely terminated, forming 7*S*DNA. Instead, upon depletion of mtRNAP, 7S DNA is no longer formed, thus, promoting full-length replication (*81*). At the same time, mtRNAP is absolutely essential for expression of mitochondrial genome and cell viability and knockout results in embryonic lethality (*81*). Therefore, it is questionable whether complete depletion of mtRNAP would be physiologically relevant.

The majority of transcription initiation events at LSP are terminated near CSBII due to formation of G-quadruplex structures in RNA. These terminated transcripts remain associated with DNA, forming an R-loop structure between the nascent RNA and the nontemplate strand of DNA (*71, 72*). It is likely that the terminated transcripts may be used as primers for initiating replication of the mtDNA

H-strand as CSBII coincides with O_H (*46*). However, initiation of mtDNA replication at O_H has not been reconstituted *in vitro*. That event may be regulated by unknown factors involving helicase or nuclease activities that would facilitate RNA-DNA transitions and replication of mtDNA (*43*).

Recent evidence suggests that TAS at the end of D-loop region could serve a crucial regulation function between abortive and processive replication. MtDNA immunoprecipitation experiments demonstrated an increased occupancy for mtDNAP and TWINKLE in the D-loop region (*50*). The decreased TWINKLE occupancy at the TAS region upon normal conditions correlated with the high levels of 7S DNA in mitochondria (*82*). This suggests an increase in the replication of 7S DNA in the D-loop due to abortive replication events. However, depletion of mtDNA resulted in increased TWINKLE occupancy and low levels of 7S DNA, indicating that TWINKLE may regulate replication initiation events at this region (*50*). In addition, genetic studies in mice support the idea of regulation of mtDNA replication by TWINKLE (*82*).

Finally, mtDNA copy number might be ~~can~~ be regulated by modulating the levels of TFAM and, thus, availability of mtDNA for replication. Small variations of TFAM levels may have dramatic effects on mtDNA transcription and replication. Down-regulation of TFAM *in vivo* coincides with reduction of mtDNA copy number (*83*) while an increase in TFAM levels results in an increase of mtDNA copy number (*84*).

4. Mitochondrial transcription.

4.1. RNA polymerase.

The mitochondrial transcription machinery is distinct from the nuclear one. It is composed of mtRNAP, two initiation factors TFAM and TFB2M, a putative elongation factor TEFM and a termination factor MTERF (*85*).

MtRNAP is single-subunit enzyme distantly related to the bacteriophage T7 RNA polymerase (T7 RNAP) and the Pol A family of polymerases (including DNA polymerases and reverse transcriptase). Recently, the crystal structure of mtRNAP revealed the precise domain composition of enzyme and structural elements involved in different transcription stages (Fig. 3) (*86*). The main structural domains of mtRNAP are the T7 RNAP-like catalytic C-terminal domain (CTD) (res. 647-1230), the N-terminal domain (NTD) (res. 368-647) that remotely resembles the promoter-binding domain of T7 RNAP, and a pentatricopeptide repeat-containing domain (PPR). The PPR domain consists of nine a-helices and two pentatricopeptide repeat (PPR) motifs known to bind RNA (*86*). The function of this domain of mtRNAP is not known.

The CTD has a characteristic thumb-palm-fingers fold similar to the Pol A family polymerases suggesting structural conservation of an active center, and likely mechanisms of substrate binding, selection and catalysis between T7 and mtRNAP (Fig. 3) (*86-88*).

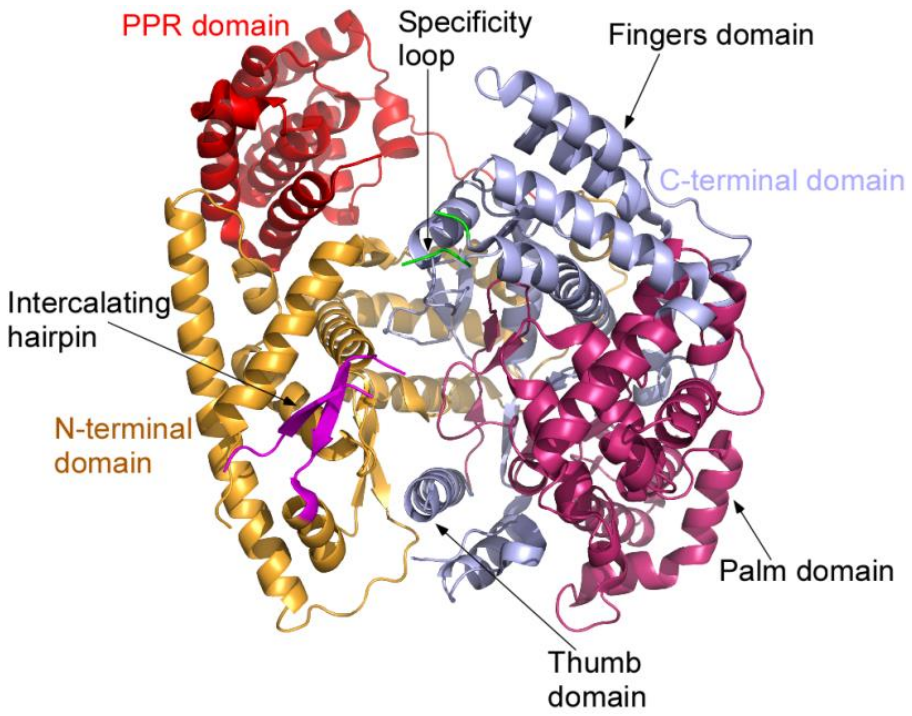

Figure 3. The structure of mtRNAP.

The most prominent structural difference between the CTD of T7 and mtRNAP is the position of the fingers subdomain that participates in catalysis of substrate incorporation (*89, 90*). The fingers subdomain in T7 RNAP contains the structural element called O-Y helices which is responsible for catalysis, NTP positioning and translocation of T7 RNAP along DNA (*91*). In mtRNAP the fingers subdomain is rotated around the O-helix and places the N-terminal portion of the Y-helix towards the NTD. In addition, part of the entrance to the active site is occluded and prevents proper positioning of the O-helix to deliver the NTP to the active site for catalysis (*86, 88, 91, 92*).

The NTD of mtRNAP shares no sequence similarity to T7 RNAP. However, it contains structural elements that correspond to the "AT-rich recognition" loop and "intercalating" hairpin in T7 RNAP (Fig.3) (*86*). The positions of the "AT-rich recognition" loop and intercalating hairpin of mtRNAP are not compatible for promoter recognition and melting as in T7 RNAP (*86*). These differences underline the need of mtRNAP for transcription factors to activate transcription. The NTD domain of T7 RNAP undergoes dramatic conformational changes upon transition from initiation to elongation that is achieved by dramatic reorganization of the promoter-binding site and the movement of the six-helix region of T7 RNAP (called the promoter-binding domain) into the position previously occupied by promoter DNA (*92*, *93*).

4.2. Initiation of mitochondrial transcription.

Initiation of mitochondrial transcription is a sequential process (Fig. 4) (*94*). The process is initiated by binding of TFAM -35 to -15 base pairs upstream of promoter (*95*). Binding of TFAM to DNA induces severe bending of the DNA (*96*). Next, TFAM recruits mtRNAP to promoter DNA. The NTD then mtRNAP makes interactions with C-terminal tail of TFAM that are important for proper positioning of

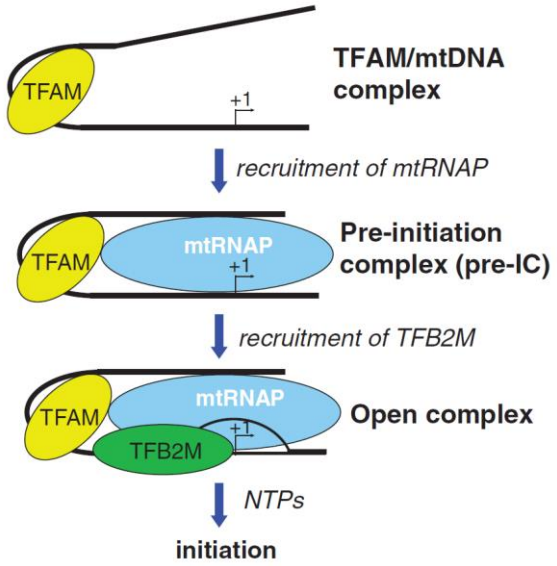

Figure 4. The assembly of mitochondrial transcription initiation complex (*1*).

mtRNAP on DNA. Thus, a pre-initiation complex (pre-IC) is formed consisting of TFAM, DNA and mtRNAP (Fig. 4) (*1*). Another factor, TFB2M, binds to the pre-IC by making interactions with the promoter DNA near the transcription start site, together with the priming NTP and the template +1 DNA base (*94*, *97*). Recruitment of TFB2M results in the transition from the pre-IC to the open initiation complex (IC) (Fig. 4).

4.3. Mitochondrial transcription elongation.

RNA synthesis occurs through repetitive cycles of substrate incorporation, which requires translocation of the active site of RNAP along the DNA template. Each nucleotide addition cycle (NAC) can be divided into four steps (Fig. 5) (*87*, *91*, *92*). First, the substrate NTP binds to RNAP in a pre-insertion state, in which it makes an incipient base-pare with the templating (+1) base of the DNA (1). At the second step of the NAC the substrate is delivered into the NTP-insertion site (or N-site) and is paired to the template nucleotide (2) whereas the 3' end of the RNA is situated in the priming site (P-site). Next, the nucleotidyl transfer reaction produces pyrophosphate (PPi) and the RNA transcript is extended by one nucleotide

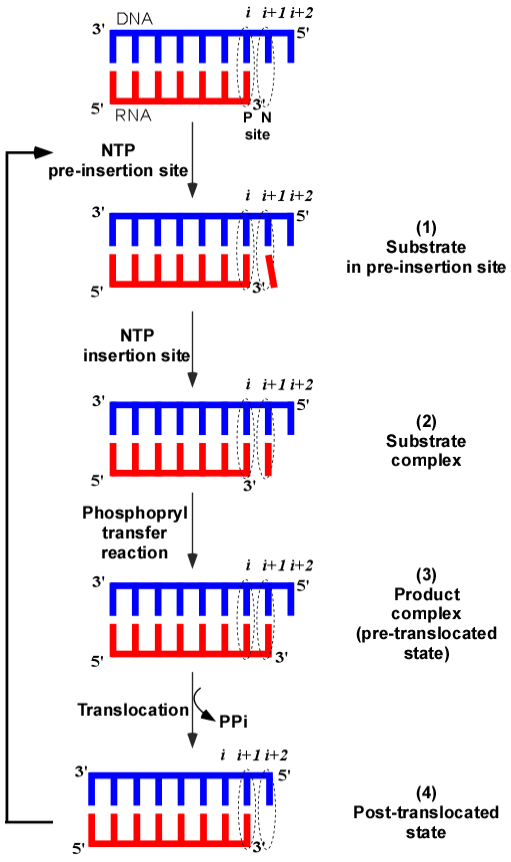

Figure 5. Nucleotide addition cycle.

that now occupies the N-site (3). Finally, due to translocation along the DNA template, the 3' end of the RNA moves from the N-site to the P-site and the free N-site may be occupied by the next incoming nucleotide (4) (*88*). The crystal structures of all four steps of nucleotide addition by T7 RNAP have illuminated the structural basis for translocation (*88, 90, 93, 98*).

Comparison of the structures of T7 RNAP in the pre-and post-translocation states shows a significant rotation of a helical subdomain in the fingers domain (*88*). Specifically, in the pre-translocation state the O helix is found in the "closed" conformation, in which the incoming nucleotide is properly positioned for base-pairing with DNA template. PPi release after nucleotidyl transfer reaction serves as a driving force providing energy required for translocation. Next, binding of NTP to pre-insertion site promotes the new cycle of nucleotide addition to the growing RNA chain.

Considering the high homology between T7 RNAP and the C-terminal domain of mtRNAP, it is likely that mtRNAP achieves single-nucleotide incorporation using the same core steps, including conformational changes before and after catalysis. It has been shown that mtRNAP complexes with 9 bp DNA-RNA hybrid are more stable than 8 bp complexes in contrast to T7 RNAP complexes(*99*). Moreover, these complexes are able to rapidly elongate RNA. Initially, mtRNAP binds to the 9 bp DNA-RNA complex rapidly followed by slower conformational change necessary for catalysis. Solvent deuterium kinetic isotope effect experiments have demonstrated that catalysis is a partially rate-limiting step for nucleotide incorporation (*100*). Pulse-chase-quench and pulse-quench experiments were used to

identify other rate-limiting steps during nucleotide addition cycle. These experiments showed that more product is formed under pulse-chase-quench indicating the presence of a rate-limiting conformational changes prior chemistry. Similar experiments demonstrated the presence of rate-limiting conformational change after chemistry. Thus, catalysis and two conformational rearrangements represent three rate-limiting steps for nucleotide incorporation by mtRNAP (*99*).

Unlike T7 RNAP, which does not require any additional components, the elongation stage of mtRNAP involves a recently described elongation factor, TEFM, which appears to be tightly associated with mtRNAP during purification (*101*). TEFM is thought to promote transcription elongation according to several lines of evidence. It has been shown that knockdown of TEFM in human cell lines decrease the abundance of promoter-distal mtDNA transcripts. Immunostaining revealed that TEFM co-localizes with the newly synthesized RNA in nucleoids (*101*).

Using pull-downs with different variants of mtRNAP, it was shown that TEFM interacts with C-terminally truncated mtRNAP, indicating that this region is required for binding. TEFM promotes the processivity of mtRNAP *in vitro* (*101*). These processivity experiments were performed using non-promoter templates (*101*) and thus the role of this transcription factor in transcription using natural promoters has not been demonstrated.

5. Termination of mitochondrial transcription.

5.1. Factor-dependent termination of mitochondrial transcription.

A mitochondrial transcription termination factor (mTERF1) has been identified based on transcription termination activity from mitochondrial lysates (*102*). Later, a 39 kDa protein was isolated by affinity chromatography and its termination function confirmed transcription assays. MTERF1 interacts with a 22-bp sequence at 3' end of $tRNA^{Leu(UUR)}$. The crystal structure revealed the mode of termination by mTERF1 (*103, 104*). The binding mechanism involves the establishment of sequence-specific interactions followed by DNA unwinding. DNA unwinding further destabilizes base pairing in the duplex resulting in base flipping. The flipped nucleotides are stabilized by stacking interactions with three amino acids (*103, 104*). Therefore, base flipping stabilizes mTERF1 on the DNA substrate. In general, mTERF1 acts as a "roadblock" preventing or interfering with elongation of mtRNAP (*105*). MTERF1 exhibits polarity towards L strand termination (*103, 106*), consistent with genetic data indicating that loss of mTERF1 did not decrease the levels of different RNAs but reduced the anti-sense transcription of rRNA genes (*107*).

Other members of the MTERF family have been identified based on their sequence similarity with mTERF1: mTERF2, mTERF3, mTERF4 (*108*). MTERF2 binds DNA non-specifically and is present in nucleoids (*109*). Knockout of the *mterf2* gene is not lethal (*43*). MTERF3 has been shown to be a transcriptional repressor as knockout of the gene results in up-regulation of mtDNA transcription initiation. Chromatin immunoprecipitation experiments have localized the binding of mTERF3

to the promoter region of mtDNA. However, the same protein interacts with 16S rRNA and promotes assembly of the large ribosomal subunit (*110, 111*). Thus, mTEFF3 may coordinate the regulation of transcription and translation; however the precise mechanisms are not known. MTERF4 forms a heterodimer with NSUN4 that methylates cytosine 911 on 12S rRNA. That heterodimer further helps to assemble the small and large subunits into a functional monosome. The loss of MTERF4 impairs translation and causes an increase in mitochondrial transcription (*112*, *113*). More work is required to elucidate the role of MTERF proteins 2-4 in transcription regulation.

5.2. Mitochondrial Transcription Termination by DNA sequence elements.

Termination of mitochondrial transcription can also be caused by a structural element encoded in mtDNA, such as conserved structural block II (CSBII) (*70*) CSBII is a G-rich sequence that is conserved in all vertebrates, located about 100 nt downstream of the LSP.

The majority of transcription initiation events are prematurely terminated downstream of CSBII between positions 300-282 of mtDNA (*70*). The 3' ends of terminated transcripts coincide with RNA-DNA transition points near O_H (*43*). Thus, CSBII presents a major regulatory point during the initial stages of transcription. The molecular mechanisms of transcription termination at CSBII are believed to include the formation of a G-quadruplex structure in the nascent RNA or a hybrid G-quadruplex structure in both RNA and DNA (*71, 72*). However, termination at CSBII

is dramatically reduced in the presence of 7-deaza-GTP, which is known to destabilize G-quadruplexes, arguing in favor of G-quadruplex formation in RNA only (*72*). K^+ ions stabilize the G-quadruplex structure and termination due to charge coordination (*71*). Every one of 12-15 G residues is important for effective termination at CSBII, as mutations decrease the percent of premature termination events (*71, 114*). It has been proposed that formation of the G-quadruplex structure results in removal of the 3' end of RNA from the active site of RNAP, - thus causing termination and negatively regulating replication initiation events (*72*). It has been proposed that additional factors may be required for an effective utilization of RNA primer for replication (*70, 71*). In conclusion, mitochondria are unique biological system where transcription termination by G-quadruplex at CSBII exists.

Recent studies indicate the existence of another transcription termination event downstream of the D-loop region at the termination-associated sequence (TAS) element. The majority of H-strand transcription has been shown to terminate near the TAS sequence, preventing further progression of mtRNAP into the D-loop region (*50*). Transcription termination at this region coincides with a reduced number of mtDNA replication events. The TAS does not appear to form a secondary structure and thus is unlikely to terminate transcription and replication on its own and would require a protein factor (*50*). Mitochondrial chromatin immunoprecipitation experiments showed that mtRNAP and TWINKLE helicase are enriched at TAS (*50*). Under normal conditions, TWINKLE levels are low at that site, but mild depletion of mtDNA results in increased TWINKLE occupancy at TAS and decreased levels of 7S DNA. This indicates that TWINKLE can be reloaded to reinitiate mtDNA replication

at the end of D-loop (*43, 50*) . As discussed above, TAS is implicated in regulation of replication and therefore this region may be an important regulatory point affecting mtDNA copy number in cells.

Rationale

Mitochondria are ubiquitous organelles present in all nucleated cells. These organelles have been involved in a number of cellular functions, including: signal transduction, innate immunity, and fatty acid and nucleic acid metabolism. They contain their own multi-copy genome that encodes components of respiratory chain complexes necessary for ATP production. Regulation of mtDNA copy number is critical to developmental processes such as spermatogenesis, embryogenesis, and differentiation, etc. Mutations in mtDNA can lead to aberrant protein synthesis and result in mitochondrial dysfunctions. Deregulation of mtDNA copy number has been linked to a number of pathologies, including: mtDNA depletion syndrome, Parkinson's disease, and different types of cancer. Understanding the molecular mechanisms of mitochondrial transcription and regulation of mtDNA copy number is thus important for understanding the processes involved in human diseases.

Although the mitochondrial transcription initiation system is well characterized, the mechanisms of transcription elongation remain obscure. Mitochondrial RNA polymerase (mtRNAP) can produce either long polycistronic transcripts or short primers required for mtDNA replication; however, it is not known how these processes are regulated. In mitochondria, all major steps of gene expression - replication, transcription and protein expression – occur in the same cellular compartment. Our knowledge about the interplay between transcription and replication and transcription and protein synthesis is very limited. The goal of this work is to characterize the molecular mechanisms of mitochondrial transcription

elongation and understand how replication and transcription are coordinated to regulate mtDNA copy number.

Materials and methods

Proteins.

Human mtRNAP, TFAM and TFB2M were purified as described previously (*1*). Yeast mtRNAP was purified as described by (*115*). T7 RNAP was purified according to (*116*). C-terminal His-6 TEFM (mature form, residues 36-360) was cloned into pET22b vector and expressed in BLR (DE3, recA-) cells (Novagen) or in Rosetta2 cells. The overnight cell culture (5 ml) was used to inoculate 1L of LB and the cell culture was incubated for 3-4 h at 37 ˚C until OD at 600 nm reached 0.4 units. The flasks were then transferred to 16°C and incubated for additional 40 min prior to addition of IPTG (0.2 mM). The cells were harvested after 18 h of incubation and disrupted by sonication. TEFM was first purified by affinity chromatography on Ni-agarose beads (Qiagen) followed by heparin-sepharose purification in 250-1500 mM gradient of NaCl. The protein was concentrated to 3-10 μM concentration, diluted 2 times with glycerol, aliquoted and stored at -70 °C.

Construction of TEFM and mtRNAP variants.

TFEM mutants were obtained by site-directed mutagenesis (Quick-change, Agilent) starting with the plasmid containing C-terminal His-6 WT TFEM (see above). . Plasmids containing the N-terminal deletion variants of TEFM Δ132 (deletion of residues 1-132) and Δ144 (deletion of residues 1-144) TEFM were provided by Hauke Hillen (Max Planck Institute for Biophysical Chemistry, Gottingen).

Preparation of PCR templates for transcription assays.

Fragment of human mtDNA (202-481) was cloned into pT7 blue vector according to manufacturer's instructions (Novagen). Templates for the transcription assays involving CSBII were amplified by PCR from a plasmid containing region 202-481 of human mtDNA. Note that the reference human mtDNA (Cambridge) contains a rare polymorphism in the CSBII region (311-315) (*117*) and therefore a more commonly occurring in humans CSBII sequence (G6AG8) was used in this study. A variant of a PCR template ("del71 CSBII") with the deletion of 71 bps region between the LSP promoter and CSBII (bps 388-318) was used for transcription and cross-linking assays. The templates having modifications in the CSBII region are listed in Table VII (appendix). To prepare templates producing 500 nt or 1000 nt run-off products we PCR-amplified regions of pT7blue plasmid containing -70 to + 70 region of the LSP promoter (region 339-478 in human mtDNA) using forward primer "500" and "1000", respectively, and the reverse primer "U19" (see Table III in the Supplement). These templates have been also used for promoter-dependent initiation assays. Plasmid pT7blue14SRO78 encoding yeast 14S rRNA promoter was kindly provided by Dr. Dmitry Markov. The template producing a 500 nt run-off product on this plasmid was generated as described above.

Transcription assays.

Promoter-dependent transcription and anti-termination assays

Transcription reactions were carried out using DNA templates (50 nM), mtRNAP (150 nM), TFAM (200 nM), TFB2M (150 nM), TEFM (300 nM) in

transcription buffer containing 40 mM Tris (pH=7.9), 10 mM $MgCl_2$ and 10 mM DTT in the presence of ATP (0.3 mM), GTP (0.3 mM), CTP (0.3 mM), UTP (0.01 mM) and 0.3 μCi [α-^{32}P] UTP (800 Ci/mmol). Reactions were carried out at 35 °C for the time indicated in figure legends and stopped by addition of an equal volume of 95% formamide/0.05M EDTA. The products were resolved by 10-20% PAGE containing 6 M urea and visualized by PhosphoImager (GE Healthcare).

Assembly of elongation complexes using nucleic acid scaffolds.

To prepare nucleic acid scaffold templates, we radioactively labeled RNA primer by T4 polynucleotide kinase for 30 min at 37 °C. The labeled RNA was annealed to template (T) and non-template (NT) DNA strands taken at the equimolar concentrations for 7 min at 75 °C followed by gradual cooling to room temperature. MtRNAP was per-incubated with scaffold complexes for 5-10 min RT prior to addition of the substrate NTPs. RNA was extended by addition of 100 μM NTPs for 5 min at RT or 35 °C.

Halted complex stability assay

The elongation complexes (ECs) having 35 nt RNA were formed using LSP promoter template (Table III, supplements) in the presence of limited NTP set: 0.3 mM each GTP, and ATP, 0.01 mM UTP and 0.3 μCi [α-^{32}P] UTP (800 Ci/mmol). The ECs were halted for 10-40 min in the presence or absence of TEFM. The ECs were then chased with 0.3 mM CTP and UTP and the amount of a halted complex left after chase was determined by autoradiography.

RNA-protein and DNA-protein cross-linking.

To probe interactions between TEFM and RNA in a scaffold EC, the photo cross-linking probe 4-thio UMP was incorporated into 15 nt RNA oligo (5'-A/4-thioU/GUCUGCGGCGCGC, Dharmacon). DNA-TEFM cross-linking was probed using DNA oligos (Midland Scientific) with 4-thio dTMP incorporated into the template DNA strand at position "+7" (3' GTACCCCATCGCCGCGCGTGCGG/4-thio-dT/CTGC-5') or into the non-template DNA strand, at position "-2" (5'-CATGGGGTATTTATTT/4-thio-dT/GACGCCAGACG-3'). The non-modified oligos used to prepare scaffolds were as described previously (*118*). The photo-reactive oligos were 5' ^{32}P-radiolabeled and the scaffolds were annealed as described above. The ECs were assembled using 1 μM scaffold, mtRNAP (1μM) and TEFM (2 μM) for 10 min and UV-irradiated for 10 min at room temperature. The cross-linked species were resolved using 4-12% SDS PAGE (Invitrogen) and visualized by PhosphoImager (GE Healthcare). To probe interactions of TEFM with the nascent RNA that encompasses CSBII, the promoter PCR template (Δ71CSBII, see above) was used. The mtRNAP was allowed to make a start-up complex and "walked" along the DNA to incorporate a photo reactive analog of GMP, 6-thio GMP (Axxora). The cross-linking was performed as described above.

Measurements of mtRNAP elongation rates.

Measurements using halted elongation complexes. Halted elongation complexes of RNAPs were obtained using limited NTPs set on promoter templates

(Fig. 7). We used PCR-amplified regions of pT7 blue plasmid containing -70 to + 70 region of the LSP promoter, 14S rRNA promoter for *S. cerevisiae* and T7 promoter as templates (50 nM). To generate start-up complexes, reactions were performed for 10 min at RT using mtRNAP (150 nM), TFAM (200 nM), TFB2M (150 nM) in a transcription buffer containing 40 mM Tris (pH=7.9), 10 mM $MgCl_2$ and 10 mM DTT and limited set of NTPs: ATP (0.3 mM), GTP (0.3 mM), UTP (0.01 mM) and 0.3 μCi [α-32P] UTP (800 Ci/mmol). The halted ECs were then chased with CTP (0.3 mM) and UTP (0.3 mM) to detect single round transcription elongation events (Fig. 7). To measure elongation rates we used a histogram with bars representing the fraction of molecules that reached run-off at particular rate interval (see. Fig. 8). Intensity of run-off product at specific time intervals was utilized as a measure of fraction of polymerases. The values were normalized to the highest intensity of run-off product within time intervals. The bars with largest fraction of molecules represented the predominant elongation rates.

Measurements using primer extension assay. Transcription reactions were performed using radioactively labeled R2-8/T1/NT1 (250 nM) scaffold template and mtRNAP variants (500 nM) in the presence of GTP (0.2 mM) and UTP (0.2 mM) – extension by two NTPs in different time intervals (0.1-2 min). Reactions were carried out at RT. Under these conditions, the primer was extended by two nts, and the rate of accumulation of the 12 nt RNA product was measured within the time intervals from 0.1 to 2 min.

Rapid quench flow (RQF) assay to measure elongation rates. The RQF protocol shown in Fig. 9 was done with the KinTek RQF-3 (3-syringe) rapid chemical

quench flow instrument that is capable to measure a reaction within 2-ms (*119*). Small sample loops with 15 µl volume were used. Transcription buffer was as described above. ECs in 15 µl (scaled-up to the number of reactions) including mtRNAP variants (250 nM) and R2-8/T1/NT1 template (125 nM) were assembled for 10 min at RT and loaded into left sample loop (Fig. 9). The instrument was set up with transcription buffer in the left and right syringes and quench solution (1 M HCl) in the middle syringe (Fig. 9). 15 µl (scaled-up to the number of reactions) of NTP substrate for elongation were loaded at twice its final concentration (0.6 mM) in the right sample loop and the valves were moved to the ‘‘fire” position. Then NTPs were added to R2-8 mtRNAP ECs to advance to the R11-14 positions. The rapid mixing program was initiated, and the sample collected in time intervals (2-120 msec). Equal volume mixing of reagents between the left and right sample loops is assumed to give working concentrations of reagents.

To prevent RNA hydrolysis, the sample collection tube contained 300 µl of Tris–base solution to neutralize HCl, so that the RNA is not exposed to low pH for a long period. The samples were collected in the tubes containing 2 µl of 20 mg/ml oyster glycogen, 150 µl 2M Tris–HCl (pH 9.0), and 30 µl 3 M sodium acetate (pH 7.0). One milliliter of ethanol was added to precipitate the RNA. The precipitated RNA was resuspended in gel loading buffer and subjected to electrophoresis.

Experimental results

Chapter 1. Molecular mechanisms of RNA synthesis by mtRNAP.

1. Elongation complexes of mtRNAP.

Elongation complexes (ECs) of mtRNAP can be obtained by two approaches: promoter-dependent transcription and assembly using RNA-DNA scaffolds (Fig. 6). To obtain ECs halted on promoter-containing templates at a defined position we performed *in vitro* initiation experiments in the presence of two initiation transcription factors, TFAM and TFB2M, and a limited set of NTPs. In the absence of CTP, mtRNAP synthesizes an 18-19 nt RNA on this template (Figure 6, A). We found, that the halted ECs could be efficiently "chased" to the run-off suggesting their full functional activity (Fig. 6A). In an alternative approach, we first prepared nucleic

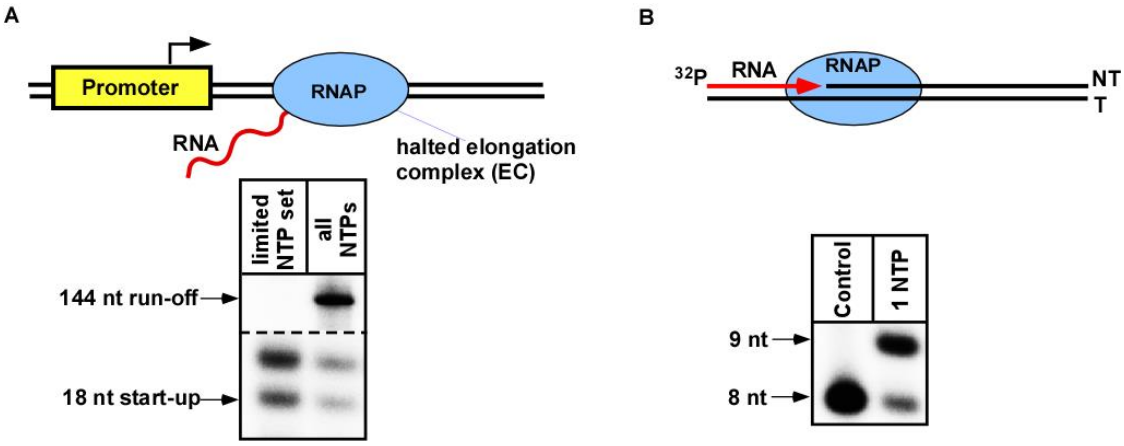

Figure 6. Formation and assembly of mtRNAP elongation complexes. A. Assembly of halted ECs on promoter templates. B. Assembly of mtRNAP ECs using RNA-DNA scaffolds.

acid scaffold templates by labeling an RNA primer with T4 polynucleotide kinase and annealing it with template (T) and non-template (NT) DNA strands at equimolar concentrations (Fig. 6 B). The ECs were assembled by incubation of the scaffold with mtRNAP for 10-20 min at room temperature. We found that mtRNAP efficiently extended RNA in the presence of GTP, suggesting that the complexes had assembled in a functionally competent conformation (Fig. 6 B). While promoter-originated ECs more closely represent the naturally occurring mitochondrial transcription complexes, the primer extension approach allows greater flexibility in assembling elongation complexes with varying lengths of the RNA-DNA hybrid and the sequence context.

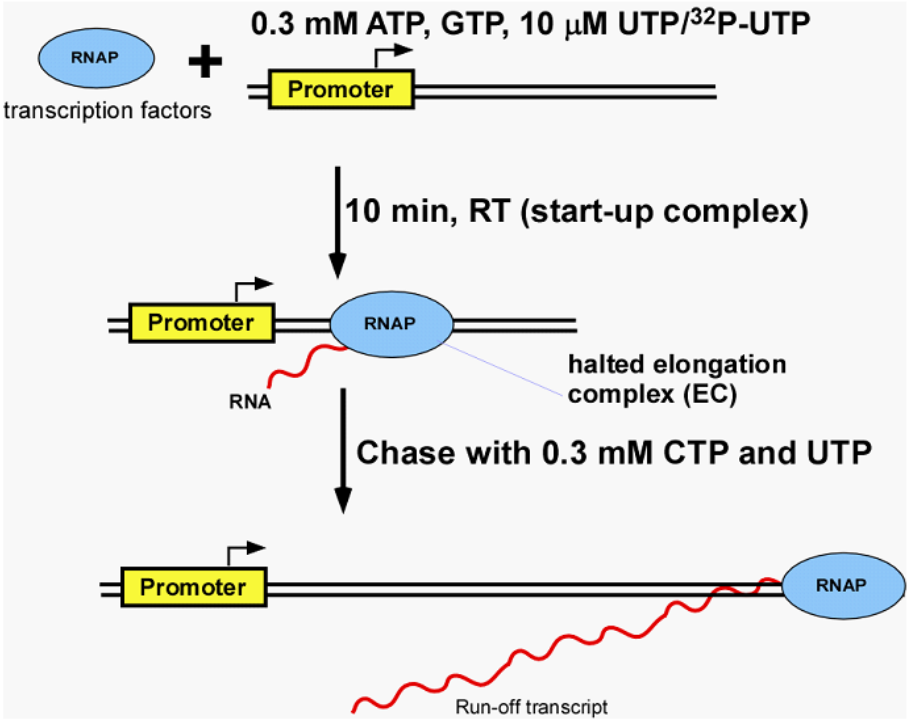

Figure 7. The experimental approach to measure the elongation rates of mtRNAP

2. Elongation rates of mitochondrial transcription.

The life cycle of bacteriophage T7 is just 17 min at 37 °C (*120*). During this time T7 RNAP must overcompete the host RNAP by producing large amounts of transcripts to ensure a preferential synthesis of phage proteins. This requires fast transcription elongation rates of T7 RNAP, which may reach 200-400 nucleotides per second *in vitro* under optimal conditions (*120-122*). In contrast, human mtRNAP has to adjust transcription rates with the rates of RNA processing and translation suggesting much lower speed of elongation. To check this hypothesis, we measured the elongation rates of different mtRNAPs and compared them to T7 RNAP rates. We halted ECs on promoter-containing templates in the presence of a limited NTP set (including radioactive UTP but omitting CTP) allowing for "synchronization" of the ECs and formation of a labeled startup RNA (Fig. 7). The resulting radiolabeled start-up ECs were then chased in the presence of CTP and "cold" UTP. Under these conditions, possible re-initiation events produced unlabeled transcripts ensuring selective detection of the products synthesized only in a single round of transcription (Fig. 7).

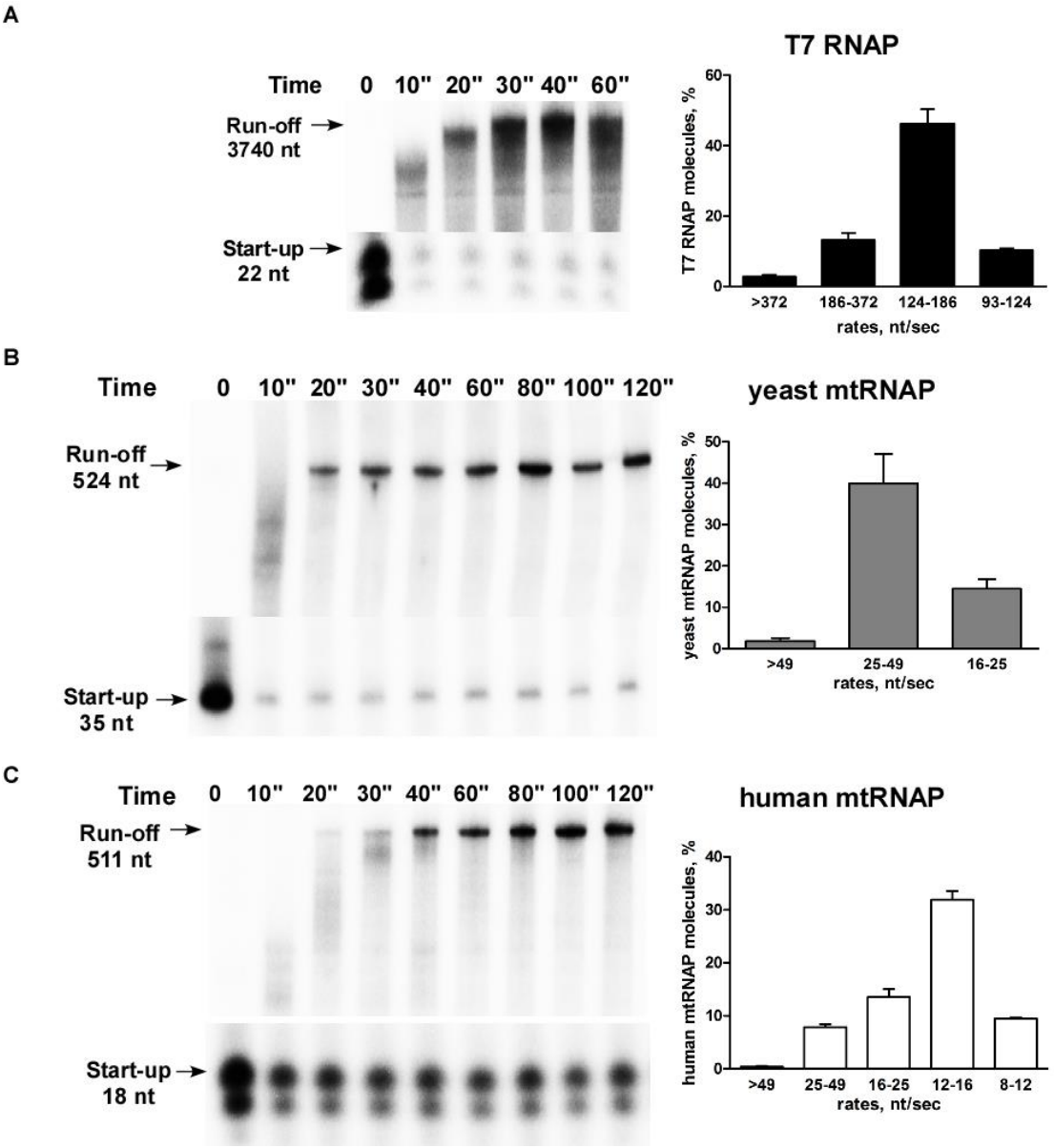

Figure 8. Human mitochondrial RNAP is a slow enzyme.

A. Elongation rates of T7 RNAP. Left panel, the T7 RNAP ECs were halted 22 nt downstream of T7 promoter and chased by CTP. Right, bar histogram depicting the fraction of T7 RNAP molecules at specific elongation rates..

B. The elongation rates of *S. c.* mtRNAP. *S. c.* mtRNAP start-up complexes were halted 35 nt downstream of 14S rRNA promoter and chased as in A.

C. The elongation rates of *H. s.* mtRNAP. *H. s.* mtRNAP start-up complexes were halted 18 nt downstream of LSP and chased as in A

We first measured the elongation rates for T7 RNAP. To assemble T7 RNAP ECs, we used linearized plasmid pT7blue (4 kbp) containing T7 promoter and formed a 22-nt start-up complex (Fig. 8 A). The complexes were chased by incubation with CTP and UTP for the time indicated. The amounts of the run-off products were determined by autoradiography after separation of the products by denaturing PAGE. These values were used as a measure of fraction of molecules that reached the end of template at specific time intervals. We determined the distribution of T7 RNAP molecules exhibiting particular elongation rates (Fig. 8 A). The majority of T7 RNAP molecules exhibited elongation rates in range of 124-186 nt/sec that is in good agreement with previously published data (*121*, *122*). We measured the elongation rates of S.c. mtRNAP using a PCR template containing the mitochondrial 14S rRNA promoter in pT7blue plasmid, by halting the ECs 35 nt downstream of transcription start site (Fig. 8 B). Template with 14S rRNA generated 500 nt run-off product and was prepared similarly to LSP-containing plasmids for human mtRNAPs (see Materials and Methods and Appendix). Similarly to T7 RNAP, we generated bar graph histogram depicting fraction of yeast mtRNAP molecules at

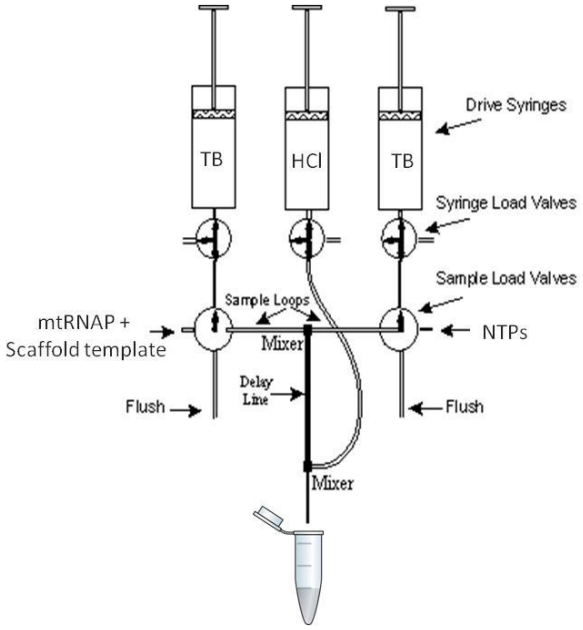

Figure 9. The schematics of Kintek RQF-3 instrument.

particular elongation speeds. The most of S.c. mtRNAPs showed elongation rates in range of 25-49 nt/sec, on average about 5 times lower than that of the majority of T7 RNAP molecules. To measure H.s. mtRNAP elongation rates we used a PCR-amplified template containing LSP to form an 18 nt halted complex (Fig. 8 C). The elongation rates of the majority of H.s. mtRNAP molecules were found to be in range of 12-16 nt/s, which is about 10 times lower than the rates measured for T7 RNAP.

3. Elongation rates of mtRNAP determined by RQF.

To test whether elongation rates of mtRNAP measured on promoter templates by chase experiments (Fig. 8) are in agreement with data obtained in other published assays (*119*), we used millisecond kinetic analysis or rapid chemical quench flow (RQF). The RQF experiments were performed using a KinTek RQF-3 (3-syringe) instrument, which is capable of measuring the rate of extension reaction within 2 ms (Fig. 9). In terms of RQF, the processive nucleotide addition is the interval between

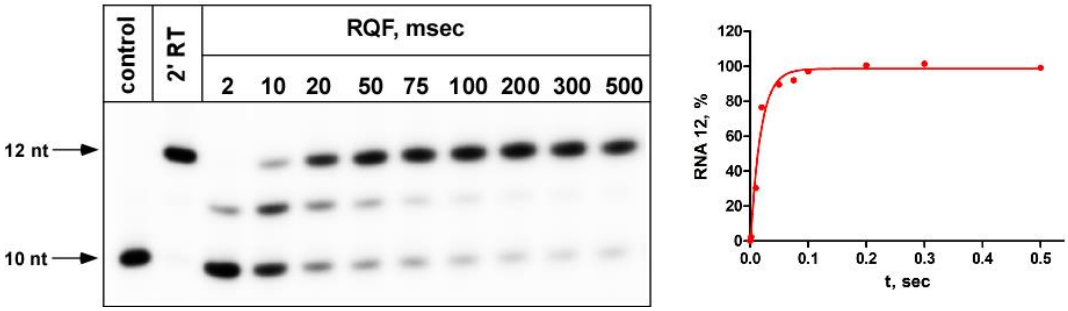

Fig. 10. Measurement of elongation rates of WT human mtRNAP by RQF. R2-8/T1/NT1 scaffold was used for extension. MtRNAP ECs were incubated 10 min at RT before extension. Control – RNA-DNA scaffold, lane 2 - extension by GTP and UTP for 2 min using Kintek RQF-3. Left. Plot of two NTPs extension of WT human mtRNAP determined by exponential curve fitting.

formation of one phosphodiester bond and loading of NTP for the next round of synthesis. During that process, translocation and/or pyrophosphate release can in principle limit the overall elongation rate of the enzyme. Stopping the reaction with 1

M HCl, which instantly inactivates RNAP, allows identification of the kinetics of phosphodiester bond formation after translocation has already occurred (*119*). ECs (0.5 µM) were assembled using the R2-8/T1/NT1 RNA-DNA scaffold (see Fig. 11) and extended with GTP. Extension of RNA by mtRNAP in the RQF assay followed exponential kinetics (Fig. 10). Based on exponential curve fitting, we determined the rate constant (units of s^{-1}) of mtRNAP under these conditions. We found that apparent rates of mtRNAP rates were 56.6 s^{-1} (Fig. 10), reflecting fast catalysis of enzyme in the absence of rate-limiting translocation step.

In conclusion, we determined that the elongation rates of mtRNAPs are significantly lower than the ones for T7 RNAP, likely reflecting the coordination of mitochondrial gene expression in mitochondria that is not required during phage expression.

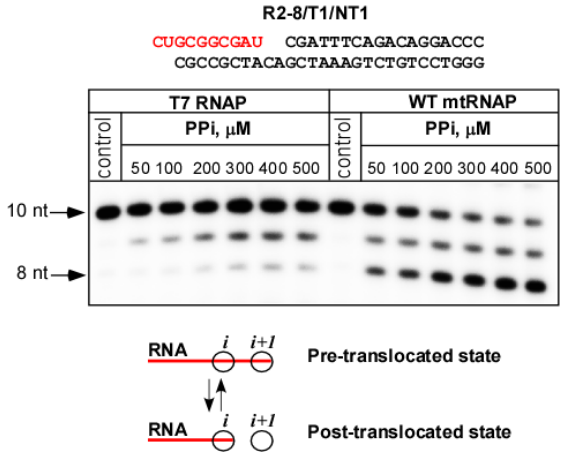

Figure 11. Pyrophosphorolytic assay of T7 and mtRNAP ECs. T7 and mtRNAP ECs were incubated with the increasing concentrations of PPi.

4. Translocation states of mtRNAP ECs.

The mechanisms of substrate incorporation and translocation of RNAPs along DNA have been studied extensively using T7 RNAP ECs (*87, 89, 123*). To examine the conformational state of mtRNAP in the assembled ECs, we performed pyrophosphorolytic assays.

Pyrophosphate (PPi) binds only to ECs having RNAP in a pre-translocated conformation and drives a reaction that is the reverse of RNA synthesis. We assembled ECs of T7 RNAP and mtRNAP on a scaffold having an 8 bps RNA-DNA hybrid and incubated the complexes with increasing concentrations of pyrophosphate (Fig. 11). We found, that while T7 ECs were more resistant to PPi, RNA in mtRNAP ECs was readily shortened to 8 nts, suggesting pyrophosphorolytic activity. We therefore conclude, that mtRNAP ECs, in contrast to T7 RNAP, were predominantly in the pre-translocated conformation (*88, 123*).

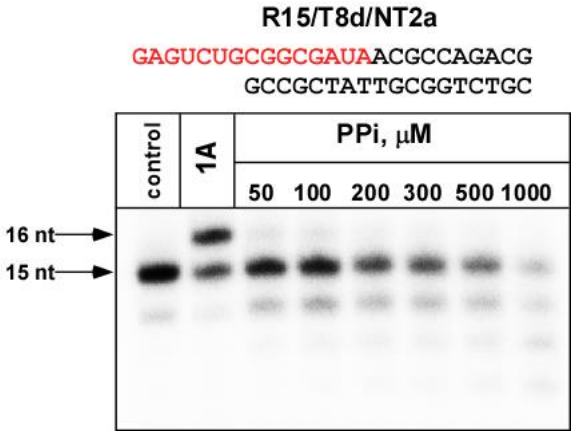

Figure 12. Post-translocated mtRNAP ECs. MtRNAP ECs were incubated in the presence of PPi for 30 min at 35 °C.

Studies of bacterial RNAP have determined that resistance of ECs to PPi treatment depends on the nature of the 3' nucleotide of RNA favoring the pre-translocated state in the order U>C>A>G (*124, 125*). We checked if mtRNAP can also form a post-translocated EC if the sequence of the 3' end nucleotides of RNA is different. Indeed, when we assembled mtRNAP ECs on an RNA-DNA scaffold containing RNA with "–UA" at the 3' end, we observed that the complexes were PPi resistant and, therefore, likely in the post-translocated state (Fig. 12). The effect of different RNA sequences on PPi-sensitivity of mtRNAP ECs is summarized in Table I (appendix). We conclude, that, similarly to the bacterial RNAPs, mtRNAP ECs can be assembled in a defined translocation state depending on RNA sequence. However,

hybrid length and 5'-tail length of RNAs in scaffolds may also reflect the transitions between tranlsocated states of polymerase.

5. What affects elongation rates in the EC?

During the nucleotide addition cycle (Fig. 5), the overall rate of elongation can be affected at several stages of the reaction. First, binding and selection of the incoming substrate in the pre-insertion and insertion sites may serve as a rate-limiting step. (*87*). Second, catalysis can directly affect the elongation rates. Mutations of the catalytic residues of T7 RNAP are known to decrease the elongation rates of this enzyme (*126*). Finally, translocation of mtRNAP along DNA template can be the rate-limiting event. Transitions between pre- and post-translocated states play a significant role in rates of nucleotide addition. Stabilization of the pre-translocated state is a prerequisite for transcriptional pausing, whereas RNAP in the post-translocated state can bind incoming NTP and initiate new cycles of NTP incorporation (*88, 91, 127, 128*). Studies of mtRNAP single nucleotide addition cycles have demonstrated that chemistry as well as conformational rearrangements before and after catalysis (i.e. translocation of mtRNAP along the DNA template) represent co-rate-limiting steps during RNA synthesis in transcription elongation (*99, 100*). We therefore employed mutational analysis to examine what structural elements in mtRNAP, as compared to T7 RNAP, might affect elongation dynamics and characteristics of the ECs.

6. Mutational analysis of structural elements involved in catalysis.

Based on previous studies (*99*), we hypothesized that the differences observed in the elongation rates between T7 and mtRNAP could be due to structural elements involved either in catalysis or translocation. The available structures of T7 RNAP and mtRNAP ECs (*86, 118*) provided an opportunity for a detailed comparative analysis of the major elements of RNAPs involved in catalysis and translocation.

The C-terminal domains of both polymerases are highly homologous (*118*). We first analyzed the differences

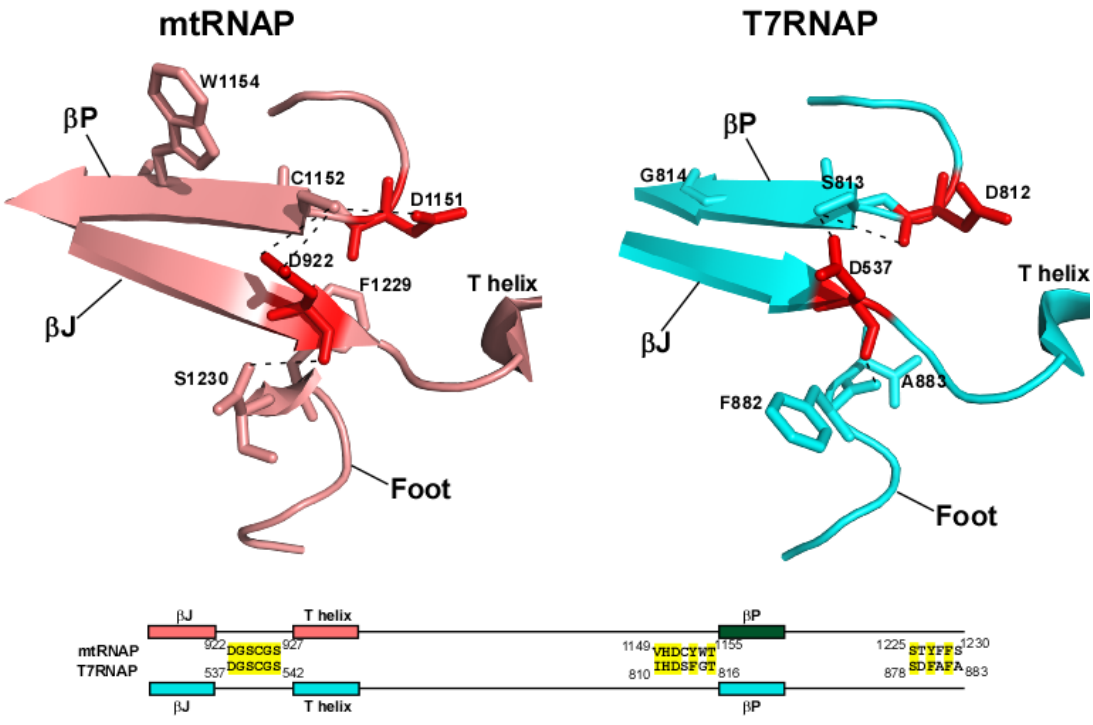

Figure 13. The sequence and structure analysis of catalytic centers of T7 and mtRNAPs. Top. Structural composition of catalytic centers of T7 and mtRNAPs. MtRNAP elements are shown in salmon, T7 RNAP – in cyan. Catalytic Asp are shown in red. The residues involved in hydrogen bond interactions are shown by sticks. Bottom. Sequence alignment of shown structural elements. Homologous residues are highlighted in yellow.

between T7 and mtRNAP in the vicinity of the active sites, within a distance of 10Å from the catalytic aspartate residues. The active sites of these RNAPs are composed

of two β-strands (βP and βJ) carrying catalytic aspartates, a part of the O helix containing residues involved in substrate binding and selection, and a C-terminal

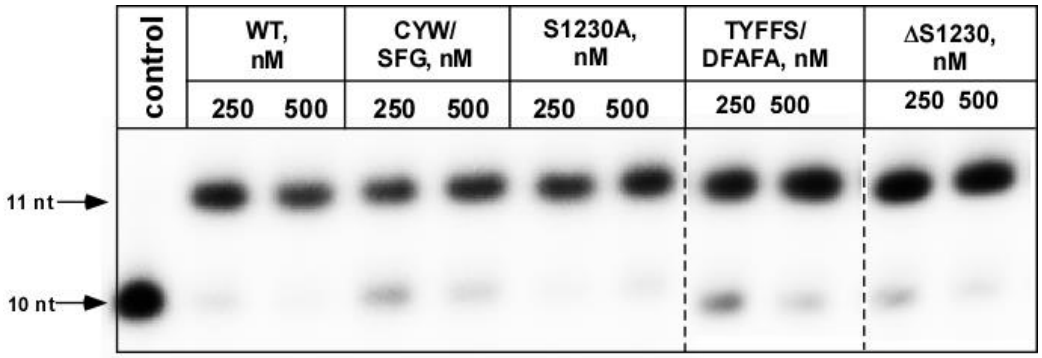

Figure 14. Primer extension assay using the catalytic center mtRNAP variants. MtRNAP ECs were assembled on linear scaffolds R2-8/T1/NT1. The complexes were pre-incubated for 5 min at RT before extension by GTP. Extension was carried out for 7-10 min at RT.

element called the "foot" (Fig. 13). The residues that compose these structural elements display even higher homology level as most of the them are identical (*118*). This includes two catalytic aspartates D1151 and D922, and residues involved in substrate binding and selection such as K853 (K472 in T7 RNAP), Y956 (Y571), R987 (R627), K991 (K631), Y999 (Y639) (*87*, *118*). There are only two small regions that are different between these two RNAPs in the active site. (Differences between the O/Y helices contributing to the active site structure are discussed separately). They involve residues

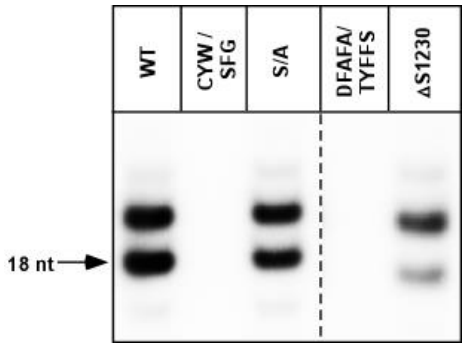

Figure 15. Promoter-dependent initiation assay of catalytic center mtRNAP variants. Reactions were performed in the presence of mtRNAP and initiation factors.

of the βP strand and the foot loop of mtRNAP (Fig. 13). Thus, while the residues adjacent to the catalytic D922 (D537) are identical in both RNAPs (Fig. 13), residues of the other βP strand that flank the catalytic D1151 (D812) are less conserved

between the two RNAPs. The SH- group of C1152 in mtRNAP (S813 in T7) is in position to make a hydrogen bond with the peptide backbone oxygen of the catalytic D922 residue and with the carboxyl group of D1151. In the same region, the hydrophobic residue W1154 in mtRNAP is changed to G815 in T7 RNAP (Fig. 13). Finally, the sequence at the end of the foot loop that projects into the active site of RNAP contains residues that also differ notably between T7 and mtRNAP and serve as a characteristic signature for these RNAPs - "YFFS" in mtRNAPs vs "FAFA" in T3/T7 family of phage RNAPs. Hydrophobic F1229 in mtRNAP is substituted to non-polar A881 in T7 RNAP. The C-terminal S1230 in mtRNAP is placed in the same orientation relative to the active

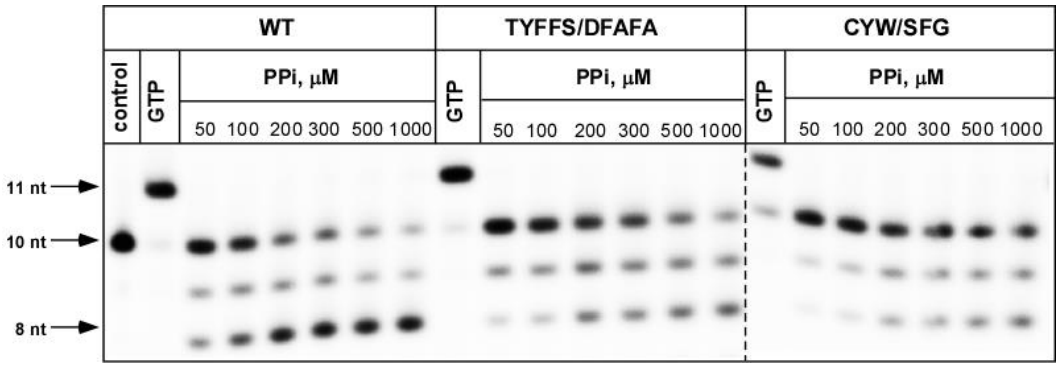

Figure 16. Pyrophosphorolytic activity of catalytic center mtRNAP variants. MtRNAP ECs (R2-8/T1/NT1 – scaffold template) were incubated with indicated concentrations of PPi for 30 min.

center as F882 of the T7 enzyme, which results in formation of an additional hydrogen bond with the peptide backbone oxygen of D922 (Fig. 13). Numerous hydrophobic aromatic amino acids stabilize the active center of mtRNAP by stacking interactions such as H920-W1154, Y1153-F1229, F1228-F1229, absent in T7 RNAP.

To address how the described differences affect elongation of mtRNAP, we constructed mutants containing substitutions of mtRNAP residues in the vicinity of

the active site to the corresponding residues in T7 RNAP. These included: substitution of residues 1152-1154 in βP strand (CYW to SFG), substitution of residues 1225-1230 in the foot loop (TYFFS to DFAFA), S1230A and ΔS1230 (Table II, appendix).

We first examined the activity of mtRNAPs variants in a primer extension assay on an RNA-DNA scaffold template R2-8/T1/NT1. All mtRNAP variants were able to extend the RNA primer suggesting that their catalytic activity in this assay was not perturbed (Fig. 14, Table II, appendix). To analyze the ability of mtRNAP

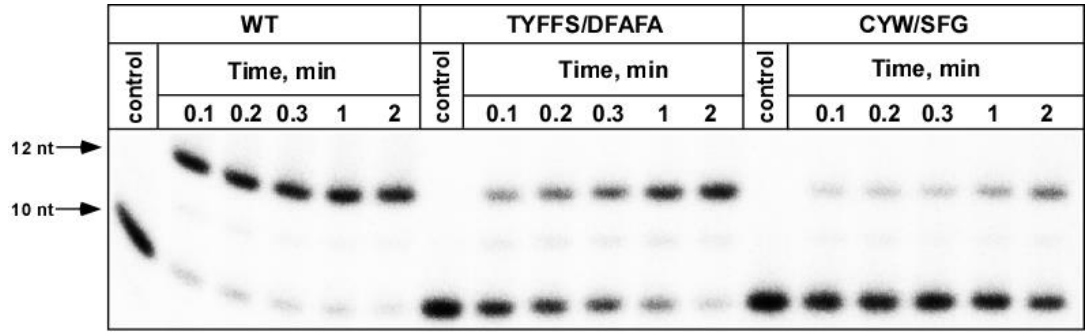

Figure 17. Primer extension assay to measure elongation rates of WT, TYFFS/DFAFA, CYW/SFG mtRNAPs at RT. R2-8/T1/NT1 was used as a template. MtRNAP ECs were incubated 10 min at RT before extension to 2 NTPs at different time intervals.

variants to initiate transcription, we used an *in vitro* transcription initiation assay on an LSP promoter template. Only S1230A mtRNAP was fully active in promoter-dependent transcription, whereas CYW/SFG and TYFFS/DFAFA mtRNAPs could not initiate transcription and ΔS1230 mtRNAP was partially defective (Fig. 15). This suggested that the catalytic center residues and the foot element of mtRNAP may be important for stabilization of RNAP-promoter interactions and initiation of transcription.

In order to test if the conformation of mtRNAP was affected in the mutants, we took advantage of the pyrophosphorolysis assay using the R2-8/T1/NT1 scaffold, on which WT mtRNAP assembles in a pre-translocated state. PPi-induced cleavage was assessed by accumulation of the 8-mer RNA, a product of the pyrophosphorolysis reaction. We found that mtRNAP variants with substitutions in the catalytic center showed variability in their sensitivity to pyrophosphate (PPi). Similar to the WT mtRNAP ECs, ECs formed with ΔS1230 and S1230A variants were PPi-sensitive (Table II, appendix) and, thus, likely in the pre-translocated state. TYFFS/DFAFA mtRNAP ECs possessed intermediate sensitivity to PPi, whereas CYW/SFG was significantly more PPi-resistant than the WT mtRNAP, and, therefore, mostly in the post-translocated state (Fig. 16).

7. Elongation rates of mtRNAP variants having substitutions in the catalytic center.

Stabilization of the post-translocated states in the EC has been suggested to increase the elongation rates of mtRNAP (*92, 124*). We therefore examined the rates of mtRNAP variants TYFFS/DFAFA (1225-1230) and CYW/SFG (1152-1154), which showed decreased pyrophosphorolysis and were mostly in the post-translocated state. To measure elongation rates we used a primer extension assay in which the RNA primer was extended by 2 nucleotides with UTP and GTP (Fig. 17). While WT mtRNAP extended the RNA within 10 sec, at least 1 min was required for primer extension by TYFFS/DFAFA mtRNAP (Fig. 17). The CYW/SFG mtRNAP required more than 4 min to complete primer extension. Thus, elongation rates of

these mtRNAP variants were very slow, lower than 1 nt/sec. We also analyzed these mutants using RQF assays in experiments involving primer extension by a single nucleotide (data not shown). Similar to the experiments above, both mutants were considerably slower, than the WT mtRNAP. This might reflect defects in the catalytic functions of these mtRNAPs that are more profound on a shorter time scale as 10 min extension of enzymes did not exhibit any abnormalities (Fig. 14).

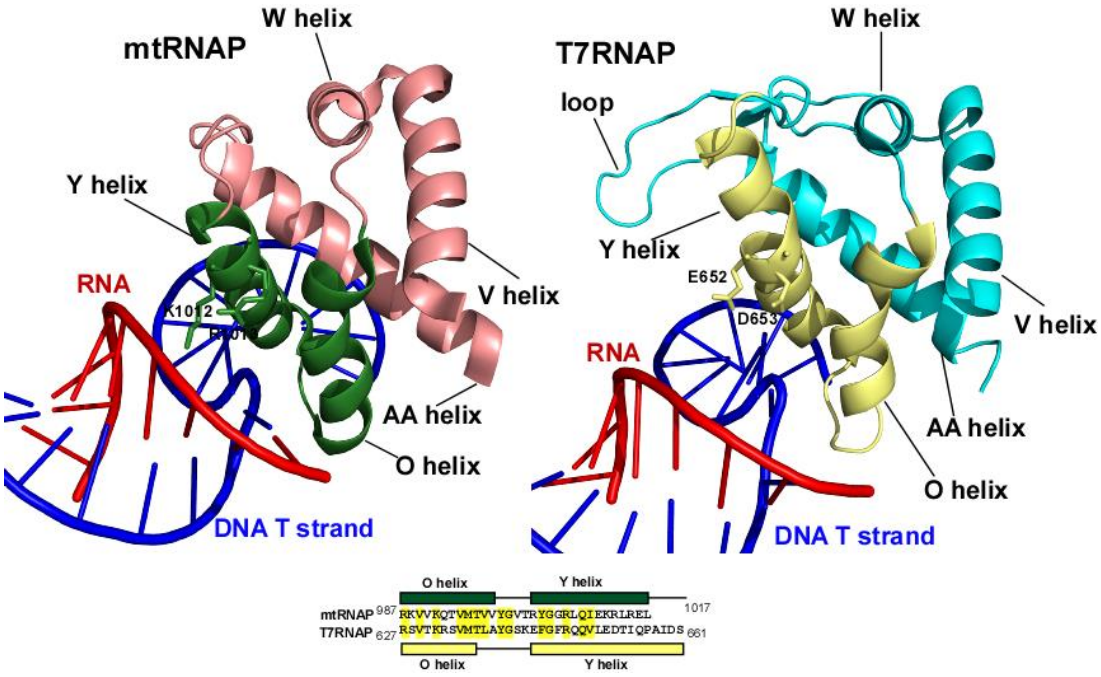

Figure 18. The sequence and structure of fingers domain of T7 and mtRNAP. Top. Structural comparison of fingers domain of T7 and mtRNAP. O/Y helices are shown by green and light orange in mtRNAP and T7 RNAP, respectively. Other helices are shown in salmon and cyan for mtRNAP and T7 RNAP, respectively. Residues interacting with RNA are shown by sticks. Bottom. Sequence alignment of O/Y helices of T7 and mtRNAP. Homologous residues are highlighted in yellow.

8. Mutational analysis of structural elements involved in translocation.

Translocation is an another step during transcription elongation that was found to be rate-limiting (*99*). Each step during the NAC is accompanied by movement of the O/Y helices in the 6-helix bundle, called fingers domain in T7 RNAP (*88, 91*). According to the EC structure, the fingers domain of mtRNAP may also be implicated in translocation (*118*). We compared the sequence and structure of the fingers domains in two RNAPs (Fig. 18). MtRNAP contains numerous amino acid substitutions in this region and lacks the "fingers insertion" loop of T7 between helices V and W (Fig. 18). Strikingly, the crystal structure of mtRNAP EC reveals that positively charged residues K1012 and R1013, conserved in all mammalian mtRNAPs, are in position to make hydrogen bonds with the RNA phosphate backbone in the pre-translocated state (*118*). While no post-translocated structure of mtRNAP is yet available, movement of the Y helix (and therefore loss of h-bonding with RNA) will be required for mtRNAP to incorporate

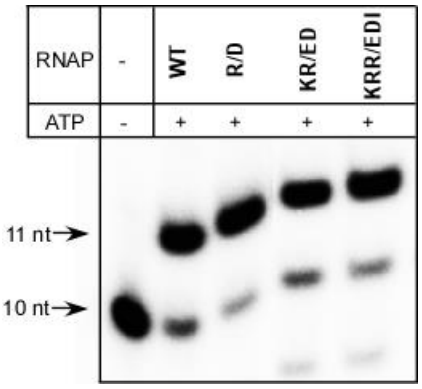

Figure 19. Primer extension assay of fingers domain mtRNAP variants. MtRNAP ECs were assembled on linear scaffolds R2-8/T1/NT1 as in Fig. 14.

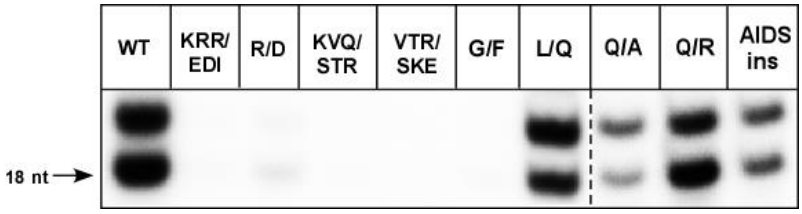

Figure 20 Initiation assay of mtRNAP variants on LSP template. Reactions were performed in the presence of mtRNAP and initiation factors TFAM and TFB2M for 30 min at 35 °C. 18 nt RNA is observed with limited NTP sets.

substrate NTPs. In all phage RNAPs the corresponding residues are negatively charged (E652 and D653 in T7 RNAP). In addition, we noted that the Y-helix of mtRNAP is one turn shorter than the corresponding structural element in T7 RNAP. We speculate that stabilization of the pre-translocated state of mtRNAP can have an effect on its elongation rates. We therefore constructed a series of mtRNAP variants by substituting residues in O/Y helices with the corresponding residues of T7 RNAP (Table III, appendix). This includes substitution of the positive residues of the Y helix with the negative residues of T7 RNAP (K1012E/R1012D/R1015I, K1012E/R1013D, R1013D), substitution of the entire fingers domain, and substitution of three-amino acids clusters with in O/Y helices (K988S/V990T/Q992R, VTR(1001-1003)/SKE), and insertion of 4 amino acids into the Y helix (insAIDS1017-1021).

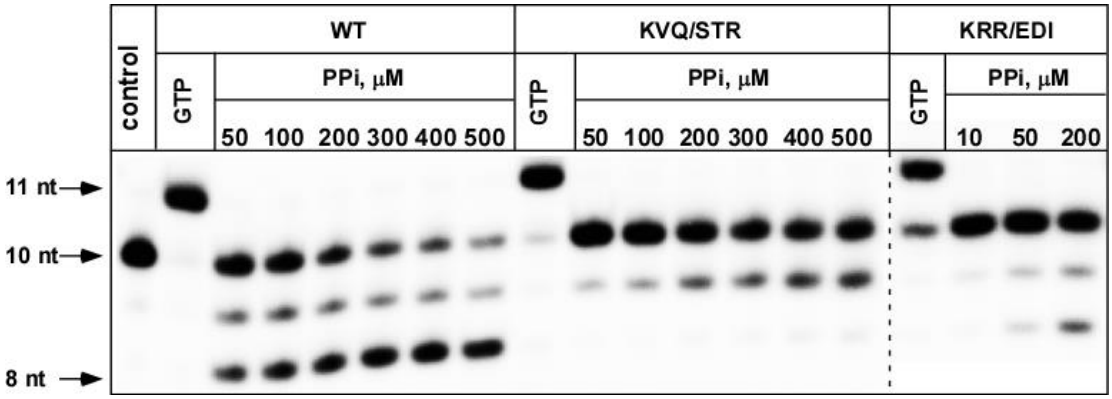

Figure 21. Pyrophosphorolytic activity of mtRNAP variants with substitutions in O/Y helices. MtRNAP ECs were incubated with corresponding concentrations of PPi for 30 min at RT as described in Materials and Methods. Single nucleotide extension of R2-8/T1/NT1 by GTP serves as control for activity.

To examine the catalytic activity of the mtRNAP variants, we used a primer extension assay. All mtRNAP variants were able to extend RNA, suggesting that the ability to extend RNA by one nucleotide was preserved (Fig. 19 and Table III,

appendix). To determine the ability of mtRNAP variants to initiate transcription, we analyzed them in *in vitro* transcription assay on LSP promoter templates. Interestingly, most of mtRNAP variants were defective in initiation and could not

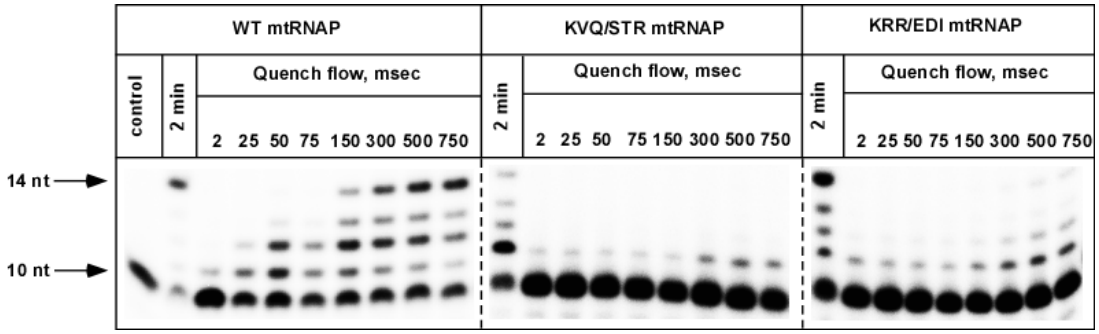

Figure 22. RQF assays to measure elongation rates of mtRNAP variants with substitutions in O/Y helices. R2-8/T1/NT1 was used as a template. MtRNAP ECs were incubated 10 min at RT before extension to 4 NTPs. Control – extension to 4 NTP for 2 min using Kintek RQF-3.

support transcription (Fig. 20). Only two single mutants – L1008Q and Q992R mtRNAP – were able to initiate transcription from LSP (Fig. 20). This suggests the importance of O/Y helices for stabilization of RNAP-promoter interactions.

To probe the conformational state of mtRNAP variants in the ECs, we utilized the PPi assay. Substitution of the positively charged residues in the Y helix (K1012E/R1012D/R1015I) resulted in mtRNAPs variants more resistant to PPi (Fig. 21), suggesting that these residues may, indeed, be important for stabilization of the pre-translocated conformation of the enzyme, as was observed in the crystal structure of mtRNAP (*118*). We also found that mutation of residues 989, 991, 993 (KVQ to

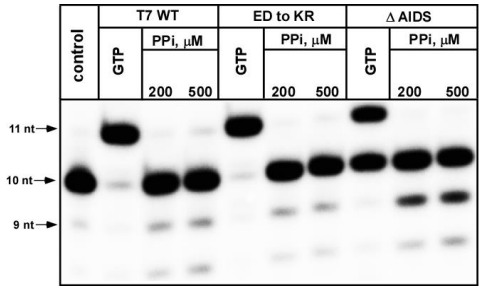

Figure 23. Pyrophosphorolyitc activity of T7 RNAP variants. T7 RNAP ECs were incubated with corresponding concentrations of PPi for 30 min at 37 °C.

STR) in the O-helix renders mtRNAP EC more resistant to PPi. Other mtRNAP variants that included substitutions of Q992R, VTR/SKE (residues 1001-1003), L1008Q and insertion of 1017-1021 were PPi-sensitive, similar to WT mtRNAP, and, thus, mostly in the pre-translocated conformation (Table III, appendix).

9. Elongation rates of mtRNAP variants having substitutions in O/Y helices.

Using the PPi assay we identified several mtRNAP variants with substitutions in O/Y helices that formed PPi-resistant ECs, suggesting that these mutations stabilized the post-translocated state of mtRNAP. These mutations included KVQ/STR (989, 991, 993) and KRR/EDI (1012-1013, 1015). Since the post-translocated state is required for the translocation cycle and increased elongation rates (*92, 124*), we asked whether these substitutions would

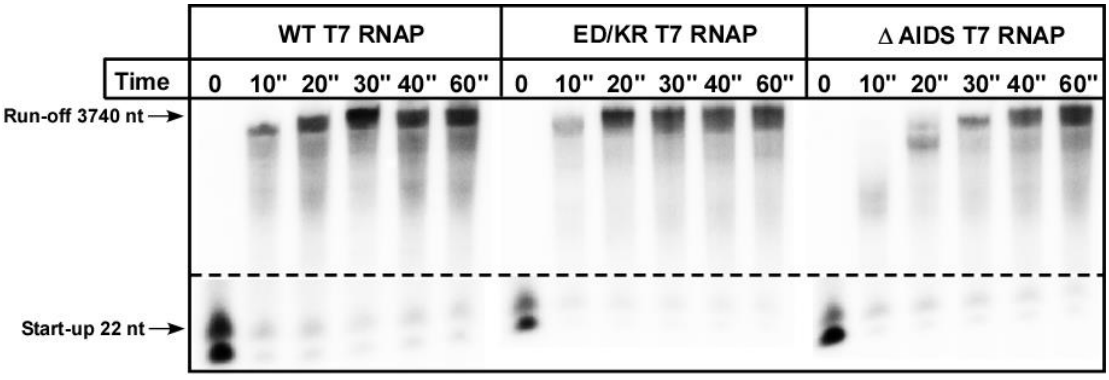

Figure 24. The elongation rates of WT T7, ED/KR, ΔAIDS RNAPs. The T7 RNAP start-up complexes were assembled 22 nt downstream of T7 promoter and chased by complete NTP set and the run-off products were resolved on 20% PAGE-urea gel.

affect elongation rates. Because these mtRNAP variants were defective in transcription initiation, and unlike mutants of the active site of mtRNAP, had no elongation defects in a scale of seconds, we measured elongation rates using the RQF protocol (Fig. 9). RQF experiments were performed as described in Materials and Methods and section 1.3 (Fig. 9). To detect the effect of translocation on elongation rates, we monitored the incorporation of 4 nucleotides on the R2-8/T1/NT1 scaffold. The elongation rates of KRR/EDI (1012-13, 1015) and KVQ/STR (989, 991, 993) mtRNAP variants were extremely slow, lower than 1 nt/sec (Fig. 22).

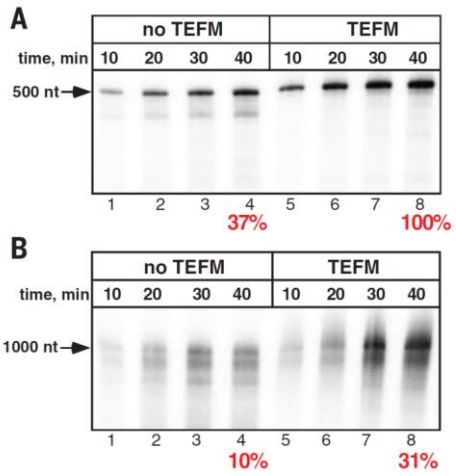

Figure 25. TEFM increases the processivity of mtRNAP ECs. Transcription was performed using LSP templates lacking CSBII sequence to generate 500-nt (A) or 1000-nt (B) runoff products. Efficiency of synthesis of 500 nt of RNA in the presence of TEFM is taken as 100%.

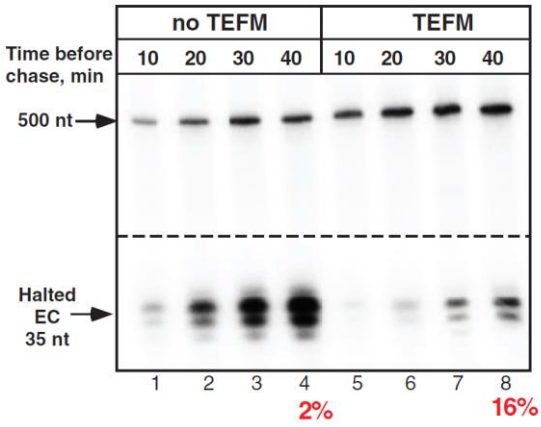

Figure 26. TEFM increases the stability of mtRNAP ECs. Complexes halted at +35 in the presence or absence of TEFM were incubated for the times indicated and chased with CTP. Numbers below indicate efficiency of a halted EC extension (%).

These results may indicate that stability of ECs of these mtRNAPs was affected and, therefore, the enzymes could not efficiently incorporate nucleotides. Overall, mtRNAP variants with substitutions in O/Y helices displayed low elongation rates, reflecting the importance of O/Y helices in the stabilization of mtRNAP EC during

translocation cycle.

10. Mutational analysis of the Fingers domain of T7 RNAP.

In a complementary approach, we performed mutagenic analysis of the T7 RNAP fingers domain. We substituted the conserved E652 and D653 residues of the Y helix of T7 RNAP to the positively charged (K1012 and R1013) residues found in mtRNAP. The Y-helix of T7 RNAP is one turn longer than the corresponding helix in mtRNAP. To test whether the length of the Y helix affects elongation by T7 RNAP, we shortened it by 4 amino acids (ΔAIDS, residues 638-641). All mutants were active, as evident from their activity in primer extension (Fig 23) and transcription initiation assays (Fig. 24). In order to define the conformational states of T7 RNAP variants, we performed PPi-sensitivity

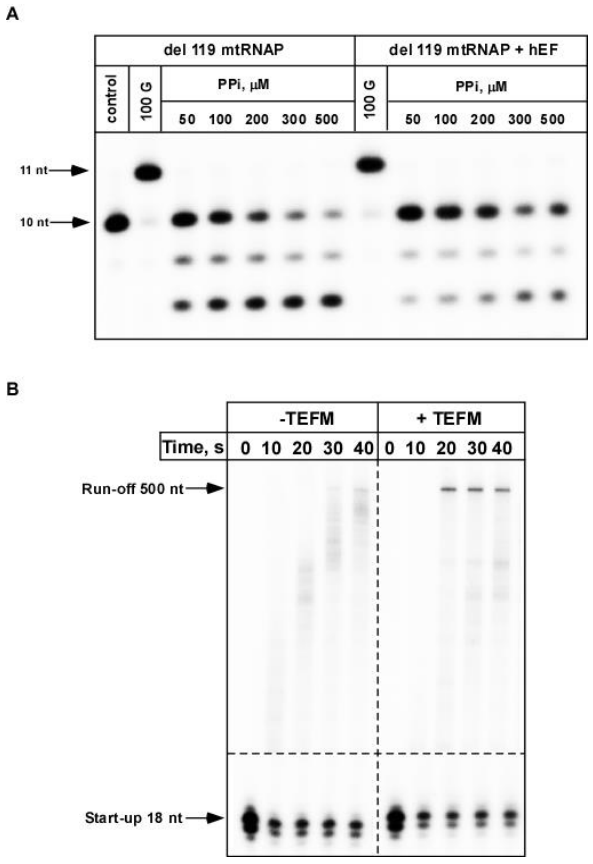

Figure 27. TEFM increases the elongation rates of mtRNAP. A. Pyrophosphorolytic assay of mtRNAP ECs in the presence and absence of TEFM. ECs were incubated with PPi 30 min at 33 °C. **B. Elongation rates of mtRNAP in the presence and absence of TEFM.** MtRNAP start-up complexes were assembled 18 nt downstream of LSP and chased by complete NTP set and the run-off products were resolved on 20% PAGE-urea gel.

assays. We found that similar to WT, the E652K/D653R T7 RNAP was more PPi-resistant, whereas ΔAIDS T7 RNAP showed an sensitivity to PPi, as evident by accumulation of a 9-mer RNA cleavage product (Fig. 23).

As both T7 RNAP variants retained transcription initiation activity, we analyzed their elongation rates using the halted complex chase assay. We found that while the rates for the ED/KR mutant were similar to the WT T7 RNAP, ΔAIDS mutant exhibited approximately three times lower speed of elongation (Fig. 24). This is consistent with the observation that this mutant T7 RNAP forms ECs that are more PPi-sensitive and can be found in pre-translocated conformation (Fig. 23).

Chapter 2. Regulation of mtRNAP elongation by elongation factor TEFM.

1. Activities of TEFM.

1.1. TFEM increases processivity of mtRNAP.

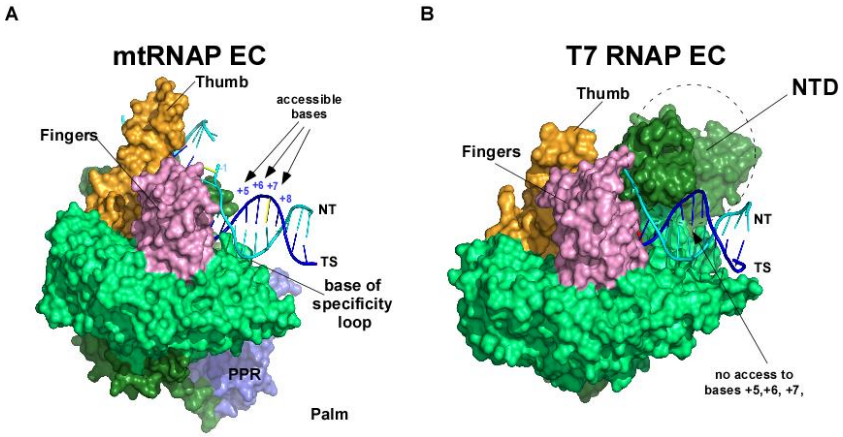

Figure 28. Downstream DNA region in mtRNAP EC is accessible for interactions with TEFM. (A and B) Elongation complexes of mtRNAP (PDB ID 3BOC) and T7 RNAP (PDB ID 1SVO) are aligned with respect to their conserved palm subdomains.

As a general elongation factor, TEFM should increase the processivity of mtRNAP to generate long transcripts. To test that possibility, we performed *in vitro* transcription assay using promoter templates having different lengths of transcribed region. We found that TEFM had no dramatic effect on the efficiency of transcription of a 500 nt RNA by mtRNAP (Fig. 25 A). However, when a longer template (1000 nt) was transcribed, mtRNAP could not efficiently complete the synthesis of RNA (Fig. 25 B). The processivity of mtRNAP was therefore substantially lower than that of T7 RNAP, which is known to transcribe large (up to 10,000) fragments of DNA (*121, 122*). Such a low processivity of mtRNAP would not be sufficient to transcribe the 16,000bp-long mtDNA. In the presence of TEFM, the efficiency of synthesis of long RNA was significantly

increased (Fig. 25 A and B), confirming that TEFM acts as a processivity factor, as originally proposed (*101*).

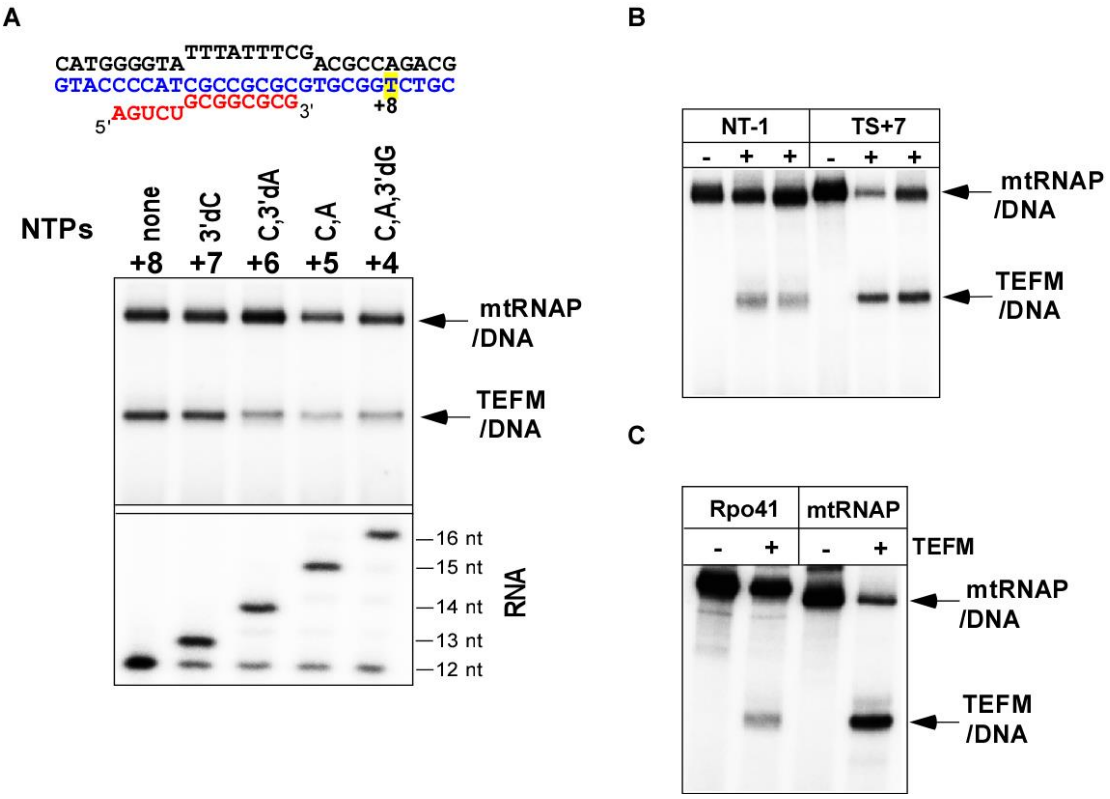

Figure 29. TEFM interacts with the downstream of DNA in mtRNAP ECs. A Template strand DNA-TEFM cross-linking. DNA-protein cross-linking of ECs using scaffold containing 4-thio dTMP at position +7 of DNA template strand. MtRNAP was walked along the template by incorporation of corresponding NTPs (see the text). Below is the 20 % PAAG gel showing RNA extension. Crosslinking was carried out, and crosslinked species are shown above. **B. TEFM cross-link to non-template strand of DNA**. The ECs were assembled using scaffold containing 4-thio dTMP at position +7 (template strand, lanes 4-5) or -1 (non-template strand, lanes 1-3) of DNA. Different concentrations of mtRNAP were tested in this experiment. **C. Specificity of TEFM-DNA cross-link.** ECs were assembled using either yeast mtRNAP (RPO41, lanes 1-2) or human mtRNAP (lanes 3-4) using scaffold containing 4-thio dTMP at position +7 in the template strand. + indicates addition of TEFM to the reaction.

1.2. TEFM increases the stability of the elongation complex.

Transcription elongation factors are known to increase processivity of RNAPs by stabilizing ECs. We therefore probed whether TEFM contributes to the stability of

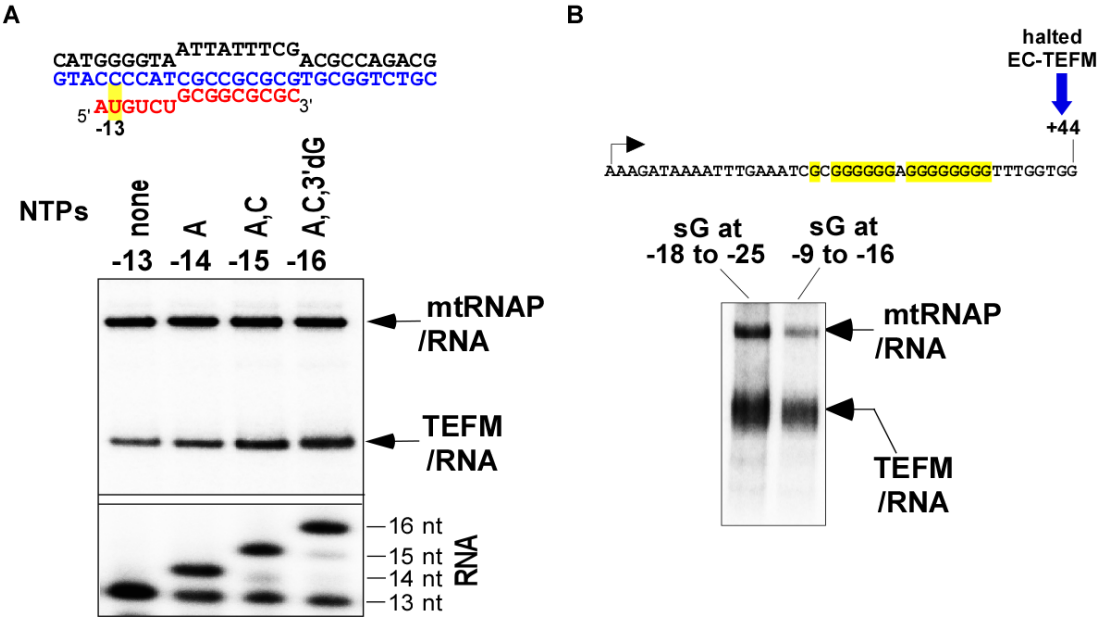

Figure 30. TEFM interacts with RNA in mtRNAP ECs. A. TEFM-RNA cross-linking on scaffold template. RNA-protein cross-linking of ECs using scaffold containing 4-thio dUMP at position -13 of RNA. MtRNAP walked along the template by incorporation of corresponding NTPs. Below is the 20 % PAAG gel showing RNA extension. **B. TEFM-RNA cross-linking on promoter template.** The EC44 was obtained by walking of mtRNAP along LSP containing template on Ni-beads. The complex was UV-irradiated, eluted from Ni-agarose with 0.25 mM Imidazol and analyzed using 4-12% SDS-PAGE followed by autoradiography.

the elongating mtRNAP. To determine this, we halted the EC 35 bp downstream of the promoter start site by omitting CTP in the presence or absence of TEFM (Fig. 26). After incubation for 10-40 min, the complexes were chased to allow synthesis of a runoff product. In the absence of TEFM, we observed accumulation of a 35-mer RNA product that could not be efficiently chased and therefore generated only small amounts of run-off products. This suggests that ECs halted in the absence of TEFM, were unstable, rapidly dissociated and re-initiated, and, as a result, were not efficiently chased (Fig. 26). In addition, ECs incubated did not retain 35-mer RNA product and exhibited increased turnover of RNA (2 times) over DNA template (data

not shown). In contrast, in the presence of TEFM, only small amount of 35-nt RNA product accumulated, indicating that, under these conditions, the ECs were stable and were efficiently chased. We therefore conclude that TEFM increases the stability of mtRNAP ECs.

1.3. TEFM increases the elongation rates of mtRNAP.

It has been suggested that transcription elongation factors are responsible for stabilization of the post-translocated states of various RNAPs (*127-129*). To test the effect of TEFM on mtRNAP conformation we performed PPi-sensitivity assays. We found that while the nascent RNA was readily cleaved by mtRNAP ECs, less cleavage was observed when TEFM was added to the reaction. This may indicate that TEFM stabilizes predominantly the post-translocated conformation of mtRNAP in the ECs (Fig. 27 A).

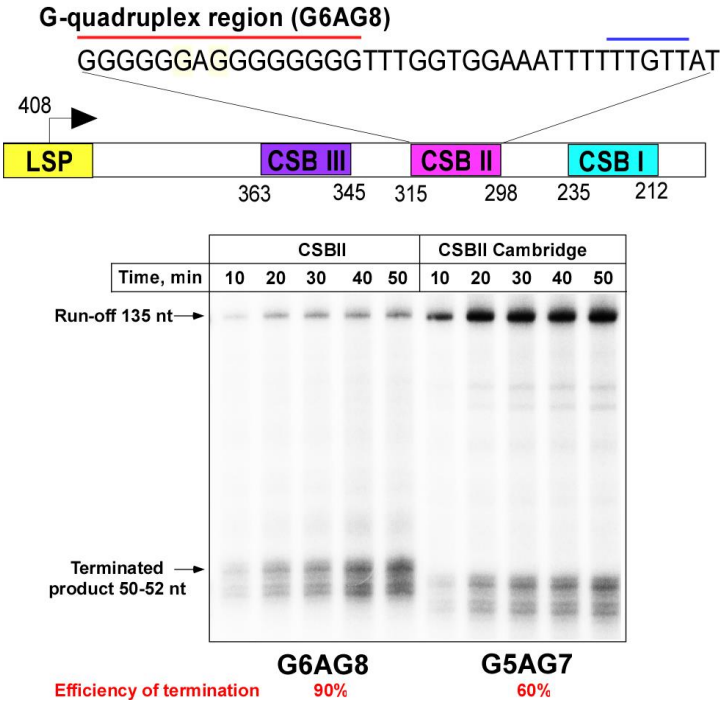

Figure 31. Transcription termination at CSBII by mtRNAP. Top. Schematic illustration of the human mtDNA region downstream of the LSP promoter. The numbers below the scheme correspond to the reference human mtDNA; the transcribed sequence of CSBII is shown at the top. Bottom. Transcription termination of mtRNAP using templates with polymorphic CSBII. Termination efficiency (%) indicated below the panel.

To investigate if the observed facilitation of post-translocated state by TEFM also led to an increase in elongation rates, we used halted EC chase experiments (Fig. 7). We found, that in the presence of TEFM, the elongation rates of mtRNAP on promoter templates were significantly increased and estimated to be approximately 40 nt/sec (Fig. 27 B). This indicates that TEFM may facilitate expression of mitochondrial genome by speeding up mtRNAP transcription.

1.4. TEFM interacts with DNA in the EC.

The catalytic domains of mtRNAP and T7 RNAP share high sequence and structural homology (*118*). However, there appears to be one striking difference regarding the overall EC organization – mtRNAP makes very few contacts with the downstream DNA duplex (Fig 28). In contrast, extensive interactions between the entire downstream duplex and N-terminus of T7 RNAP are observed (Fig. 28). These extensive interactions with the downstream DNA can contribute to the high processivity of T7 RNAP (*93, 121*). Because TEFM interacts with the C-terminus of mtRNAP (*101*), we hypothesized

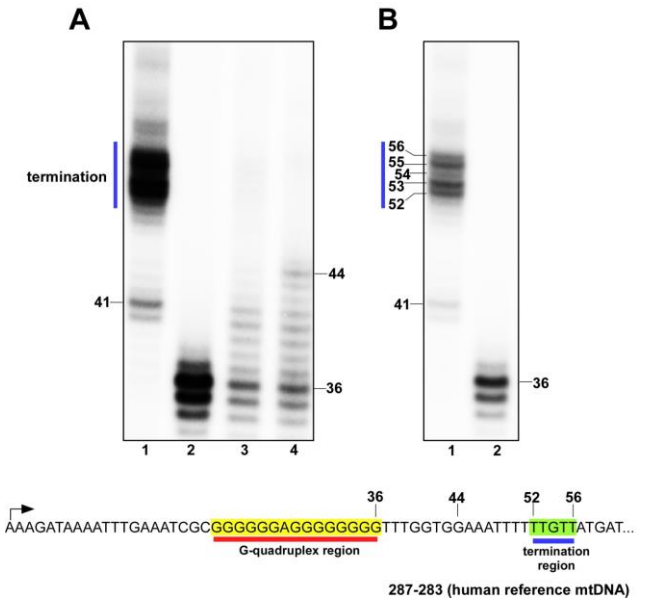

Figure 32. Mapping of the termination site at CSBII. A. Transcription reactions were performed using Δ71CSBII template (transcribed sequence is indicated below the panel). In the absence of TEFM, mtRNAP efficiently terminates at CSBII (lane 1). RNA markers (36 nt and 44 nt, lanes 2-4) were prepared by walking mtRNAP using Ni-agarose. B. Lighter exposure of the autoradiogram shown in A (lane 1 and 2 only).

that it may also bind the downstream DNA to compensate for the lack of mtRNAP-DNA interactions. To evaluate this, we tested the ability of photo-reactive probes introduced in downstream DNA to photo cross-link to TEFM in EC. We incorporated 4-thio-2'-deoxythymidine-5'-monophosphate (4-thio dTMP) into the DNA template in the downstream region and assembled the EC using RNA-DNA scaffolds. We found that TEFM efficiently cross-linked to the +7 and +8 bases in the DNA template strand (Fig. 29 A). In a control experiment, we used a scaffold EC in which the photo reactive probe was placed at position -1 in the non-template strand of DNA. While crosslinking to the RNAP was still observed in this case, no efficient DNA-TEFM crosslink was detected, indicating specificity of TEFM interactions with the downstream DNA region. We also found that the DNA-TEFM cross-linking was species-specific, as no efficient cross-linking was detected when yeast mtRNAP was used to assemble the ECs (Fig. 29 B). Thus, TEFM interacts with downstream of DNA duplex, which likely results in stabilization of the EC and contributes to mtRNAP processivity.

1.5. TEFM interacts with RNA in the EC.

Next, we investigated whether TEFM can interact with the nascent transcript. We assembled ECs on a nucleic acid scaffold containing a photo-reactive analog of uridine (4-thio-uridine) 13 nt upstream from the 3' end of RNA in the scaffold. We "walked" mtRNAP along the DNA template by providing different mixtures of NTPs to allow placing the photo-reactive probe at positions: -14, -15 and -16 by RNA extension (Fig. 30 A). We observed efficient cross-linking between TEFM and RNA

when the photoreactive base was 15 to 16 nt away from the 3' end of RNA. Using a template DNA containing the LSP promoter, we incorporated a photo reactive 6-thio-GMP into RNA transcript (Fig. 30 B). The start-up complexes were formed to allow mtRNAP to synthesize a 18-mer RNA in the presence of TEFM. These complexes were immobilized on Ni-agarose beads. The photo reactive analog, 6-thio GMP (sG, 50 μM) was incorporated during the transcription cycles on Ni-agarose beads and the ECs were halted 44 bps away from the promoter start site. We detected cross-link of TEFM to the RNA bases 9-25 nt away from the 3' end of RNA confirming our data with the scaffold (Fig. 30 B). Combining cross-linking data on scaffold and promoter template, this suggests that TEFM interacts with the nascent transcript after it emerges from the RNA exit pore in EC. Thus, TEFM may contribute to formation of the RNA exit channel in mtRNAP and thus contribute to the overall stability of the ECs

2. Transcription termination at CSB II.

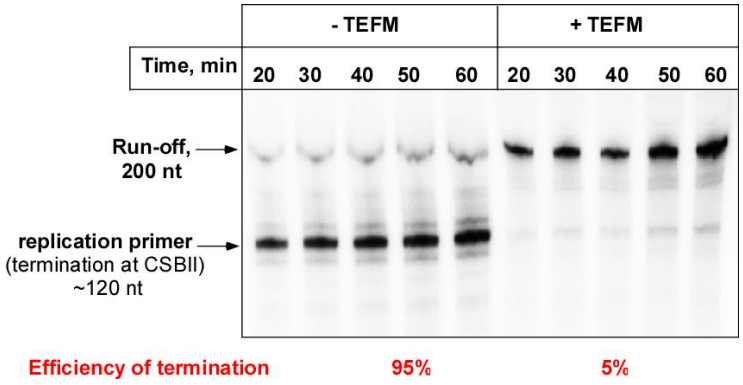

Figure 33. MtRNAP does not terminate at CSBII in the presence of TEFM. Termination efficiency (%) is indicated below the panel.

Replication and transcription are interconnected processes in human mitochondria. Human mitochondria lack primase activity

necessary for initiation of DNA replication. It has been shown that human mtRNAP serves as a sole enzyme responsible for the priming of replication events (*41, 43, 81*). MtRNAP generates the RNA replication primer by initiating transcription from LSP and terminating at a G-rich sequence about 120 nt downstream of the promoter (*41, 43, 70*). The G-rich sequence, called conserved sequence block II (CSBII) results in the formation of a G-quadruplex structure in nascent RNA (Fig. 2). It has been suggested that the G-quadruplex may also transform into a hybrid R-loop between nascent RNA and the non-template DNA strand (*71, 72*). The termination event occurs near the replication origin O_H, thus, generating a the replication primer that is used by replisome to drive the ensuing replication events (*46*).

To examine this, we performed *in vitro* transcription assays on a template containing the region of mitochondrial DNA encompassing LSP and CSBII (Fig. 31). MtRNAP terminated downstream of LSP at CSBII resulting in generation of ~120 nt replication primer. Only small population of mtRNAP molecules was able to produce a 200 nt run-off transcript. In order to determine the location of the termination signal at CSBII, we prepared a template that lacked 71 nucleotides between LSP and CSB II (Δ 71 CSBII template). Use of this template allowed us to locate the termination signal using high-resolution sequencing gels and RNA markers of different sizes (Fig. 32). MtRNAP terminated 16-20 nt downstream of the G-quadruplex region at CSBII near the end of U6 run (Fig. 32). At this point, the RNA-DNA hybrid formed in mtRNAP EC consists of A-U and T-A pairs and, therefore, is extremely weak, likely resulting in overall instability of the ECs. This

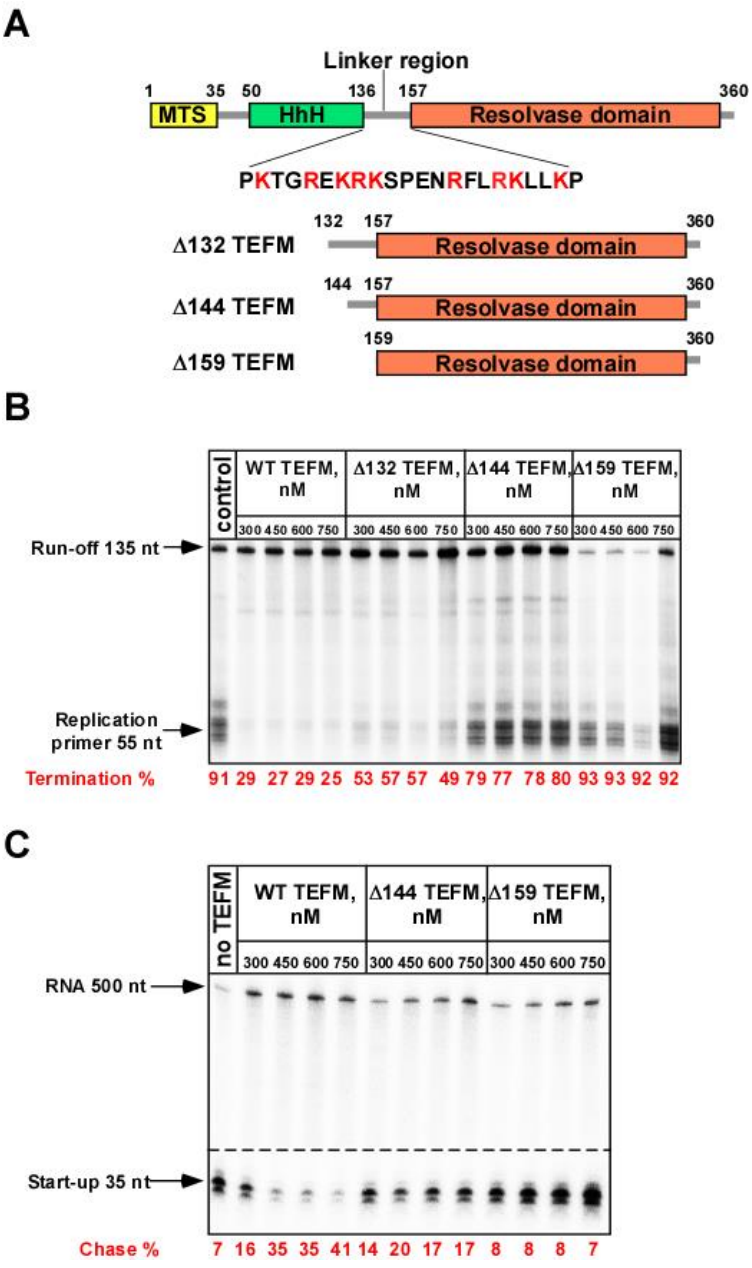

Figure 34. The linker region of TEFM is essential for anti-termination activity of TEFM and stability of ECs. A. The schematics of TEFM with protein domains. The sequence of linker region with highlighted positively charged residues is shown. Variants of TEFM with different deletions in the linker region are depicted. **B. Anti-termination activity of TEFM variants with deletions in the linker region.** Conditions as in Fig. 31, 33. **C. Halted complex stability assay of TEFM variants with deletions in the linker region** (see Fig. 26).

situation is similar to intrinsic termination signals – hairpins that are responsible for termination of prokaryotic transcription. Thus, bacterial RNAPs terminate due to the formation of an RNA hairpin that disrupts the upstream region of the RNA-DNA hybrid including the run of six to eight uridine 5'-monophosphate residues (*71, 72, 130*).

Human mtDNA is highly polymorphic in the CSBII region. Significantly, the commonly used

reference genome (Cambridge) contains a rare polymorphism in the G-quadruplex region – G5AG7. – while the majority of human mtDNA haplotypes include two additional G residues at CSBII (G6AG8) (*117*). We demonstrated that the termination efficiency of mtRNAP was dependent on G-quadruplex region sequence. We found that the efficiency of transcription termination was substantially higher when the sequence corresponded to the major polymorphism in mtDNA (G6AG8) (Fig. 31) was used, suggesting an effect of G run length on quadruplex formation. These data indicate that different polymorphisms in this region may directly affect efficiencies of termination by mtRNAP and result in various transcription elongation profiles.

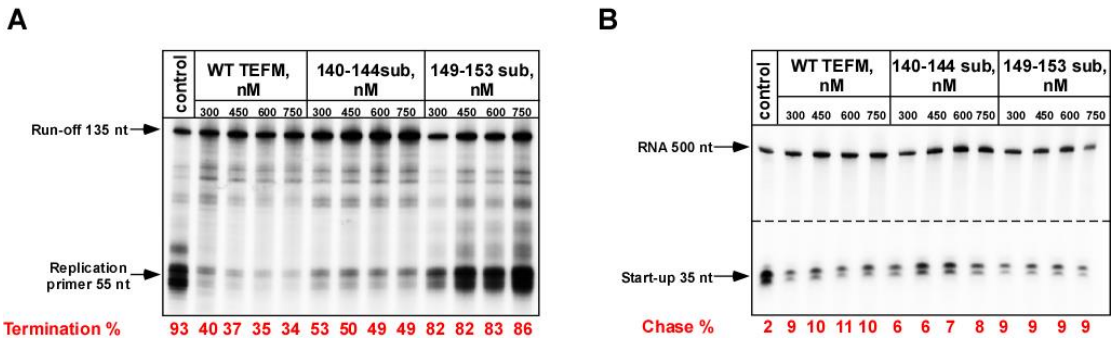

Figure 35. Positively charged residues in the linker region of TEFM are responsible for anti-termination activity. A. Anti-termination activity of TEFM variants with substitutions in the linker region (140-44 and 149-53 substitutions). Reaction conditions as in Fig. 31 and 33. **B. Halted complex stability assay of TEFM variants with substitutions in the linker region.** Reaction conditions as in Fig. 26.

For a long time, termination at CSBII by mtRNAP just 120 nt downstream of LSP has been a puzzling phenomenon, as a number of essential genes are encoded downstream of the G-quadruplex signal, including one protein gene and 8 tRNA genes. Premature termination of mtRNAP at this point would not allow efficient expression of these genes. We therefore tested whether TEFM affected mtRNAP termination at CSBII. We found that in the presence of TEFM, mtRNAP efficiently transcribed through CSBII generating a full size run-off product (Fig. 33). Thus,

TEFM acts as an anti-termination factor that suppresses termination at CSBII and synthesis of a replication primer for mtDNAP.

3. Functional analysis of TEFM domains.

TEFM shares sequence homology to the group of prokaryotic and eukaryotic transcription factors that contain helix-hairpin-helix (HhH) domains (*101*). TEFM contains two tandemly repeated HhH domains at the N-terminal portion of the protein. The C-terminal domain of TEFM consists of RNaseH fold characteristic for RuvC resolvases. However, TEFM lacks any detectable resolvase activity (*101*). In addition to HhH and RNaseH folds, TEFM also contains a 20 amino acid linker region rich in Arg and Lys residues (Fig. 34 A). The large number of positively charged residues in the linker region may potentially be important for interactions with RNA, contributing to the anti-termination activity of TEFM and stabilization of ECs. To understand the functional roles of TEFM subdomains, we generated variants of TEFM that lack NTD (deletion of residues 35-132 – Δ132 TEFM), NTD and part of the linker (deletion of residues 35-144 – Δ144 TEFM) and NTD together with the linker (deletion of residues 35-159 – Δ159 TEFM or CTD TEFM). We tested these variants in *in vitro* transcription and halted complex stability assays (Fig. 34 B, C). Δ144 and Δ159 TEFM possessed dramatically diminished anti-termination activity as can be seen by accumulation of the termination product. At the same time, deletion of the NTD did not result in noticeable reduction of anti-termination activity (Fig. 34 B). The Δ132 NTD deletion did not affect the ability of TEFM to stabilize the EC in the halted complex stability assay (Fig. 34 C). In contrast, we found that Δ144 TEFM

provided only moderate stability to ECs and Δ159 TEFM did not stabilize the ECs at all (Fig. 34 B). Thus, we conclude that the CTD and the linker region (residues 136-156) are essential for anti-termination and stabilization activities of TEFM, while the NTD is dispensable for the *in vitro* activities.

We investigated the linker subdomain in more detail by scanning mutagenesis within the region containing amino acids 136-156. We substituted the charged residues in the first (residues 140-144) or the second (residues 149-153) half of the linker with the sequence AAGAA, thus generating sub140-144AAGAA and sub 149-153AAGAA TEFM variants. In addition, we made the RK152,153AA substitution in this region (Table V, appendix). The constructed mutants were tested in *in vitro* transcription and halted complex stability assays. We found that TEFM sub 149-153 and RK/AA (152-153) variants had the most severe defect in anti-termination activity (Fig. 35 A, Table V, appendix). However, both TEFM variants stabilized mtRNAP ECs as efficiently as WT TEFM (Fig. 35 B, Table V, appendix). We conclude that region 149-153 of the linker subdomain is necessary for anti-termination activity but is dispensable for stabilization of the EC. In this region, residues R152 and K153 were the most critical contributors to the anti-termination activity of TEFM.

4. Replication and transcription are mutually exclusive processes in human mitochondria.

Our results suggest that in the absence of TEFM, mtRNAP terminates transcription from LSP at CSBII, likely due to formation of a hybrid G-quadruplex

structure, and generates a 120 nt replication primer. At the same time, mtRNAP is not sufficiently processive to generate long polycistronic transcripts from HSP. Thus, the absence of TEFM is more favorable for mtDNA replication. In the presence of TEFM, termination of mtRNAP at CSBII is suppressed as TEFM interferes with the formation of the G-quadruplex structure, allowing the polymerase to transcribe past this region. Increased processivity of mtRNAP facilitates the production of long polycistronic transcripts from HSP. This situation is beneficial for expression of the mitochondrial genome. Thus, we propose that TEFM serves as a regulatory component of a molecular switch that determines whether to replicate mtDNA or increase transcription rates (Fig. 36).

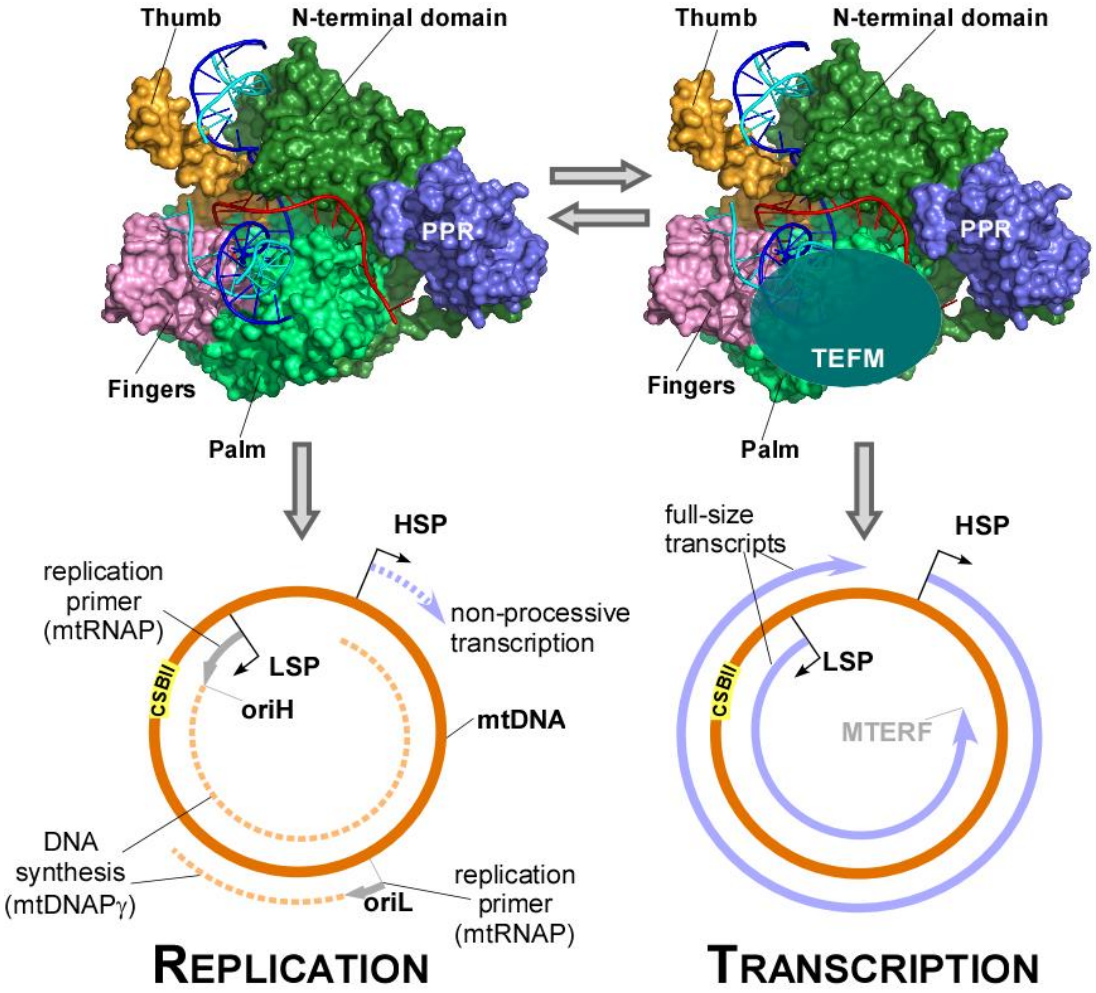

Figure 36. Model for a replication-transcription switch in mitochondria. In the absence of TEFM, mtRNAP initiates transcription at the LSP promoter but terminates at CSBII and produces a 120-nt replication primer at the replication origin oriH. The primer is used by the replisome for replication of the heavy strand of mtDNA, during which a separate initiation event involving mtRNAP generates a replication primer at oriL. Because of low processivity, transcription from the HSP does not produce a full-size transcript (bottom left). When TEFM is bound (bottom right), transcription from HSP produces genome-length transcripts. At the same time, LSP-driven transcription is not terminated at CSBII and proceeds to the site of factor-dependent termination (MTERF).

Discussion.

1. Mechanisms of transcription elongation in human mitochondria.

A comparison of the crystal structures of T7 and mtRNAPs showed high similarity in C-terminal domains of RNAPs, suggesting conservation of the mechanisms for substrate selection, binding and catalysis. T7 RNAP ECs have been studied extensively and several crystal structures that outline different stages of nucleotide incorporation cycles are now available (*87, 88, 93, 123*). In agreement with the crystal structures, biochemical studies reveal that most mtRNAP ECs are in the pre-translocated state, whereas all T7 RNAP ECs are in the post-translocated state. Stabilization of the pre-translocated state is a prerequisite for transcriptional pausing, whereas ECs in the post-translocated state can rapidly bind incoming NTP and initiate new cycles of NTP incorporation (*88, 91, 127, 128*). Thus, it has been suggested that the conformational states of polymerases can affect their elongation rates. We therefore investigated the elongation rates of yeast and human mtRNAPs and compared them to T7. We found that yeast and human mtRNAPs were 10-20 times slower than T7 RNAP, which is consistent with the finding that the pre-translocated state is stabilized in the mtRNAP ECs.

The elongation phase of transcription is a highly regulated step in many systems. T7 bacteriophage has a very fast life cycle, lasting around 10-20 minutes (*120*). The physiological function of T7 RNAP is to produce a large amount of its transcripts to outcompete bacterial RNA for ribosome occupancy, requiring fast

elongation rates as measured in this and other studies (*121, 122*). However, mitochondrial transcription needs to be coordinated with RNA processing, degradation, and translation to prevent the accumulation of potentially deleterious free RNAs. Indeed, the elongation rates of mtRNAP are comparable with the rates for the highly regulated multi-subunit bacterial and eukaryotic RNAPs. Bacterial RNAPs transcribe at a rate of about 50–100 nt/sec measured *in vivo (131)*, and 10–35 nt/sec *in vitro (132)*. In eukaryotic cells mRNA synthesis proceeds at a rate of 20-30 nt/sec (*133*). *In vitro,* purified RNAP II elongates at rates of 1-10 nt/sec (*127*). Transcript elongation by multi-subunit polymerases is a dynamic process that does not occur at a constant rate but on average, is dramatically slower than T7 rates. Despite the structural similarities between the mtRNAP and T7 RNAP, its elongation rates indicate a functional similarity to multi-subunit polymerases.

According to recent studies, the rate-limiting steps of transcription elongation by mtRNAP are likely to be catalysis and translocation (*99*). We compared the elongation structures of T7 and mtRNAPs and characterized elements involved in catalysis and translocation. Despite high conservation between the active sites of mtRNAP and T7 RNAP, we noticed changes in the elements that could affect catalysis or translocation. We made substations in mtRNAP or T7RNAP that reflect these differences and examined the effects of these substitutions (Table II, III, appendix). Residues C1152 and W1154 flanking the catalytic Asp 1151, which are involved in numerous hydrogen bond interactions that stabilize the active center of mtRNAP, were substituted with S813 and G815, respectively, in T7. In addition, the C-terminal residues TYFFS, conserved in all mammalian mtRNAPs, were substituted

to residues DFAFA of T7 which are completely conserved in all phage polymerases. Positively charged K1012 and R1013 of Y helix in mtRNAP interacting with RNA in EC were substituted to E652 and D653 in T7. Another group of differences in O-helix and loop connecting O/Y helices are prominent between the two polymerases (Fig. 18). Mutations in catalytic center and O/Y helices had different sensitivity to PPi and, thus, stabilized alternate conformational states of mtRNAP. Most prominently, several substitutions in O/Y helices (K1012E/R1013D/R1015I, K988S/V990T/Q992R) as well as in catalytic center (CYW/SFG (1152-1154), TYFFS/DFAFA (1225-1230)) were PPi-resistant and mostly in post-translocated conformation. Thus, we were able to manipulate the conformational states of mtRNAP by site-directed mutagenesis of residues to the ones found in T7 RNAP.

Stabilization of the post-tranlsocated state is necessary to increase elongation rates, as observed for T7 RNAP. However, all mtRNAP variants having a post-translocated conformation were defective in a promoter-initiation assay, suggesting that the altered residues may be involved in stabilization of mtRNAP-promoter interactions. Measurements of elongation rates using a primer extension assay demonstrated that these variants were dramatically slower than WT mtRNAP and could not efficiently incorporate nucleotides during translocation cycles, implicating impaired EC stability. We analyzed the residues of T7 RNAP (in Y-helix) involved in translocation by generating substitutions to the residues found in mtRNAP: E652D/K653R and ΔAIDS (658-661). In contrast to mtRNAP, substitutions in the Y helix of T7 RNAP variants had no effect on promoter initiation. Interestingly, shortening of the Y helix in T7 RNAP (ΔAIDS) stabilized the pre-translocated

conformation of the EC and resulted in two fold reduction of the elongation rates of T7 RNAP.

Overall, amino acid substitutions in the structural elements of mtRNAP responsible for catalysis and translocation to the corresponding residues in T7 RNAP did not result in generation of a "fast" mtRNAP, though a reciprocal approach allowed generation of a "slow" version of T7 RNAP. It has been suggested that T7 RNAP may be an evolutionary precursor of mtRNAPs (*134*). This hypothesis was proposed based upon the high homology between mtRNAP and T7 and the absence of bacterial-like multi subunit RNAP in mitochondria. Our data argue against this hypothesis. We reason that since no single point mutation or even a "cluster" of mutations in various structural elements in mtRNAP resulted in higher elongation rates, mtRNAP and T7 RNAP diverged from their common single-subunit ancestor very early in evolution. We suggest that modern mtRNAPs are much closer in their structure to the ancestral RNAP than to T7 because the elongation rates of mitochondrial enzymes should reflect their biological function. We therefore propose that the "missing" RNAP from the early bacterial endosymbiont was a single-subunit RNAP that served as a house-keeping RNAP in pre-endosymbiotic bacteria, and gave rise to modern single-subunit RNAPs from phages, mitochondria and plastids.

2. Regulation of mtRNAP by TEFM.

The processivity and fast elongation rates of T7 RNAP elongation are largely associated with the stabilization of a post-translocated conformation of the enzyme (*92, 121, 122*). Factor-dependent prokaryotic and eukaryotic RNAPs are also known

to oscillate between post-tranlsocated and pre-translocated states during elongation. Numerous transcription elongation factors are responsible for stabilization of post-translocated states of those polymerases (*128, 129*). Moreover, in yeast mitochondria putative elongation factor, Mss116 has been shown to promote stability of yeast mtRNAP and stabilize post-tranlsocated state of the enzyme (*135*). Therefore, it was reasonable to propose that newly identified TEFM would perform similar functions for mtRNAP in human mitochondria. First, we confirmed that, as it was suggested (but never demonstrated) (*101*), TEFM increases the processivity of mtRNAP on long DNA templates, which is necessary for transcription of 16 000 bp-long mtDNA. Next, we demonstrated that the higher processivity of mtRNAP/TEFM complex is, likely, explained by the increased stability of ECs in the presence of TEFM. Finally, we determined that TEFM enabled stabilization of the post-translocated conformation of mtRNAP ECs. As a consequence, elongation rates of mtRNAP are increased about four times in the presence of TEFM. Thus, TEFM, similarly to transcription factors in bacteria, increases transcription rates of mtRNAP by stabilizing a post-translocated state of polymerase and increasing the overall stability of ECs (*128, 129*).

In bacteria, Gre A, Gre B, Mfd reactivate arrested or stalled RNAP ECs by cleaving the nascent transcript (*129*). Nus G and Nus G-like factors stimulate transcription by increasing the processivity and elongation rates of bacterial RNAP (*136*). In eukaryotic cells, an entire class of general elongation factors suppresses pausing and stimulates productive elongation primarily involving P-TEFb and SII complex. These elongation factors bind to RNAP II EC and provide it with the ability to transcribe at steady rate of about 3.8 kb/min (*128*). The similar activity in

productive transcription elongation is characteristic for large group of factors comprising super elongation complexes (*137*). These factors increase the elongation rate of RNAP II by suppressing transient pausing and augmenting processivity and stability of polymerase, analogous to mtRNAP-TEFM complex.

3. Transcription termination and mechanisms of anti-termination.

We demonstrated that mtRNAP efficiently terminated at CSBII. The termination efficiency of mtRNAP was dependent on the region that encodes a G-quadruplex sequence in CSBII. Thus, the G6AG8 sequence, found in the majority of human haplotypes, is a more efficient terminator than the G5AG7 sequence found in the reference (Cambridge) mitochondrial genome. Other polymorphisms in this region such as G5AG8 and G7AG6 decrease or have no effect on termination (Table VII, appendix). The occurrence of different pathologies including cancers and neuro-degenerative diseases (*138, 139*) correlates with polymorphisms in the G-quadruplex region. Sequence composition of the G-quadruplex region can therefore regulate the efficiency of termination, and, thus, positively or negatively regulate mitochondrial gene expression.

We determined the region of termination in CSBII to be 16-18 nt downstream of the G-quadruplex, near the end of a U6 run. At this point, the 9-bp RNA-DNA hybrid consists of weak A-U and T-A pairs those likely results in instability of the ECs. This situation is similar to the intrinsic termination signals – hairpins - that are responsible for termination of prokaryotic transcription. The hairpin terminator

consists of two components: a GC-rich sequence that folds into hairpin structure in the emerging transcript and a U-rich region that weakens RNA-DNA hybrid (*129*). Together these two elements cause destabilization of the EC and termination of transcription in bacteria. Other classes of termination events also destabilize EC and release transcripts but operate using different mechanisms. The Rho-dependent termination depends on the termination factor (Rho for *E. coli*) that binds to the emerging transcript and pulls it out of the EC (*140*). Mfd-dependent-termination relies on the transcription repair coupling factor Mfd that destabilizes RNAP-DNA interactions (*141*). Mitochondrial transcription termination, which involves the formation of a G-quadruplex, is a unique event not described in any other biological system. The requirement for a G-quadruplex structure instead of a stem loop as in bacteria may be related to the other functions of that structure. It has been shown that G-quadruplex can tether RNA to non-template strand of DNA, forming a stable R-loop(*72*). Moreover, the G-quadruplex may have a role in stabilizing a triple-stranded structure in the D-loop by preventing re-annealing of template and non-template strands of mtDNA.

One of the major finding of this work is identification of the anti-termination activity of TEFM within the CSBII region. We showed that TEFM interacts with the downstream DNA duplex in the mtRNAP EC. Moreover, we detected interactions between TEFM and the emerging transcript as well as the G-quadruplex region of RNA. Based on interactions with DNA and RNA, we proposed that the putative binding site of TEFM should be near the palm region of mtRNAP. This would allow TEFM to be located close to the PPR domain and interact with the emerging

transcript in RNA exit channel. It is possible that TEFM may even form a part of the RNA exit channel, thereby increasing the stability of mtRNAP ECs and helping to retain the post-translocated state of enzyme. Thus, TEFM likely interferes with the formation of G-quadruplex structure in RNA exit channel of mtRNAP. Such a mechanism is reminiscent of bacterial anti-termination by bacteriophage proteins λN and λQ that cause bacterial RNAP to transform into a termination-resistant form. Besides λN and λQ proteins, the bacterial anti-termination complex typically includes an RNA component, *nut*, and bacterial accessory factors belonging to the Nus family, primarily Nus A and Nus G (*129*). N or Q-modification of bacterial RNAP prevents termination by directly interfering with the formation of the hairpin at an intrinsic terminator (*142, 143*) analogous to proposed mechanism of action of TEFM.

4. Regulation of the replication-transcription switch in mitochondria.

We proposed a model for a mitochondrial replication-transcription switch in which TEFM serves as a molecular component that determines whether mtDNA will be replicated or transcribed. This switch may be important for a number of developmental processes such as embryogenesis, oogenesis and differentiation where mtDNA copy number is strictly regulated at specific stages of differentiation (*56, 57*). We suggest that replication and transcription are mutually exclusive processes, as simultaneous functioning of the corresponding machineries can result in head-on collisions with detrimental consequences for the genome, particularly at replication forks leading to replication arrest, premature transcription termination and DNA

breaks (*144*). Recent studies demonstrated that collisions also trigger duplication/deletions and base substitutions at promoter regions in bacteria (*145*). Separation of replication and transcription in time would imply that there are different populations of mitochondrial nucleoids. While one group of nucleoids could be replicating (in the absence of TEFM), another population would be transcribing (in the presence of TEFM). This notion is supported by immunostaining of TEFM and newly synthesized RNA in mitochondrial nucleoids. Nucleoids with newly synthesized RNA also contained TEFM, whereas TEFM was absent in transcription-negative nucleoids consisting only of mtDNA (*101, 146*). Another piece of supporting evidence comes from the observation that only 50% of mtDNA genomes have a D-loop (*47*). The mtDNA molecules that contain a D-loop likely represent the population of replication intermediates containing the 7S nascent DNA. It is clear that any transcription event would displace the 7S DNA, suggesting that the D-loop containing nucleoids are reserved for future replication events.

According to our model, the levels of functional TEFM should be regulated in order to adjust to different cellular conditions requiring increase of transcription or replication events in mitochondria. Regulation of TEFM activity may occur at different stages: transcriptional, translational or post-translational. To date, thorough investigation of TEFM gene expression in various cells and tissues is lacking. However, given the fact that different organs require specific amounts of mitochondrial function, it is plausible that the levels of TEFM may be regulated at the stage of gene expression. Accordingly, tissues with high mitochondrial demand such as heart, liver, brain could expresses greater amounts of TEFM compared to others.

Regulation at the translational level may include regulation of stability and translation rate of TEFM mRNA. TEFM as a protein molecule may be stabilized by post-translational modifications or interaction with other factors. The most universal post-translational modification in cells is phosphorylation. Phosphorylation and dephosphorylation processes could quickly regulate the activity of protein depending on cellular needs. This scenario is plausible considering the role of TEFM in the replication-transcription switch. Phosphorylation of TEFM may function analogously to phosphorylation of another mitochondrial transcription initiation factor, TFAM that is degraded by Lon protease (*147*). Another important area of research is the interaction and modulation of TEFM activity by other mitochondrial proteins. TEFM has been shown to interact with RNA helicase DHX30 and the PPR protein PTCD3 (*101*). These proteins are presumably involved in RNA processing and translation events, respectively, and, thus, could assist TEFM in transcribing mtDNA. However, the direct role of these proteins in TEFM interaction has not been demonstrated. Further studies are needed to provide new insights into mechanisms of TEFM regulation in mitochondria.

Summary and conclusions.

1. Transcription elongation rates of mtRNAPs are 10-20 times slower that T7 RNAP, likely reflecting the need to coordinate mitochondrial gene expression with RNA processing and translation.

2. Unlike T7 RNAP, mtRNAP can be found in both pre- and post-translocated conformations in elongation complexes. Positively charged residues in the Y helix interacting with RNA appear to stabilize the post-translocated conformation of mtRNAP. Substitution of the residues in the structural elements, responsible for mtRNAP translocation (O/Y helices) and catalysis, results in mtRNAP variants with the stabilized post-translocated conformation.

3. Unlike the situation in T7 RNAP, mutation of residues in the O/Y helices results in mtRNAP variants defective in transcription initiation. This suggests the importance of this structural element for promoter binding and recognition in mtRNAP.
4. No single point mutation or even a "cluster" of mutations in various structural elements in mtRNAP resulted in higher elongation rates. Mutations of the corresponding non-conserved elements of T7 RNAP resulted only in modest changes of elongation rates. This suggests mtRNAP and T7 RNAP diverged

from a common ancestor single-subunit RNAP very early during evolutionary history.

5. TEFM is required for processivity of mtRNAP on long templates and its stability. TEFM increases the elongation rates of mtRNAP by stabilization of the post-translocated state of mtRNAP ECs.

6. TEFM likely prevents termination of mtRNAP at CSBII by interfering with the formation of the RNA G-quadruplex. In the elongation complex, TEFM interacts with downstream DNA and the 5' end of nascent RNA, likely contributing to the formation of RNA exit channel.

7. TEFM serves as a component of the molecular switch that allows mitochondria to either replicate mtDNA or to elevate its transcription rates. In the absence of TEFM, transcription from HSP is not processive and mtRNAP primes replication of mtDNA. In the presence of TEFM, mtRNAP does not terminate at CSBII and produces large polycistronic transcripts.

8. Replication and transcription appear to be mutually exclusive processes in mitochondria. This likely serves to prevent the detrimental effects of head-on collisions of the replication and transcription machineries on circular mtDNA.

9. Human mitochondria likely contain different nucleoids. Some nucleoids are involved in active transcription, some undergo replication (no TEFM), and some are replication and transcription-negative as they are tightly packed and not accessible for protein machineries.

References

1. **Y. I. Morozov *et al.*, A novel intermediate in transcription initiation by human mitochondrial RNA polymerase. *Nucleic acids research* 42, 3884 (Apr, 2014).**

2. **V. Jayashankar, I. A. Mueller, S. M. Rafelski, Shaping the multi-scale architecture of mitochondria. *Current opinion in cell biology* 38, 45 (Feb, 2016).**

3. **F. Vogel, C. Bornhovd, W. Neupert, A. S. Reichert, Dynamic subcompartmentalization of the mitochondrial inner membrane. *The Journal of cell biology* 175, 237 (Oct 23, 2006).**

4. **S. Hoppins *et al.*, A mitochondrial-focused genetic interaction map reveals a scaffold-like complex required for inner membrane organization in mitochondria. *The Journal of cell biology* 195, 323 (Oct 17, 2011).**

5. **M. Harner *et al.*, The mitochondrial contact site complex, a determinant of mitochondrial architecture. *The EMBO journal* 30, 4356 (Nov 2, 2011).**

6. **Y. Chaban, E. J. Boekema, N. V. Dudkina, Structures of mitochondrial oxidative phosphorylation supercomplexes and mechanisms for their stabilisation. *Biochimica et biophysica acta* 1837, 418 (Apr, 2014).**

7. **P. Mitchell, Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191, 144 (Jul 8, 1961).**

8. **M. J. Gresser, J. A. Myers, P. D. Boyer, Catalytic site cooperativity of beef heart mitochondrial F1 adenosine triphosphatase. Correlations of initial velocity, bound intermediate, and oxygen exchange measurements with an alternating three-site model. *The Journal of biological chemistry* 257, 12030 (Oct 25, 1982).**

9. **H. A. Krebs, The history of the tricarboxylic acid cycle. *Perspectives in biology and medicine* 14, 154 (Autumn, 1970).**

10. **N. S. Chandel, Mitochondria as signaling organelles. *BMC biology* 12, 34 (2014).**

11. **D. P. Barupala, S. P. Dzul, P. J. Riggs-Gelasco, T. L. Stemmler, Synthesis, delivery and regulation of eukaryotic heme and Fe-S cluster cofactors. *Archives of biochemistry and biophysics* 592, 60 (Feb 15, 2016).**

12. **R. Lill *et al.*, The role of mitochondria in cellular iron-sulfur protein biogenesis and iron metabolism. *Biochimica et biophysica acta* 1823, 1491 (Sep, 2012).**

13. **T. E. Gunter, L. Buntinas, G. C. Sparagna, K. K. Gunter, The Ca2+ transport mechanisms of mitochondria and Ca2+ uptake from physiological-type Ca2+ transients. *Biochimica et biophysica acta* 1366, 5 (Aug 10, 1998).**

14. **J. G. McCormack, R. L. Daniel, N. J. Osbaldeston, G. A. Rutter, R. M. Denton, Mitochondrial Ca2+ transport and the role of matrix Ca2+ in mammalian tissues. *Biochemical Society transactions* 20, 153 (Feb, 1992).**

15. **D. A. East, M. Campanella, Ca2+ in quality control: an unresolved riddle critical to autophagy and mitophagy. *Autophagy* 9, 1710 (Nov 1, 2013).**

16. **B. Glancy, R. S. Balaban, Role of mitochondrial Ca2+ in the regulation of cellular energetics. *Biochemistry* 51, 2959 (Apr 10, 2012).**

17. **D. G. Nicholls, Mitochondria and calcium signaling. *Cell calcium* 38, 311 (Sep-Oct, 2005).**

18. **G. Szabadkai, M. R. Duchen, Mitochondria: the hub of cellular Ca2+ signaling. *Physiology* 23, 84 (Apr, 2008).**

19. P. Pizzo, T. Pozzan, Mitochondria-endoplasmic reticulum choreography: structure and signaling dynamics. *Trends in cell biology* 17, 511 (Oct, 2007).

20. A. Murley, J. Nunnari, The Emerging Network of Mitochondria-Organelle Contacts. *Molecular cell* 61, 648 (Mar 3, 2016).

21. C. Lopez-Crisosto *et al.*, ER-to-mitochondria miscommunication and metabolic diseases. *Biochimica et biophysica acta* 1852, 2096 (Oct, 2015).

22. B. Kornmann, P. Walter, ERMES-mediated ER-mitochondria contacts: molecular hubs for the regulation of mitochondrial biology. *Journal of cell science* 123, 1389 (May 1, 2010).

23. S. Marchi, S. Patergnani, P. Pinton, The endoplasmic reticulum-mitochondria connection: one touch, multiple functions. *Biochimica et biophysica acta* 1837, 461 (Apr, 2014).

24. R. Bravo-Sagua *et al.*, Organelle communication: signaling crossroads between homeostasis and disease. *The international journal of biochemistry & cell biology* 50, 55 (May, 2014).

25. X. Liu, C. N. Kim, J. Yang, R. Jemmerson, X. Wang, Induction of apoptotic program in cell-free extracts: requirement for dATP and cytochrome c. *Cell* 86, 147 (Jul 12, 1996).

26. A. Gross, BCL-2 family proteins as regulators of mitochondria metabolism. *Biochimica et biophysica acta*, (Jan 29, 2016).

27. P. D. Bhola, A. Letai, Mitochondria-Judges and Executioners of Cell Death Sentences. *Molecular cell* 61, 695 (Mar 3, 2016).

28. M. K. Shigenaga, T. M. Hagen, B. N. Ames, Oxidative damage and mitochondrial decay in aging. *Proceedings of the National Academy of Sciences of the United States of America* 91, 10771 (Nov 8, 1994).

29. **N. S. Chandel *et al.*, Mitochondrial reactive oxygen species trigger hypoxia-induced transcription. *Proceedings of the National Academy of Sciences of the United States of America* 95, 11715 (Sep 29, 1998).**

30. **A. P. West *et al.*, TLR signalling augments macrophage bactericidal activity through mitochondrial ROS. *Nature* 472, 476 (Apr 28, 2011).**

31. **S. E. Weinberg, L. A. Sena, N. S. Chandel, Mitochondria in the regulation of innate and adaptive immunity. *Immunity* 42, 406 (Mar 17, 2015).**

32. **R. B. Seth, L. Sun, C. K. Ea, Z. J. Chen, Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* 122, 669 (Sep 9, 2005).**

33. **A. Carlucci, L. Lignitto, A. Feliciello, Control of mitochondria dynamics and oxidative metabolism by cAMP, AKAPs and the proteasome. *Trends in cell biology* 18, 604 (Dec, 2008).**

34. **U. Topf, L. Wrobel, A. Chacinska, Chatty Mitochondria: Keeping Balance in Cellular Protein Homeostasis. *Trends in cell biology*, (Mar 19, 2016).**

35. **Y. F. Lin, C. M. Haynes, Metabolism and the UPR(mt). *Molecular cell* 61, 677 (Mar 3, 2016).**

36. **A. B. Harbauer, R. P. Zahedi, A. Sickmann, N. Pfanner, C. Meisinger, The protein import machinery of mitochondria-a regulatory hub in metabolism, stress, and disease. *Cell metabolism* 19, 357 (Mar 4, 2014).**

37. **L. S. Wenz, L. Opalinski, N. Wiedemann, T. Becker, Cooperation of protein machineries in mitochondrial protein sorting. *Biochimica et biophysica acta* 1853, 1119 (May, 2015).**

38. **A. Chacinska, C. M. Koehler, D. Milenkovic, T. Lithgow, N. Pfanner, Importing mitochondrial proteins: machineries and mechanisms. *Cell* 138, 628 (Aug 21, 2009).**

39. M. J. Baker, A. E. Frazier, J. M. Gulbis, M. T. Ryan, Mitochondrial protein-import machinery: correlating structure with function. *Trends in cell biology* 17, 456 (Sep, 2007).

40. P. Mishra, D. C. Chan, Metabolic regulation of mitochondrial dynamics. *The Journal of cell biology* 212, 379 (Feb 15, 2016).

41. M. Falkenberg, N. G. Larsson, C. M. Gustafsson, DNA replication and transcription in mammalian mitochondria. *Annual review of biochemistry* 76, 679 (2007).

42. N. G. Larsson, Somatic mitochondrial DNA mutations in mammalian aging. *Annual review of biochemistry* 79, 683 (2010).

43. C. M. Gustafsson, M. Falkenberg, N. G. Larsson, Maintenance and Expression of Mammalian Mitochondrial DNA. *Annual review of biochemistry*, (Mar 24, 2016).

44. J. P. Uhler, M. Falkenberg, Primer removal during mammalian mitochondrial DNA replication. *DNA repair* 34, 28 (Oct, 2015).

45. J. Fish, N. Raule, G. Attardi, Discovery of a major D-loop replication origin reveals two modes of human mtDNA synthesis. *Science* 306, 2098 (Dec 17, 2004).

46. D. Kang, K. Miyako, Y. Kai, T. Irie, K. Takeshige, In vivo determination of replication origins of human mitochondrial DNA by ligation-mediated polymerase chain reaction. *The Journal of biological chemistry* 272, 15275 (Jun 13, 1997).

47. T. J. Nicholls, M. Minczuk, In D-loop: 40 years of mitochondrial 7S DNA. *Experimental gerontology* 56, 175 (Aug, 2014).

48. B. Xu, D. A. Clayton, A persistent RNA-DNA hybrid is formed during transcription at a phylogenetically conserved mitochondrial DNA sequence. *Molecular and cellular biology* 15, 580 (Jan, 1995).

49. D. A. Clayton, Replication of animal mitochondrial DNA. *Cell* 28, 693 (Apr, 1982).

50. E. Jemt *et al.*, Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. *Nucleic acids research* 43, 9262 (Oct 30, 2015).

51. C. Kukat *et al.*, Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proceedings of the National Academy of Sciences of the United States of America* 108, 13534 (Aug 16, 2011).

52. D. F. Bogenhagen, D. Rousseau, S. Burke, The layered structure of human mitochondrial DNA nucleoids. *The Journal of biological chemistry* 283, 3665 (Feb 8, 2008).

53. N. G. Larsson *et al.*, Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nature genetics* 18, 231 (Mar, 1998).

54. D. F. Bogenhagen, Mitochondrial DNA nucleoid structure. *Biochimica et biophysica acta* 1819, 914 (Sep-Oct, 2012).

55. C. Kukat *et al.*, Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proceedings of the National Academy of Sciences of the United States of America* 112, 11288 (Sep 8, 2015).

56. J. St John, The control of mtDNA replication during differentiation and development. *Biochimica et biophysica acta* 1840, 1345 (Apr, 2014).

57. J. C. St John, Mitochondrial DNA copy number and replication in reprogramming and differentiation. *Seminars in cell & developmental biology* 52, 93 (Apr, 2016).

58. **T. Wai, D. Teoli, E. A. Shoubridge, The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nature genetics* 40, 1484 (Dec, 2008).**

59. **J. M. Facucho-Oliveira, J. C. St John, The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem cell reviews* 5, 140 (Jun, 2009).**

60. **J. M. Facucho-Oliveira, J. Alderson, E. C. Spikings, S. Egginton, J. C. St John, Mitochondrial DNA replication during differentiation of murine embryonic stem cells. *Journal of cell science* 120, 4025 (Nov 15, 2007).**

61. **H. A. Tuppen, E. L. Blakely, D. M. Turnbull, R. W. Taylor, Mitochondrial DNA mutations and human disease. *Biochimica et biophysica acta* 1797, 113 (Feb, 2010).**

62. **A. Grunewald *et al.*, Mitochondrial DNA Depletion in Respiratory Chain-Deficient Parkinson Disease Neurons. *Annals of neurology* 79, 366 (Mar, 2016).**

63. **Y. C. Chae *et al.*, Mitochondrial Akt Regulation of Hypoxic Tumor Reprogramming. *Cancer cell* 30, 257 (Aug 8, 2016).**

64. **N. E. Scharping *et al.*, The Tumor Microenvironment Represses T Cell Mitochondrial Biogenesis to Drive Intratumoral T Cell Metabolic Insufficiency and Dysfunction. *Immunity*, (Aug 2, 2016).**

65. **E. Reznik *et al.*, Mitochondrial DNA copy number variation across human cancers. *eLife* 5, (2016).**

66. **K. Szczepanowska, A. Trifunovic, Different faces of mitochondrial DNA mutators. *Biochimica et biophysica acta* 1847, 1362 (Nov, 2015).**

67. **I. J. Holt, A. Reyes, Human mitochondrial DNA replication. *Cold Spring Harbor perspectives in biology* 4, (Dec, 2012).**

68. **T. A. Brown, C. Cecconi, A. N. Tkachuk, C. Bustamante, D. A. Clayton, Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. *Genes & development* 19, 2466 (Oct 15, 2005).**

69. **T. A. Brown *et al.*, Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction. *Molecular and cellular biology* 31, 4994 (Dec, 2011).**

70. **X. H. Pham *et al.*, Conserved sequence box II directs transcription termination and primer formation in mitochondria. *The Journal of biological chemistry* 281, 24647 (Aug 25, 2006).**

71. **P. H. Wanrooij, J. P. Uhler, T. Simonsson, M. Falkenberg, C. M. Gustafsson, 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* 107, 16072 (Sep 14, 2010).**

72. **P. H. Wanrooij *et al.*, A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop. *Nucleic acids research* 40, 10334 (Nov 1, 2012).**

73. **S. Wanrooij *et al.*, Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 105, 11122 (Aug 12, 2008).**

74. **J. M. Fuste *et al.*, Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. *Molecular cell* 37, 67 (Jan 15, 2010).**

75. **S. Wanrooij *et al.*, In vivo mutagenesis reveals that OriL is essential for mitochondrial DNA replication. *EMBO reports* 13, 1130 (Dec, 2012).**

76. **A. J. Berk, D. A. Clayton, Mechanism of mitochondrial DNA replication in mouse L-cells: asynchronous replication of strands, segregation of**

circular daughter molecules, aspects of topology and turnover of an initiation sequence. *Journal of molecular biology* 86, 801 (Jul 15, 1974).

77. D. F. Bogenhagen, D. A. Clayton, The mitochondrial DNA replication bubble has not burst. *Trends in biochemical sciences* 28, 357 (Jul, 2003).

78. I. J. Holt, H. E. Lorimer, H. T. Jacobs, Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell* 100, 515 (Mar 3, 2000).

79. A. Reyes *et al.*, Mitochondrial DNA replication proceeds via a 'bootlace' mechanism involving the incorporation of processed transcripts. *Nucleic acids research* 41, 5837 (Jun, 2013).

80. J. L. Pohjoismaki *et al.*, Alterations to the expression level of mitochondrial transcription factor A, TFAM, modify the mode of mitochondrial DNA replication in cultured human cells. *Nucleic acids research* 34, 5815 (2006).

81. I. Kuhl *et al.*, POLRMT regulates the switch between replication primer formation and gene expression of mammalian mtDNA. *Science advances* 2, e1600963 (Aug, 2016).

82. D. Milenkovic *et al.*, TWINKLE is an essential mitochondrial helicase required for synthesis of nascent D-loop strands and complete mtDNA replication. *Human molecular genetics* 22, 1983 (May 15, 2013).

83. A. Rantanen, N. G. Larsson, Regulation of mitochondrial DNA copy number during spermatogenesis. *Human reproduction* 15 Suppl 2, 86 (Jul, 2000).

84. M. I. Ekstrand *et al.*, Mitochondrial transcription factor A regulates mtDNA copy number in mammals. *Human molecular genetics* 13, 935 (May 1, 2004).

85. **B. M. Hallberg, N. G. Larsson, Making proteins in the powerhouse. *Cell metabolism* 20, 226 (Aug 5, 2014).**

86. **R. Ringel *et al.*, Structure of human mitochondrial RNA polymerase. *Nature* 478, 269 (Oct 13, 2011).**

87. **D. Temiakov *et al.*, Structural basis for substrate selection by t7 RNA polymerase. *Cell* 116, 381 (Feb 6, 2004).**

88. **Y. W. Yin, T. A. Steitz, The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* 116, 393 (Feb 6, 2004).**

89. **D. Temiakov *et al.*, The specificity loop of T7 RNA polymerase interacts first with the promoter and then with the elongating transcript, suggesting a mechanism for promoter clearance. *Proceedings of the National Academy of Sciences of the United States of America* 97, 14109 (Dec 19, 2000).**

90. **Y. W. Yin, T. A. Steitz, Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science* 298, 1387 (Nov 15, 2002).**

91. **T. A. Steitz, The structural changes of T7 RNA polymerase from transcription initiation to elongation. *Current opinion in structural biology* 19, 683 (Dec, 2009).**

92. **T. A. Steitz, The structural basis of the transition from initiation to elongation phases of transcription, as well as translocation and strand separation, by T7 RNA polymerase. *Current opinion in structural biology* 14, 4 (Feb, 2004).**

93. **T. H. Tahirov *et al.*, Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. *Nature* 420, 43 (Nov 7, 2002).**

94. **Y. I. Morozov *et al.*, A model for transcription initiation in human mitochondria. *Nucleic acids research* 43, 3726 (Apr 20, 2015).**

95. **M. Gaspari, M. Falkenberg, N. G. Larsson, C. M. Gustafsson, The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. *The EMBO journal* 23, 4606 (Nov 24, 2004).**

96. **H. B. Ngo, J. T. Kaiser, D. C. Chan, The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nature structural & molecular biology* 18, 1290 (Nov, 2011).**

97. **M. Sologub, D. Litonin, M. Anikin, A. Mustaev, D. Temiakov, TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase. *Cell* 139, 934 (Nov 25, 2009).**

98. **G. M. Cheetham, T. A. Steitz, Structure of a transcribing T7 RNA polymerase initiation complex. *Science* 286, 2305 (Dec 17, 1999).**

99. **E. D. Smidansky, J. J. Arnold, S. L. Reynolds, C. E. Cameron, Human mitochondrial RNA polymerase: evaluation of the single-nucleotide-addition cycle on synthetic RNA/DNA scaffolds. *Biochemistry* 50, 5016 (Jun 7, 2011).**

100. **J. J. Arnold, E. D. Smidansky, I. M. Moustafa, C. E. Cameron, Human mitochondrial RNA polymerase: structure-function, mechanism and inhibition. *Biochimica et biophysica acta* 1819, 948 (Sep-Oct, 2012).**

101. **M. Minczuk *et al.*, TEFM (c17orf42) is necessary for transcription of human mtDNA. *Nucleic acids research* 39, 4284 (May, 2011).**

102. **B. Kruse, N. Narasimhan, G. Attardi, Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. *Cell* 58, 391 (Jul 28, 1989).**

103. **E. Yakubovskaya, E. Mejia, J. Byrnes, E. Hambardjieva, M. Garcia-Diaz, Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. *Cell* 141, 982 (Jun 11, 2010).**

104. **N. Jimenez-Menendez *et al.*, Human mitochondrial mTERF wraps around DNA through a left-handed superhelical tandem repeat. *Nature structural & molecular biology* 17, 891 (Jul, 2010).**

105. **K. E. Guja, M. Garcia-Diaz, Hitting the brakes: termination of mitochondrial transcription. *Biochimica et biophysica acta* 1819, 939 (Sep-Oct, 2012).**

106. **J. Shang, D. A. Clayton, Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro. *The Journal of biological chemistry* 269, 29112 (Nov 18, 1994).**

107. **M. Terzioglu *et al.*, MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation. *Cell metabolism* 17, 618 (Apr 2, 2013).**

108. **T. Linder *et al.*, A family of putative transcription termination factors shared amongst metazoans and plants. *Current genetics* 48, 265 (Oct, 2005).**

109. **M. Pellegrini *et al.*, MTERF2 is a nucleoid component in mammalian mitochondria. *Biochimica et biophysica acta* 1787, 296 (May, 2009).**

110. **A. Wredenberg *et al.*, MTERF3 regulates mitochondrial ribosome biogenesis in invertebrates and mammals. *PLoS genetics* 9, e1003178 (2013).**

111. **C. B. Park *et al.*, MTERF3 is a negative regulator of mammalian mtDNA transcription. *Cell* 130, 273 (Jul 27, 2007).**

112. **M. D. Metodiev *et al.*, NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. *PLoS genetics* 10, e1004110 (Feb, 2014).**

113. **H. Spahr, B. Habermann, C. M. Gustafsson, N. G. Larsson, B. M. Hallberg, Structure of the human MTERF4-NSUN4 protein complex that**

regulates mitochondrial ribosome biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 109, 15253 (Sep 18, 2012).

114. J. Asin-Cayuela, T. Schwend, G. Farge, C. M. Gustafsson, The human mitochondrial transcription termination factor (mTERF) is fully active in vitro in the non-phosphorylated form. *The Journal of biological chemistry* 280, 25499 (Jul 8, 2005).

115. D. A. Markov *et al.*, Identification of proteins associated with the yeast mitochondrial RNA polymerase by tandem affinity purification. *Yeast* 26, 423 (Aug, 2009).

116. B. He *et al.*, Rapid mutagenesis and purification of phage RNA polymerases. *Protein expression and purification* 9, 142 (Feb, 1997).

117. R. M. Andrews *et al.*, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature genetics* 23, 147 (Oct, 1999).

118. K. Schwinghammer *et al.*, Structure of human mitochondrial RNA polymerase elongation complex. *Nature structural & molecular biology* 20, 1298 (Nov, 2013).

119. M. Kireeva *et al.*, Millisecond phase kinetic analysis of elongation catalyzed by human, yeast, and Escherichia coli RNA polymerase. *Methods* 48, 333 (Aug, 2009).

120. F. W. Studier, Bacteriophage T7. *Science* 176, 367 (Apr 28, 1972).

121. V. S. Anand, S. S. Patel, Transient state kinetics of transcription elongation by T7 RNA polymerase. *The Journal of biological chemistry* 281, 35677 (Nov 24, 2006).

122. M. Golomb, M. Chamberlin, Characterization of T7-specific ribonucleic acid polymerase. IV. Resolution of the major in vitro transcripts by gel

electrophoresis. *The Journal of biological chemistry* 249, 2858 (May 10, 1974).

123. D. Temiakov, M. Anikin, W. T. McAllister, Characterization of T7 RNA polymerase transcription complexes assembled on nucleic acid scaffolds. *The Journal of biological chemistry* 277, 47035 (Dec 6, 2002).

124. P. P. Hein, M. Palangat, R. Landick, RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. *Biochemistry* 50, 7002 (Aug 16, 2011).

125. E. Kashkina *et al.*, Elongation complexes of Thermus thermophilus RNA polymerase that possess distinct translocation conformations. *Nucleic acids research* 34, 4036 (2006).

126. G. Bonner, E. M. Lafer, R. Sousa, Characterization of a set of T7 RNA polymerase active site mutants. *The Journal of biological chemistry* 269, 25120 (Oct 7, 1994).

127. S. M. Uptain, C. M. Kane, M. J. Chamberlin, Basic mechanisms of transcript elongation and its regulation. *Annual review of biochemistry* 66, 117 (1997).

128. Q. Zhou, T. Li, D. H. Price, RNA polymerase II elongation control. *Annual review of biochemistry* 81, 119 (2012).

129. J. W. Roberts, S. Shankar, J. J. Filter, RNA polymerase elongation factors. *Annual review of microbiology* 62, 211 (2008).

130. V. Epshtein, C. J. Cardinale, A. E. Ruckenstein, S. Borukhov, E. Nudler, An allosteric path to transcription termination. *Molecular cell* 28, 991 (Dec 28, 2007).

131. P. P. Dennis, M. Ehrenberg, D. Fange, H. Bremer, Varying rate of RNA chain elongation during rrn transcription in Escherichia coli. *Journal of bacteriology* 191, 3740 (Jun, 2009).

132. S. F. Tolic-Norrelykke, A. M. Engh, R. Landick, J. Gelles, Diversity in the rates of transcript elongation by single RNA polymerase molecules. *The Journal of biological chemistry* 279, 3292 (Jan 30, 2004).

133. P. Maiuri *et al.*, Fast transcription rates of RNA polymerase II in human cells. *EMBO reports* 12, 1280 (Dec, 2011).

134. T. E. Shutt, M. W. Gray, Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends in genetics : TIG* 22, 90 (Feb, 2006).

135. D. A. Markov, I. D. Wojtas, K. Tessitore, S. Henderson, W. T. McAllister, Yeast DEAD box protein Mss116p is a transcription elongation factor that modulates the activity of mitochondrial RNA polymerase. *Molecular and cellular biology* 34, 2360 (Jul, 2014).

136. R. S. Washburn, M. E. Gottesman, Regulation of transcription elongation and termination. *Biomolecules* 5, 1063 (2015).

137. A. Shilatifard, W. S. Lane, K. W. Jackson, R. C. Conaway, J. W. Conaway, An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* 271, 1873 (Mar 29, 1996).

138. S. K. Bharti *et al.*, DNA sequences proximal to human mitochondrial DNA deletion breakpoints prevalent in human disease form G-quadruplexes, a class of DNA structures inefficiently unwound by the mitochondrial replicative Twinkle helicase. *The Journal of biological chemistry* 289, 29975 (Oct 24, 2014).

139. D. W. Dong *et al.*, Association of G-quadruplex forming sequences with human mtDNA deletion breakpoints. *BMC genomics* 15, 677 (2014).

140. M. S. Ciampi, Rho-dependent terminators and transcription termination. *Microbiology* 152, 2515 (Sep, 2006).

141. A. M. Deaconescu *et al.*, Structural basis for bacterial transcription-coupled DNA repair. *Cell* 124, 507 (Feb 10, 2006).

142. I. Gusarov, E. Nudler, Control of intrinsic transcription termination by N and NusA: the basic mechanisms. *Cell* 107, 437 (Nov 16, 2001).

143. J. W. Roberts *et al.*, Antitermination by bacteriophage lambda Q protein. *Cold Spring Harbor symposia on quantitative biology* 63, 319 (1998).

144. D. Dutta, K. Shatalin, V. Epshtein, M. E. Gottesman, E. Nudler, Linking RNA polymerase backtracking to genome instability in E. coli. *Cell* 146, 533 (Aug 19, 2011).

145. T. S. Sankar, B. D. Wastuwidyaningtyas, Y. Dong, S. A. Lewis, J. D. Wang, The nature of mutations induced by replication-transcription collisions. *Nature* 535, 178 (Jul 7, 2016).

146. L. Chatre, M. Ricchetti, Large heterogeneity of mitochondrial DNA transcription and initiation of replication exposed by single-cell imaging. *Journal of cell science* 126, 914 (Feb 15, 2013).

147. B. Lu *et al.*, Phosphorylation of human TFAM in mitochondria impairs DNA binding and promotes degradation by the AAA+ Lon protease. *Molecular cell* 49, 121 (Jan 10, 2013).

Appendix, Abbreviations list

Table I

The effect of RNA sequence on PPi-sensitivity of the mtRNAP ECs

Scaffold name	**RNA sequence (5'-3)**	**PPi sensitivity**
R2-8/T1/NT1	CUGCGGCGAU	++
R15/T8d/NT2a	GAGUCUGCGGCGAUA	-
R8mod/T42U/NT42U	CGGCGAUA	-
R16/T390/NT380	GAGUCUGCGGCGAUA	-
R14/T02/NT02	AGUCUGCGGCGCGC	++

Table II

Activity of catalytic center variants of mtRNAP

Variant	**Activity on scaffold**	**Activity on promoter**	**PPi sensitivity**
CYW to SFG (1152-1154)	100 %	1%	-
DFAFA to TYFFS (1225-1230)	100 %	1%	+-
S1230A	100 %	100%	++
ΔS1230	100 %	1%	++

Table III. Activity of fingers domain variants of mtRNAP

Variant	Activity/ RNA extension on scaffold	Activity/ promoter-dependent transcription	PPi sensitivity
WT	100 %	100%	++
R1013D	100 %	1%	-
K1012E/R1013D	100 %	1%	-
K1012E/R1013D/R1015I	100 %	1%	-
K988S/V990T/Q992R	100 %	1%	-
VTR(1001-1003)SKE	100 %	1%	++
G1006F	100 %	1%	++
L1008Q	100 %	100%	++
K1012E	100 %	100%	++
K988S	100 %	100%	++
E889A	100 %	100%	++
Q992A	100 %	100%	++
Q992R	100 %	100%	++
insAIDS(1017-1021)	100 %	100%	++
T7 fingers domain swap (628-671 region from T7 RNAP to 957-1044 of mtRNAP)	100 %	1%	++

Table IV

Elongation rates of RNAP variants used in experiments.

Variant	**RNAP**	**Elongation rates**
WT mtRNAP human	Δ119 mtRNAP	10 nt/sec
Δ43 mtRNAP	-	10 nt/sec
Δ368 mtRNAP	-	10 nt/sec
Q992A	Δ119 mtRNAP	10 nt/sec
Q992R	Δ119 mtRNAP	10 nt/sec
K988S/V990T/Q992R	Δ119 mtRNAP	~1 nt/sec
VTR(1001-1003)SKE	Δ119 mtRNAP	-
R1013D	Δ119 mtRNAP	-
K1012E/R1013D/R1015I	Δ119 mtRNAP	~1 nt/sec
G1006F	Δ119 mtRNAP	-
L1008K	Δ119 mtRNAP	10 nt/sec
insAIDS(1017-1021)	Δ119 mtRNAP	-
T7 fingers domain swap (628-671 region from T7 RNAP to 957-1044 of mtRNAP)	Δ119 mtRNAP	10 nt/sec
CYW to SFG (1152-1154)	Δ119 mtRNAP	~1 nt/sec
DFAFA to TYFFS (1225-1230)	Δ119 mtRNAP	~1 nt/sec
S1230A	Δ119 mtRNAP	10 nt/sec
ΔS1230	Δ119 mtRNAP	~1 nt/sec
WT Rpo41 yeast	-	20 nt/sec
WT T7 RNAP bacteriophage	-	213 nt/sec
E652K/D653R	T7 RNAP	~200 nt/sec
ΔAIDS(658-661)	T7 RNAP	~100 nt/sec

Table V

Antitermination activity of TEFM variants.

Variant	**Anti-Termination efficiency, %**	**Halted complex chase, %**
WT TEFM	5	16
Δ50 TEFM	5	16
Δ132 TEFM	50	-
Δ144 TEFM	80	20
Δ159 TEFM	93	8
140-144 subst. (AAGAA)	50	7
149-153 subst. (AAGAA)	85	9
R152A/K153A	75	-

Table VI

Primers used to amplify PCR templates for transcription.

Primer name	**Sequence**	**Plasmid**
Primer 500 bp For	5’ CATGTTCTTTCCTGCGTTATC 3’	pT7 blue LSP -70 +70, pT7 blue LSP 35 and 45 nt start-ups
Primer 1000 bp For	5’ – GTCCTTCTAGTGTAGCCG – 3’	pT7 blue LSP -70 +70
U-19 mer Rev	5′ GTTTTCCCAGTCACGACGT 3′	pT7 blue LSP -70 +70, pT7 blue LSP 35 and 45 nt start-ups
T7 promoter For	5’ CTAATACGACTCACTATAGGG 3’	pT7 blue 14s rRNA yeast, pT7 blue LSP -70 +70
Y14S 500r For	5’ CCCCAAAAAACTTGATTAGGG 3’	pT7 blue 14s rRNA yeast

CSB2 for	5’ AGTGTGTTAATTAATTAATGCT TGTAGGAC 3’	pT7 blue human mtDNA 202-481 and Δ71 mtDNA 202-481
CSB2 rev	5’ GATTAGTAGTATGGGAGTGGG AG 3’	pT7 blue human mtDNA 202-481 and Δ71 mtDNA 202-481

Table VII

Templates containing modification of CSBII.

Modification	**Plasmid name**	**mtRNAP termination efficiency, %**
WT CSBII (G6AG8)	pT7 blue human mtDNA 202-481	95
Δ61 CSBII (G6AG8)	pT7 blue human mtDNA Δ61 mtDNA 202-481	95
Δ71 CSBII (G6AG8)	pT7 blue human mtDNA Δ71 mtDNA 202-481	95
Δ87 CSBII (G6AG8)	pT7 blue human mtDNA Δ87 mtDNA 202-481	95
Δ71 CSBII (G5AG7)	pT7 blue human mtDNA Δ71 mtDNA 202-481	60
Δ71 CSBII (G7AG6)	pT7 blue human mtDNA Δ71 mtDNA 202-481	90
WT CSBII (G5AG8)	pT7 blue human mtDNA 202-481	60
WT CSBII (G6AG7)	pT7 blue human mtDNA 202-481	-
WT CSBII	pT7 blue human mtDNA 202-	95

(G6AG8) First G to A	481	
WT CSBII (G5AG8) First G to A	pT7 blue human mtDNA 202-481	-

Table VIII

Primers for mutagenesis.

Mutation	**Sequence**
Q992A mtRNAP	5’ GCAAGGTGGTGAAGGCGACGGTGATGACGG 3’
Q992A mtRNAP	5’ CAAGGTGGTGAAGCGTACGGTGATGACGG 3’
K988S/V990T/Q992RmtRNAP	5’ GTTTCATCACCCGCAGCGTGACGAAGCGGACGGTG ATGACGGTG 3’
VTR(1001-1003)SKEmtRNAP	5’ GACGGTGGTGTACGGGTCCAAGGAATATGGCGGGC GCCTGC 3’
R1013D mtRNAP	5’ CTGCAGATTGAGAAGGACCTCCGGGAGCTG 3’
K1012E/R1013D/R1015I mtRNAP	5’ CCTGCAGATTGAGGAGGACCTCATCGAGCTGAGC GACTTTCCCC 3’
G1006F mtRNAP	5’ GTCACGCGCTATGGCTTCCGCCTGCAGATTGAG 3’
L1008K mtRNAP	5’ CTATGGCGGGCGCCAGCAGATTGAGAAGC 3’
AIDS insertion (1017-1020) mtRNAP	5’ GAAGCGCCTCCGGGAGCTGGCGATCGACTCCAGC GACTTTCCCCAGG 3’

mtRNAP deletion Fingers domain (957-1044)	5’ CGGATGTGCCGCAGGACGTGTACTCGGGGACCCG G GCCATCCAGC 3’
T7 RNAP Fingers domain swap in mtRNAP megaprimer For	5' CGGATGTGCCGCAGGACGTGTACGGGATTGTTGC TAAGAAAGTCAACG 3’
T7 RNAP Fingers domain swap in mtRNAP megaprimer Rev	5’ GCTGGATGGCCCGGGTCCCCGACACCGTCACGC TCACAGATTC 3’
CYW to SFG (1152-1154) mtRNAP	5’ GTCTCTGTGCACGACAGTTTCGGGACTCACGCA GCTG 3’
TYFFS to DFAFA (1225-1230) mtRNAP	5’ GGAGCAGGTGAAGCGTTCCGACTTCGCCTTCGCCT GATAACTCGAGGCATGCG 3’
S1230A mtRNAP	5’ CCACCTACTTCTTCGCCTGATAACTCGAGGC 3’
S1230stop mtRNAP	5’ CCACCTACTTCTTCTAGTGATAACTCGAGGC 3’
E652K/D653R T7 RNAP	5’ CGTCAACAAGTGCTGAAACGTACCATTCAGCC AGC 3’
AIDS deletion (658-661) T7 RNAP	5’ GATACCATTCAGCCAGGCAAGGGTCTGATG 3’
Δ50 TEFM	5’ GAAGGAGATATACATATGGACGAAAATGCAAAA GAACC 3’
Δ132 TEFM	
Δ144 TEFM	
Δ159 TEFM	5’ GAAGGAGATATACATATGGAACGTGAACGTCTGA

	AAGCCG 3’
140-144 substitution AAGAA in TEFM	5’ GTGTCCGAAAACCGGTGCAGCAGGTGCAGCAAGT C CGGAAAATCG 3’
149-53 substitution AAGAA in TEFM	5’ CGTAAAAGTCCGGAAAATGCAGCAGGTGCAGCAC T GCTGAAACCGG 3’
R152A/K153A in TEFM	5’ CCGGAAAATCGTTTTCTGGCCGCACTGCTGAAACC GGATATTG 3’
Δ71 CSBII	5’ GCCAAAAGATAAAATTTGAAATCGCGGGGGGAGG GGGGGGTTTGG 3’

Scaffold templates used in transcription and cross-linking assays.

R2-8/T1/NT1

5’ CUGCGGCGAU CGATTTCAGACAGGACCC 3’

CGCCGCTACAGCTAAAGTCTGTCCTGGG

R15/T8d/NT2a

5’GAGUCUGCGGCGAUAACGCCAGACG 3’

GCCGCTATTGCGGTCTGC

R14/T2sU+7/NT2

TTTATTTCG

CATGGGGTA ACGCCAGACG

GTACCCCATCGCCGCGCGTGCGGTCTGC

GCGGCGCG 3’ +8

5’AGUCU

R14/T2/NT2sU-2 -2

TTTATTTCG

CATGGGGTA ACGCCAGACG

GTACCCCATCGCCGCGCGTGCGGTCTGC

GCGGCGCG 3’

5’AGUCU

R14sU-13/NT2/NT2

TTTATTTCG

CATGGGGTA ACGCCAGACG

GTACCCCATCGCCGCGCGTGCGGTCTGC

GCGGCGCG 3’

5’AGUCU

-13

bp (bps) – base pairs

CSB – conserved sequence box

CTD – C-terminal domain

D-loop – displacement loop

EC – elongation complex

H – heavy

HSP – heavy strand promoter

L – light

LSP – light strand promoter

mtDNA – mitochondrial DNA

mtRNAP – mitochondrial RNA polymerase

nt (nts) – nucleotides

NT – nontemplate

NTD – N-terminal domain

NTP – nucleotides

PPi – pyrophosphate

RQF – rapid chemical quench flow

T – template

TEFM – mitochondrial transcription elongation factor

T7 RNAP – T7 bacteriophage RNA polymerase

Attributes

All experiments presented in figures and tables were performed by Karen Agaronyan, except experiments in Fig. 30, 32 (Dr. Dmitry Temiakov). Fig. 28 and Fig. 36 (with modifications by Karen Agaronyan) were prepared by Dr. Dmitry Temiakov.