3726–3735 *Nucleic Acids Research, 2015, Vol. 43, No. 7* *Published online 23 March 2015*
doi: 10.1093/nar/gkv235

A model for transcription initiation in human mitochondria

Yaroslav I. Morozov[1,†], Andrey V. Parshin[1,†], Karen Agaronyan[1], Alan C. M. Cheung[2], Michael Anikin[1], Patrick Cramer[3] and Dmitry Temiakov[1,*]

[1]Department of Cell Biology, School of Osteopathic Medicine, Rowan University, 2 Medical Center Dr., Stratford, NJ 08084, USA, [2]Institute of Structural and Molecular Biology, University College London, Gower Street, London WC1E 6BT, UK and [3]Max Planck Institute for Biophysical Chemistry, Department of Molecular Biology, Am Fassberg 11, 37077 Gottingen, Germany

Received January 07, 2015; Revised February 21, 2015; Accepted March 08, 2015

ABSTRACT

Regulation of transcription of mtDNA is thought to be crucial for maintenance of redox potential and vitality of the cell but is poorly understood at the molecular level. In this study we mapped the binding sites of the core transcription initiation factors TFAM and TFB2M on human mitochondrial RNA polymerase, and interactions of the latter with promoter DNA. This allowed us to construct a detailed structural model, which displays a remarkable level of interaction between the components of the initiation complex (IC). The architecture of the mitochondrial IC suggests mechanisms of promoter binding and recognition that are distinct from the mechanisms found in RNAPs operating in all domains of life, and illuminates strategies of transcription regulation developed at the very early stages of evolution of gene expression.

INTRODUCTION

The remnants of the ancient single-subunit RNA polymerases of the Pol A family can be found in modern bacteriophages and DNA-maintaining organelles, such as plastids and mitochondria. In the latter, these RNAPs are charged with responsibility to synthesize mRNA, tRNA and rRNA, and to make RNA primers for replication (1–3). Even though some mitochondrial genomes are small (e.g. human mtDNA), regulation of mitochondrial gene expression is an elaborate process that occurs at various stages and involves many auxiliary factors and DNA and RNA modifying enzymes (4–8). Numerous mitochondrial dysfunctions are associated with defects in expression of mitochondrial genes and contribute to aging and severe pathologies and dysfunctions (9). At the beginning of the gene expression process, transcription of human mitochondrial DNA requires assembly of an initiation complex (IC) composed of mitochondrial RNA polymerase (mtRNAP) and two core transcription factors: TFAM and TFB2M (10–12). Recent studies demonstrated that mtRNAP is recruited to the promoter by formation of direct interactions with the nucleoid protein, TFAM (Figure 1A) (13,14). The resulting complex, called the pre-IC, lacks specificity toward DNA and cannot initiate transcription unless another transcription factor, TFB2M, is bound (13). Upon binding, the N-terminus of TFB2M reaches the active site of mtRNAP where it interacts with the priming substrate and assists in promoter melting (15). However, neither TFB2M nor TFAM binding sites on mtRNAP have been identified and thus the overall architecture of the IC as well as the pre-IC remains obscure. In this work, using biochemical, genetic and structural data we build a comprehensive map of interactions between all components of the IC. These data allowed us to construct a model of transcription initiation which is essential for understanding of molecular mechanisms of promoter binding and recognition and future studies of regulation of gene expression in human mitochondria.

MATERIALS AND METHODS

Expression and purification of the components of human mitochondrial transcription

TFB2M (res 21–396) carrying 6-His tag at the C-terminus was obtained by deletion of the intein region from pTYB11-TFB2M construct described previously (15). For cross-linking experiments a variant of TFB2M having an engineered protein kinase (PKA) site at the C-terminus (...RRASVHHHHHH) was used. TFAM, TFB2M and mtRNAP variants were obtained by site-directed mutagenesis (QuikChange, Agilent) as described (13) and purified as in (15).

*To whom correspondence should be addressed. Tel: +1 856 566 6327; Fax: +1 856 566 2881; Email: d.temiakov@rowan.edu
†These authors contributed equally to this work.

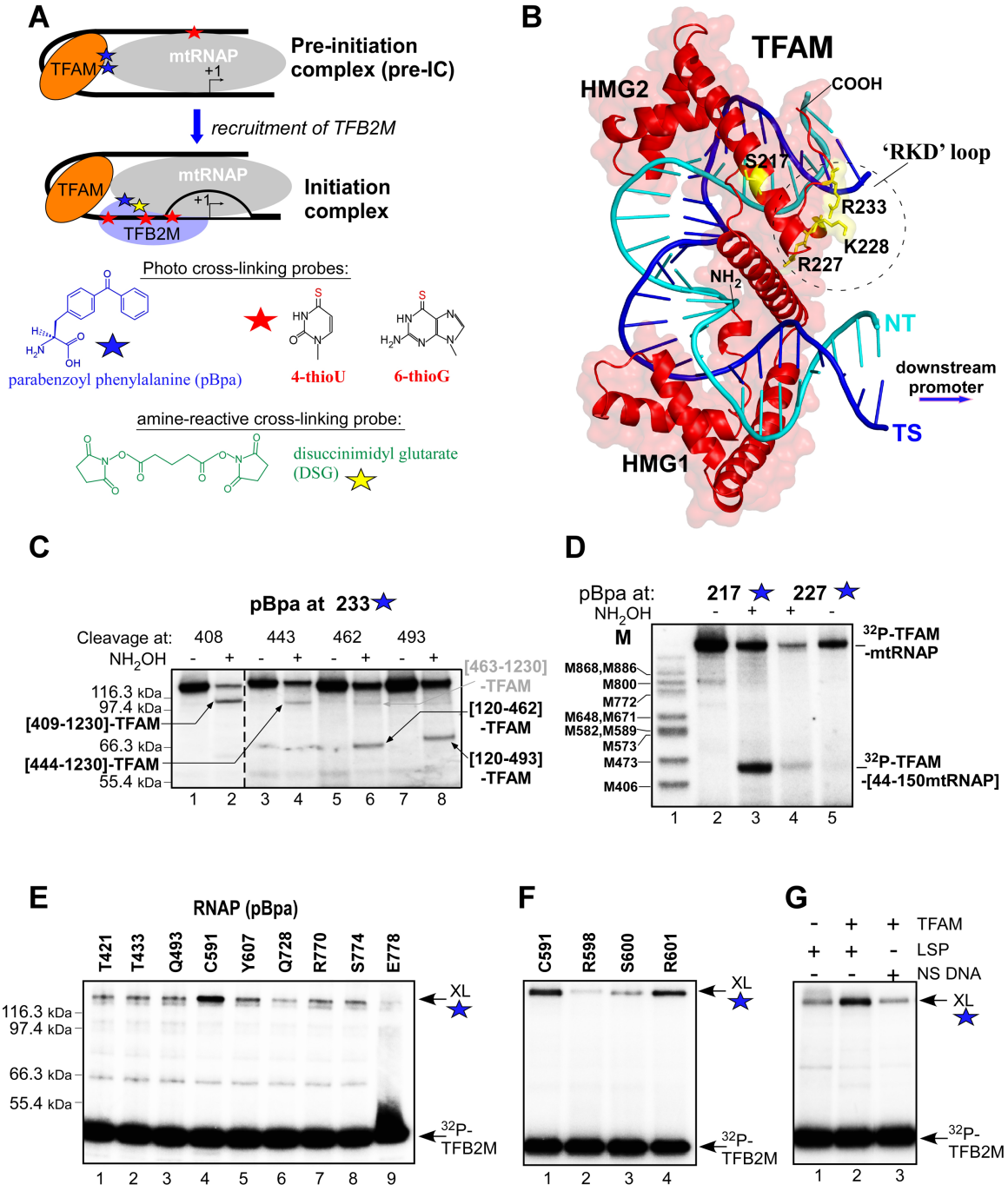

Figure 1. Identification of the key interactions between mtRNAP and transcription factors. **(A)** Schematic model of transcription initiation in mitochondria. TFAM recruits mtRNAP to promoter to form a pre-initiation complex (pre-IC). TFB2M binding to the pre-IC results in promoter melting and formation of an 'open' IC. Protein–protein interactions in pre-IC and IC were probed by pBpa cross-linking (blue stars); interactions between TFB2M variants and mtRNAP—with DSG cross-linker (yellow star); DNA–protein interactions—with photo reactive base analogs, 4-thioUMP and 6-thioGMP (red stars). (**B**) Structure of TFAM showing major cross-link sites in the C-terminus. Conserved 'RKD loop' in the C-terminal domain of TFAM (PDB ID 3TMM) is shown illustrating location of residues which substitution to pBpa resulted in cross-link with mtRNAP. **(C)** Mapping of TFAM-233pBpa cross-link to mtRNAP. The pre-ICs were assembled using mutant mtRNAPs having a single asparagine-glycine (NG) pair at position 408, 443, 462 or 493 and ^{32}P-labeled 233pBpa-TFAM, UV-irradiated and treated with hydroxylamine (lanes 2, 4, 6 and 8). Cleavage pattern is consistent with location of the major cross-linking site in the region 444–462 of the D helix of mtRNAP. **(D)** Mapping of TFAM-227pBpa cross-link to mtRNAP. The pre-ICs were assembled using mutant mtRNAPs having a single NG pair at position 150 and ^{32}P-labeled 217pBpa-TFAM or 227pBpa-TFAM, UV-irradiated and treated with hydroxylamine (lanes 3 and 4). In both reactions the N-terminal mtRNAP fragments were labeled (lanes 3 and 4), suggesting that the cross-link is to region 44–150 in mtRNAP. These data, taken together with the finding that Δ119 mtRNAP efficiently cross-links to the 227pBpaTFAM (13), suggest that the cross-linking is to the interval 120–150. Lanes 1–3 represent essential controls and have been published previously (13). **(E, F)** Scanning cross-linking of pBpa-containing mtRNAP and TFB2M. The ICs were assembled using ^{32}P-labeled TFB2M, LSP, TFAM and mtRNAP having pBpa at the position indicated. **(G)** Pre-IC interacts with TFB2M when assembled on promoter DNA. The ICs were assembled as above using 591pBpa-mtRNAP and the LSP promoter (lane 2) or non-specific DNA (NS, lane 3).

Promoter templates for transcription assays

Templates for transcription assays were prepared by polymerase chain reaction (PCR) amplification of pT7blue plasmid containing −70 to +70 of native light strand promoter (LSP) sequence. PCR template with the HSP1 promoter contained native HSP1 sequence from −70 to +5 and a downstream sequence (from +6 to +70) identical to that of the LSP promoter. Thus, both LSP and HSP1 templates contained identical initially transcribed sequence. Sequences of the synthetic templates with LSP and HSP1 promoter used in cross-linking are shown in the Supplemental Material.

Transcription assays

Standard transcription reactions were carried out using synthetic or PCR DNA templates (50 nM), mtRNAP (50 nM), TFAM (50 nM), TFB2M (50 nM) in a 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), UTP (0.01 mM) and 0.3 μCi[α-^{32}P] UTP (800 Ci/mmol) to produce 17–18 mer RNA products. Reactions were carried out at 35°C for the 30 min and stopped by addition of an equal volume of 95% formamide/0.05M EDTA. The products were resolved by 20% PAGE containing 6 M urea and visualized by PhosphorImager (GE Health).

Protein–protein cross-linking using artificial photo reactive amino acid (pBpa)

The amber codon was introduced to TFAM or Δ119 mtRNAP gene using Quik-Change site directed mutagenesis kit (Agilent) as described above. Expression of pBpa-containing proteins was performed as described previously (13). The ICs were assembled using ^{32}P-labeled TFB2M (100 nM), TFAM (100 nM), pBpa-mtRNAP (300 nM) and the LSP promoter (300 nM) or non-specific DNA (NS, 300 nM). The cross-linking was activated by UV irradiation at 312 mn for 15 min at room temperature.

Protein–DNA photo cross-linking

Promoter-mtRNAP cross-link with 4-thio UTP at base −49 was generated as describe previously (13) (see also Supplemental Material and Methods). To probe interactions of the pre-IC and the IC with bases +1, −5, −9 and −10 synthetic LSP promoter templates were used as shown in Supplementary Figure S1. The oligonucleotides containing the photo reactive probes (either 4-thio UMP or 6-thio dGMP) were from Midland Scientific. Transcription complexes (250 nM) were formed as described above and UV irradiated (312 nm) for 15 min at room temperature in the presence of non-specific oligonucleotides (10 μM). Cross-linking products were resolved using a 4–12% Bis-Tris NuPAGE gel (Invitrogen) and visualized by PhosphorImager (GE Health).

Cross-linking using disuccinimidyl glutarate

The ICs (0.5 μM in 40 mM Na-HEPES, pH = 7.5) assembled using ^{32}P-labeled TFB2M were treated with a mixture of disuccinimidyl glutarate (DSG) (0.6 mM DMSO solution, ProteoChem) and the products of the reaction separated using 7% Tris-glycine sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS PAGE).

Mapping of the cross-linking sites in mtRNAP and TFB2M

Hydroxylamine (NH_2OH) cleavage of pre-IC was performed as in (13). MtRNAP mutants used in photo cross-linking and mapping experiments were constructed using NG-less Δ119 mtRNAP (residues 120–1230) described previously (16) by introducing a single asparagine-glycine pair at position 408 (L408N/A409G), 926 (S926N, restores natural site) or 1117 (S1117G, restores natural site). NG443 (K443N/A444G) and NG150 (S150N) variants were constructed using mature NG-less mtRNAP (res 44–1230) described previously (15).

To map the cross-link site in TFB2M, the ICs were assembled using ^{32}P-labeled TFB2M (100 nM), C591pBpa-mtRNAP, (500 nM), TFAM (600 nM) and LSP (500 nM) in a transcription buffer containing 0.1 mg/ml BSA and 70 mM NaCl, UV irradiated for 15 min, and the cross-linking species separated using 7% Tris-glycine SDS PAGE. The TFB2M–mtRNAP cross-link was excised from the gel, eluted with 0.2% SDS for 1 h at 4°C and precipitated with 80% acetone. The dried material (2 μl) was dissolved in solution containing 50 mM CNBr/ 0.2M HCl and incubated at 37°C for 10–30 min.

Construction of the TFB2M model

TFB2M model was built based on TFB1M coordinates (PDB ID 4GC5) using Swiss-Model Server (http://swissmodel.expasy.org/).

RESULTS AND DISCUSSION

Binding sites of TFAM and TFB2M on mtRNAP

The C-terminus of TFAM was found to interact with the N-terminal extension region of mtRNAP (residues 120–134, Figures 1B and 3) (13). This region, however, is represented only by an unassigned helix attached to the pentatricopeptide repeat (PPR) domain in the crystal structure of mtRNAP, thus precluding the modeling of the pre-IC. When analyzing cross-linking between mtRNAP and TFAM having photo reactive amino acid pBpa (parabenzoyl phenylalanine) inserted at positions 217 and 228 (or 233) we noticed a markedly different mobility of the cross-linked species, which was indicative of an additional site of interaction with mtRNAP (Supplementary Figure S1A). Using chemical peptide mapping we determined that pBpa228-TFAM and pBpa233-TFAM interact with region 444–462/444–473 of the D-helix of mtRNAP N-terminal domain (NTD) (Figure 1C and Supplementary Figure S1). Further, we mapped the interaction site of the pBpa227-TFAM between residues 120–150 in the N-terminal extension region of mtRNAP, suggesting that the C-terminus region of TFAM (the RKD loop) is located in close proximity to both identified structural elements in mtRNAP (Figure 1D). In agreement with these data, deletion of 15 or more C-terminal TFAM residues or substitution of the charged residues in

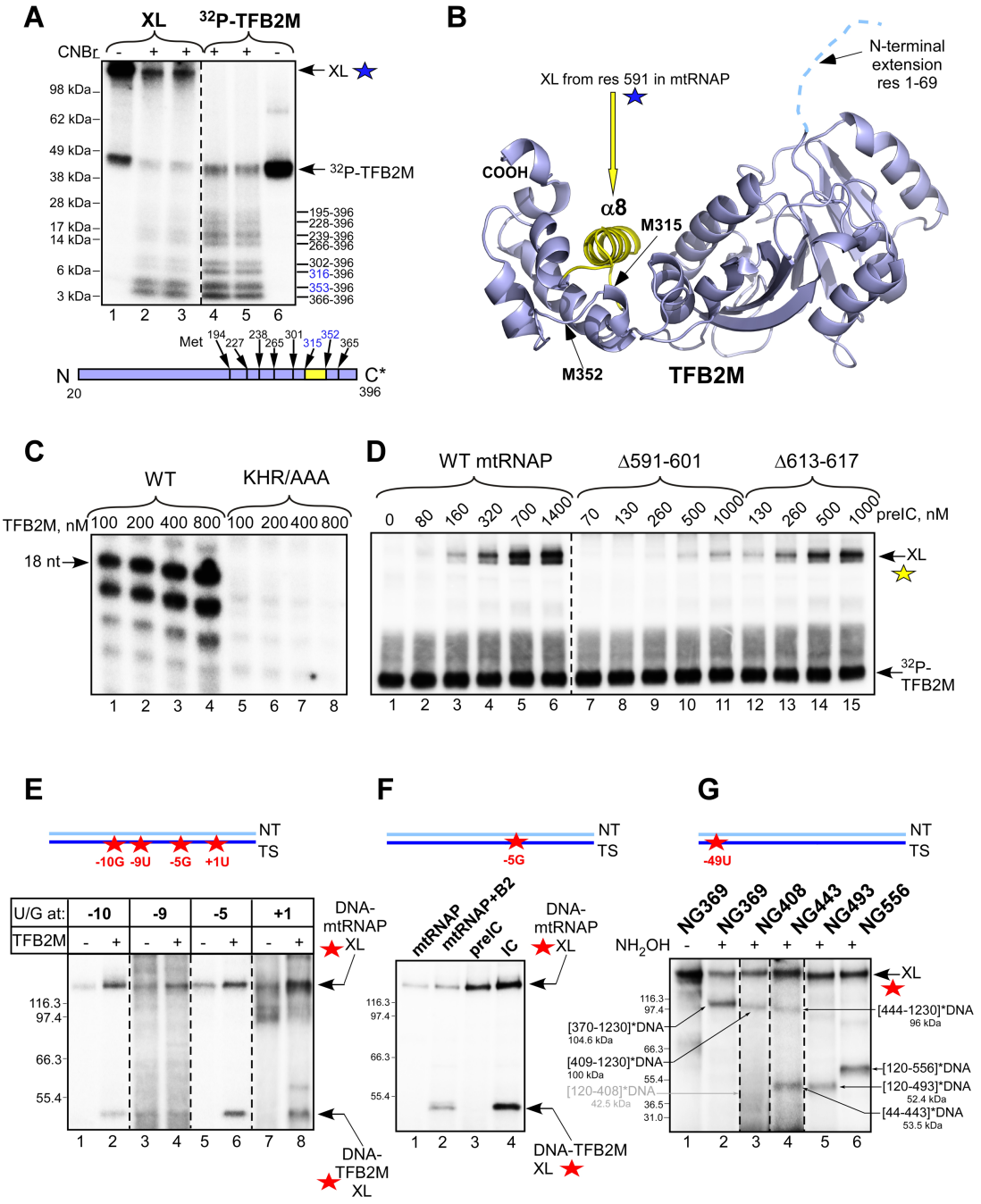

Figure 2. Identification of the binding sites in mtRNAP and TFB2M. **(A, B)** Mapping of mtRNAP–TFB2M cross-link. MtRNAP–TFB2M cross-link (lanes 1–3) and ^{32}P-TFB2M (lanes 4–6) were treated with CNBr for the time indicated. Methionine cleavage pattern reveals two band representing labeled peptides 352–396 and 365–396 in TFB2M that were not shifted by the cross-linking to mtRNAP and is consistent with cross-link location between M315 and M352 in the α8 helix of TFB2M (lower panel and Supplementary Figure S3). **(C)** TFB2M mutant having substitutions in the α8 helix cannot support transcription. Transcription assay was performed with the wild-type (WT) (lanes 1–4) and mutant TFB2M (KHR/AAA, lanes 5–8). **(D)** Deletion of the B-loop of mtRNAP results in decrease of IC assembly efficiency. The pre-ICs were assembled using WT (lanes 1–6), B-loop deletion (lanes 7–11) or intercalating hairpin deletion (lanes 12–15) mtRNAP and incubated with TFB2M in the presence of DSG. **(E)** DNA–mtRNAP interactions in pre-IC and IC. The complexes were assembled on the LSP promoter containing photo reactive 4-thio UMP or 6-thio dGMP at the positions indicated and UV-irradiated. **(F)** Both TFB2M and mtRNAP interact with the −5 template base of promoter. The LSP promoter template containing 6-thio dGMP was incubated with the proteins indicated and UV irradiated. Note the increase of DNA–mtRNAP cross-link upon addition of TFB2M to the pre-IC (lane 4). **(G)** Mapping of mtRNAP interaction with the −49 template base in the pre-IC. The cross-linking was performed using Δ119 mtRNAP having NG pair at position 369 (lanes 1 and 2), 408 (lane 3), 443 (lane 4), 493 (lane 5) and 556 (lane 6). Cleavage of the NG369 and NG408 mtRNAP mutants generates a single labeled fragment corresponding to the C-terminal region of mtRNAP (lanes 2 and 3). The NG443 mtRNAP cleavage produces two fragments. One fragment corresponds to the N-terminal region (residues 44–443, 80% efficiency), the other fragment corresponds to the C-terminal fragment (residues 444–1230, 20% efficiency), suggesting that the major DNA cross-linking site is located between residues 409–443 (lane 4). Consistent with this, cleavage of NG493 and NG556 mtRNAPs verified that the location of the major cross-linking site was N-terminal to these cleavage site positions (lanes 5 and 6). Mol. weights of the protein markers (Mark12, Invitrogen) are indicated to the left of the panel (kDa).

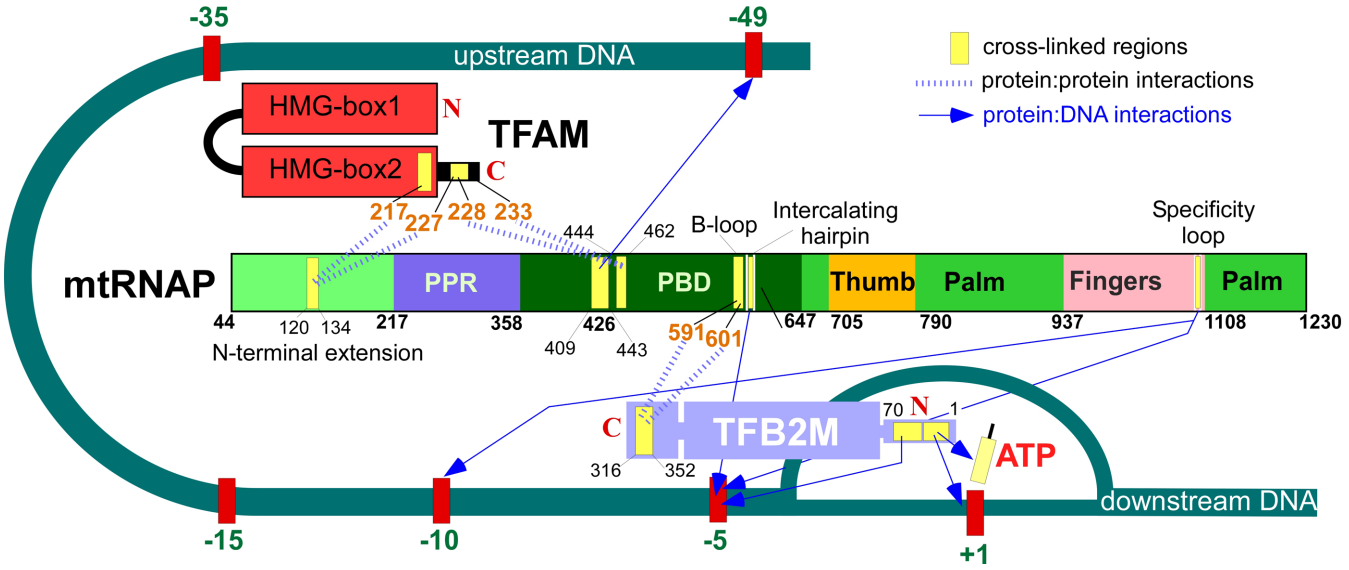

Figure 3. Structural model of human mitochondrial transcription IC. The model combines data obtained by mapping of protein–protein and DNA–protein interactions (this work) as well as mapping of TFB2M interactions with the priming substrate and +1 template DNA base (13). Large horizontal bars denote mtRNAP (with major domain and structural elements indicated), TFB2M (blue) and TFAM (red). Lines and arrows connect contacts identified by cross-links from specific nucleotide positions or amino acids and protein segments (yellow shaded boxes) of mtRNAP or transcription factors. The positions of nucleotide residues are numbered relative to the promoter start site, which is denoted +1.The extent of promoter melting (region −4 to +3) is shown according to pre-melted promoter template assay (13). TFAM makes interactions with the −15/−35 region of promoter (not shown).

the RKD loop with uncharged residues resulted in a significant loss of transcription activity (Supplementary Figure S2 and Supplementary Table S1) (17).

We next probed interaction between mtRNAP and TFB2M in the IC using photo cross-linking. Based on the previously generated model of the IC (18), we made a series of mtRNAP variants having pBpa incorporated at the desired position (Figure 1E and F). Specific cross-linking was detected when the probes were located in the loop (termed here the 'B loop', res 588–604) adjacent to the intercalating hairpin of mtRNAP. The B loop is present in all mitochondrial RNAPs but is not found in phage RNAPs, suggesting a conserved role of TFB2M-like transcription factors in mitochondria. We found that deletion of this region results in inability of mtRNAP to support transcription initiation (18).

In a previous study, we demonstrated that the pre-IC could be assembled on non-specific DNA (13). Such complexes are likely transient, and dissociate if TFB2M is not present. In contrast, the IC could only form on a promoter template, as evidenced by cross-linking between TFB2M and mtRNAP (Figure 1G, lanes 2 and 3). In the absence of TFAM, TFB2M does not bind efficiently to mtRNAP (Figure 1G, lane 1) as the specific cross-link was detected only when the pre-IC was formed prior to addition of TFB2M. These findings further confirm the sequential model of transcription initiation in human mitochondria (13) (Figure 1A).

Using chemical mapping with CNBr under 'single hit' (19,20) conditions, we mapped the region in TFB2M that interacts with mtRNAP to the interval flanked by residues 315–352 (Figure 2A and Supplementary Figure S3). This region corresponds to the α8 helix in the model of TFB2M, which was built based on the structure of its homolog, TFB1M (21) (Figure 2B). Despite a relatively low-sequence conservation between TFB1M (which is not a transcription factor) and yeast homolog of TFB2M, Mtf1 (transcription factor) these proteins exhibit high structural conservation (21,22). The very N-terminal domain of TFB2M (res 1–69) is implicated in DNA binding and interaction with the priming ATP (15), however, does not share any sequence homology with TFB1M and is illustrated as a dashed line in this model (Figure 2B).

To confirm the role of α8 helix of TFB2M in binding to mtRNAP, we made substitutions of the solvent exposed charged residues in this region—K325, H326, R330 (KHR to AAA mutation) (Figure 2B). This mutant TFB2M was unable to support transcription (Figure 2B) further confirming the mapping data.

Finally, to confirm the TFB2M binding site on mtRNAP we analyzed the ability of deletion mutants of mtRNAP to form the IC. We employed a DSG cross-linker (Figure 1A) that allows linking of lysine residues located within ∼8 Å distance (23), and observed efficient cross-linking with wild-type mtRNAP and ^{32}P-labeled TFB2M (Figure 2D). While the intercalating hairpin deletion mutant (deletion of res. 613–617) was also able to cross-link to TFB2M when assembled in an IC, deletion of the B-loop resulted in a dramatically lower cross-linking efficiency, confirming this structural element as a primary binding site for TFB2M (Figure 2D).

MtRNAP–DNA contacts become prominent upon recruitment of TFB2M to the pre-IC

Using the photo reactive base analogs 4-thio UMP and 6-thio dGMP (Figure 1A), we probed interactions between mtRNAP and TFB2M with promoter DNA. We found that

at most positions analyzed the cross-link between DNA and mtRNAP in the pre-IC was relatively weak (Figure 2E). Addition of TFB2M and formation of the IC resulted in a noticeable increase of the efficiency of DNA–mtRNAP cross-linking. We mapped the most efficient cross-links in the IC at positions −5 and −10 (Supplementary Figure S4). The −5G template strand base was found to interact with the intercalating hairpin (res. 604–614) and the (base of) specificity loop of mtRNAP, consistent with the location and function of these elements in T7 RNAP structure (24). Unexpectedly, we detected a strong cross-link between the −5G template base and TFB2M (Figure 2F, lane 4). This interaction was mapped to the N-terminal region of TFB2M (res. 1–59) (Supplementary Figure S4C) supporting previous observations that this functionally important region is inserted into the RNA–DNA hybrid binding cavity of mtRNAP (15). Further upstream of the promoter start site, the −10G base was found to interact with the specificity loop of mtRNAP (Supplementary Figure S4D and E). Consistent with these data, deletion of the tip of the loop (res. 1087–1099) results in a variant of mtRNAP that is active on RNA–DNA scaffold but unable to support promoter-driven transcription (Supplementary Figure S4F).

Finally, we mapped interactions between the far upstream region of promoter DNA (at base -49) and mtRNAP in the pre-IC (Figure 2G). This cross-link has been previously shown to be TFAM dependent and suggested that mtRNAP is sandwiched between the upstream and downstream promoter regions (13). Using site-specific protein mapping employing a hydroxylamine cleavage reaction, the region of interaction with the -49 base was mapped to the solvent-exposed loop (residues 415–426) between the D helix and the β-strand A in the N-terminal domain of mtRNAP (Figures 2G and 3).

Spatial organization of the human mitochondrial pre-IC

The large body of cross-linking and functional data indicates that TFAM binds mtRNAP only in the presence of DNA and that this process is TFB2M-independent. To obtain a structural model of the pre-IC, we combined two experimentally determined models: the high-resolution structures of human mtRNAP (18) and the TFAM/DNA complex (25,26), using the cross-linking data as constraints. For the purposes of modeling we assumed that only minor conformational changes would occur in the individual protein molecules upon their interaction. We also used a related T7 RNAP promoter complex (24) as a reference for location of the downstream DNA region. First, we extended the DNA of the TFAM/DNA complex at both ends to include the mitochondrial promoter region from −14 to −5, which is within the mtRNAP footprint (15,27). This region was aligned in the same orientation as the upstream DNA of the promoter complex of T7 RNAP (24), which was previously suggested by homology modeling (18). Second, TFAM region 227–232 was positioned in proximity to the mapped region 444–462 in the D helix of mtRNAP, according to the cross-linking data. Finally, the TFAM region centered near residue 217 was placed in proximity to the N-terminal extension of mtRNAP, of which only one unassigned helix is visible in a crystal structure (18) (Figures 3 and 4 and Supplementary Figure S5).

In this model, the DNA bent by TFAM is wrapped around mtRNAP and the promoter is directed toward the intercalating hairpin and the specificity loop of mtRNAP which have been implicated in promoter melting and recognition, respectively (18,28,29) and found to interact with the DNA in cross-linking studies (this work). The region of the D helix distal to TFAM contains a large number of positively charged residues such as R431, K435, K443, R477 and R481 that may be involved in interaction with DNA upstream of TFAM (Figure 4). The upstream portion of the promoter DNA (region −35/−60) contours the mtRNAP molecule along its N-terminal domain and required a bend of ∼15–20° to avoid clashes with the protein body. In agreement with the model, the −49 base of promoter DNA is within interacting distance of the loop (415–426) mapped in cross-linking experiments, further validating the model of the pre-IC (Figure 4).

The modeling places TFAM residue 217 in proximity to the PPR domain and an unassigned helix in the N-terminal extension region of mtRNAP. In the model, the residues in TFAM that cross-link to mtRNAP (R227, K228 and R233) face mtRNAP and fall within the theoretical interaction range of the cross-linkers used (13). Consistent with this, TFAM residues G226 and T234 which did not produce an efficient cross-link with mtRNAP when substituted with pBpa face away from the mtRNAP surface (Figure 4).

Model of transcription initiation in human mitochondria

The cross-linking data indicate that in the absence of the promoter TFB2M is unable to bind to mtRNAP efficiently suggesting that binding of TFAM and promoter DNA causes structural changes in mtRNAP. Since the N-terminus of TFB2M is directed toward the active site of mtRNAP, where it interacts with both the priming ATP and a templating (+1) DNA base, such changes should provide a passage within mtRNAP in order to accommodate TFB2M. The opening of a nucleic acid binding cavity and movement of the N-terminal domain relative the palm domain was observed in the structure of mtRNAP elongation complex (16). We therefore used this structure to model interactions within the IC (Figure 5). To model the melted promoter DNA region proximal to the start site, we used homology modeling with the T7 RNAP IC (24). The DNA duplex in this complex is melted in the region from +3 to −4, which is consistent with the data for mtRNAP promoter complex (15). TFB2M was docked manually to the B loop of mtRNAP so it would come in a close proximity to the α8 helix of TFB2M. We further considered that, since the N-terminal extension region of TFB2M interacts with both −5 and +1 bases of the template strand of promoter DNA, the only way for TFB2M to reach the active site of mtRNAP would be via the opening between the thumb (res 720–760) and the intercalating hairpin of mtRNAP. This restraint sets the orientation of TFB2M relative to mtRNAP and results in minimal clashes between these proteins (Figure 5). As we noted previously when modeling the melted promoter region into mtRNAP structure, the intercalating hairpin clashes with the promoter DNA indicating that its location is not fixed in the apo enzyme (18). In phage

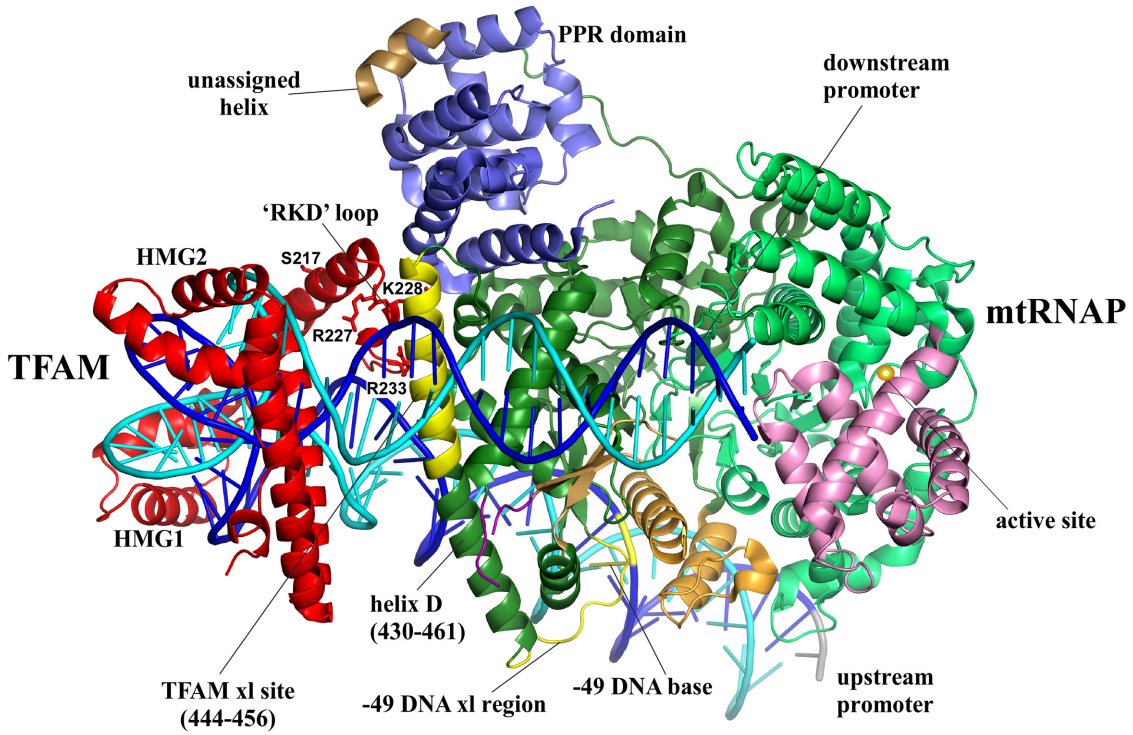

Figure 4. Model of the pre-IC. Two experimental data sets (mtRNAP, PDB ID 3SPA and DNA/TFAM complex, PDB ID 3TMM) were used to model the pre-IC based on biochemical data. The upstream and downstream promoter DNA regions (depicted as ribbons, T-strand, blue, NT-strand, cyan) were extended to emulate trajectory of the nucleic acid in the pre-IC. MtRNAP is depicted as a ribbon (color codes as in Figure 3). A Mg^{2+} ion (orange) was placed according to a T7 RNAP structure. TFAM is shown in red (ribbon). The TFAM and −49 DNA cross-linking sites in mtRNAP are indicated in yellow.

RNAPs, this structural element is inserted between DNA stands and maintains the trailing edge of the transcription bubble during initiation of transcription (30,31). Our cross-linking data indicate that the intercalating hairpin interacts with the −5 base of promoter only when TFB2M is present (Figure 2F and Supplementary Figure S4). We therefore propose that binding of TFB2M to the adjacent B-loop pushes the intercalating hairpin toward the DNA duplex and forces it to assume a position identical to the position of the corresponding loop in T7 RNAP IC. The repositioning of the intercalating loop may open the passage between the N-terminal domain and the thumb allowing access of TFB2M to the active site of mtRNAP.

Mechanism of promoter binding and recognition and implications for transcription regulations

Our data suggest an elaborate mechanism of promoter binding and recognition by mtRNAP in which both transcription initiation factors are required for the specificity of promoter interactions. While TFAM is an abundant mitochondrial protein covering the entire mtDNA, interactions with this factor and the upstream promoter region are essential for the proper geometry of the pre-IC. In the absence of TFB2M, mtRNAP appears to have very few (if any) specific contacts with promoter DNA. In the IC, mtRNAP interacts the trailing edge of the transcription bubble via intercalating hairpin to maintain separation of the DNA strands and TFB2M is required for these interactions (15). However, TFB2M likely has an additional role during initiation of transcription. The major specificity element in phage RNAPs—the specificity loop—is a β hairpin that inserts into the major groove of promoter and makes multiple interactions with DNA bases at position −7 to −11 in T7 promoter (31,32). While essential for promoter-driven transcription in human mitochondria (Supplementary Figure S4), this element does not interact with promoter in the pre-IC and becomes engaged in DNA binding only when TFB2M is bound. This suggests that TFB2M can reposition not only the intercalating hairpin but also the specificity loop of mtRNAP so it can bind and recognize promoter.

In all living systems that rely upon multi-subunit RNAPs, regulation of transcription initiation is achieved during promoter binding and melting steps (33–35). Our model of a sequential assembly of the IC suggests that this approach is also used by human mtRNAP, a member of arguably the oldest family of polymerases—Pol A—that includes phage RNAPs, RT polymerase, bacterial DNA pol I, and single-subunit RNAPs from mitochondria and plastids. Notably, RNAPs of the Pol A family lack subunits dedicated for sequence-specific promoter recognition such as sigma in bacterial RNAPs. While inherently limited in ability to recognize many different promoters and regulatory signals, mtRNAP employs a set of transcription factors with specialized functions to reposition structural elements that are involved in promoter binding and recognition, and conserved between the members of Pol A family of RNAPs.

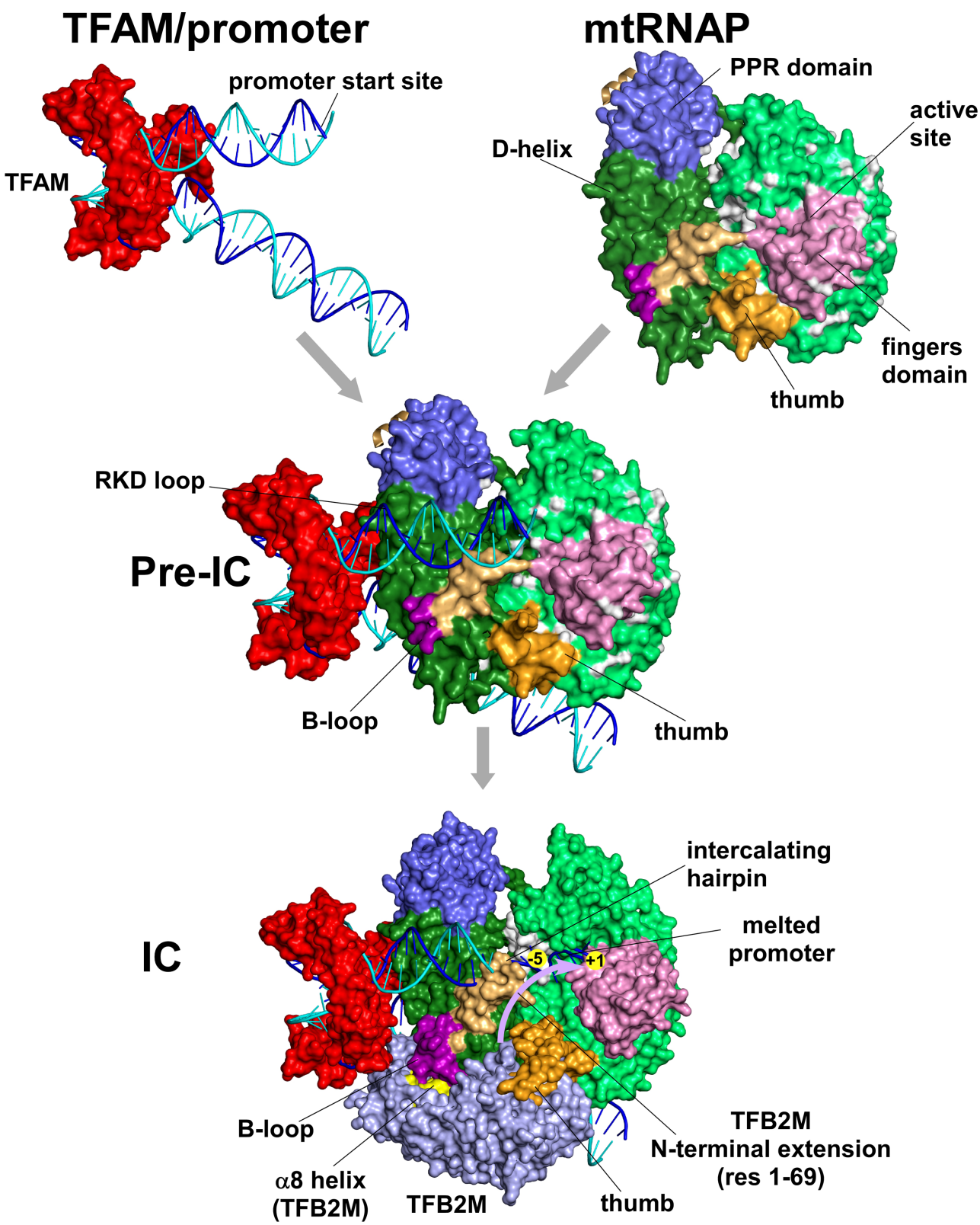

Figure 5. A model of transcription initiation process in human mitochondria. The models of the pre-IC and IC are shown as a surface representation. The IC model was generated using mtRNAP 'open' conformation found in the elongation complex (PDB ID 4BOC). MtRNAP subdomains are colored according to Figure 3, TFB2M is in light blue. The model suggests that the N-terminus of TFB2M descents into the active site of mtRNAP by passing through the opening between the thumb subdomain and the intercalating hairpin.

The most dramatic example of such a modification involves the ability of TFB2M to reach the active site of mtRNAP upon repositioning of other elements (15). While the ability of auxiliary initiation factors such as sigma in bacteria and TFIIB in eukaryotes to participate in catalytic functions of RNAP have been documented (36,37), only TFB2M has been found to make direct contact with the priming substrate. The apparent advantage of this modified approach is an opportunity to regulate replication and expression of mitochondrial genome at various stages of the cell cycle or in response to the energetic state of the cell (15,38). Studies of the extent of this regulation and the mechanisms behind it are crucial to further our understanding of mitochondrial functions.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are thankful to W. McAllister and A. Mustaev for critical reading of the manuscript. S. Borukhov, D. Markov, H. Hillen and K. Schwinghammer are acknowledged for fruitful discussions and helpful suggestions.
Author Contributions: Y.M., A.P. and K.A. performed transcription assays, mutagenesis and mapped protein–protein and DNA–protein interactions. A.C., P.C. and D.T. built structural models. Y.M., A.P., A.K., M.A., A.C., P.C. and D.T. designed and analyzed the experiments; D.T. prepared the manuscript.

FUNDING

NIH RO1GM104231 and Foundation of UMDNJ (PC88 to D.T.); Wellcome Trust Fellowship 102535/Z/13/Z (to A.C.); Deutsche Forschungsgemeinschaft (SFB646, SFB960, TR5, GRK1721, CIPSM, NIM, QBM); BioImaging Network, an ERC Advanced Grant, the Jung-Stiftung and the Vallee Foundation (to P.C.). Funding for open access charge: NIH RO1GM104231.
Conflict of interest statement. None declared.

REFERENCES

1. Asin-Cayuela,J. and Gustafsson,C.M. (2007) Mitochondrial transcription and its regulation in mammalian cells. *Trends Biochem. Sci.*, **32**, 111–117.
2. Fuste,J.M., Wanrooij,S., Jemt,E., Granycome,C.E., Cluett,T.J., Shi,Y., Atanassova,N., Holt,I.J., Gustafsson,C.M. and Falkenberg,M. (2010) Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. *Mol. Cell*, **37**, 67–78.
3. Bonawitz,N.D., Clayton,D.A. and Shadel,G.S. (2006) Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol. Cell*, **24**, 813–825.
4. Minczuk,M., He,J., Duch,A.M., Ettema,T.J., Chlebowski,A., Dzionek,K., Nijtmans,L.G., Huynen,M.A. and Holt,I.J. (2011) TEFM (c17orf42) is necessary for transcription of human mtDNA. *Nucleic Acids Res.*, **39**, 4284–4299.
5. Wanrooij,P.H., Uhler,J.P., Shi,Y., Westerlund,F., Falkenberg,M. and Gustafsson,C.M. (2012)A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop. *Nucleic Acids Res.*, **40**, 10334–10344.
6. Yakubovskaya,E., Mejia,E., Byrnes,J., Hambardjieva,E. and Garcia-Diaz,M. (2010) Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. *Cell*, **141**, 982–993.
7. Bestwick,M.L. and Shadel,G.S. (2013) Accessorizing the human mitochondrial transcription machinery. *Trends Biochem. Sci.*, **38**, 283–291.
8. Hallberg,B.M. and Larsson,N.G. (2014) Making proteins in the powerhouse. *Cell Metab.*, **20**, 226–240.
9. Shutt,T.E. and Shadel,G.S. (2010) A compendium of human mitochondrial gene expression machinery with links to disease. *Environ. Mol. Mutagen.*, **51**, 360–379.
10. Litonin,D., Sologub,M., Shi,Y., Savkina,M., Anikin,M., Falkenberg,M., Gustafsson,C.M. and Temiakov,D. (2010) Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. *J. Biol. Chem.*, **285**, 18129–18133.
11. Shi,Y., Dierckx,A., Wanrooij,P.H., Wanrooij,S., Larsson,N.G., Wilhelmsson,L.M., Falkenberg,M. and Gustafsson,C.M. (2012) Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proc. Natl Acad. Sci. U.S.A.*, **109**, 16510–16515.
12. Yakubovskaya,E., Guja,K.E., Eng,E.T., Choi,W.S., Mejia,E., Beglov,D., Lukin,M., Kozakov,D. and Garcia-Diaz,M. (2014) Organization of the human mitochondrial transcription initiation complex. *Nucleic Acids Res.*, **42**, 4100–4112.
13. Morozov,Y.I., Agaronyan,K., Cheung,A.C., Anikin,M., Cramer,P. and Temiakov,D. (2014) A novel intermediate in transcription initiation by human mitochondrial RNA polymerase. *Nucleic Acids Res.*, **42**, 3884–3893.
14. Posse,V., Hoberg,E., Dierckx,A., Shahzad,S., Koolmeister,C., Larsson,N.G., Wilhelmsson,L.M., Hallberg,B.M. and Gustafsson,C.M. (2014) The amino terminal extension of mammalian mitochondrial RNA polymerase ensures promoter specific transcription initiation. *Nucleic Acids Res.*, **42**, 3638–3647.
15. Sologub,M., Litonin,D., Anikin,M., Mustaev,A. and Temiakov,D. (2009) TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase. *Cell*, **139**, 934–944.
16. Schwinghammer,K., Cheung,A., Morozov,A.I., Agaronyan,K., Temiakov,D. and Cramer,P. (2013) Structure of mitochondrial RNA polymerase elongation complex. *Nat. Struct. Mol. Biol.*, **20**, 1298–1303.
17. Dairaghi,D.J., Shadel,G.S. and Clayton,D.A. (1995) 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.
18. Ringel,R., Sologub,M., Morozov,Y.I., Litonin,D., Cramer,P. and Temiakov,D. (2011) Structure of human mitochondrial RNA polymerase. *Nature*, **478**, 269–273.
19. Grachev,M.A., Kolocheva,T.I., Lukhtanov,E.A. and Mustaev,A.A. (1987) Studies on the functional topography of *Escherichia coli* RNA polymerase. *Eur. J. Biochem.*, **163**, 113–121.
20. Korzheva,N. and Mustaev,A. (2001) Transcription elongation complex: structure and function. *Curr. Opin. Microbiol.*, **4**, 119–125.
21. Guja,K.E., Venkataraman,K., Yakubovskaya,E., Shi,H., Mejia,E., Hambardjieva,E., Karzai,A.W. and Garcia-Diaz,M. (2013) Structural basis for S-adenosylmethionine binding and methyltransferase activity by mitochondrial transcription factor B1. *Nucleic Acids Res.*, **41**, 7947–7959.
22. Schubot,F.D., Chen,C.J., Rose,J.P., Dailey,T.A., Dailey,H.A. and Wang,B.C. (2001) Crystal structure of the transcription factor sc-mtTFB offers insights into mitochondrial transcription. *Protein Sci.*, **10**, 1980–1988.
23. Waugh,S.M., DiBella,E.E. and Pilch,P.F. (1989) Isolation of a proteolytically derived domain of the insulin receptor containing the major site of cross-linking/binding. *Biochemistry*, **28**, 3448–3455.
24. Cheetham,G.M., Jeruzalmi,D. and Steitz,T.A. (1999) Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature*, **399**, 80–83.
25. Ngo,H.B., Kaiser,J.T. and Chan,D.C. (2011) The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat. Struct. Mol. Biol.*, **18**, 1290–1296.
26. Rubio-Cosials,A., Sidow,J.F., Jimenez-Menendez,N., Fernandez-Millan,P., Montoya,J., Jacobs,H.T., Coll,M., Bernado,P. and Sola,M. (2011) Human mitochondrial transcription factor A

induces a U-turn structure in the light strand promoter. *Nat. Struct. Mol. Biol.*, **18**, 1281–1289.
27. Gaspari,M., Falkenberg,M., Larsson,N.G. and Gustafsson,C.M. (2004) The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. *EMBO J.*, **23**, 4606–4614.
28. Velazquez,G., Guo,Q., Wang,L., Brieba,L.G. and Sousa,R. (2012) Conservation of promoter melting mechanisms in divergent regions of the single-subunit RNA polymerases. *Biochemistry*, **51**, 3901–3910.
29. Nayak,D., Guo,Q. and Sousa,R. (2009) A promoter recognition mechanism common to yeast mitochondrial and phage T7 RNA polymerases. *J. Biol. Chem.*, **284**, 13641–13647.
30. Gleghorn,M.L., Davydova,E.K., Rothman-Denes,L.B. and Murakami,K.S. (2008) Structural basis for DNA-hairpin promoter recognition by the bacteriophage N4 virion RNA polymerase. *Mol. Cell*, **32**, 707–717.
31. Cheetham,G.M. and Steitz,T.A. (1999) Structure of a transcribing T7 RNA polymerase initiation complex. *Science*, **286**, 2305–2309.
32. Rong,M., He,B., McAllister,W.T. and Durbin,R.K. (1998) Promoter specificity determinants of T7 RNA polymerase. *Proc. Natl Acad. Sci. U.S.A.*, **95**, 515–519.
33. Kornberg,R.D. (2007) The molecular basis of eukaryotic transcription. *Proc. Natl Acad. Sci. U.S.A.*, **104**, 12955–12961.
34. Vannini,A. and Cramer,P. (2012) Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell*, **45**, 439–446.
35. Borukhov,S. and Nudler,E. (2008) RNA polymerase: the vehicle of transcription. *Trends Microbiol.*, **16**, 126–134.
36. Sainsbury,S., Niesser,J. and Cramer,P. (2013) Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. *Nature*, **493**, 437–440.
37. Pupov,D., Kuzin,I., Bass,I. and Kulbachinskiy,A. (2014) Distinct functions of the RNA polymerase sigma subunit region 3.2 in RNA priming and promoter escape. *Nucleic Acids Res.*, **42**, 4494–4504.
38. Amiott,E.A. and Jaehning,J.A. (2006) Mitochondrial transcription is regulated via an ATP 'sensing' mechanism that couples RNA abundance to respiration. *Mol. Cell*, **22**, 329–338.