Published in final edited form as:

Nat Struct Mol Biol. 2018 September ; 25(9): 754–765. doi:10.1038/s41594-018-0122-9.

Structural basis of mitochondrial transcription

Hauke S. Hillen[*,1], **Dmitry Temiakov**[2], and **Patrick Cramer**[*,1]

[1]Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Am Fassberg 11, 37077 Göttingen, Germany.

[2]Department of Cell Biology, Rowan University, School of Osteopathic Medicine, 2 Medical Center Dr, Stratford, NJ 08084, USA

Summary

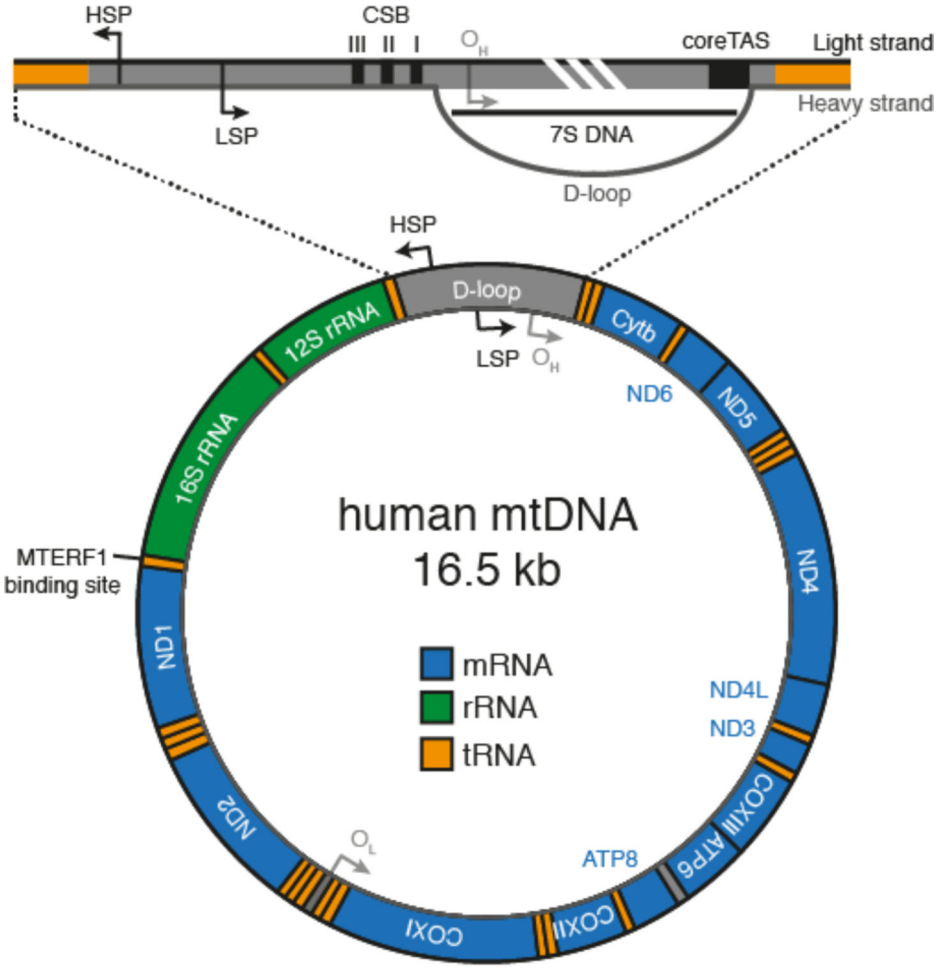
The mitochondrial genome is transcribed by a single-subunit DNA-dependent RNA polymerase (mtRNAP) and its auxiliary factors. Structural studies have elucidated how the auxiliary factors cooperate with mtRNAP to bring about transcription. Initiation factors TFAM and TFB2M assist mtRNAP in promoter DNA binding and opening, respectively, whereas the elongation factor TEFM renders mtRNAP processive, thereby enabling synthesis of the long polycistronic RNA chains. Here we summarize a large body of structural and functional work that has accumulated on human mitochondrial transcription, provide a molecular movie that can be used for teaching purposes, and define open questions and future directions.

Overview

[*]**Corresponding authors:** Hauke S. Hillen (hauke.hillen@mpibpc.mpg.de) and Patrick Cramer (pcramer@mpibpc.mpg.de).

Author contributions

H.S.H. prepared figures and the movie. H.S.H., D.T. and P.C. wrote the manuscript.



Overview figure. Schematic representation of the human mitochondrial genome

The circular human mitochondrial genome is depicted schematically with mRNA coding regions in blue, rRNA coding regions in green, tRNA coding regions in orange and non-coding regions in grey. LSP: light strand promoter; HSP: heavy strand promoter; CSB: conserved sequence block; $O_{H/L}$: origin of replication of the heavy/light strand; Cytb: cytochrome b; ND1/2/3/4/5/6: NADH dehydrogenase subunit 1/2/3/4/5/6; COXI/II/III: cytochrome c oxidase subunit I/II/III; ATP6/8: ATP synthase F0 subunit 6/8. The regulatory region containing the two promoters is shown enlarged above with important elements discussed in the text indicated. Parts of this region frequently form the D-loop, a triple stranded region caused by the presence of a third DNA strand, the 7S DNA, annealed to the light strand.

Mammalian mitochondria contain a small circular genome that encodes for 13 essential proteins of the respiratory chain, 22 transfer RNAs and 2 ribosomal RNAs[1]. The two strands of this genome are distinguished as the heavy and light strands based on their asymmetric purine/pyrimidine distribution. Transcription of the mitochondrial genome is carried out by a dedicated RNA polymerase (mtRNAP) that is related to the single-subunit RNA polymerases in bacteriophages and chloroplasts, and the pol I family of DNA polymerases[2]. Each strand of mtDNA contains a promoter, the heavy or light strand promoter (HSP or LSP,

respectively), which are located within a non-coding regulatory region of mtDNA[3] and drive transcription of polycistronic transcripts spanning nearly the entire genome length[4,5]. This region also contains the origin of replication for the heavy strand and can be partially triple-stranded due to the presence of a third DNA strand called 7S DNA[6]. Unlike other single-subunit RNA polymerases, mtRNAP requires the assistance of protein factors for each step of the transcription cycle[7]. To initiate transcription, mtRNAP assembles with the initiation factors TFAM and TFB2M at the promoter to form the initiation complex. Following initial RNA synthesis, initiation factors are lost and the elongation factor TEFM is recruited as mtRNAP engages in the elongation phase. The termination of transcripts initiated at LSP involves the road block protein MTERF1, which binds at a distinct site downstream of the rRNA genes. However, the precise mechanisms of transcription termination remain unknown.

Introduction

Mitochondria are organelles that carry out oxidative phosphorylation to produce the energy required by eukaryotic cells[8]. In addition, they are involved in numerous cellular processes such as signaling[9], ion homeostasis[10], apoptosis[11] and aging[12]. Mitochondrial dysfunction is associated with various pathological phenotypes in humans[13]. In accordance with their evolutionary origin from an endosymbiontic event[14], mitochondria contain a separate genome and utilize a dedicated, nuclear-encoded molecular machinery for its expression. In humans, the mitochondrial genome comprises 16.5 kb of circular DNA encoding for 13 essential polypeptides, 22 transfer RNAs (tRNAs) and 2 ribosomal RNAs (rRNAs) that are required for the formation of the dedicated mitochondrial ribosome[1].

Transcription of the mammalian mitochondrial genome is carried out by a single-subunit mitochondrial RNA polymerase (mtRNAP or POLRMT) that partly resembles RNA polymerases (RNAPs) from T7-like bacteriophages, but is not related to the multi-subunit RNAPs found in bacteria or the eukaryotic cell nucleus[2]. Transcription initiates from two promoters (light- and heavy-strand promoters or LSP and HSP, respectively) located in close vicinity to each other within the non-coding regulatory region of the mitochondrial genome[3] and results in long poly-cistronic transcripts that are processed to yield mature mRNAs and tRNAs[4,15,16]. This regulatory region frequently contains a triple-stranded structure termed the ‘D-loop’, which is caused by the presence of an additional DNA strand, the 7S DNA[6]. Although its precise role remains obscure, this DNA stretch is thought to originate from abortive or paused DNA replication. Remarkably, the RNA primers required for DNA replication also originate from transcription by mtRNAP[17], placing this enzyme at the heart of both gene expression and genome maintenance in mitochondria[18–20].

In contrast to its self-sufficient bacteriophage relatives, mtRNAP requires the assistance of additional protein factors to initiate transcription, to productively elongate the RNA chain and to terminate transcription[7]. Compared to other transcription systems, the mechanistic and structural basis of mitochondrial transcription remained poorly understood. In recent years, however, numerous structural studies have elucidated various components and diverse stages of the mitochondrial transcription cycle (see Table 1). This review summarizes these

studies and our current understanding of the mechanism of mitochondrial transcription that has resulted from them.

RNA polymerase

Before structural data on mtRNAP became available, sequence analysis suggested that it may share structural similarity with T-odd bacteriophage RNAPs, the most well-studied of which is T7 RNA polymerase[2]. Similar to the distantly related pol A family of DNA polymerases[21,22], the C-terminal catalytic domain (CTD) of these RNAPs resembles a right hand formed by the thumb, palm and fingers domains[23]. The active site is located within the palm domain and utilizes a canonical two-metal dependent reaction mechanism for nucleic acid polymerization[24,25]. The fingers domain harbors the O-helix, a mobile element involved in catalytic functions and in strand separation at the downstream edge of the transcription bubble[26,27]. The O-helix corresponds topologically to the bridge helix found in multi-subunit RNAPs[28]. Promoter binding is mediated by the N-terminal domain (NTD), which contains two nucleic acid-interacting elements, the AT-rich recognition loop and the intercalating hairpin[29,30]. While the former is involved in promoter DNA recognition, the latter separates the DNA strands to form the transcription bubble during initiation[29]. A third element involved in promoter binding, the specificity loop, is located within the CTD and stabilizes the initially melted transcription bubble by binding the major groove upstream of the point of strand separation and guiding the template strand to the active site[29].

The first structure of human mtRNAP revealed that its catalytic domain is indeed similar to bacteriophage T7 RNAP, suggesting a conserved catalytic mechanism[31] (Figure 1). As in T7 RNAP, the conserved fingers, palm and thumb domains of the CTD adopt the canonical right-hand fold[31]. In this apo mtRNAP structure, however, the fingers domain appears rotated to a ‘clenched’ conformation[31]. The N-terminal domain (NTD) is partially similar to the promoter-binding domain of T7 RNA polymerase, but the AT-rich recognition loop differs substantially. Whereas this element contacts the minor groove of promoter DNA in the initiation complex of T7 RNAP[29,30], it is sequestered by intra-protein interactions in mtRNAP and charged DNA-interacting residues are not conserved[31]. In addition to the bacteriophage-like core, mtRNAP contains an N-terminal pentatricopeptide repeat (PPR) domain and a largely flexible N-terminal extension (NTE) which was not resolved in the apo mtRNAP structure although the construct used for crystallization lacked only the first 104 amino acids, 43 of which constitute the mitochondrial targeting signal[31]. PPR domains can be found in many mitochondrial and chloroplast proteins, where they most often mediate protein-RNA interactions[32,33]. The PPR domain of mtRNAP is connected to the N-terminal domain by a proline-rich linker, which potentially confers some degree of mobility to this domain relative to the enzyme core[31].

The other element involved in promoter recognition by T7 RNAP, the specificity loop, was not entirely resolved in the mtRNAP structure[31]. Furthermore, comparison with the T7 initiation complex structure revealed that the intercalating hairpin is positioned differently in the mtRNAP structure[31]. However, mutational analysis showed that both the intercalating hairpin as well as the preceding sequence region (the ‘lever loop’), which is found only in mtRNAPs, are required for promoter-specific initiation[31]. This suggested that initiation

factors may reposition the intercalating hairpin such that it could fulfill the function it has in T7 RNA polymerase, namely to separate the DNA strands upstream of the transcription start site[31].

Taken together, the structure of human apo mtRNAP revealed the architecture of this enzyme, confirmed the principal resemblance of its CTD to bacteriophage RNAPs and provided first clues for the structural basis of factor dependence of transcription initiation. However, it remained unclear how the polymerase would interact with nucleic acids and what the role of its mitochondria-specific features would be.

Transcribing elongation complex

Questions on the basic mechanism of DNA-dependent RNA synthesis by mtRNAP could be addressed with the structure of human mtRNAP in the transcribing state (Figure 2a). This elongation complex (EC) consisted of mtRNAP and a RNA-DNA scaffold mimicking a transcription bubble with the RNA transcript[34]. The mtRNAP variant used in this study lacked part of the N-terminal extension (residues 1–150), which was shown to be dispensable for transcript elongation[34]. Compared to the apo structure, the fingers and palm domains appear rotated, resulting in a conformation that is highly similar to that observed for the T7 RNA polymerase EC[26,27]. The polymerase active center cleft tightly embeds a 9-base pair A-form DNA-RNA hybrid, with the downstream duplex emerging at a perpendicular angle[34]. This arrangement shows a striking resemblance to both T7 RNAP and the structurally unrelated multi-subunit RNAPs in eukaryotes and bacteria[35,36]. The hybrid is stabilized by interactions with the thumb domain, which became fully visible in the EC structure. As in T7 RNAP, separation of the downstream DNA strands is achieved by the O-helix in the fingers domain, which wedges apart the +1 and +2 nucleotides of the template strand[26,27,34,37] (Figure 2b).

The mitochondrial EC structure revealed the pre-translocated state of the nucleotide addition cycle, in which the 3’ end of the nascent transcript occupies the NTP-binding site, thus occluding this site from substrate binding[34] (Figure 2b). Despite the overall similarity to T7 RNA polymerase in the catalytic domain, the mtRNAP EC structure also brought to light tangible mechanistic differences between these enzymes. Most notably, T7 RNAP undergoes a major refolding and re-arrangement of the N-terminal promoter-binding domain during the transition from initiation to elongation[26,27], moving the intercalating hairpin away from the nucleic acids. MtRNAP, however, does not undergo such a conformational change. Instead, the intercalating hairpin remains close to the nucleic acids and peels off the 5’ end of the nascent transcript from the template strand at the upstream edge of the transcription bubble[34]. Although the 5’ end of the nascent RNA was not clearly visible in the EC structure, difference density and crosslinking experiments indicated that it runs along the mobile specificity loop towards the PPR domain[34].

In summary, the EC structure revealed how nucleic acids bind to the polymerase during RNA chain elongation, argued for a conserved catalytic mechanism and revealed mitochondria-specific mechanistic features. The lack of rearrangement of the NTD during the transition from initiation to elongation suggested that initiation factors prime mtRNAP

for initiation and elongation[34]. However, in the absence of structural information on the initiation complex, the mechanism of transcription initiation in mitochondria remained obscure.

Initiation factor TFAM

To initiate transcription, mtRNAP is recruited to the DNA by the promoter-binding protein TFAM to form the closed pre-initiation complex (preIC)[38,39]. Binding of the initiation factor TFB2M then facilitates DNA opening and formation of the open initiation complex (IC)[38–43]. Due to a lack of homology to initiation factors found in other transcription systems, it was not obvious how these proteins would facilitate transcription. This problem was addressed by first solving structures of the two initiation factors TFAM and TFB2M and then also of the initiation complex comprising mtRNAP, TFAM, TFB2M and promoter DNA.

The mammalian mitochondrial transcription factor TFAM belongs to the high mobility group box (HMG box) family of proteins and contains two HMG domains interspersed by a linker region[44]. The ability of TFAM to facilitate transcription initiation is dependent on its C-terminal tail region[40,45]. The binding sites of TFAM at the two promoters have been mapped between DNA bases −10 and −35 upstream of the transcription start site (TSS)[38,46]. In addition to specifically binding to these sites, TFAM also binds double-stranded DNA in a non-specific fashion[47] and this function is thought to mediate packaging of the mitochondrial genome[48–51]. Although a TFAM homolog exists in yeast (Abf2)[52,53], this protein lacks the C-terminal tail found in TFAM and seems to function solely as a DNA packaging factor[45].

The crystal structure of TFAM was solved in complex with a DNA segment encompassing its binding site at the LSP promoter by two groups independently[54,55] (Figure 3a). These structures demonstrated that each of its HMG-box domains induces a ~90° bend in the DNA, suggesting how it can compact the mitochondrial genome into higher-order structures called nucleoids[48–51]. Further structures of TFAM bound to other DNA segments revealed that bending seems to occur independent of the bound DNA sequence[56]. Notably, this study also suggested that TFAM may bind at HSP in a reversed orientation as compared to LSP, thus implying structurally distinct initiation complexes at the two promoters. However, the DNA sequence used only comprised part of the TFAM binding site at the HSP and the proposed model is inconsistent with biochemical and crosslinking studies, which indicate an identical topology for the IC at both promoters[57].

Initiation factor TFB2M

The second mammalian initiation factor, TFB2M, and its paralog TFB1M were originally identified based on sequence homology to the yeast mitochondrial transcription initiation factor Mtf1[58,59]. In yeast, transcription initiation requires only mtRNAP (Rpo41) and Mtf1[60], which was initially hypothesized to resemble the bacterial sigma factor based on partial sequence homology[61]. However, mutational studies did not support a functional homology to sigma factor proteins [62]. The structure of *S. cerevisiae* Mtf1 demonstrated that

it instead shows structural similarity to ErmC, a methyltransferase involved in the maturation of ribosomal RNA[63] (Figure 3b). This surprising finding led to the hypothesis that promoter recognition may reside largely within mtRNAP itself and that potential RNA-binding capabilities of Mtf1 may play a role during transcription initiation[63].

Indeed, sequence and structural homology to rRNA methyltransferases appears to be conserved across mitochondrial initiation factors from different species, as it can also be found in the mammalian TFB-type proteins. While TFB1M appears to have retained its function as a mitochondrial rRNA methyltransferase[64,65], TFB2M is a *bona fide* transcription initiation factor[66]. Recently, the structure of human TFB2M was reported, completing the set of structures of proteins of the initiation machinery[67] (Figure 3b). As in Mtf1, the N-terminal part of TFB2M resembles rRNA methyltransferases, whereas the C-terminal part forms a helical domain with an extended flexible tail[67]. Deletion of a flexible loop specific to mitochondrial transcriptional activator proteins showed no effect on the activity of TFB2M[67], although residues within this loop had been implied in the Mtf1-Rpo41 interaction[68]. Comparison of TFB2M to the structure of mouse TFB1M, which had been reported earlier[69], revealed an overall high structural similarity, but also suggested differences which may account for their diverging functions[67]. Most prominently, TFB2M contains a loop insertion in its N-terminal domain which is absent in both TFB1M and Mtf1 and contributes to an extended positively charged surface patch on TFB2M. Moreover, a number of positively charged residues constituting this charged groove are conserved in TFB2M between humans and mouse, but not in TFB1M[67].

Taken together, these studies revealed the structures of the two initiation factors TFAM and TFB2M but did not explain how these factors interact with mtRNAP and with nucleic acids to assemble the initiation complex. Therefore, the mechanism of mitochondrial transcription initiation remained unclear.

Initiation complex

The recent structure of the mitochondrial transcription initiation complex revealed the interactions between TFAM, TFB2M, mtRNAP and DNA, showing how the two initiation factors cooperate with the polymerase to recruit, position and melt promoter DNA[67] (Figure 4). Consistent with biochemical data[39,70], the structure demonstrated that mtRNAP utilizes its N-terminal extension to stack against the distal HMG-Box domain of TFAM in the IC, which is brought in position by the TFAM-induced DNA bending[67]. This interaction is mediated by a previously unresolved structural element within the N-terminal extension of mtRNAP, the ‘tether helix’[67], and seems to position the active site of mtRNAP at the region of initial DNA melting around the TSS[71]. Tethering of the N-terminal extension of mtRNAP to TFAM explains how TFAM binding may compensate for the absence of the AT-rich recognition loop in mtRNAP, as suggested previously[30,31]. In T7 RNAP, this element contacts the promoter DNA region 13 −17 bp upstream of the TSS. In the mitochondrial IC, the DNA region between bases −10 and −15 is positioned in proximity to the PPR domain, where DNA backbone interactions may contribute to the DNA bending observed in this region[67]. The C-terminal tail of TFAM is positioned in immediate vicinity of mtRNAP, where it mediates additional interactions with the ‘D helix’ and/or parts of the N-terminal

extension of mtRNAP[40,67]. Contrary to an earlier model based on low-resolution electron microscopic data[72] TFB2M interacts with the DNA around the point of strand separation with its C-terminal helical domain[67]. Comparison of IC structures assembled on DNA scaffolds representing the LSP and HSP promoter sequences confirmed the identical topology of the IC at both promoters[57,67].

The conformation of mtRNAP in the open IC resembles that observed for T7 RNAP in its respective initiation complex[30]. In particular, similarly to the T7 RNAP IC, the intercalating hairpin is positioned to separate the DNA strands at position −4 relative to the TSS. This conformation is stabilized by the interaction of TFB2M with the polymerase, likely locking it in the melting-competent state[67]. This seems to be achieved by an interaction between the C-terminal helical domain of TFB2M and a mtRNAP-specific ‘lever loop’ adjacent to the intercalating hairpin, explaining why a deletion mutant lacking this region is incapable of promoter-specific initiation[31]. In addition, the flexible C-terminal tail of TFB2M buttresses the intercalating hairpin and appears partially ordered in the IC structure[67]. The rRNA methyltransferase-like domain of TFB2M is positioned along the trajectory of the single-stranded non-template strand, embedding it in an extensive positively charged surface[67]. Thus, it seems that the basic surface groove that binds nucleic acid in the structurally related rRNA methyltransferases is utilized by TFB2M to stabilize the initially opened DNA.

This mechanism of non-template DNA binding by TFB2M is somewhat reminiscent to that of bacterial sigma factor, although structural similarity is lacking. Sigma interacts with the non-template strand both non-specifically and via a base-specific flipping mechanism to stabilize the initially unwound DNA[73–75]. Whether TFB2M also interacts with the nucleic acid in a base-specific fashion remains to be seen, as the reported resolution of the IC structure was not sufficient to detect such interactions. This also holds true for potential interactions of the specificity loop with promoter DNA, which are mediated by extensive base contacts in the case of the T7 RNAP IC[29]. In the mitochondrial IC, this element is positioned similarly along the major groove of the DNA just upstream of the point of strand separation[67]. However, the structure could not elucidate whether base-specific contacts are formed. Thus, while it is clear that the specificity loop is required for promoter-directed initiation by mtRNAP[40,76], it remains open whether it plays a significant role in promoter recognition.

In conclusion, the structure of the IC with open promoter DNA revealed how the initiation factors facilitate transcription initiation and provided rationale for the factor-dependence of mtRNAP for this step. Furthermore, it revealed the role of the N-terminal extension of mtRNAP. In conjunction with previous studies this has led to a basic mechanistic understanding of mammalian mitochondrial transcription initiation.

Elongation factor TEFM

Following initiation of RNA synthesis, the initiation factors are released during promoter escape and the elongation factor TEFM must be recruited. This factor was originally identified by its sequence homology with RuvC-like Holliday junction resolvases and a limited sequence relationship with the eukaryotic nuclear transcription factor Spt6[77]. TEFM

is required for processive synthesis of the near genome-length mitochondrial transcripts[77–79]. In the absence of this factor, mtRNAP prematurely terminates at a conserved G-quadruplex-forming sequence downstream of the LSP (the conserved sequence block II, CSBII; see also 2YHUYLHZ.) This yields the short transcripts that were suggested to serve as primers for replication of the heavy strand of mtDNA[17,18,80,81]. Thus, it appears that transcription and replication primer formation are mutually exclusive processes and that TEFM may be involved in switching between the two[20,78].

The structure of TEFM confirmed the predicted combination of a C-terminal Holliday junction resolvase-like homo-dimerisation domain with a N-terminal helix-hairpin-helix domain[82] (Figure 5a). However, it also revealed that TEFM is a pseudonuclease, as it does not possess endonucleolytic activity and the active site appears to be degenerate[82]. Thus, the Holliday junction resolvase fold seems to be utilized for a function in mitochondrial transcription. Interestingly, the closest structural homolog of TEFM, the mitochondrial Holliday junction resolvase Cce1 in yeast, is a functional nuclease required for replication of mitochondrial genomes[83,84]. Biochemical characterization of TEFM demonstrated that the C-terminal pseudonuclease domain and a short, basic linker connecting it to the N-terminal domain are sufficient to prevent termination at the CSBII-site and to enhance mtRNAP processivity[82]. Thus, the N-terminal domain seems dispensable for all observed activities of TEFM *in vitro*[82].

The structure of TEFM revealed that this protein forms a homodimer via its C-terminal pseudonuclease domains, whereas the two N-terminal domains are connected to it via a flexible, basic linker. However, these structural data could not explain how TEFM exerts its profound effect on the transcriptional activity of mtRNAP.

Processive elongation complex

The crystal structure of the processive elongation complex, also called anti-termination complex, consists of mtRNAP and the transcriptionally active core of TEFM[82]. The structure revealed that TEFM forms extensive contacts with the nucleic acids in the transcription bubble[82] (Figure 5b). Notably, these interactions distantly resemble those observed between Holliday junction resolvases and their cruciform substrates[85], suggesting that this fold was re-purposed to bind the transcription bubble. Recruitment of TEFM to the EC encloses the downstream DNA duplex in a sliding clamp between TEFM and the polymerase and stabilizes the single-stranded non-template strand, reminiscent of elongation factors from other transcription systems such as NusG (bacteria)[86] or Spt4/5 and DSIF (RNA polymerase II system in yeast and human, respectively)[87,88]. This explains how TEFM enhances general processivity of the polymerase. Notably, mtRNAP was found to be in the post-translocated state in the processive EC structure, further suggesting that TEFM may stabilize this conformation (Figure 5c)[82].

In addition, TEFM binds to the intercalating hairpin of mtRNAP[82], which separates the nascent RNA from the template strand in the EC[34], and cross-linking indicated that it may also interact with the specificity loop[82] (Figure 5d). This suggests how TEFM can exert its striking effect on transcription through the CSBII sequence. By ensuring maintenance of the

upstream edge of the transcription bubble and creating a narrow RNA exit path along mtRNAP, binding of TEFM is likely to prevent formation of bulky secondary structures, such as G-quadruplexes, in the nascent RNA, which would otherwise disrupt the EC and lead to termination of transcription[82].

In summary, the structures of TEFM and of the anti-termination complex provided the structural and mechanistic basis for processive transcription elongation. Furthermore, they illustrated how TEFM enables mtRNAP to transcribe through the CSBII region, driving gene expression over replication primer formation.

The transition from initiation to elongation

The availability of structures of the IC and the processive EC provided insights into the structural changes that must occur during the initiation-elongation transition. First, the initiation factor TFB2M and the elongation factor TEFM occupy overlapping binding sites on mtRNAP, and thus TFB2M must dissociate before TEFM can bind[67,82]. Second, the upstream DNA duplex must undergo a substantial repositioning from its location in the IC to the position observed in the EC[34,67]. This is likely to occur with concurrent loss of TFB2M, thus liberating the binding site for TEFM. Third, the interaction of TFAM with the tether helix must be broken, but it remains unclear when this occurs. Taken together, the initiation-elongation transition involves an exchange of factors that use overlapping binding sites on the polymerase, resembling the situation in multi-subunit polymerase systems[88–91].

Transcription termination

At the end of each transcription cycle, mtRNAP is expected to cease RNA synthesis and dissociate from mtDNA. One termination site has been identified within the $tRNA^{Leu}$ gene[92], which serves as a binding site for the 39 kDa protein MTERF1[93]. This factor exhibits strong polarity in terminating transcription initiated at the LSP but not at the HSP promoter[94–97]. Indeed, the light strand does not encode any genes downstream of the MTERF1 binding site. Moreover, knockout of MTERF1 results in impaired transcription from LSP but not from HSP [98]. This suggests that MTERF1 may act to prevent interference of the transcribing complexes with initiation within the control region[98]. In addition, MTERF1 also seems to affect mtDNA replication, as it has been shown to act as a ‘contrahelicase’ causing the mitochondrial replisome to pause[99]. This may help to avoid collisions between the transcription and replication machineries[99].

MTERF1 belongs to a highly conserved family of proteins that localize to mitochondria and have been implicated in DNA binding[100]. The structure of MTERF1 bound to the $tRNA^{Leu}$ termination sequence revealed a modular structure of this protein, which harbors 8 motifs, each composed of two α-helices followed by a 3_{10} helix[94,101] (Figure 6). MTERF1 binds along the major groove of DNA, unwinds and partially melts it by flipping out three nucleotides. Base flipping is essential for stable binding and sequence-specific recognition of DNA and explains how MTERF1 acts as a roadblock for the transcribing polymerase, thereby mediating transcription termination[94,101]. The structure of MTERF1 bound to DNA

also provided a structural rational for two known pathogenic mutations in the mitochondrial genome, which apparently impair transcription termination[94].

In contrast to the well-defined mechanism of LSP transcript termination by MTERF1, the mechanism of heavy strand transcription termination remains unknown. It has been suggested that, similar to the case with LSP, termination involves a roadblock protein that binds near the 3' end of the 7S DNA[102]. This region, termed coreTAS, does not form a secondary structure but shares sequence similarity with a conserved sequence in the regulatory D-loop region called CSBI. As of today, no protein factors that would mediate termination of H-strand transcription have been identified.

In mammals, several other MTERF proteins (MTERF2–4) exist which are all involved in mitochondrial gene expression[100], but do not appear to play a direct role in transcription termination[103,104,105,106].

A movie of mitochondrial transcription

Over the last decade, the structural basis of mitochondrial transcription has emerged. Together with biochemical data[3,38–43,45–47,57,58,62,70,71,76,77,79,80,93,94,107–110], the structural studies reviewed here have led to a model of the mammalian mitochondrial transcription cycle (Figure 7). Based on these insights we prepared a molecular movie that illustrates the key known steps in the mitochondrial transcription cycle (Supplemental Movie 1). To initiate transcription, TFAM binds to the promoter upstream of the TSS, which may involve sliding of TFAM along the mitochondrial genome, scanning for promoter DNA[111]. MtRNAP is then recruited to form the closed promoter pre-initiation complex, which, with the help of TFB2M, transitions to the open promoter complex[39]. Following initial RNA synthesis, the initiation factors are released and TEFM is recruited to the transcribing polymerase. With the help of TEFM, mtRNAP can then synthesize complete RNA transcripts. Termination of transcription is mediated by MTERF1 and potentially other, yet unidentified factors. The fact that mtRNAP has evolved to depend on these factors likely reflects the need to regulate mitochondrial gene expression. The recent structural studies provide a framework to analyze such factor-dependent regulation.

Comparison with other single-subunit RNA polymerase systems

The available structures also shed light on the mechanistic differences between mtRNAP and the related single-subunit phage RNA polymerases. In conjunction with biochemical data, it becomes evident that promoter recognition in the mitochondrial system is not based on the structural elements shared with T7 RNAP but instead relies on recognition of a ~50–55 bp promoter region by cooperative interaction of initiation factors and mtRNAP. The structures of the initiation factors and of the initiation complex demonstrate how these factors substitute for the lack of T7 RNAP-like elements in mtRNAP and how mitochondria-specific elements such as the lever loop and the N-terminal region are involved in this process[67].

Unlike T7 RNAP, which undergoes significant refolding during the transition from initiation to elongation, progression of mtRNAP to elongation involves loss of initiation factors and a

rearrangement of the upstream DNA duplex to the position occupied by TFB2M in the initiation complex[67]. The identification and characterization of the elongation factor TEFM[77–79,82] revealed another striking difference between mtRNAP and T7 RNAP, which does not rely on processivity factors. The structure of the processive elongation complex suggests that TEFM renders the EC more stable by stabilizing the RNA exit path[82], which may be particularly crucial for transcription of the mitochondrial genome because it encodes highly structured tRNAs and rRNAs. This may also in part explain the dependence of mtRNAP on roadblock termination factors, as it tolerates less bulky secondary structures such as hairpin terminators [82].

Evolutionary considerations

Taken together, a picture has emerged that illuminates how a single-subunit RNAP has evolved to depend on protein factors for its function, many of which show some degree of similarity in their mode of action to transcription factors from the multi-subunit RNAP transcription systems. However, these mitochondrial factors do not share sequence homology with their functional counterparts in other transcription systems. Instead, they appear to utilize originally unrelated protein folds, which may have evolved for a role in transcription through gene duplication events[112]. Thus, the mitochondrial transcription system appears to have evolutionarily converged on mechanistic principles conserved in all DNA-dependent RNA synthesizing machineries.

The presence of a phage-like mitochondrial RNAP is virtually ubiquitous among eukaryotic organisms, suggesting that this transcription system was adopted very early in eukaryotic evolution[113,114]. However, whether this happened before or after the endosymbiotic event remains enigmatic[115]. Strikingly, the mitochondrial genomes of some jakobids encode all four core subunits of a eubacterial multi-subunit RNAP[116,117]. It is unclear, however, whether this eubacterial RNAP is employed to transcribe the mtDNA[117]. Chloroplasts contain both a phage-like polymerase and a eubacterial-like RNAP, which both seem to be utilized in organellar gene expression[118]. Notably, the replicative apparatus in mitochondria is also related to the corresponding machinery from T-odd bacteriophages[119–121].

Open questions & future directions

Despite the structural advances described here, our understanding of the mitochondrial transcription system remains incomplete. Further structural and biochemical studies will be necessary to obtain a detailed mechanistic view of this evolutionarily unique system.

One open question regards the precise mechanism of promoter recognition and melting. The open initiation complex structure revealed substantial distortions in the DNA[67], which may facilitate initial duplex melting as in multi-subunit RNAPs[122,123], but it is unclear at what stage they are induced. It further remains unclear whether the interaction of TFAM with promoter DNA is the sole determinant of promoter recognition. Biochemical data suggest that mtRNAP may engage in sequence-specific interactions within the initially unwound DNA region[38] and that the N-terminus of TFB2M interacts with the priming nucleotide[43], but these interactions were not resolved in the IC structure. Higher resolution structures of

the closed pre-initiation complex and of an initially transcribing complex (ITC) will be necessary to elucidate all details of transcription initiation.

It also remains open how the unique PPR domain of mtRNAP functions. During initiation, this domain is involved in upstream promoter contacts, albeit likely non-specifically[67]. During elongation, it has been suggested that the growing RNA chain may interact with the PPR domain[34]. However, no direct structural evidence has been obtained for such an interaction. Another interesting possibility is that the PPR domain may be involved in protein-protein rather than in protein-RNA contacts, as suggested recently for the mitochondrial protein LRPPRC and its interaction partner SLIRP[124]. Thus, the PPR domain of mtRNAP may serve as a platform to recruit factors involved in co-transcriptional processes.

Whereas the structure of MTERF1 suggests how this protein facilitates termination of L-strand transcription, the mechanism of H-strand termination remains elusive. Future efforts should be directed at identifying the factors involved in H-strand termination and developing a unified model for transcription termination in mammalian mitochondria.

Finally, it will be necessary to investigate how transcription is coordinated with RNA processing and translation. Mitochondrial ribosome biogenesis, for example, has recently been shown to occur co-transcriptionally[125]. Likewise, we need to study how transcription relates to mtDNA replication. The finding that mtRNAP acts as the primase at both origins of replication[17,18,81] suggests that some sort of crosstalk exists between the transcription and replication machineries. Although transcription termination at CSBII seems to be responsible for replication primer formation[80,126] and TEFM efficiently prevents such termination *in vitro*[78,79,127], it remains unclear how the primer is processed and handed over to the replication machinery *in vivo*. Further, it remains open how TEFM is regulated and whether additional proteins are involved in the transcription-replication switch *in vivo*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We wish to thank current and past members of the Cramer and Temiakov laboratories for critical discussions and valuable comments. H.S.H. was supported by a Boehringer Ingelheim Fonds PhD fellowship. P.C. was supported by the Deutsche Forschungsgemeinschaft (SFB860, SPP1935), the European Research Council Advanced Investigator Grant TRANSREGULON (grant agreement No 693023) and the Volkswagen Foundation. D.T. was supported by NIH RO1 GM104231 and R01 GM118941.

References

1. Falkenberg M, Larsson N-G & Gustafsson CM DNA replication and transcription in mammalian mitochondria. Annu. Rev. Biochem 76, 679–699 (2007). [PubMed: 17408359]
2. Masters BS, Stohl LL & Clayton DA Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. Cell 51, 89–99 (1987). [PubMed: 3308116]
3. Chang DD & Clayton DA Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell 36, 635–643 (1984). [PubMed: 6697390]

4. Aloni Y & Attardi G Expression of the mitochondrial genome in HeLa cells. II. Evidence for complete transcription of mitochondrial DNA. J. Mol. Biol 55, 251–267 (1971). [PubMed: 5548607]
5. Murphy WI, Attardi B, Tu C & Attardi G Evidence for complete symmetrical transcription in vivo of mitochondrial DNA in HeLa cells. J. Mol. Biol 99, 809–814 (1975). [PubMed: 1214305]
6. Kasamatsu H, Robberson DL & Vinograd J A novel closed-circular mitochondrial DNA with properties of a replicating intermediate. PNAS 68, 2252–2257 (1971). [PubMed: 5289384]
7. Gustafsson CM, Falkenberg M & Larsson N-G Maintenance and Expression of Mammalian Mitochondrial DNA. Annu. Rev. Biochem 85, 133–160 (2016). [PubMed: 27023847]
8. Ernster L & Schatz G Mitochondria: a historical review. J. Cell Biol 91, 227s–255s (1981). [PubMed: 7033239]
9. Hamanaka RB & Chandel NS Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes. Trends in Biochemical Sciences 35, 505–513 (2010). [PubMed: 20430626]
10. Pozzan T & Rizzuto R High tide of calcium in mitochondria. Nat. Cell Biol 2, E25–7 (2000). [PubMed: 10655598]
11. Kroemer G & Reed JC Mitochondrial control of cell death. Nat. Med 6, 513–519 (2000). [PubMed: 10802706]
12. Sun N, Youle RJ & Finkel T The Mitochondrial Basis of Aging. Molecular Cell 61, 654–666 (2016). [PubMed: 26942670]
13. Nunnari J & Suomalainen A Mitochondria: in sickness and in health. Cell 148, 1145–1159 (2012). [PubMed: 22424226]
14. Gray MW, Burger G & Lang BF Mitochondrial evolution. Science 283, 1476–1481 (1999). [PubMed: 10066161]
15. Aloni Y & Attardi G Symmetrical in vivo transcription of mitochondrial DNA in HeLa cells. Proc. Natl. Acad. Sci. U.S.A 68, 1757–1761 (1971). [PubMed: 5288761]
16. Ojala D, Montoya J & Attardi G tRNA punctuation model of RNA processing in human mitochondria. Nature 290, 470–474 (1981). [PubMed: 7219536]
17. Chang DD & Clayton DA Priming of human mitochondrial DNA replication occurs at the light-strand promoter. Proc. Natl. Acad. Sci. U.S.A 82, 351–355 (1985). [PubMed: 2982153]
18. Wanrooij S et al. Human mitochondrial RNA polymerase primes laggingstrand DNA synthesis in vitro. Proc. Natl. Acad. Sci. U.S.A 105, 11122–11127 (2008). [PubMed: 18685103]
19. Fusté JM et al. Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. Molecular Cell 37, 67–78 (2010). [PubMed: 20129056]
20. Kühl I et al. POLRMT regulates the switch between replication primer formation and gene expression of mammalian mtDNA. Sci Adv 2, e1600963–e1600963 (2016). [PubMed: 27532055]
21. Sousa R Structural and mechanistic relationships between nucleic acid polymerases. Trends in Biochemical Sciences 21, 186–190 (1996). [PubMed: 8871404]
22. McAllister WT & Raskin CA The phage RNA polymerases are related to DNA polymerases and reverse transcriptases. Molecular Microbiology 10, 1–6 (1993). [PubMed: 7526118]
23. Sousa R, Chung YJ, Rose JP & Wang BC Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution. Nature 364, 593–599 (1993). [PubMed: 7688864]
24. Steitz TA, Smerdon SJ, Jäger J & Joyce CM A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. Science 266, 2022–2025 (1994). [PubMed: 7528445]
25. Sosunov V et al. Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. EMBO J 22, 2234–2244 (2003). [PubMed: 12727889]
26. Tahirov TH et al. Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. Nature 420, 43–50 (2002). [PubMed: 12422209]
27. Yin YW & Steitz TA Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science 298, 1387–1395 (2002). [PubMed: 12242451]
28. Cramer P Common structural features of nucleic acid polymerases. Bioessays 24, 724–729 (2002). [PubMed: 12210533]

29. Cheetham GM, Jeruzalmi D & Steitz TA Structural basis for initiation of transcription from an RNA polymerase-promoter complex. Nature 399, 80–83 (1999). [PubMed: 10331394]
30. Cheetham GM & Steitz TA Structure of a transcribing T7 RNA polymerase initiation complex. Science 286, 2305–2309 (1999). [PubMed: 10600732]
31. Ringel R et al. Structure of human mitochondrial RNA polymerase. Nature 478, 269–273 (2011). [PubMed: 21947009] The crystal structure of human mitochondrial RNA polymerase demonstrated its resemblance to phage RNAPs and highlighted differences that may explain factor dependence of mtRNAP for initiation.
32. Lightowlers RN & Chrzanowska-Lightowlers ZMA PPR (pentatricopeptide repeat) proteins in mammals: important aids to mitochondrial gene expression. Biochem. J 416, e5–6 (2008). [PubMed: 18939947]
33. Filipovska A & Rackham O Pentatricopeptide repeats: modular blocks for building RNA-binding proteins. RNA Biol 10, 1426–1432 (2013). [PubMed: 23635770]
34. Schwinghammer K et al. Structure of human mitochondrial RNA polymerase elongation complex. Nat. Struct. Mol. Biol 20, 1298–1303 (2013). [PubMed: 24096365] The structure of mtRNAP bound to nucleic acids revealed the overall architecture of the elongation complex and highlighted conserved catalytic mechanisms and striking differences to phage RNAPs.
35. Gnatt AL, Cramer P, Fu J, Bushnell DA & Kornberg RD Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science 292, 1876–1882 (2001). [PubMed: 11313499]
36. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH & Artsimovitch I Structural basis for transcription elongation by bacterial RNA polymerase. Nature 448, 157–162 (2007). [PubMed: 17581590]
37. Temiakov D et al. Structural basis for substrate selection by t7 RNA polymerase. Cell 116, 381–391 (2004). [PubMed: 15016373]
38. Gaspari M, Falkenberg M, Larsson N-G & Gustafsson CM The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. EMBO J 23, 4606–4614 (2004). [PubMed: 15526033]
39. Morozov YI et al. A novel intermediate in transcription initiation by human mitochondrial RNA polymerase. Nucleic acids research 42, 3884–3893 (2014). [PubMed: 24393772] This study provided the first evidence of a direct interaction between mtRNAP and TFAM and described the pre-initiation complex.
40. Morozov YI et al. A model for transcription initiation in human mitochondria. Nucleic acids research 43, 3726–3735 (2015). [PubMed: 25800739] Biochemical study mapping the interactions between mtRNAP and initiation factors in the IC.
41. Posse V & Gustafsson CM Human Mitochondrial Transcription Factor B2 is Required for Promoter Melting During Initiation of Transcription. J. Biol. Chem 292, jbc.M116.751008–2645 (2016).This study provided biochemical evidence that TFB2M is required for the initial melting of promoter DNA.
42. Ramachandran A, Basu U, Sultana S, Nandakumar D & Patel SS Human mitochondrial transcription factors TFAM and TFB2M work synergistically in promoter melting during transcription initiation. Nucleic acids research 45, 861–874 (2016). [PubMed: 27903899]
43. Sologub M, Litonin D, Anikin M, Mustaev A & Temiakov D TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase. Cell 139, 934–944 (2009). [PubMed: 19945377]
44. Parisi MA & Clayton DA Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science 252, 965–969 (1991). [PubMed: 2035027]
45. Dairaghi DJ, Shadel GS & Clayton DA Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. J. Mol. Biol 249, 11–28 (1995). [PubMed: 7776365] Elegant biochemical demonstration that the C-terminal tail of TFAM is required for transcription activation.
46. Fisher RP, Topper JN & Clayton DA Promoter selection in human mitochondria involves binding of a transcription factor to orientationindependent upstream regulatory elements. Cell 50, 247–258 (1987). [PubMed: 3594571]

47. Fisher RP & Clayton DA Purification and characterization of human mitochondrial transcription factor 1. Mol. Cell. Biol 8, 3496–3509 (1988). [PubMed: 3211148]
48. Alam TI et al. Human mitochondrial DNA is packaged with TFAM. Nucleic acids research 31, 1640–1645 (2003). [PubMed: 12626705]
49. Kanki T et al. Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA. Mol. Cell. Biol 24, 9823–9834 (2004). [PubMed: 15509786]
50. Kaufman BA et al. The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. Mol. Biol. Cell 18, 3225–3236 (2007). [PubMed: 17581862]
51. Kukat C et al. Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. Proc. Natl. Acad. Sci. U.S.A 112, 11288–11293 (2015). [PubMed: 26305956]
52. Diffley JF & Stillman B A close relative of the nuclear, chromosomal highmobility group protein HMG1 in yeast mitochondria. PNAS 88, 7864–7868 (1991). [PubMed: 1881919]
53. Fisher RP, Lisowsky T, Breen GA & Clayton DA A rapid, efficient method for purifying DNA-binding proteins. Denaturation-renaturation chromatography of human and yeast mitochondrial extracts. J. Biol. Chem 266, 9153–9160 (1991). [PubMed: 1851168]
54. Rubio-Cosials A et al. Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. Nat. Struct. Mol. Biol 18, 1281–1289 (2011). [PubMed: 22037172] The structure of human TFAM bound to a DNA stretch encompassing ist binding site at LSP reveal that it induces a 180° turn in the DNA (see also ref. 55).
55. Ngo HB, Kaiser JT & Chan DC The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. Nat. Struct. Mol. Biol 18, 1290–1296 (2011). [PubMed: 22037171] The structure of human TFAM bound to a DNA stretch encompassing ist binding site at LSP reveal that it induces a 180° turn in the DNA (see also ref. 54).
56. Ngo HB, Lovely GA, Phillips R & Chan DC Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. Nat Comms 5, (2014).
57. Morozov YI & Temiakov D Human Mitochondrial Transcription Initiation Complexes Have Similar Topology on the Light and Heavy Strand Promoters. J. Biol. Chem 291, 13432–13435 (2016). [PubMed: 27226527]
58. Falkenberg M et al. Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nat. Genet 31, 289–294 (2002). [PubMed: 12068295]
59. McCulloch V, Seidel-Rogol BL & Shadel GS A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. Mol. Cell. Biol 22, 1116–1125 (2002). [PubMed: 11809803]
60. Xu B & Clayton DA Assignment of a yeast protein necessary for mitochondrial transcription initiation. Nucleic acids research 20, 1053–1059 (1992). [PubMed: 1549466]
61. Jang SH & Jaehning JA The yeast mitochondrial RNA polymerase specificity factor, MTF1, is similar to bacterial sigma factors. J. Biol. Chem 266, 22671–22677 (1991). [PubMed: 1939277]
62. Shadel GS & Clayton DA A Saccharomyces cerevisiae mitochondrial transcription factor, sc-mtTFB, shares features with sigma factors but is functionally distinct. Mol. Cell. Biol 15, 2101–2108 (1995). [PubMed: 7891705]
63. Schubot FD et al. Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. Protein Sci 10, 1980–1988 (2001). [PubMed: 11567089] The structure of the yeast mitochondrial transcription initiation factor Mtf1 demonstrated it similarity to bacterial rRNA methyltransferases.
64. Seidel-Rogol BL, McCulloch V & Shadel GS Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. Nat. Genet 33, 23–24 (2003). [PubMed: 12496758]
65. Metodiev MD et al. Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome. Cell Metabolism 9, 386–397 (2009). [PubMed: 19356719]
66. Litonin D et al. Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. J. Biol. Chem 285, 18129–18133 (2010). [PubMed: 20410300]

67. Hillen HS, Morozov YI, Sarfallah A, Temiakov D & Cramer P Structural Basis of Mitochondrial Transcription Initiation. Cell 171, 1072–1081.e10 (2017). [PubMed: 29149603] Structures of TFB2M and the mitochondrial transcription initiation complex elucidated how the intiation factors interact with mtRNAP to facilitate transcription initiation.
68. Cliften PF, Park JY, Davis BP, Jang SH & Jaehning JA Identification of three regions essential for interaction between a sigma-like factor and core RNA polymerase. Genes Dev 11, 2897–2909 (1997). [PubMed: 9353258]
69. Guja KE et al. Structural basis for S-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1. Nucleic acids research 41, 7947–7959 (2013). [PubMed: 23804760]
70. Posse V et al. The amino terminal extension of mammalian mitochondrial RNA polymerase ensures promoter specific transcription initiation. Nucleic acids research 1–10 (2014).
71. Dairaghi DJ, Shadel GS & Clayton DA Human mitochondrial transcription factor A and promoter spacing integrity are required for transcription initiation. Biochim. Biophys. Acta 1271, 127–134 (1995). [PubMed: 7599198]
72. Yakubovskaya, E; Organization of the human mitochondrial transcription initiation complex.. Nucleic acids research.
73. Feklistov A & Darst SA Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. Cell 147, 1257–1269 (2011). [PubMed: 22136875]
74. Zhang Y et al. Structural Basis of Transcription Initiation. Science 338, 1076–1080 (2012). [PubMed: 23086998]
75. Murakami KS & Darst SA Bacterial RNA polymerases: the wholo story. Curr. Opin. Struct. Biol 13, 31–39 (2003). [PubMed: 12581657]
76. Nayak D, Guo Q & Sousa R A promoter recognition mechanism common to yeast mitochondrial and phage t7 RNA polymerases. J. Biol. Chem 284, 13641–13647 (2009). [PubMed: 19307179]
77. Minczuk M et al. TEFM (c17orf42) is necessary for transcription of human mtDNA. Nucleic acids research 39, 4284–4299 (2011). [PubMed: 21278163] This study identified TEFM as a mitochondrial transcription elongation factor required for processive transcription of mtDNA.
78. Agaronyan K, Morozov YI, Anikin M & Temiakov D Replicationtranscription switch in human mitochondria. Science 347, 548–551 (2015). [PubMed: 25635099]
79. Posse V, Shahzad S, Falkenberg M, Hallberg BM & Gustafsson CMTEFM is a potent stimulator of mitochondrial transcription elongation in vitro. Nucleic acids research (2015). doi:10.1093/nar/gkv105
80. Wanrooij PH, Uhler JP, Simonsson T, Falkenberg M & Gustafsson CM G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. Proc. Natl. Acad. Sci. U.S.A 107, 16072–16077 (2010). [PubMed: 20798345] This study provided in vitro evidence that transcription termination at CSBII is mediated by formation of a G-quadruplex in the nascent RNA.
81. Kang D, Miyako K, Kai Y, Irie T & Takeshige K In vivo determination of replication origins of human mitochondrial DNA by ligation-mediated polymerase chain reaction. J. Biol. Chem 272, 15275–15279 (1997). [PubMed: 9182553]
82. Hillen HS et al. Mechanism of Transcription Anti-termination in Human Mitochondria. Cell 171, 1082–1093.e13 (2017). [PubMed: 29033127] Biochemical and structural characterization of TEFM domains and of the processive anti-termination complex revealed how TEFM confers processivity to mtRNAP.
83. Lockshon D et al. A role for recombination junctions in the segregation of mitochondrial DNA in yeast. Cell 81, 947–955 (1995). [PubMed: 7781070]
84. Ceschini S Crystal structure of the fission yeast mitochondrial Holliday junction resolvase Ydc2. EMBO J 20, 6601–6611 (2001). [PubMed: 11726496]
85. Górecka KM, Komorowska W & Nowotny M Crystal structure of RuvC resolvase in complex with Holliday junction substrate. Nucleic acids research 41, 9945–9955 (2013). [PubMed: 23980027]
86. Liu B & Steitz TA Structural insights into NusG regulating transcription elongation. Nucleic acids research 45, 968–974 (2017). [PubMed: 27899640]

87. Martinez-Rucobo FW, Sainsbury S, Cheung ACM & Cramer P Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. EMBO J 30, 1302–1310 (2011). [PubMed: 21386817]

88. Bernecky C, Plitzko JM & Cramer P Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp. Nature Publishing Group 24, 809–815 (2017).

89. Werner F, Werner F 2012 A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. J. Mol. Biol 417, 13–27 (2012). [PubMed: 22306403]

90. Grohmann D et al. The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. Molecular Cell 43, 263–274 (2011). [PubMed: 21777815]

91. Plaschka C et al. Transcription initiation complex structures elucidate DNA opening. Nature 533, 353–358 (2016). [PubMed: 27193681]

92. Montoya J, Gaines GL & Attardi G The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. Cell 34, 151–159 (1983). [PubMed: 6883508]

93. Kruse B, Narasimhan N & Attardi G Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. Cell 58, 391–397 (1989). [PubMed: 2752429]

94. Yakubovskaya E, Mejia E, Byrnes J, Hambardjieva E & Garcia-Diaz M Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. Cell 141, 982–993 (2010). [PubMed: 20550934] The crystal structure of MTERF1 bound to a sequence from the tRNALeu gene revealed how it wraps around the DNA and flips out bases from the DNA duplex (see also ref. 101).

95. Asin-Cayuela J, Schwend T, Farge G & Gustafsson CM The human mitochondrial transcription termination factor (mTERF) is fully active in vitro in the non-phosphorylated form. J. Biol. Chem 280, 25499–25505 (2005). [PubMed: 15899902]

96. Shang J & Clayton DA Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro. J. Biol. Chem 269, 29112–29120 (1994). [PubMed: 7525579]

97. Christianson TW & Clayton DA In vitro transcription of human mitochondrial DNA: accurate termination requires a region of DNA sequence that can function bidirectionally. PNAS 83, 6277–6281 (1986). [PubMed: 3018722]

98. Terzioglu M et al. MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation. Cell Metabolism 17, 618–626 (2013). [PubMed: 23562081]

99. Shi Y et al. Mitochondrial transcription termination factor 1 directs polar replication fork pausing. Nucleic acids research 44, 5732–5742 (2016). [PubMed: 27112570]

100. Linder T et al. A family of putative transcription termination factors shared amongst metazoans and plants. Curr. Genet 48, 265–269 (2005). [PubMed: 16193327]

101. Jiménez-Menéndez N et al. Human mitochondrial mTERF wraps around DNA through a left-handed superhelical tandem repeat. Nat. Struct. Mol. Biol 17, 891–893 (2010). [PubMed: 20543826] Crystal structure of MTERF1 bound to a double-stranded DNA segment (see also ref. 94).

102. Jemt E 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 gkv804 (2015). doi:10.1093/nar/gkv804

103. Park CB et al. MTERF3 is a negative regulator of mammalian mtDNA transcription. Cell 130, 273–285 (2007). [PubMed: 17662942]

104. Wredenberg A et al. MTERF3 regulates mitochondrial ribosome biogenesis in invertebrates and mammals. PLoS Genet 9, e1003178 (2013). [PubMed: 23300484]

105. Cámara Y et al. MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome. Cell Metabolism 13, 527–539 (2011). [PubMed: 21531335]

106. Pellegrini M et al. MTERF2 is a nucleoid component in mammalian mitochondria. Biochim. Biophys. Acta 1787, 296–302 (2009). [PubMed: 19366608]

107. Hixson JE & Clayton DA Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. Proc. Natl. Acad. Sci. U.S.A 82, 2660–2664 (1985). [PubMed: 3857608]
108. Fisher RP & Clayton DA A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. J. Biol. Chem 260, 11330–11338 (1985). [PubMed: 4030791]
109. Topper JN & Clayton DA Identification of transcriptional regulatory elements in human mitochondrial DNA by linker substitution analysis. Mol. Cell. Biol 9, 1200–1211 (1989). [PubMed: 2725494]
110. Wang Y & Shadel GS Stability of the mitochondrial genome requires an amino-terminal domain of yeast mitochondrial RNA polymerase. Proc. Natl. Acad. Sci. U.S.A 96, 8046–8051 (1999). [PubMed: 10393945]
111. Farge G et al. Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. Nat Comms 3, 1013 (2012).
112. Cotney J & Shadel GS Evidence for an early gene duplication event in the evolution of the mitochondrial transcription factor B family and maintenance of rRNA methyltransferase activity in human mtTFB1 and mtTFB2. J. Mol. Evol 63, 707–717 (2006). [PubMed: 17031457]
113. Cermakian N, Ikeda TM, Cedergren R & Gray MW Sequences homologous to yeast mitochondrial and bacteriophage T3 and T7 RNA polymerases are widespread throughout the eukaryotic lineage. Nucleic acids research 24, 648–654 (1996). [PubMed: 8604305]
114. Cermakian N et al. On the evolution of the single-subunit RNA polymerases. J. Mol. Evol 45, 671–681 (1997). [PubMed: 9419244]
115. Shutt TE & Gray MW Bacteriophage origins of mitochondrial replication and transcription proteins. Trends Genet 22, 90–95 (2006). [PubMed: 16364493]
116. Lang BF et al. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387, 493–497 (1997). [PubMed: 9168110]
117. Burger G, Gray MW, Forget L & Lang BF Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5, 418–438 (2013). [PubMed: 23335123]
118. Zhelyazkova P et al. The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. Plant Cell 24, 123–136 (2012). [PubMed: 22267485]
119. Ropp PA & Copeland WC Cloning and characterization of the human mitochondrial DNA polymerase, DNA polymerase gamma. Genomics 36, 449–458 (1996). [PubMed: 8884268]
120. Korhonen JA, Gaspari M & Falkenberg M TWINKLE Has 5’ -> 3’ DNA helicase activity and is specifically stimulated by mitochondrial singlestranded DNA-binding protein. J. Biol. Chem 278, 48627–48632 (2003). [PubMed: 12975372]
121. Maier D et al. Mitochondrial single-stranded DNA-binding protein is required for mitochondrial DNA replication and development in Drosophila melanogaster. Mol. Biol. Cell 12, 821–830 (2001). [PubMed: 11294889]
122. Schilbach S et al. Structures of transcription pre-initiation complex with TFIIH and Mediator. Nature 21, 327–209 (2017).
123. Engel C et al. Structural Basis of RNA Polymerase I Transcription Initiation. Cell 169, 120–131.e22 (2017). [PubMed: 28340337]
124. Spåhr H et al. SLIRP stabilizes LRPPRC via an RRM-PPR protein interface. Nucleic acids research 44, 6868–6882 (2016). [PubMed: 27353330]
125. Rackham O et al. Hierarchical RNA Processing Is Required for Mitochondrial Ribosome Assembly. Cell Reports 16, 1874–1890 (2016). [PubMed: 27498866]
126. Pham XH et al. Conserved sequence box II directs transcription termination and primer formation in mitochondria. J. Biol. Chem 281, 24647–24652 (2006). [PubMed: 16790426]
127. Tan BG, Wellesley FC, Savery NJ & Szczelkun MD Length heterogeneity at conserved sequence block 2 in human mitochondrial DNA acts as a rheostat for RNA polymerase POLRMT activity. Nucleic acids research 44, gkw648–7829 (2016).

128. Gangelhoff TA, Mungalachetty PS, Nix JC & Churchill MEA Structural analysis and DNA binding of the HMG domains of the human mitochondrial transcription factor A. Nucleic acids research 37, 3153–3164 (2009). [PubMed: 19304746]

129. Spåhr H, Samuelsson T, Hällberg BM & Gustafsson CM Structure of mitochondrial transcription termination factor 3 reveals a novel nucleic acidbinding domain. Biochemical and biophysical research communications 397, 386–390 (2010). [PubMed: 20430012]

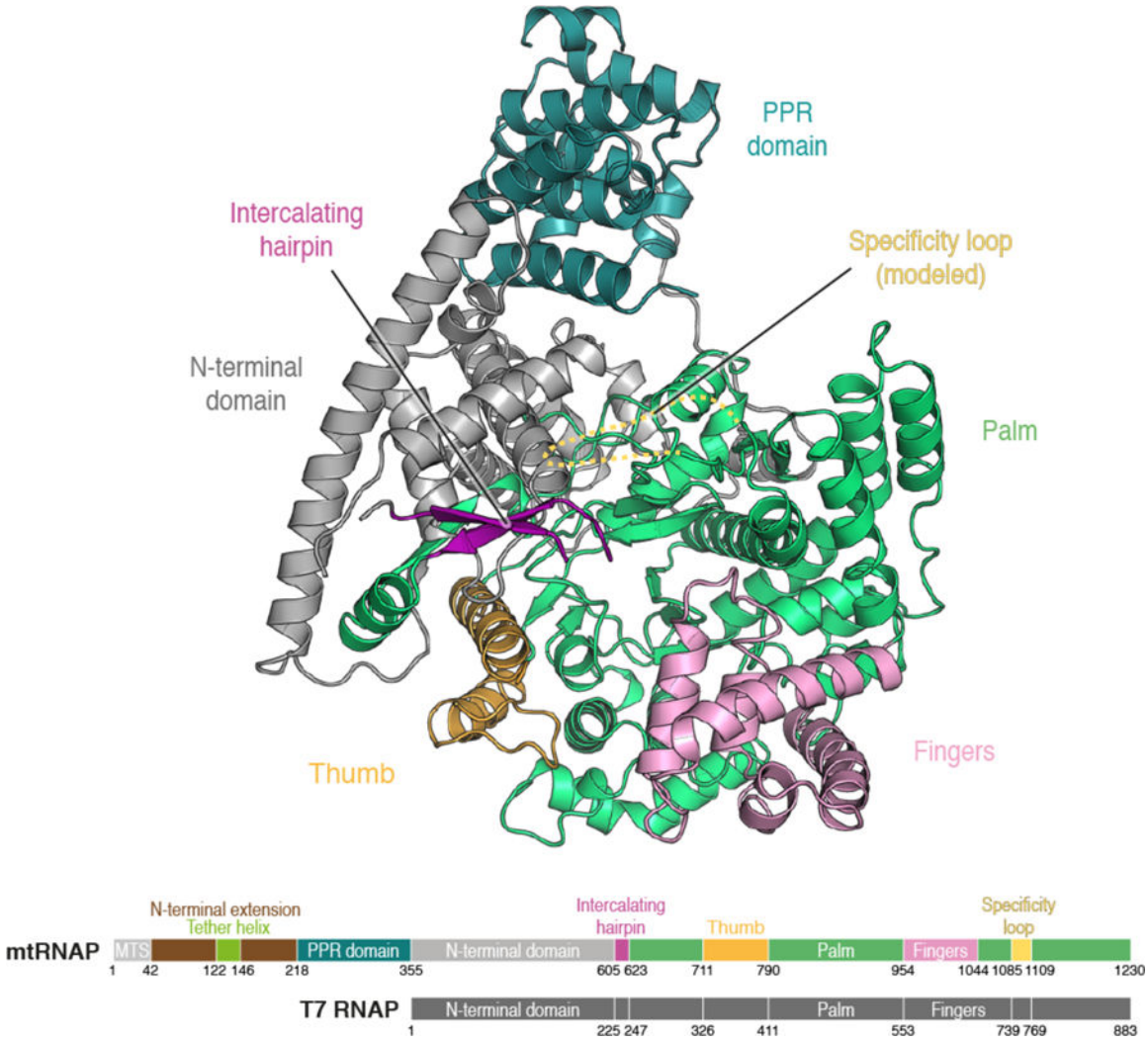


Figure 1. Structure of mitochondrial RNA polymerase (mtRNAP).

Ribbon depiction of human mitochondrial RNA polymerase (PDB ID: 3SPA) colored according to domains: PPR domain: deepteal; N-terminal domain: grey; Palm: green; Fingers: pink; intercalating hairpin: purple, Thumb: orange. Important structural elements are indicated. Below is a schematic depiction of mtRNAP and T7 RNAP domains and important structural elements. Structurally homologous regions between mtRNAP and T7 RNAP are aligned.

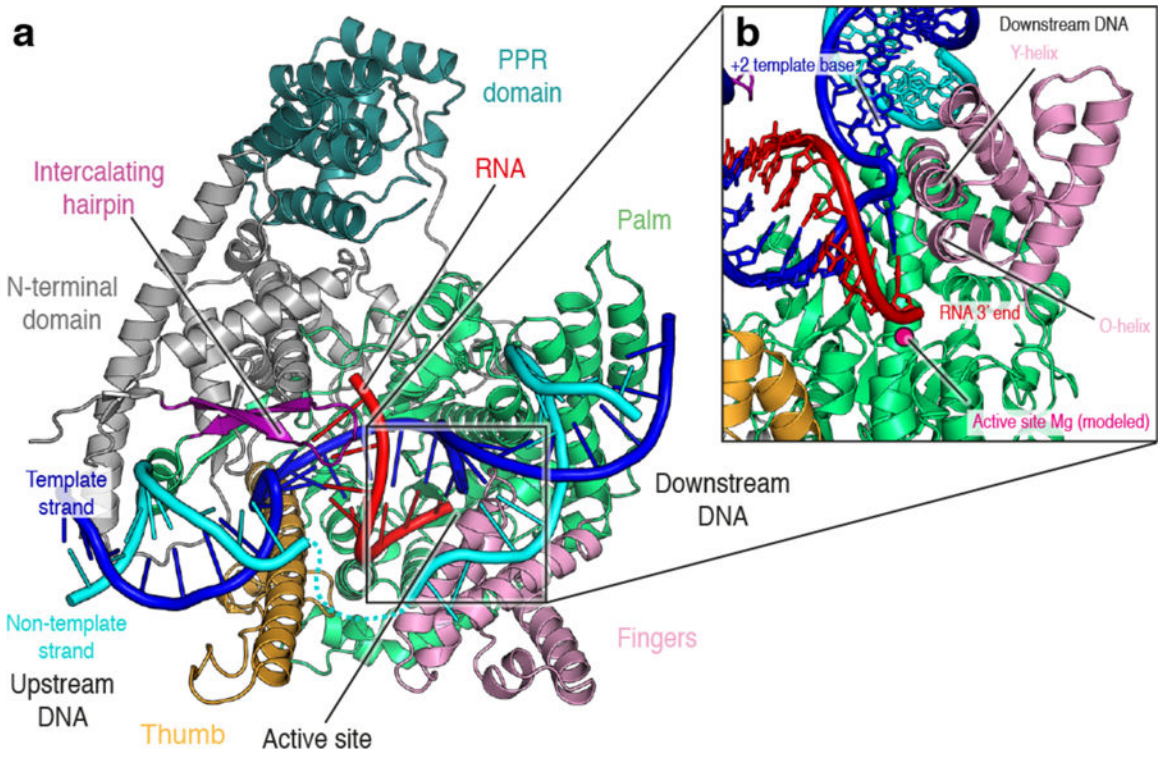


Figure 2. Structure of transcribing mtRNAP (elongation complex)

(**a**) Ribbon depiction of the human mitochondrial transcription elongation complex (PDB ID: 4BOC). MtRNAP is colored according to domains as in Figure 1. Important structural elements are indicated. Non-template DNA is depicted in cyan, template DNA in blue and RNA in red. The upstream DNA duplex is located close to the thumb domain and the fingers domain is involved in separation of the downstream DNA duplex, as in phage RNAPs. The nascent RNA is separated from template DNA by the intercalating hairpin and runs toward the PPR domain. (**b**) Close-up view of the active site in the mtRNAP elongation complex. The active site Magnesium is shown modeled and functional helices conserved between mtRNAP and phage RNAPs are indicated. The elongation complex is in the pre-translocated state, with the next templating base still base-paired to the non-template strand in the downstream DNA duplex.

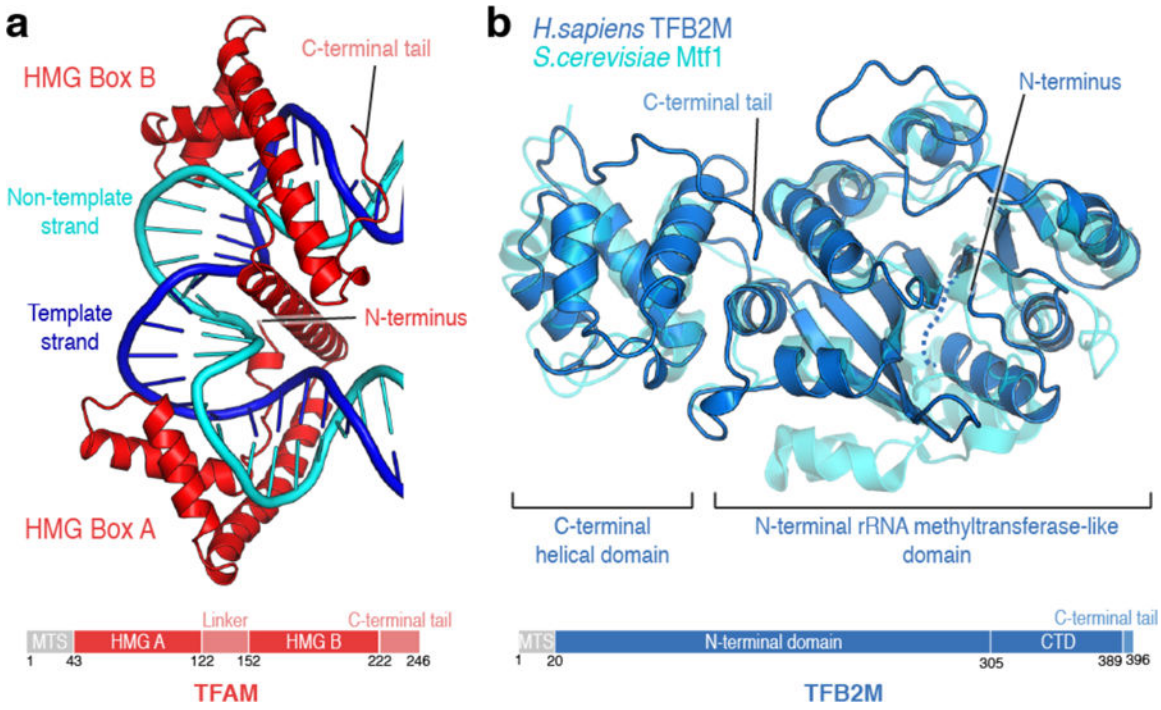


Figure 3. Structures of initiation factors TFAM and TFB2M

(**a**) Ribbon depiction of human TFAM bound to the LSP promoter sequence (PDB ID: 3TMM) colored in red. DNA coloring as in Figure 2. Important structural elements are indicated. A schematic depiction of the domain architecture of TFAM is shown below. MTS: Mitochondrial targeting sequence. Each of the HMG box domains of TFAM induces an approximately 90° bend to the duplex DNA by intercalating a hydrophobic residue, resulting in a 180° turn. (**b**) Ribbon depiction of human TFB2M (PDB ID: 6ERO) colored in marine blue with the structure of the yeast mitochondrial transcription initiation factor Mtf1 (PDB ID: 1I4W) overlaid transparently in cyan. A schematic depiction of the domain architecture of TFB2M is shown below. MTS: Mitochondrial targeting sequence. The N-terminal part of TFB2M and Mtf1 adopt a rRNA methyltransferase-like fold, while the C-terminal part is a globular, all-helical domain. TFB2M has a flexible tail which is important for its function in transcription.

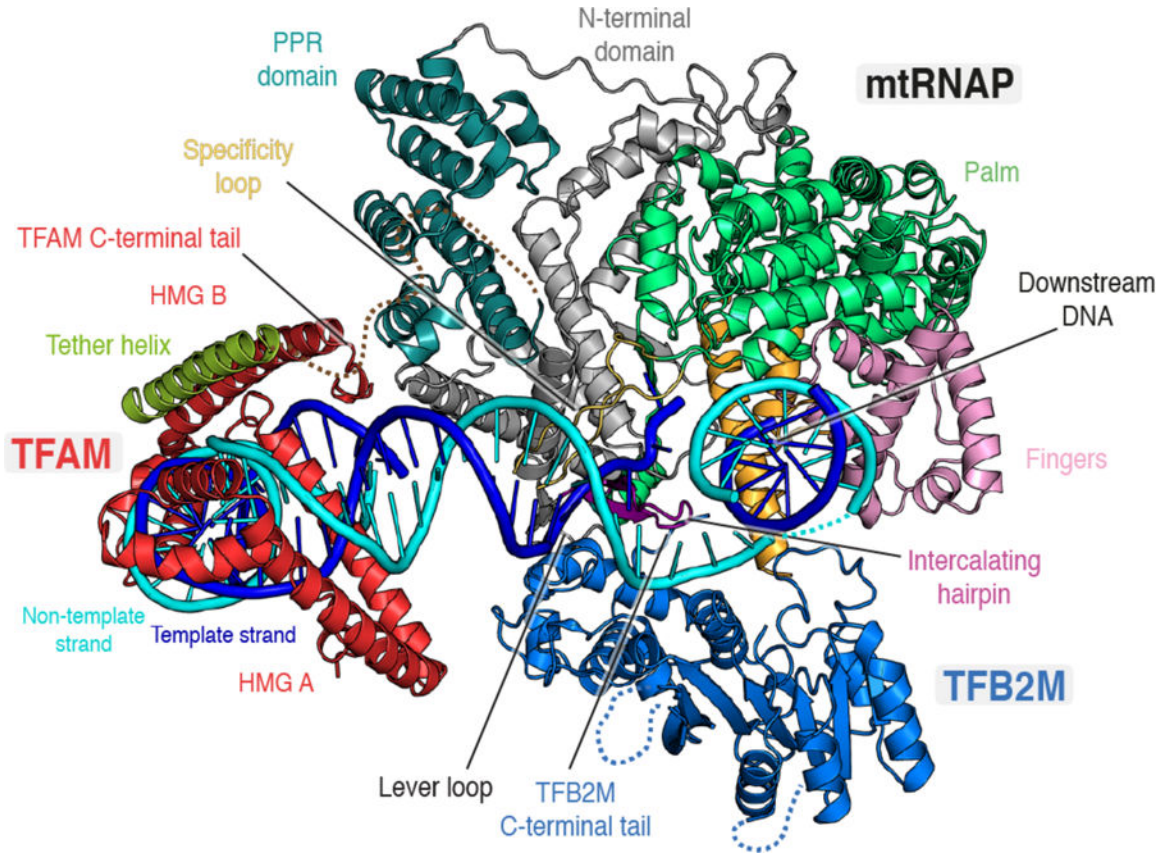


Figure 4. Structure of the mitochondrial transcription initiation complex

Ribbon depiction of the human mitochondrial transcription initiation complex (PDB ID 6ERP). The RNA polymerase is colored according to domains as in Figure 1. TFAM binds and bends the DNA upstream of the transcription start site. MtRNAP binds to TFAM via the tether helix located in the N-terminal extension and TFB2M binds around the site of DNA strand separation. The melted non-template strand runs along the positively charged surface of TFB2M. The intercalating hairpin of mtRNAP separates the DNA strands and TFB2M interacts with the adjacent lever loop, apparently stabilizing this conformation. Parts of the PPR domain of mtRNAP are in close vicinity to the DNA duplex upstream of the point of strand separation. The downstream DNA duplex emerges at a sharp angle.

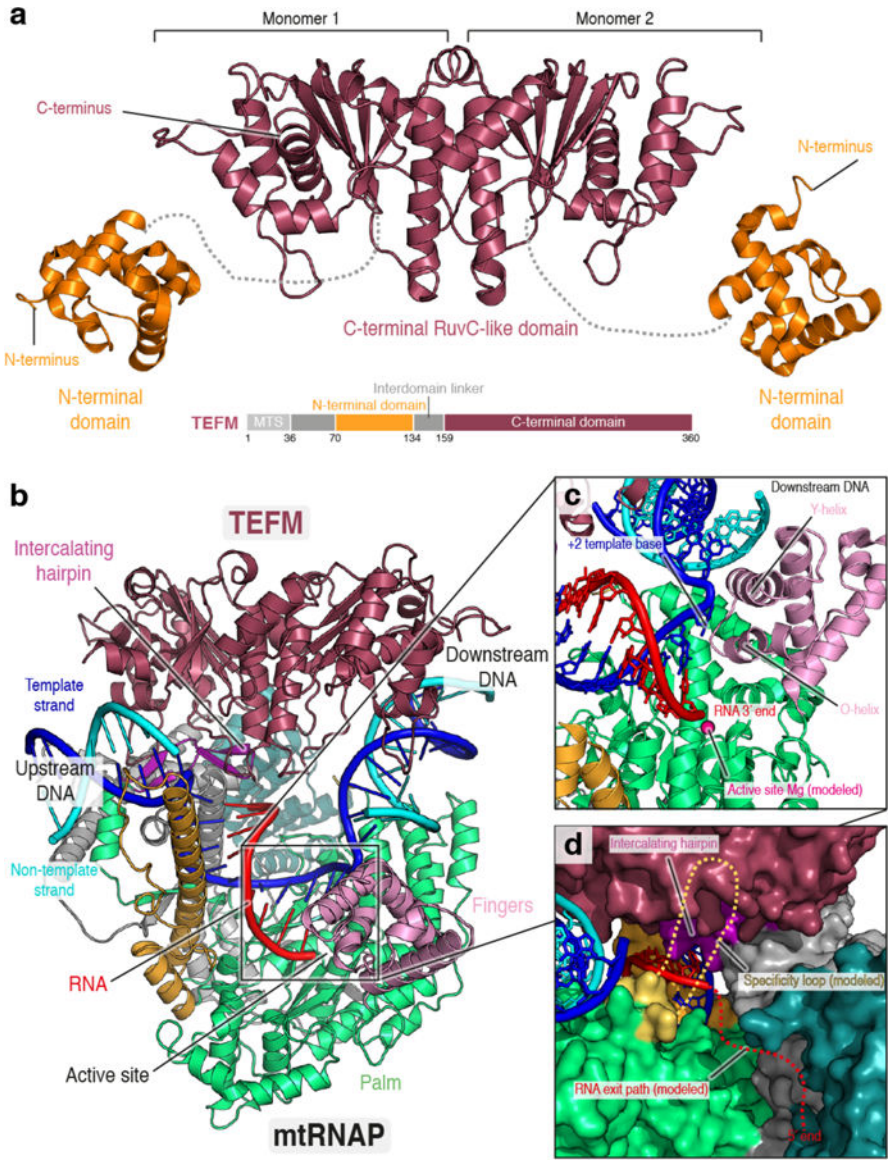


Figure 5. Structure of elongation factor TEFM and the processive antitermination complex

(**a**) Ribbon depiction of the human mitochondrial transcription elongation factor TEFM. The CTD (PDB ID: 5OL8) is colored in raspberry and the NTD (PDB ID: 5OL9) in orange. The unresolved linker region is shown as dashed line. A schematic depiction of the domain architecture of TEFM is shown below. MTS: Mitochondrial targeting sequence. The C-terminal domain of TEFM forms a homodimer in solution and resembles RuvC-like endonucleases involved in resolution of Holliday Junctions. (**b**) Ribbon depiction of the human mitochondrial processive anti-termination complex consisting of an elongation complex with the TEFM C-terminal domain and interdomain linker bound (PDB ID 5OLA). Coloring as in (a) and Figure 1. TEFM interacts with the nucleic acid in the transcription bubble and stabilizes the downstream DNA duplex. In addition, it interacts with the intercalating hairpin and specificity loop of mtRNAP and may stabilize the post-translocated state. (**c**) Close-up view of the active site of the anti-termination complex. The active site

Magnesium is shown modeled and functional helices conserved between mtRNAP and phage RNAPs are indicated. The anti-termination complex is in the post-translocated state, as evident by the next templating base flipped out towards the O-helix. (**d**) Close-up view of the RNA exit path in the anti-termination complex as viewed from the back of the orientation depicted in (b). Proteins are shown in surface representation with coloring as in (b). TEFM stabilizes the intercalating hairpin, which separates the nascent RNA from the template DNA and forms a tight RNA exit channel that is thought to prevent formation of disruptive secondary structures within the EC.

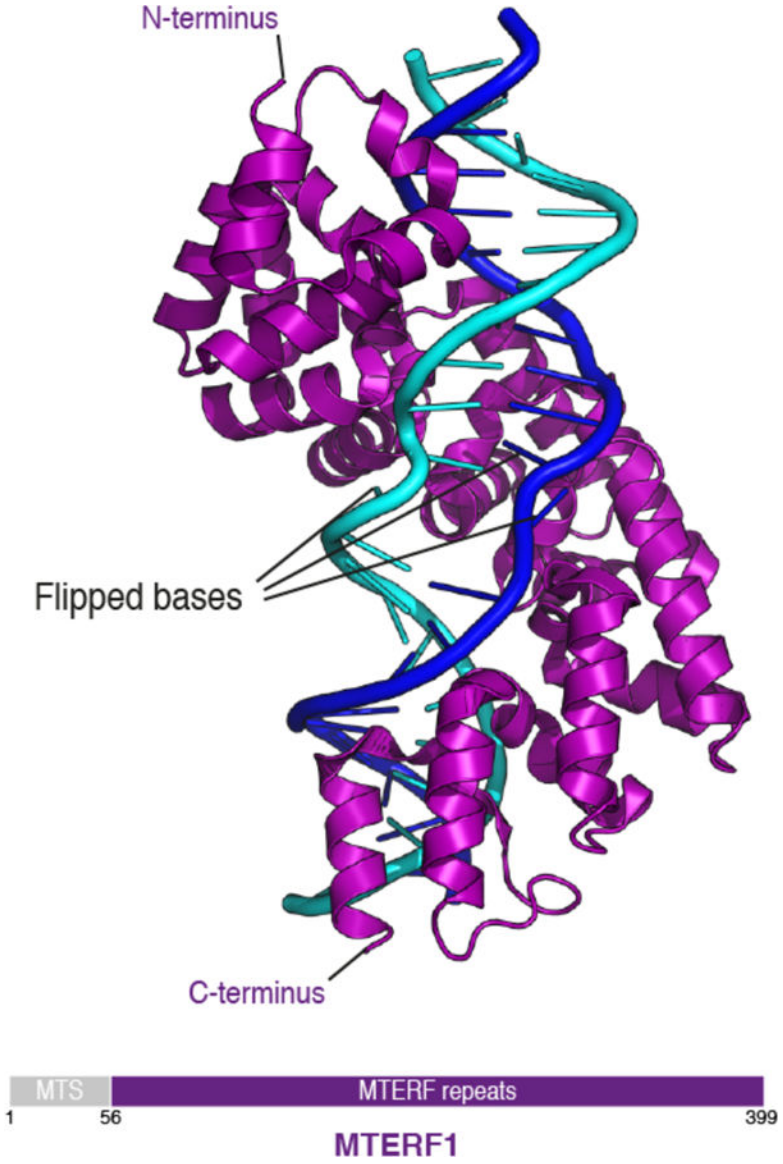


Figure 6. Structure of MTERF1

Ribbon depiction of human MTERF1 (PDB ID: 3MVA) in purple bound to the $tRNA^{Leu}$ termination sequence. DNA coloring as in Figure 2. A schematic depiction of MTERF1 domains is shown below. MTS: Mitochondrial targeting sequence. MTERF1 binds to a specific site in the mitochondrial genome and flips out multiple bases from the double stranded DNA. This binding is thought to act as a roadblock for the transcribing RNA polymerase, thus causing transcription termination.

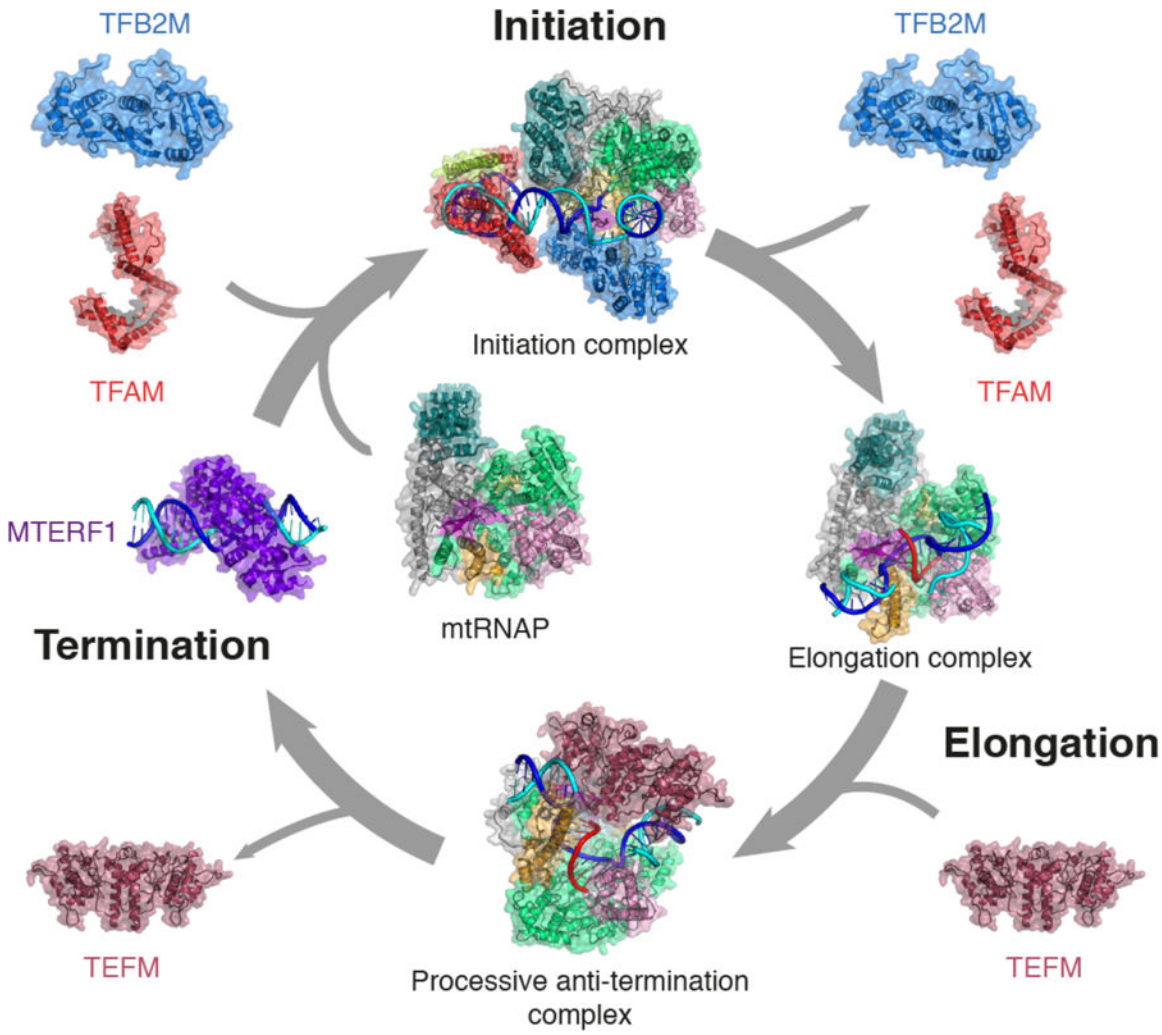


Figure 7. A structural view of the human mitochondrial transcription cycle

Schematic depiction of the human mitochondrial transcription cycle with known factors involved. Structures discussed in this review are depicted as ribbon representation with transparent surface overlaid. Coloring as in Figures 1–6. PDB IDs used: mtRNAP: 3SPA; TFAM: 3TMM; TFB2M: 6ERO; Initiation complex: 6ERP, Elongation complex: 4BOC; TEFM C-terminal domain: 5OL8; Processive anti-termination complex: 5OLA; MTERF1: 3MVA.

Table 1:

Structures of proteins involved in human mitochondrial transcription

Protein or Complex	Nucleic acid type	Resolution (Å)	PDB code	Reference
TFAM	TFAM binding site at LSP −39 to −12	2.5	3TMM	Ngo et al., 2011[55]
TFAM	TFAM binding site at LSP promoter −39 to −12	2.45	3TQ6	Rubio-Cosials et al., 2011[54]
TFAM	random sequence	2.81	4NNU	Ngo et al., 2014[56]
TFAM	Partial TFAM binding site at HSP −29 to −8	2.9	4NOD	Ngo et al., 2014[56]
TFAM HMG Box 2	-	1.35	3FGH	Gangelhoff et al., 2009[128]
TFB2M	-	1.75	6ERO	Hillen et al., 2017[67]
TEFM (N-terminal domain)	-	1.3	5OL9	Hillen et al., 2017[82]
TEFM (C-terminal domain)	-	1.9	5OL8	Hillen et al., 2017[82]
Apo mtRNAP	-	2.5	3SPA	Ringel et al., 2011[31]
Open initiation complex (OC)	LSP promoter −39 to +11	4.5	6ERP	Hillen et al., 2017[67]
Open initiation complex (OC)	HSP promoter −39 to +11	4.5	6ERQ	Hillen et al., 2017[67]
Elongation complex (EC)	DNA/RNA hybrid scaffold	2.65	4BOC	Schwinghammer et al., 2013[34]
Processive EC with TEFM (Antitermination complex)	DNA/RNA hybrid scaffold	3.9	5OLA	Hillen et al., 2017[82]
mTERF1	MTERF1 binding site in $tRNA^{Leu}$ DNA	2.2	3MVA	Yakubovskaya et al., 2010[94]
mTERF1 RFY mutant	MTERF1 binding site in $tRNA^{Leu}$ DNA	2.79	3MVB	Yakubovskaya et al., 2010[94]
mTERF1	12-mer $tRNA^{Leu}$(UUR) binding sequence	2.4	3N7Q	Jimenez-Menendez et al, 2010[101].
mTERF1	15-mer $tRNA^{Leu}$(UUR) binding sequence	3.1	3N6S	Jimenez-Menendez et al, 2010[101].
mTERF3	-	1.6	3M66	Spahr et al., 2010[129]