Structure and Function of Human Mitochondrial RNA Polymerase Elongation Complex

Kathrin Schwinghammer
aus München
2014

Structure and Function of Human Mitochondrial RNA Polymerase Elongation Complex

Kathrin Schwinghammer
aus München
2014

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28.November 2011 von Herrn Prof. Dr. Patrick Cramer betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde selbstständig und ohne unerlaubte Hilfe erarbeitet.

München, den 13.02.2014

_____

Kathrin Schwinghammer

Dissertation eingereicht am   13.02.2014

1. Gutachter   Prof. Dr. Patrick Cramer

2. Gutachter   PD Dr. Dietmar Martin

Mündliche Prüfung am   24.03.2014

Acknowledgements

First of all I want to thank my supervisor Prof. Dr. Patrick Cramer for giving me the opportunity to work in such an inspiring atmosphere and for all the personal support and trust during the last three years. He is an exceptional group leader that combines organization, leadership and motivation in such a perfect way that is not often found elsewhere. Dear Patrick, I wish you a good start and a successful era in Göttingen!

I especially want to thank Prof. Dmitry Temiakov for being a great stepsupervisor, collaborator and discussion partner. The times we met in the US and in Germany were always very inspiring. Let me know whenever you come to Munich, it will be a pleasure to meet up again.

Dear Elisabeth, thank you for being my mentor and friend not only during my Bachelor studies but especially during the first year of my PhD.

Many thanks to you Alan, for all the crystallographic advice and for sharing the first moments with the crystal structure. I'll never forget the magic moment when we saw the first initial nucleic acid density in the polymerase that was "so impressive because it's in all of us"! At the same time I also want to thank Sarah for introducing me to the Synchrotron in Switzerland and for making long nights at the beamline more entertaining. I'm not sure whether your theory of "the lucky meal" is true, but it is always important to consider all critical parameters.

I am also very thankful to Yarik and Karen, my US analogs, for making this huge number of mutants and biochemical assays that complemented this work in an exceptional manner.

I also want to thank Claudia and Stefan for keeping the lab running and making it a place I always liked working. Successful work always includes nice people around you that contribute to professional topics as well as to the personal sense of being. Therefore, I want to thank Larissa, Merle, Simon and Tobi for great lunch times including Maultaschen, a lot of pasta, yoghurt, fruits and the obligatory espresso. Dear Carlo, Christoph and Daniel, thank you for your extraordinary sense of humor, it is appreciated. With Hauke the Munich mtRNAP group finally became a two-men show, at least for half a year. Dear Hauke, I wish you an exciting and successful time in the Cramer laboratory, I'm glad the mtRNAP project is in very good hands now! Of course I want to thank all other members of the lab for many celebrations, cake sessions,

retreats and the great working atmosphere in general. Furthermore, I want to thank my Bachelor student Kilian for his valuable contributions to the project during his time in the lab.

I want to thank my student fellows Vroni and Chris who spent so many endless learning sessions, Weißwurst-Lunches and Topmodel evenings with me. Even though we spread over all four branches of chemistry we have a lot in common. Never forget that she "likes eating Gewürzgurken"!

Liebe Mama, lieber Papa, lieber Flo und liebe Anni, euch gebührt mein ganz besonderes "Dankeschön". Ohne euch wäre ich nicht zu dem Menschen geworden, der ich heute bin! Vielen Dank, dass ihr auf meinem bisherigen Weg stets zu mir gestanden seid und mich bei jedem Vorhaben bedingungslos unterstützt habt.

Lieber Martin, zuletzt möchte ich vorallem dir für deine Loyalität, dein Verständnis und deine Unterstützung danken. Vielen Dank, dass du mich auch in schwierigen Zeiten immer daran erinnerst was wirklich wichtig ist im Leben!

Summary

Mitochondria are often described as molecular power stations of the cell as they generate most of the energy that drives cellular processes. Mitochondria are eukaryotic organelles with bacterial origin that contain an extra-nuclear source of genetic information. Although most mitochondrial proteins are encoded in the nucleus, the mitochondrial genome still encodes key components of the oxidative phosphorylation machinery that is the major source for cellular adenosine 5'-triphosphate (ATP). The mitochondrial genome is transcribed by a singlesubunit DNA-dependent RNA polymerase (RNAP) that is distantly related to the RNAP of bacteriophage T7. Unlike its T7 homolog, mitochondrial RNA polymerase (mtRNAP) relies on two transcription factors, TFAM and TFB2M, to initiate transcription. The previously solved structure of free mtRNAP has revealed a unique pentatricopeptide repeat (PPR) domain, a N-terminal domain (NTD) that resembles the promoter-binding domain of T7 RNAP and a C-terminal catalytic domain (CTD) that is highly conserved in T7 RNAP. The CTD adopts the canonical right-hand fold of polymerases of the pol A family, in which its 'thumb', 'palm' and 'fingers' subdomains flank the active center. Since the structure represents an inactive "clenched" conformation with a partially closed active center, only limited functional insights into the mitochondrial transcription cycle have been possible so far.

This work reports the first crystal structure of the functional human mtRNAP elongation complex, determined at 2.65 Å resolution. The structure reveals a 9-base pair DNA-RNA hybrid formed between the DNA template and the RNA transcript and one turn of DNA both upstream and downstream of the hybrid. Comparisons with the distantly related T7 RNAP indicate conserved mechanisms for substrate binding and nucleotide incorporation, but also strong mechanistic differences. Whereas T7 RNAP refolds during the transition from initiation to elongation, mtRNAP adopts an intermediary conformation that is capable of elongation without NTD refolding. The intercalating hairpin that melts DNA during mtRNAP and T7 RNAP initiation additionally promotes separation of RNA from DNA during mtRNAP elongation.

The structure of the mtRNAP elongation complex (this work) and free mtRNAP (previously published) demonstrate that mtRNAP represents an evolutionary intermediate between singlesubunit and multisubunit polymerases. Furthermore, it illustrates the adaption of a phage-like RNAP to a new role in mitochondrial gene expression.

Publications

Parts of this work have been published.

Schwinghammer, K., Cheung, A.C.M., Morozov, Y.I., Agaronyan, K., Temiakov, D. and Cramer, P. **Structure of human mitochondrial RNA polymerase elongation complex**. Nat Struct Mol Biol. 20(11), 1298-303 (2013).

Author contributions:

K.A. and Y.I.M. cloned mtRNAP variants and performed biochemical assays. D.T. and K.S. performed RNAP purification and prepared crystals. K.S. and A.C.M.C. performed structure determination and modelling. D.T. and P.C. designed and supervised research. K.S., A.C.M.C., D.T. and P.C. wrote the manuscript.

Table of contents

1 Introduction

1.1 Gene transcription

Genetic information is fundamental for all life and is universally stored in form of deoxyribonucleic acid (DNA). In 1958 Francis Crick described the directional flow of genetic information from DNA via ribonucleic acid (RNA) to proteins as the "central dogma" of molecular biology (Crick, 1970). Here, transcription is the process in which RNA is synthesized from a DNA template by DNA-dependent RNA polymerases (RNAPs) (Weiss and Gladstone, 1959). Based on structural homology, RNAPs can be grouped into two classes, multisubunit and singlesubunit polymerases, that are the product of convergent evolution (Cramer, 2002a).

1.1.1 Multisubunit RNA polymerases

Gene transcription by multisubunit RNA polymerases is found over all three kingdoms of life. Whereas bacteria and archaea rely on a single multisubunit polymerase to transcribe their entire genome, eukaryotes have three multisubunit polymerases that synthesize different kinds of RNA from their nuclear genome (Roeder and Rutter, 1969). RNAP I is located in the nucleoli and transcribes the precursor of 18S, 5.8S and 28S ribosomal RNA (rRNA) (Grummt, 2003). RNAP II is located in the nucleoplasm and transcribes messenger RNA (mRNA) from all protein coding genes, small nucleolar RNAs (snoRNAs) and some small nuclear RNAs (snRNAs) (Wyers et al., 2005). Also located in the nucleoplasm, RNAP III transcribes 5S rRNA and all transfer RNAs (tRNAs) (Weinmann and Roeder, 1974; Zylber and Penman, 1971). Recently, two additional, but non-essential plant-specific RNAPs, RNAP IV and RNAP V, have been described to be involved in the formation and maintenance of heterochromatin by RNA interference (Lahmy et al., 2010; Pontier et al., 2005).

Even though multisubunit polymerases differ in their subunit composition, they all share the general overall structure of a crab claw consisting of up to 17 polypeptide subunits (Cramer, 2002b). The highly conserved active center cleft indicates a general

catalytic mechanism for all multisubunit polymerases. Variations are commonly found in peripheral subunits and accessory factors essential for transcriptional regulation (Cramer et al., 2008). Whereas RNAP initiation in bacteria relies on a single regulatory factor, the sigma factor, for promoter recognition and enzyme recruitment, archaea employ two factors, TFB and the TATA-binding protein (TBP) for transcription initiation (Geiduschek and Ouhammouch, 2005; Mooney et al., 2005). The much bigger eukaryotic RNAP I, II and III utilize a large set of regulatory factors to fulfill the cellular needs for transcription regulation (Roeder, 1996).

1.1.2 Singlesubunit RNA polymerases

Singlesubunit RNAPs are found in bacteriophages (e.g. T7 phage) and eukaryotic cell organelles (e.g. mitochondria) (Masters et al., 1987; Tiranti et al., 1997). The respective enzymes consist of only one polypeptide chain and adapt the canonical architecture of a right-hand including a palm, fingers and thumb subdomain similar to DNA polymerases (DNAPs) (Cheetham et al., 1999; Ringel et al., 2011).

The best-characterized singlesubunit RNA polymerase is the bacteriophage T7 RNAP. Over the last 18 years, several structures illuminated T7 RNAP in its initiation state (Cheetham and Steitz, 1999), the transition state from initiation to elongation phase (Yin and Steitz, 2002), the four different steps of the nucleotide addition cycle during elongation (Cheetham et al., 1999; Durniak et al., 2008; Jeruzalmi and Steitz, 1998; Tahirov et al., 2002; Temiakov et al., 2004; Yin and Steitz, 2004) and an inhibitory state in which with T7 RNAP is complexed with T7 lysozyme (Jeruzalmi and Steitz, 1998). In eukaryotes, the singlesubunit mitochondrial DNA-dependent RNA polymerase (mtRNAP) transcribes a small mitochondrial genome that encodes rRNAs, tRNAs and a few subunits of respiratory chain complexes that are involved in cellular ATP production (Sologub et al., 2009).

Despite their high degree of structural conservation, singlesubunit RNAPs serve distinct biological roles. In T7-like phages, singlesubunit RNAPs are optimized to produce large quantities of mRNA transcripts to compete the host RNAP (Studier,

1972). In contrast, mitochondrial and plastid RNAPs synthesize diverse types of RNA and must coordinate transcription with processing, editing and translation in context of the changing needs of the cell (Asin-Cayuela and Gustafsson, 2007; Yin et al., 2010). Although all these singlesubunit RNAPs are evolutionary conserved and contain a highly conserved catalytic core (Masters et al., 1987), they achieve their specific roles by using different strategies. T7 RNAP is a self-sufficient polymerase that is highly specific for its promoters (Cheetham et al., 1999). Promoter initiation is factor-independent and the transition into elongation phase is achieved by a major domain rearrangement of the N-terminal domain (NTD) (Tahirov et al., 2002; Yin and Steitz, 2002). In human mtRNAP, structural alterations observed in the promoter binding domain require the enzyme to recruit two transcription initiation factors for promoter specificity, binding and melting (Litonin et al., 2010; Ringel et al., 2011). Release of these factors marks the transition to the elongation phase of transcription, a mechanism commonly employed by multisubunit RNAPs (Borukhov and Nudler, 2008).

1.1.3 Evolution of DNA-dependent RNA polymerases

Increased genetic complexity in higher organisms does not necessarily correlate with an enlarged number of genes but rather with an increased need for gene expression and regulation (Levine and Tjian, 2003). This circumstance is reflected by the varying sequence and structure compositions of RNAPs (Levine and Tjian, 2003). Since multisubunit polymerases and bacteriophage-like singlesubunit polymerases do not share structural similarities it is likely that they have evolved from separate ancestors (Cermakian et al., 1997; Werner and Grohmann, 2011).

Multisubunit polymerases comprise a common subunit architecture including the central cleft with its three catalytic aspartate residues (Cramer et al., 2008). According to the 'RNA world hypothesis' postulated by Steitz in 1993, this enzyme class evolved from an ancient homodimeric ribozyme without any catalytic activity (Steitz and Steitz, 1993). It was suggested that during evolution the homodimeric architecture converted into a heterodimeric core, RNA components were lost and polymerase activity was acquired (Iyer et al., 2003). Through an increasing recruitment

of regulatory factors, the subunit complexity of multisubunit polymerases rises from bacteria to archaea and eukaryotes (Carter and Drouin, 2010).

Although singlesubunit polymerases do not show significant homologies with their multisubunit relatives, they provide a strong sequence and structure conservation within their class (Cermakian et al., 1997). It was postulated that they evolved from ancient DNAPs or reverse transcriptases (Cermakian et al., 1997; Delarue et al., 1990; Steitz et al., 1994). Among the six families of singlesubunit DNAPs (A, B, C, D, F, X, Y) singlesubunit RNAPs are most similar to the pol A Klenow fragment of *Escherichia coli* (*E.coli*) DNAP I (Cermakian et al., 1997; Sousa, 1996). From the phylogenetic point of view it needs to be further investigated at which stage of evolution the ancestor singlesubunit RNAP gene was acquired (Cermakian et al., 1997).

According to the widely accepted endosymbiotc theory, mitochondria evolved from an ancient bacteria that was engulfed by a primitive eukaryotic cell (Gray, 2012). A striking argument herefore is the ancestry of key components of the mitochondrial transcription and replication machinery with T7 bacteriophages (Shutt and Gray, 2006). Since phage-like genes were found in bacterial genomes, it seems likely that the mtRNAP gene was acquired as part of the endosymbiotic genome instead of a direct attendence of a phage-like entity (Shutt and Gray, 2006). Initially functioning as a primase during DNA replication, mtRNAP later acquired the ability to transcribe genes encoded in the mitochondrial genome (Shutt and Gray, 2006). With this central role in mitochondrial gene expression, mtRNAP replaced the bacterial-like multisubunit RNAP that was originally acquired from the protobacterial genome into the eukaryotic cell (Shutt and Gray, 2006).

1.1.4 The nucleotide addition cycle

Even though there are many structural and functional aspects that distinguish singlesubunit polymerases from multisubunit polymerases, they both share the conserved mechanism of nucleotide addition (Sousa, 1996; Temiakov et al., 2000).
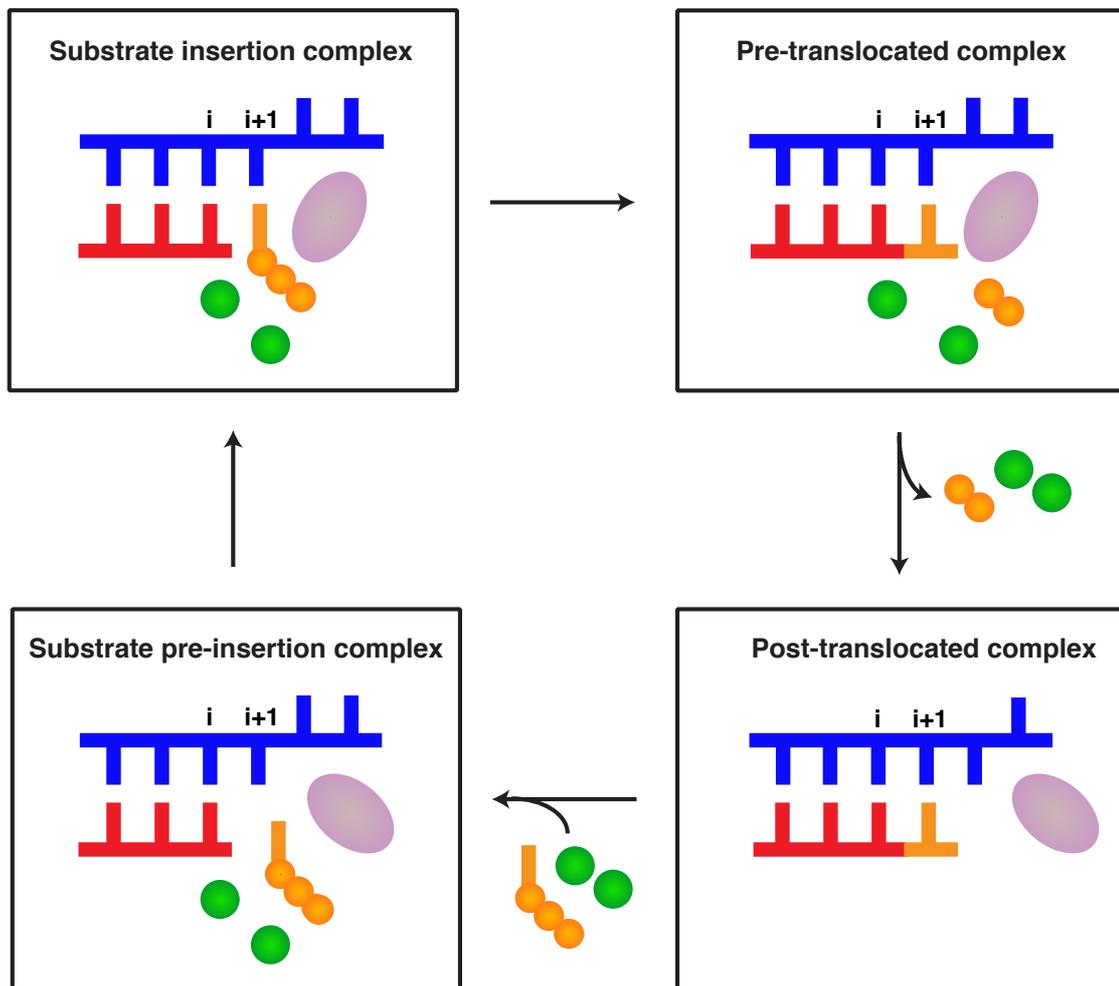
Figure 1 - Scheme of nucleotide addition cycle of RNAPs during elongation.
Nucleic acids are shown as lines (DNA, blue; RNA, red), Mg^{2+} ions (green) and the O helix of the fingers domain (pink) as spheres, nucleoside triphosphate (NTP) as line with three spheres (orange). An incoming NTP binds to the pre-insertion complex of the post-translocated RNAP (lower left). Upon a conformational change of the O helix in the RNAP fingers domain, the NTP is properly positioned for later insertion (upper left). A Mg^{2+} catalyzed phosphoryl transfer reaction results in the incorporation of the NTP at the 3'-end of the RNA, extending it by +1 and coordinating pyrophosphate (PP_i) by metal ions (upper right). The release of the PP_i and the Mg^{2+} ions is accompanied by a translocation step, enabling RNAP to bind another NTP in the insertion site again (lower right). (Scheme adapted from (Yin and Steitz, 2004)).

During recent years T7 RNAP became the best characterized singlesubunit polymerase with many functional states visualized in crystal structures (Steitz, 2009). As exemplarily shown for the T7 system, elongation can be divided into four stages

termed nucleotide addition cycle (Fig. 1). An incoming nucleoside triphosphate (NTP) approaches the active center of the post-translocated polymerase causing an open conformation due to initial interactions between the substrate phosphate backbone and two O helix residues (substrate pre-insertion complex) (Temiakov et al., 2004). A rotation of the fingers subdomain causes the active center to close and to properly position the substrate NTP for the insertion reaction (substrate insertion complex) (Yin and Steitz, 2004). A Mg^{2+} catalyzed phosphoryl transfer reaction results in the extension of the nascent RNA chain by one nucleotide. The pyrophosphate (PP_i) forms an ionic cross-link with both a metal ion and the protein (pre-translocated complex) (Yin and Steitz, 2004). Dissociation of PP_i and Mg^{2+} ions is accompanied by the formation of an open complex and the translocation of the DNA-RNA hybrid (post-translocated complex) (Yin and Steitz, 2004). Another conformational change in the fingers subdomain causes the unwinding of the downstream DNA duplex by one base pair. DNA backtracking is avoided by a stacking interaction of a tyrosine residue into the insertion site of the post-translocated complex until another NTP is bound for the next round of the nucleotide addition cycle (Sousa, 1996).

Due to the high sequence and structure homology between mtRNAP and T7 RNAP it was suggested that the nucleotide addition cycle in mitochondria is conserved (Masters et al., 1987; Ringel et al., 2011).

1.2   Origin and function of mitochondria

Mitochondria are eukaryotic dual-membrane organelles that contain their own genome. The outer membrane separates the organelle from the cellular cytosol, whereas the inner membrane forms inward foldings called cistrae. Mitochondria are the power stations of the cell since they are responsible for adenosine 5'-triphosphate (ATP) synthesis through their oxidative phosphorylation system (OXPHOS) (Hatefi, 1985). Beside its role in energy production, the mitochondrion is the stage for a variety of other important metabolic processes, such as the regulation of apoptosis, nucleotide biosynthesis, control of cytosolic calcium concentration, cellular differentiation and fatty acid metabolism (Brookes et al., 2002; Carafoli, 1970; Chen et al., 2012; Green and Reed, 1998; Ott et al., 2007). Remarkably, only genes involved in OXPHOS are encoded in the mitochondrial genome itself (Bonawitz et al., 2006).

The origin of mitochondria is still highly debated. The maintenance of its own genome is the most striking evidence that mitochondria are derived from ancient bacteria (Gray and Doolittle, 1982). The generally accepted endosymbiotic hypothesis suggests that the mitochondrion was inherited from an α-proteobacterium that developed a symbiotic relationship with a primitive eukaryotic cell over two billion years ago (Martin and Muller, 1998). Phylogenic data suggests that this partnership enabled them to use increasing amounts of oxygen in the atmosphere in a non-toxic way (Andersson et al., 2003). Over time, bacterial genes were either lost or transferred from the mitochondrial to the nuclear genome (Martin et al., 2005). Today, except for some OXPHOS genes, most proteins needed in the mitochondrion are encoded in the nuclear genome (Becker et al., 2012). There are three potential reasons why the cell still accepts the high effort of keeping some genes encoded in the mitochondrion (Adams and Palmer, 2003). First, some proteins might be too hydrophobic for being imported across the mitochondrial membrane into the organelle (Popot and de Vitry, 1990). This seems plausible since the two OXPHOS genes encoding cytochrome b and cytochrome c oxidase subunit I are two of the most hydrophobic proteins in a eukaryotic cell (Claros et al., 1995; Popot and de Vitry, 1990; von Heijne, 1986). Second, mitochondria and the nucleus might have evolved a different codon usage that makes mitochondrial genes unreadable in the nucleus and most likely stopped further gene transfer (Andersson and Kurland, 1991). Third, direct gene expression within the mitochondrion may be crucial for a metabolic control mechanism that regulates the response to energy requirements in eukaryotes (Allen, 1993). In general a small genome makes it easier to quickly respond to environmental changes (Wallace, 2007). Mitochondrial gene expression may be directly influenced by the oxidative state or the activity of the electron transport chain in mitochondria. A similar example of a rapid and direct redox control was found in chloroplasts of plants (Pfannschmidt et al., 1999).

During evolution, the mitochondrial genome may have lost some genes whose function is replaced by unrelated genes of the nucleus (Gray and Lang, 1998). One prominent example here is the substitution of the originally multisubunit bacteria-like RNA polymerase by a singlesubunit bacteriophage-like T7 RNAP responsible for mitochondrial transcription (see also chapter 1.1.3). Regardless of the several reasons for gene transfer, the crosstalk between both genomes has been maintained throughout evolution to efficiently regulate mitochondrial activities (Gray and Lang, 1998).

1.3   The mitochondrial transcription machinery

1.3.1   The mitochondrial genome

The mitochondrial DNA (mtDNA) is a double-stranded, circular genome that represents the only extra-nuclear source of DNA in mammals (Nass, 1966). In contrast to its nuclear relative, mtDNA is inherited maternally as mitochondria from sperm cells are actively eliminated during early stages of the cell development (Sutovsky et al., 1997). The mitochondrial genome is organized in histone-free structures, the so-called nucleoids (Bogenhagen et al., 2008; Bogenhagen et al., 2003). Depending on their tissue specific energy demand, cells contain between 1,000 to 10,000 copies of mtDNA (Shadel and Clayton, 1997; Taylor et al., 2005). Cells with a huge energy usage like brain, liver and muscle cells contain a higher copy number of mtDNA (Bonawitz et al., 2006).

Both strands of the mtDNA provide an uneven nucleotide content and were therefore characterized as guanine rich (heavy) and guanine poor (light) DNA strand (Anderson et al., 1981). Although the size of mtDNA varies from 16.6 kbp in human to 75 kbp in yeast *Saccharomyces cerevisiae* (*S.c.*) it always encodes for 37 genes: the heavy strand encodes for two rRNAs of mitochondrial ribosomes, 12 mRNAs of the approximately 80 key subunits of the oxidative phosphorylation machinery and 14 tRNAs essential for mitochondrial translation, whereas the light strand encodes for only one mRNA and 8 tRNAs (Fig. 2) (Anderson et al., 1981). The rest of the approximately 1,500 proteins needed for the metabolic activity of mitochondria are encoded in the nuclear genome, transcribed by nuclear RNAPs, synthesized in the cytosol and imported into mitochondria via a cleavable N-terminal mitochondrial localization signal (MLS) sequence (Mokranjac and Neupert, 2005). Similarly, the basic components of the mitochondrial transcription machinery are not encoded in the organelle itself. Consequently mitochondrial transcription regulation relies on both genomes. Another unique feature of the human mitochondrial genome is the lack of introns (Gaspari et al., 2004b). Gene sequences are so closely arranged that some even overlap. The only major non-coding region was characterized as displacement loop (D-loop) since both genomic DNA strands are displaced through a third, 500-700 bp heavy strand DNA product (7S DNA) (Shadel and Clayton, 1997).
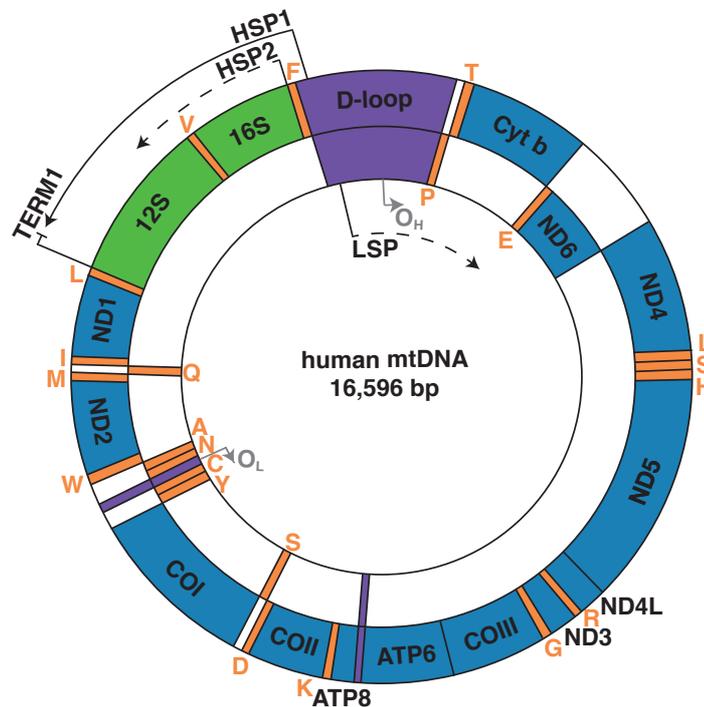
Figure 2 - Schematic map of the human mitochondrial genome.
The heavy and the light strand are depicted as the outer and inner circle respectively, comprising coding regions for mRNA (blue), rRNA (green), tRNA (orange) and non-coding regions (violet). Transcription is initiated from two promoters on the heavy strand (HSP1 and HSP2) and only one promoter on the light strand (LSP). Termination of transcripts from the HSP1 is introduced downstream of the 12S rRNA by binding of the mitochondrial transcription termination factor mTerf1 to its binding region (TERM1). Replication of mtDNA is initiated from one origin of each strand (O_H and O_L). (Scheme adapted from (Greaves et al., 2012).)

The D-loop accommodates well-conserved regulatory elements for transcription and replication (Gaspari et al., 2004b). A second non-coding element for mitochondrial replication is located in a minor non-coding region roughly 5,000 bp apart from the D-loop. Transcription in mitochondria is initiated on the strand specific promoters named light strand promoter (LSP) and heavy strand promoters 1 and 2 (HSP1 and HSP2) (Fig. 2). Transcripts generated from LSP or HSP2 have genomic length, i.e. encompass all genetic information of the respective strand, and are subsequently processed in individual species of RNA (Montoya et al., 1982). Transcription from the HSP1 is terminated after synthesis of the 12S rRNA (Clayton, 1991; Ojala et al., 1981).

9

An earlier study has shown that the transcription rate from HSP1 is more than 50 times higher than from HSP2 (Gelfand and Attardi, 1981). Therefore, the existence of two HSPs could be due to a flexible regulation of the ratio of rRNA to mRNA in respect of physiological changes (Kucej et al., 2008).

1.3.2 Mitochondrial RNA polymerase

The mitochondrial genome is transcribed by the singlesubunit polymerase mtRNAP. Unlike most other known eukaryotic polymerases, mtRNAP is not related to multisubunit polymerases in bacteria (Masters et al., 1987). Instead, mtRNAP comprises extensive sequence homology with singlesubunit RNAPs encoded by T3 and T7 bacteriophages (Cermakian et al., 1997).

Although the human mtRNAP was initially identified in 1997 (Tiranti et al., 1997), it took another 14 years to gain further insights into its structural features (Nayak et al., 2009; Ringel et al., 2011). MtRNAP comprises three major domains, characterized as the highly conserved C-terminal domain (CTD), the minor conserved NTD and an N-terminal extension domain (NED) that is missing in the coding sequence of T7 RNAP (Fig. 3).

The CTD (residues 648-1230) can also be classified as the catalytic domain, as it harbors regions that are involved in essential polymerase activities like DNA template and nucleotide binding as well as nucleotide incorporation. As shown in a recent crystal structure, the CTD adopts the canonical right-hand fold that is typical for members of the pol A family (Joyce and Steitz, 1994; Ringel et al., 2011). A 'thumb,' 'palm' and 'fingers' subdomain flank the active center (Ringel et al., 2011). Within the palm domain, the highly conserved aspartic acids, D922 and D1151, coordinate two divalent Mg^{2+} cations that are essential for catalytic activity of the polymerase (Smidansky et al., 2011). The O helix, which is part of the fingers domain, also contributes to catalysis as well as substrate selection and translocation of the nascent RNA strand (Doublie and Ellenberger, 1998; Kiefer et al., 1997; Yin and Steitz, 2002).
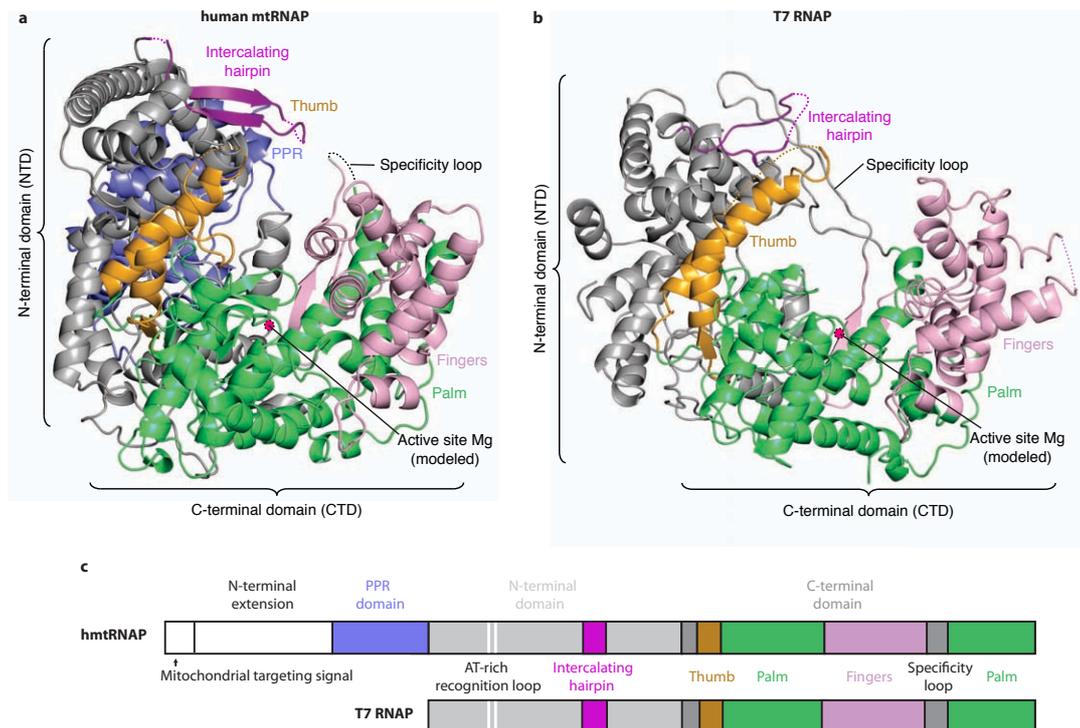
Figure 3 - Domain structure of free human mtRNAP and T7 RNAP determined by X-ray crystallography.

(a) MtRNAP (PDB code 3SPA, (Ringel et al., 2011)) is depicted as a ribbon (orange, thumb; green, palm; pink, fingers; purple, intercalating hairpin; slate, pentratricopeptide repeat (PPR). The N-terminal extension domain (NED, residues 1-217), a part of the intercalating hairpin (residues 592-602), the specificity loop (residues 1086-1105) and half of the thumb subdomain (residues 726-769) are unstructured in the crystal structure and therefore represented as dashed lines. A Mg^{2+} ion (magenta) was placed according to a T7 RNAP structure (Yin and Steitz, 2004).

(b) T7 RNAP (PDB code 1ARO, (Jeruzalmi and Steitz, 1998)) structural domains are colored as in (a). The catalytic Mg^{2+} ion was also placed according to another T7 RNAP structure (Yin and Steitz, 2004). The co-crystallized lysozyme moiety was omitted for clarity.

(c) Schematic domain comparison of mtRNAP and T7 RNAP. Structural elements are highlighted in the same color code as in (a) and (b). Beneath a highly conserved CTD and a minor conserved NTD mtRNAP comprises a PPR domain and a NED domain. (Scheme adapted from (Ringel et al., 2011)).

Since the recent crystal structure of free mtRNAP reveals an inactive 'clenched' conformation with a partially closed active center, further functional insights are restrained (Ringel et al., 2011). Another structural element of the fingers subdomain is

the specificity loop that contributes to promoter recognition and the formation of the RNA exit channel in the T7 system (Paratkar and Patel, 2010; Temiakov et al., 2000; Yin and Steitz, 2002). No structural or functional analogy could be assigned for the specificity loop in human mtRNAP. Recent studies in yeast revealed that the *S.c.* RNAP (Rpo41) utilizes similar structural elements to specifically recognize the promoter sequence in the absence of transcription factors (Matsunaga and Jaehning, 2004b; Nayak et al., 2009).

In contrast to mtRNAP, T7 RNAP possesses a short insertion in the fingers domain, termed fingers flap that interacts with the downstream DNA duplex during transcription elongation. In the mitochondrial system this function could have been overtaken by additional transcription factors (Guo et al., 2005).

The NTD (residues 369-647) comprises two loops that correspond to functional elements in T7 RNAP: the AT-rich recognition loop and the intercalating hairpin (Steitz, 2009; Temiakov et al., 2004). The AT-rich recognition loop binds promoter DNA during initiation of T7 RNAP but is sequestered by a pentatricopeptide repeat (PPR) domain in mtRNAP and not required for mtRNAP initiation (Ringel et al., 2011). In the RNAP of bacteriphage N4, the AT-rich recognition loop is capable of specifically recognizing hairpin-shaped promoters (Davydova et al., 2007). Its specific role in the mitochondrial transcription system needs to be further investigated. The intercalating hairpin is involved in promoter melting, as shown by a deletion mutant that was not able to initiate transcription from double-stranded promoter templates (Ringel et al., 2011). In the T7 system the intercalating hairpin also melts DNA during transcription initiation but is repositioned far away from the nucleic acids during the transition from initiation to elongation in which a massive NTD refolding takes place (Yin and Steitz, 2002). It is unknown whether a similar refolding of the NTD occurs in mtRNAP and what the function of the intercalating hairpin during mitochondrial transcription elongation is.

The NED (residues 1-368) shows the highest degree of sequence variability between different species (Cermakian et al., 1997; Masters et al., 1987). Again, not much is known about this region in the human system. In yeast, the NED serves as a binding platform for transcription and translation factors as well as RNA processing

proteins (Paratkar et al., 2011; Rodeheffer and Shadel, 2003). The NED is attached to the NTD via a short proline-rich linker and comprises a MLS sequence, an uncharacterized, flexible region and a PPR domain (Ringel et al., 2011). The PPR domain consists of two tandemly arranged 35 residue repeats. These domains are exclusively found in plant and mitochondrial proteins which are involved in RNA editing and processing events (Delannoy et al., 2007; Small and Peeters, 2000). The need for the PPR domain in mtRNAP of higher eukaryotes is unknown. NED deletion studies in human mtRNAP showed that this domain is required for promoter specific transcription, but not for polymerase activity itself (Ringel et al., 2011). This result, in combination with the tight association of NED with the rest of human mtRNAP, indicates the functional importance of this domain (Ringel et al., 2011).

Various studies discovered that mtRNAP provides additional, transcription-independent functions such as ribosomal biogenesis (Surovtseva and Shadel, 2013). Since yeast mtRNAP functions as an ATP-sensor, it seems likely that human mtRNAP can also adjust protein expression levels in response to fluctuations in the ATP pool of mitochondria (Amiott and Jaehning, 2006). Even though the nuclear encoded mtRNAP is usually imported into mitochondria, an alternative splicing form was observed that accumulated in the nucleus for unidentified reasons (Kravchenko et al., 2005). Taken together, mtRNAP is not only the main component of the mitochondrial transcription machinery but also functions as a bridging element to other regulatory pathways.

1.3.3 Transcription factors

In order to efficiently initiate mitochondrial transcription mtRNAP relies on two transcription factors: TFAM and TFB1M or TFB2M (Fig. 4). Hence, the basal human mitochondrial transcription machinery *in vitro* consists of mtRNAP, TFAM, TFB1M or TFB2M and a DNA template containing HSP or LSP sequence (Falkenberg et al., 2002). Both mtRNAP and Rpo41 can initiate transcription factor-independently on pre-melted promoter sequences (Litonin et al., 2010; Matsunaga and Jaehning, 2004a). This indicates that initiation factors are exclusively needed for promoter recognition, melting.
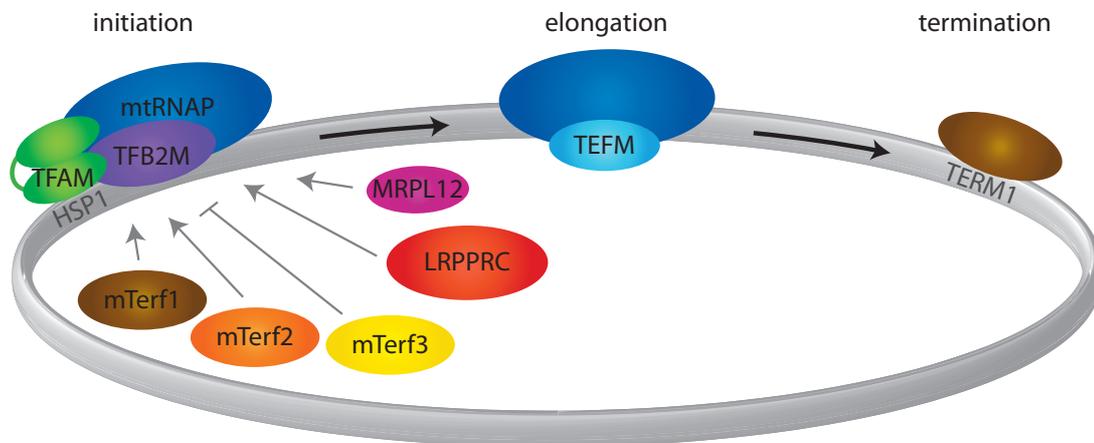
Figure 4 - Scheme of the human mitochondrial transcription machinery.
After specific TFAM (green) binding to the mitochondrial promoter DNA (e.g. HSP1), mtRNAP (dark blue) and TFB2M (purple) are recruited and form the mitochondrial initiation complex (IC). Regulatory factors that are discussed in the following chapter have been shown to have stimulating (↑) or inhibiting (T) effects on the IC. Whereas LRPPRC (red) and MRLP12 (magenta) directly interact with mtRNAP, it needs to be further investigated how members of the mTerf1 family interact with the transcription machinery (brown, mTerf1; orange, mTerf2; yellow, mTerf3). MTerf1 induces HSP1-dependent termination by binding to a 22 bp region (TERM1) on the heavy strand of the mitochondrial genome. TEFM (light blue) was identified as the mitochondrial elongation factor as it enhances mtRNAP processivity *in vitro*.

The need for transcription factors represents a major functional difference to the T7 system. Unlike mtRNAP, T7 RNAP can initiate transcription without the recruitment of additional factors (Chamberlin et al., 1983). Whereas mitochondrial transcription factors are released during the transition from initiation to elongation phase, the NTD of T7 RNAP undergoes an extensive structural rearrangement (Mangus et al., 1994; Yin and Steitz, 2002). Thereby, the contacts with the promoter sequence are lost and an RNA exit tunnel is formed by sub domain H, part of the NTD and the specificity loop (Tahirov et al., 2002; Yin and Steitz, 2002).

1.3.3.1   TFAM

TFAM was the first identified human mitochondrial factor that is recruited by mtRNAP to initiate transcription (Fisher and Clayton, 1985; Larsson et al., 1997). It is encoded in

the nuclear genome, synthesized in the cytoplasm and imported into mitochondria with the help of an N-terminal MLS sequence that is cleaved after translocation (Parisi and Clayton, 1991). The 25 kDa protein comprises the two high mobility group (HMG) boxes A and B, a 27 aa linker region and a 25 aa C-terminal domain (Fisher and Clayton, 1988). Like other members of the ubiquitous HMG box family of DNA binding proteins (Parisi and Clayton, 1991), TFAM can specifically or non-specifically bind, unwind and bend DNA. HMG box A is mainly responsible for DNA contacts, whereas HMG box B has only weak DNA affinity (Gangelhoff et al., 2009). Several deletion studies showed that the TFAM C-terminal domain is required for specific promoter binding during initiation (Gangelhoff et al., 2009). Two recent crystal structures showed that TFAM induces a U-turn in the promoter sequence (Ngo et al., 2011; Rubio-Cosials et al., 2011). Together with the linker region, each HMG domain stabilizes a kink of 90° by a series of basic amino acids that contact the negatively charged phosphate backbone of the DNA. Whether TFAM binds promoter DNA as a monomer or a dimer is still under debate (Gangelhoff et al., 2009). Recent studies indicate that TFAM binds to the NTD of mtRNAP, resulting in a promoter DNA bend around the polymerase (Morozov et al., 2014; Posse et al., 2014).

TFAM is required for transcription initiation, from LSP and HSP1 but not from HSP2 (Fisher and Clayton, 1985; Fisher et al., 1987; Litonin et al., 2010). Specific promoter selection is controlled in a TFAM concentration-dependent manner: LSP initiated transcription is activated under low TFAM concentrations, whereas transcription activity switches to HSP1 with increasing TFAM concentrations and transcription inhibition in the presence of TFAM over expression (Shutt et al., 2010). A tunable TFAM activity at different promoter regions may be needed to adjust protein synthesis to environmental changes (Rebelo et al., 2011).

In addition to its function in promoter selection and transcription activation, TFAM also contributes to mitochondrial genome compaction and mtDNA copy control (Alam et al., 2003; Kaufman et al., 2007). Due to its unspecific DNA binding ability, TFAM is, together with other proteins, involved in nucleoid formation in human mitochondria (Kang et al., 2007; Ruhanen et al., 2010; Spelbrink et al., 2001; Wang and Bogenhagen, 2006). Increasing amounts of bound TFAM correlates with a

decrease of DNA accessibility for other DNA binding proteins (Alam et al., 2003; Fisher and Brown, 1980; Rebelo et al., 2011). High TFAM concentrations were shown to destabilize mtDNA *in vivo*, suggesting the importance of TFAM in cellular homeostasis and regulation of nucleoid activity (Ekstrand et al., 2004). TFAM stability itself may be regulated via post-translational phosphorylation of the protein or other interacting factors that are not identified as such yet (Lu et al., 2013; Matsushima et al., 2010).

In general, TFAM induced conformational changes in the DNA both affect transcription and nucleoid stability, suggesting that the mitochondrial genome organization is coupled to transcription, similar to the bacterial system (Ohniwa et al., 2007).

In yeast, the TFAM homologue Abf2 also compacts mtDNA but does not have any activating contribution in transcription initiation due to the lack of the C-terminal domain (Diffley and Stillman, 1991). Therefore, the yeast mitochondrial transcription machinery is not a three-component system as found in human mitochondria, but a two-component system.

1.3.3.2 TFB2M

The third component of the human transcription machinery is TFB2M. Similar to the other components of the transcription machinery, it is encoded in the nuclear genome and translated across the mitochondrial membrane. TFB2M was originally identified together with a second protein named TFB1M (Falkenberg et al., 2002). Both proteins share a high sequence homology with an ancestral bacterial rRNA methyltransferase and are capable to dimethylate 12S rRNA of mitochondrial ribosomes *in vitro* (Cotney et al., 2009; Sologub et al., 2009). During evolution the function of TFB2M and TFB1M diverged, due to the variety of regulatory needs of mitochondria (McCulloch and Shadel, 2003). Recent studies revealed, that only TFB1M retained its rRNA methyltransferase activity and assists in the biogenesis of the small subunit of the mitochondrial ribosome (Seidel-Rogol et al., 2003).

TFB2M on the other side lost its methylransferase activity during evolution and adapted the ability to activate mitochondrial transcription initiation (Sologub et al.,

2009). Although both proteins were able to stimulate initiation *in vitro*, TFB2M was discovered to be several magnitudes more efficient than TFB1M (Falkenberg et al., 2002). In addition, its transcriptional contribution is independent of the rRNA methyltransferase domain (Cotney et al., 2009) or non-specific DNA-binding affinities (McCulloch and Shadel, 2003). Although TFB2M does not provide any promoter recognition activity, it assists in promoter melting and contributes to an open complex formation (Gaspari et al., 2004a; Sologub et al., 2009). Moreover, TFB2M facilitates binding of the priming nucleotide in the active center of mtRNAP by a transient interaction of its N-terminal domain with the +1 and +3 bases of the DNA template strand (Litonin et al., 2010; Lodeiro et al., 2010; Sologub et al., 2009). Whether the overall structure of the mitochondrial initiation complex is stabilized by a direct interaction of TFB2M with the second essential transcription factor TFAM is still under debate (McCulloch and Shadel, 2003; Morozov et al., 2014). TFB2M binding affinities for the mtRNAP were only discovered in the yeast system (Diffley and Stillman, 1991). The yeast homologue of TFB2M, the mitochondrial transcription factor 1 (Mtf1) forms an interactive two-component system with Rpo41 for mitochondrial transcription, independent of the presence of Abf2 (Paratkar et al., 2011; Paratkar and Patel, 2010).

1.3.3.3   TEFM

Although current research focuses more and more on the investigation of mtRNAP regulatory factors, the transcription elongation factor of mitochondria (TEFM) was only recently identified (Minczuk et al., 2011). Based on a sequence homology with the bacterial Holliday Junction Resolvase (HJR), TEFM was initially characterized as a putative mitochondrial HJR, which was not confirmed during later experiments (Connolly et al., 1991; Minczuk et al., 2011). Instead, there are three indications that TEFM functions as a mitochondrial elongation factor. First, TEFM provides an RnaseH fold and two tandem helix-hairpin-helix (HhH) domains which are also present in the nuclear transcription factor Spt6, and the bacterial regulator protein Tex (Ponting, 2002). Similar to Spt6, which directly interacts with RNAP II, TEFM is capable of binding to the catalytic region of mtRNAP (Minczuk et al., 2011). Second, TEFM was shown to enhance mtRNAP processivity *in vitro* (Minczuk et al., 2011). Third, TEFM co-

localizes with newly synthesized RNA and may therefore contribute to the processing of polycistronic transcripts from mitochondrial promoters (Minczuk et al., 2011).

To provide a complete picture of mtRNAP transcription, the regulatory function and interaction network of TEFM need to be further investigated in the future.

1.3.3.4 Other regulatory factors involved in mitochondrial transcription

Besides initiation and elongation, transcription termination is also a highly regulated process. In contrast to polycistronic transcripts from HSP2, transcripts from HSP1 are immediately terminated downstream of both rRNA genes (Montoya et al., 1982). HSP1-dependent termination is induced by the mitochondrial termination factor 1 (mTerf1) that specifically binds with its conserved five-arginine-motif to a 22 bp region within the tRNALeu(UUR) gene (TERM1, Fig. 4) (Kruse et al., 1989; Roberti et al., 2006). MTerf1 can simultaneously bind to both TERM1 and HSP1 itself, forming a DNA-loop that assists in recycling components of the core transcription machinery back to the promoter (Martin et al., 2005). A recent study suggests an additional field of mTerf1 activity, as it seems to be involved in modulation of replicational pausing (Hyvarinen et al., 2007).

Besides mTerf1, the prototype of the mTerf family, mTerf2 and mTerf3 also adopt roles in mitochondrial transcription and gene expression. Depletion studies showed that mTerf2 represents a positive and mTerf3 a negative regulator of transcription of the mitochondrial genome (Park et al., 2007; Wenz et al., 2009).

The mitochondrial leucine-rich pentatricopeptide repeat containing protein (LRPPRC) comprises not only two PPR domains as mtRNAP, but 22 domains (Mili and Pinol-Roma, 2003). LRPPRC is involved in multiple stages of the mitochondrial RNA metabolism (Chujo et al., 2012; Ruzzenente et al., 2012). LRPPRC stimulates transcriptional activity of mtRNAP *in vitro*, most likely through direct interactions with mtRNAP or other regulatory proteins (Liu et al., 2011; Sondheimer et al., 2010).

The mitochondrial ribosomal protein L12 (MRLP12) is a component of the large subunit of mitochondrial ribosomes (Surovtseva et al., 2011). In its ribosome-free form it was characterized as a mtRNAP interactor with transcription activating properties (Wang et al., 2007).

Although the human mitochondrial genome is relatively small, it relies on a variety of regulatory proteins with multiple activities each. It needs to be further investigated if and how all these different factors interact with the primary transcription machinery. This will help to draw a complete picture of the detailed regulatory mechanisms controlling transcription of the mitochondrial genome.

1.3.4 Mitochondrial replication

Replication of the mitochondrial genome is independent of the cell cycle or the nuclear replication processes (Bogenhagen and Clayton, 1977; Pica-Mattoccia and Attardi, 1972). The duplication of the mtDNA is carried out by the replisome that consists of exclusively nuclear-encoded proteins: the DNAP γ (Burgers et al., 2001), mitochondrial single-stranded DNA binding proteins (mtSSB) (Korhonen et al., 2004), the mitochondrial DNA helicase TWINKLE (Spelbrink et al., 2001), topoisomerases (Zhang et al., 2001) and RNaseH (Cerritelli et al., 2003).

Two models for mitochondrial replication are under current discussion. In the strand-coupled bidirectional replication model multiple replication origins cause symmetrical DNA synthesis on both the leading and the lagging strand (Holt and Jacobs, 2003; Yang et al., 2002). In the asynchronous strand-displacement model, replication of the heavy strand is initiated from the origin of replication (O_H) in the D-loop region. After DNA synthesis of the heavy strand has proceeded to two thirds of the genome it runs into the origin of replication on the light strand promoter (O_L). The disposed O_L forms a stem-loop structure that initiates replication of the light strand. DNA synthesis continuously proceeds until the full circle of the mitochondrial genome is reached (Brown et al., 2005; Tapper and Clayton, 1981; Wong and Clayton, 1985).

A unique feature of the mitochondrial replisome is the lack of primases. Instead, mtRNAP synthesizes the short RNA primers needed for replication initiation (Wanrooij et al., 2008). Transcription initiated from the LSP generates transcripts that can subsequently be processed into short-length primers essential for replication initiation at the O_H (Xu and Clayton, 1996). Although mtRNAP is highly processive on double-

stranded DNA, it is also capable to synthesize 25-27 bp long transcripts that are used as primers for DNA duplication by DNAP γ (Wanrooij et al., 2008). Therefore, activation of the second DNA strand is achieved by binding of mtRNAP to the single-stranded O_L stem-loop structure (Chang and Clayton, 1985; Fuste et al., 2010).

Another link between mitochondrial transcription and replication is indicated by the transcription factor TFAM that indirectly stimulates replication initiation (Kang and Hamasaki, 2005) and pausing (Hyvarinen et al., 2007). Even though a close interplay between transcriptional and replicational proteins is essential, this has not been shown through physical interactions.

1.4   Mitochondrial dysfunctions

DNA damage has an intrinsic effect on gene stability and gene expression. Since mitochondria are the stage for many metabolic processes, it is not surprising that they provide a high risk for disorders. Mitochondria are semi-autonomous organelles that require proteins encoded in both the nuclear and the mitochondrial genome (Holt et al., 1988; Wallace et al., 1988; Zeviani et al., 1989). Therefore, mutations in both genomes can lead to mitochondrial diseases (Larsson and Clayton, 1995). Even though only a minority of the mitochondrial proteins is encoded in the organelle itself, mtDNA underlies a higher mutation rate than the nuclear genome (Brown et al., 1979; Calvo and Mootha, 2010). This can be due to a reduced set of DNA repair mechanisms in mitochondria compared to the nuclear DNA repair pathways (Liu and Demple, 2010). Since mtRNAP was found to arrest at damaged genomic sites and TFAM may mark DNA damage by interaction with p53, mitochondria might also provide a mechanism of transcription-coupled DNA maintenance (Cline et al., 2010; Wong and Clayton, 1985; Yoshida et al., 2003).

The second reason for an increased number of mtDNA mutations is the oxidative environment of the mitochondrial matrix caused by reactive oxygen species (ROS) that are generated as a side product of OXPHOS. Tissues with a high energy

demand like brain, heart or muscle tissues are more sensitive to mitochondrial dysfunctions than others (Wallace et al., 2010). Among the over 300 observed pathogenic mtDNA mutations, defects in the ATP production represent the major cause for cellular disorders and show a wide range of phenotypes (McFarland et al., 2010; MITOMAP, 2013; Wallace et al., 2010). Dysfunctions in the respiratory chain have been linked to neurodegenerative defects, such as Alzheimer's or Parkinson's disease, (Trifunovic et al., 2004; Weissman et al., 2007) as well as an increased risk for breast and prostate cancer (Canter et al., 2005; Pedersen, 1978; Petros et al., 2005). In addition, mitochondrial dysfunctions are also involved in cell aging, as the accumulation of mtDNA mutations over time can lead to a decline of mitochondrial function (Miquel et al., 1980).

Although mitochondria enable the cell to perform a variety of essential cellular processes, defects in a single pathway can cause a severe threat for human health. The further investigation and identification of potential molecular triggers leading to mitochondrial diseases will be a major task for future research.

1.5 Aims and scope of this work

The singlesubunit mtRNAP occupies an exceptional position in the evolution of RNAPs, as it comprises properties of both, singlesubunit and multisubunit RNAPs. On the one hand, mtRNAP shares a high sequence and structure homology with the RNAP of bacteriophage T7 (Masters et al., 1987). Both polymerases are equally capable to specifically recognize promoter DNA (Matsunaga and Jaehning, 2004a). On the other hand, mtRNAP relies on additional factors to initiate and regulate mitochondrial transcription (Litonin et al., 2010). This is a common strategy of the structurally unrelated multisubunit polymerases, such as RNAP II (Gnatt et al., 2001a). Identifying more details about the molecular mechanisms in mitochondria will allow a deeper comprehension of evolutionary relationships between phages, bacteria and eukaryotes.

Although mtRNAP has been studied more extensively in recent years, detailed mechanistic insights into the mitochondrial transcription cycle are still lacking. Until today there is only one crystal structure of mtRNAP available (Ringel et al., 2011). The

herein identified 'clenched' conformation of mtRNAP is unlikely to represent a functional state during transcription. Therefore, the major intention of this work was to visualize mtRNAP in its elongating conformation and to expand the knowledge of mtRNAP activity. To gain insights into the elongation phase of mitochondrial transcription, a combination of X-ray crystallography, transcription assays and cross-linking experiments was used. Structural and mechanistically comparisons of the mitochondrial system with the T7 system were used to facilitate the understanding of the mitochondrial transcription cycle on a molecular level.

At the same time, this work represents an important step towards future attempts to investigate larger mtRNAP complexes comprising transcription initiation factors and regulatory factors.

Since mitochondrial dysfunction can cause severe disorders and cell aging, the reported molecular insights into mtRNAP elongation contribute to disease related research and anti-viral drug design.

2 Materials and Methods

2.1 Materials

2.1.1 Bacterial strains

Table 1 - Bacterial strains

Strain	Genotype	Company
XL1-blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacIqZΔM15* Tn*10* (Tetr)]	Stratagene
BL21-CodonPlus® (DE3) RIL	*E.coli* B F- ompT hsdS(r - m -) dcm+ Tetr gal endA Hte [argU ileY leuW Camr]	Stratagene

2.1.2 Plasmids

Table 2 - Plasmids

Plasmid	Insert	Type	Tag	Restriction sites
Δ150mtRNAP	residues 151-1230 of human mtRNAP, vector with mutation in NcoI cutting site, by Dmitry Temiakov	pProExHb	N-term His$_6$	NcoI, XhoI

2.1.3 Synthetic oligonucleotides

Oligonucleotides purchased from metabion (Germany) were HPLC-purified, delivered lyophilized and dissolved in TE buffer to a final concentration of 1.6 mM. Oligonucleotides purchased from IDT DNA (USA) were standard-desalted, delivered lyophilized and also dissolved in TE buffer to a final concentration of 1.6 mM. RNA Oligonucleotides purchased from Dharmacon Inc (USA) were synthesized 2'-ACE protected, standard-desalted, delivered lyophilized, deprotected and dissolved in TE buffer to a final concentration of 1.6 mM.

Table 3 - DNA oligonucleotides used for crystallization

Name	Sequence 5′→3′	Scaffold	Source
DKS01	TAG TGC ATA CCG CCA	CC2	metabion
DKS02	TCT TTT GGC GGT ATG CAC T	CC2	metabion
DKS03	TGT TAG TTG GGG GGT GAC TGT TAA AAG TGC ATA CCG CCA AAA GAT AAG G	CC1	metabion
DKS04	AAT TAT CTT TTG GCG GTA TGC ACT TTT AAC AGT CAC CCC CCA ACT AAC A	CC1	metabion
DKS05	AAA AGT GCA TAC CGC CA	CC4	metabion
DKS08	TGG CGG TAT GCA CTT TT	CC4	metabion
DKS09	TGT TAA AAG TGC ATA CCT TAT CCC GAT A	OC1	metabion
DKS10	TAT CTT TTG GCG GTA TGC ACT TTT AAC A	OC1	metabion
DKS11	AAA AGT GCA TAC CTT ATC CCG ATA AAA TT	OC2	metabion
DKS12	AAT TTT ATC TTT TGG CGG TAT GCA CTT TT	OC2	metabion
DKS13	TGT TAA AAG TGC ATA CCT TAT CCC GAT AAA ATT	OC3	metabion
DKS14	AAT TTT ATC TTT TGG CGG TAT GCA CTT TTA ACA	OC3	metabion
DKS15	CGC CAG ACA GG	EC2,3	metabion
DKS17	CCT GTC TGG CGT GCG CGC CGC	EC3	metabion
DKS18	GGG GTT GTA GCT TAT GTC GAA GTA TGG GAG	EC4	metabion
DKS19	CTC CCA TAC TAA TCT CAT CAA TAC AAC CCC	EC4	metabion
DKS20	GGG AAT GCA TGG CGC GGC	EC5	metabion
DKS21	CCT GTC TGG CGT GCG CGC CGG	EC2	metabion
DKS22	GTG CAT ACC GTA TCC CCA TAG GAT TGG	OC4	metabion
DKS23	CCA ATC CTA TCT TTT GGC GGT ATG CAC	OC4	metabion
DKS27	GGG GTA GCT TAT GTC GAA GTA TGG GAG	EC6	metabion
DKS28	CTC CCA TAC TAA TCT CAT CAA TAC CCC	EC6	metabion
DKS29	GGG GTA GCT TAT GTC GAA GTG TG	EC7	metabion
DKS30	CAC ACT AAT CTC ATC AAT ACC CC	EC7	metabion
DKS31	CATGGGGTAATTATTTCGACTGACGCAG	EC8-10	metabion
DKS32	GGG GTA ATT ATT TCG ACT GAC GCA G	EC11-13	metabion
DKS33	ATT ATT TCG ACT GAC GCA G	EC14,15,32	metabion
DKS34	ACT GAC GCA G	EC16,17,33	metabion
DKS35	GGG GTA ATT ATT TCG ACT GAC GC	EC18-20	metabion
DKS36	ATT ATT TCG ACT GAC GC	EC21,22	metabion
DKS37	ACT GAC GC	EC23,24	metabion
DKS38	GGG GTA ATT ATT TCG ACT GAC	EC25-27	metabion

Name	Sequence 5′→3′	Scaffold	Source
DKS39	ATT ATT TCG ACT GAC	EC28,29,34	metabion
DKS40	ACT GAC	EC31,32,35	metabion
DKS41	CTG CGT CAG TGC GGG CCG GTA CCC CAT G	EC8-10	metabion
DKS42	CTG CGT CAG TGC GGG CCG GTA CCC C	EC11-13	metabion
DKS43	CTG CGT CAG TGC GGG CCG G	EC14-17, 32,33	metabion
DKS44	GCG TCA GTG CGG GCC GGT ACC CC	EC18-20	metabion
DKS45	GCG TCA GTG CGG GCC GG	EC21-24	metabion
DKS46	GTC AGT GCG GGC CGG TAC CCC	EC25-27	metabion
DKS47	GTC AGT GCG GGC CGG	EC28-31, 34,35	metabion
DKS51	CATG GGG TAA TTA TTT CGA CGC CAG ACG	EC36	metabion
DKS70	CGT CTG GCG TGC GCG CCG GTA CCC CAT G	EC36	metabion
DT01	ACG CCA GAC AGG	EC1	IDT DNA
DT02	CCT GTC TGG CGT GCG GCG CCG	EC1	IDT DNA
NT02	CAT GGG GTA ATT ATT TCG ACG CCA GAC G	DT1-3,6	IDT DNA
NT03	GTC GAT TTC AGA CAG GAC CC	DT5	IDT DNA
NT06	CAT GGG GTA ATT ATT TTC ATC GCC AGA CG	DT4	IDT DNA
TS01	GGG TCC TGT CTG AAA TCG ACA TCG CCG C	DT5	IDT DNA
TS02	CGT CTG GCG TGC GCG CCG CTA CCC CAT G	DT1,3,6	IDT DNA
TS0X	CGT CTG GCG TGC GCG CCG TTA CCC CAT G	DT2	IDT DNA
TS06	CGT CTG GCG ATC GCG CCG CTA CCC CAT G	DT4	IDT DNA
TS35sU	CCT GTC TGA ATC GAU* ATC GCC GC	DT7	IDT DNA
YMNT1	GCG GCG ATC ATT CGC TTG ACA GG	DT7	IDT DNA

Table 4 - RNA oligonucleotides used for crystallization

Name	Sequence 5′→3′	Scaffold	Source
R14mt	AGU CUG CGG CGC GC	DT1,2,EC1	Dharmacon
RS11sU	GAG U*GC GGC GA	DT5	Dharmacon
R15mtsU	AU*G UCU GCG GCG CGC	DT6	Dharmacon
R20mt	GAA GAC AGU CUG CGG CGC GC	DT3	Dharmacon
mtR12G	GUC UGC GGC GCG	DT4	Dharmacon
RKS01	GUC UGC CCG GCG CGC	EC2	metabion
RKS02	GCG CGC	EC3	metabion
RKS03	UUU UUA GUU GAU GAG AU	EC4	metabion
RKS04	UUU UGC CGC GCC A	EC5	metabion

Table continued on next page

Name	Sequence 5′→3′	Scaffold	Source
RKS05	UUA GUU GAU GAG AU	EC6,7	metabion
RKS06	CUG CCC GGC CCG C	EC8,11,18,25,32-35	metabion
RKS07	CCG GCC CGC	EC9,12,14,16,19,21,23,26,28,30	metabion
RKS08	GCC CGC	EC10,13,15,17,20,22,24,27,29,31	metabion
YMRNA1	UCG CUC GAU UCA	DT7	Dharmacon

+1

EC8
5' CATGGGGTA ATTATTTCG ACTGACGCAG 3' DKS31
3' GTACCCCAT GGCCGGGGC GTGACTGCGTC 5' DKS41
5' CUGC CCGGGCCGC 3' RKS06

EC9
5' CATGGGGTA ATTATTTCG ACTGACGCAG 3' DKS31
3' GTACCCCAT GGCCGGGGC GTGACTGCGTC 5' DKS41
5' CCGGCCCGC 3' RKS07

E10
5' CATGGGGTA ATTATTTCG ACTGACGCAG 3' DKS31
3' GTACCCCAT GGCCGGGGC GTGACTGCGTC 5' DKS41
5' GCCCGC 3' RKS08

E11
5' GGGGTA ATTATTTCG ACTGACGCAG 3' DKS32
3' CCCCATGGCC GGGGC GTGACTGCGTC 5' DKS42
5' CUGC CCGGGCCGC 3' RKS06

E12
5' GGGGTA ATTATTTCG ACTGACGCAG 3' DKS32
3' CCCCATGGCC GGGGC GTGACTGCGTC 5' DKS42
5' CCGGCCCGC 3' RKS07

E13
5' GGGGTA ATTATTTCG ACTGACGCAG 3' DKS32
3' CCCCATGGCC GGGGC GTGACTGCGTC 5' DKS42
5' GCCCGC 3' RKS08

E14
5' ATTATTTCG ACTGACGCAG 3' DKS33
3' GGCCGGGGC GTGACTGCGTC 5' DKS43
5' CCGGGCCGC 3' RKS07

E15
5' ATTATTTCG ACTGACGCAG 3' DKS33
3' GGCCGGGGC GTGACTGCGTC 5' DKS43
5' GCCCGC 3' RKS08

E16
5' ACTGACGCAG 3' DKS34
3' GGCCGGGGC GTGACTGCGTC 5' DKS43
5' CCGGCCCGC 3' RKS07

E17
5' ACTGACGCAG 3' DKS34
3' GGCCGGGGC GTGACTGCGTC 5' DKS43
5' GCCCGC 3' RKS08

E18
5' GGGGTA ATTATTTCG ACTGACGC 3' DKS35
3' CCCCATGGCC GGGGC GTGACTGCG 5' DKS44
5' CUGC CCGGCCCGC 3' RKS06

EC19
5' GGGGTA ATTATTTCG ACTGACGC 3' DKS35
3' CCCCATGGCC GGGGC GTGACTGCG 5' DKS44
5' CCGGCCCGC 3' RKS07

EC20
5' GGGGTA ATTATTTCG ACTGACGC 3' DKS35
3' CCCCATGGCC GGGGC GTGACTGCG 5' DKS44
5' GCCCGC 3' RKS08

EC21
5' ATTATTTCG ACTGACGC 3' DKS36
3' GGCCGGGGC GTGACTGCG 5' DKS45
5' CCGGCCCGC 3' RKS07

Figure 5 - Schematic overview of all scaffolds used in this study.
The nucleic acid scaffolds contain template DNA (blue), non-template DNA (cyan) and RNA (red).

2.1.4 Media and additives

All chemicals used to prepare buffers or other solutions had p.a. quality and were produced by one of the following companies: Bio-Rad, Biozyme, Dianova-Jackson, Fisher Scientific, Fluka, Merck, Invitrogen, Roth, Sigma-Aldrich and VWR.

Table 5 - Media for *E.coli* cultivation

Name	Description
LB	1% (w/v) tryptone; 0.5% (w/v) yeast extract; 0.5% (w/v) NaCl; (+2% (w/v) agar for selective media plates)

Table 6 - Additives for *E.coli* cultivation

Name	Stock solution	Applied concentration
Ampicillin	10% (w/v) ampicillin	0.1% (w/v) ampicillin
Tetracycline	1.25% (w/v) tetracycline in EtOH	0.00125% (w/v) tetracycline
Chloramphenicol	30 mg/mL chloramphenicol in EtOH	30 µg/mL chloramphenicol
IPTG	1 M IPTG	0.1 mM IPTG

2.1.5 Buffers, markers, solutions and enzymes

Table 7 - General buffers and solutions

Name	Description or source	Application
100 × Protease inhibitor	60 µM leupeptin, 200 µM pepstatin A, 98 mM PMSF, 211 mM benzamidine; in EtOH	Protein purification
1 × TBE	8.9 mM Tris-HCl; 8.9 mM boric acid; 2 mM EDTA; pH 8.0 at 25°C	Agarose gel electrophoresis
6 × DNA loading dye	Fermentas	Agarose gel electrophoresis
Gene Ruler 1 kb DNA ladder (0.1 µg/µL)	Fermentas	Agarose gel electrophoresis
SYBR Safe (10,000 × in DMSO)	Invitrogen	Agarose gel electrophoresis
20 × MES SDS running buffer	50 mM MES; 50 mM Tris Base; 0.1% SDS; 1 mM EDTA; pH 7.3 at 25°C	SDS-PAGE
20 × MOPS SDS running buffer	50 mM MOPS; 50 mM Tris Base; 0.1% SDS; 1 mM EDTA; pH 7.7 at 25°C	SDS-PAGE
5x SDS sample buffer	250 mM Tris-HCl (pH 7.0 at 25°C); 50% (v/v) glycerol; 0.5% (w/v) bromophenol blue; 7.5% (w/v) SDS; 500 mM DTT	SDS-PAGE
Broad range MW marker	Bio-Rad	SDS-PAGE
Coomassie gel staining solution	50% (v/v) ethanol; 7% (v/v) acetic acid; 0.125% (w/v) Coomassie Brilliant Blue R-250	SDS-PAGE
Instant Blue	Expedeon	SDS-PAGE
Destain solution	5% (v/v) EtOH; 7.5% (v/v) acetic acid	SDS-PAGE
Instant coomassie	10 mM MOPS (pH 7.0 at 25°C); 10 mM RbCl; 75 mM $CaCl_2$; 15% (v/v) glycerol	SDS-PAGE
TFB-I	30 mM K acetate; 50 mM $MnCl_2$; 100 mM RbCl; 10 mM $CaCl_2$; 15% (v/v) glycerol; pH 5.8 at 25°C	Competent cells
TFB-II	10 mM MOPS (pH 7.0 at 25°C); 10 mM RbCl; 75 mM $CaCl_2$; 15% (v/v) glycerol	Competent cells
1 × TE	10 mM Tris-HCl (pH 8.0 at 25°C); 1 mM EDTA	Various
1000 × SYPRO Orange	Invitrogen	Thermal shift assay
Primer extension buffer	20 mM Tris (pH 7.9 at 20°C); 10 mM $MgCl_2$; 10 mM DTT; 0.05% (v/v) Tween 20	Primer extension assay
Transcription run-off buffer	40 mM Tris (pH 7.9 at 20°C); 10 mM $MgCl_2$; 10 mM DTT	Transcription run-off assay

Table 8 - Protein purification buffer

Name	Description
Buffer A	40 mM Tris-HCl (pH 8.0 at 25°C); 300 mM NaCl; 5% glycerol; 5 mM DTT
Buffer B	40 mM Tris-HCl (pH 8.0 at 25°C); 1.5 M NaCl; 5% glycerol; 5 mM DTT
Buffer C	40 mM Tris-HCl (pH 8.0 at 25°C); 1.5 M NaCl; 5% glycerol; 200 mM imidazole; 5 mM DTT
Buffer D	40 mM Tris-HCl (pH 8.0 at 25°C); 300 mM NaCl; 5% glycerol; 1 mM EDTA; 5 mM DTT
Buffer E	40 mM Tris-HCl (pH 8.0 at 25°C); 5% glycerol; 5 mM DTT
Buffer F	40 mM Tris-HCl (pH 8.0 at 25°C); 2 M NaCl; 5% glycerol; 5 mM DTT
Buffer G	100 mM Tris-HCl (pH 8.0 at 25°C); 100 mM NaCl; 5% glycerol; 0.1 mM EDTA; 5 mM DTT
Buffer H	100 mM Tris-HCl (pH 8.0 at 25°C); 200 mM NaCl; 5% glycerol; 0.1 mM EDTA; 5 mM DTT
Buffer I	100 mM Tris-HCl (pH 8.0 at 25°C); 300 mM NaCl; 5% glycerol; 0.1 mM EDTA; 5 mM DTT

Table 9 - Components used for crystallization

Name	Description or source
Endoproteinase ArgC	Sigma-Aldrich
100 mM ATP	Jena Bioscience
100 mM GTP	Jena Bioscience
100 mM 3'dATP	Jena Bioscience
100 mM 3'dGTP	Jena Bioscience
100 mM AMPCPP	Jena Bioscience
100 mM GMPCPP	Jena Bioscience
Cryo solution	10% PEG 4000; 60 mM Na acetate; 30 mM trisodium citrate; 25% glycerol; 5% ethylene glycol

Table 10 - Enzymes, buffers and components used for PCR and plasmid cloning

Name	Source
dNTP mix, 2 mM each	Fermentas
DMSO	Fermentas
Phusion High-Fidelity DNA polymerase (2 U/µL)	Finnzymes
5 × Phusion HF buffer	Finnzymes

Name	Source
Taq DNA Polymerase (recombinant)	Fermentas
10 × Taq Buffer with KCl	Fermentas
25 mM MgCl$_2$	Fermentas
NotI (10,000 U/mL)	New England Biolabs
XhoI (20,000 U/mL)	New England Biolabs
10 × NEBuffer 4	New England Biolabs
CIP	New England Biolabs
100 × BSA	New England Biolabs
T4 DNA ligase	New England Biolabs
10 × T4 DNA ligase reaction buffer	New England Biolabs
Quick T4 DNA ligase	New England Biolabs
2 × Quick ligation buffer	New England Biolabs

2.1.6 Crystallization screens

Table 11 - Crystallization screens

Name	Abbreviation	Source
Classic Lite Suite	NCL	QIAGEN
Complex screen 1	COM	Crystallization facility MPI
Complex screen 1	CO2	Crystallization facility MP
CP-PEGS-Salt screen	PSA	Crystallization facility MP
Cryos Suite	NCO	QIAGEN
Index HT	IND	Hampton Research
AMSO$_4$ Suite	NAS	QIAGEN
Cation Suite	NCA	QIAGEN
Morpheus	MFU	Crystallization facility MPI
PACT Suite	PAC	QIAGEN
Wizard I II	NWU	Crystallization facility MPI

2.2   Methods

2.2.1   Molecular cloning

2.2.1.1   DNA amplification by polymerase chain reaction (PCR)

PCR primers were designed with a 5'-overhang consisting of eight nucleotides followed by the desired restriction site and 20-25 nt complementary to the target sequence. The PCR primers had a melting temperature of 50-65°C and a GC content of 40-60% with a G or a C at their 3'-ends. 50 µL PCR mix typically contained 1-50 ng template DNA, 0.5 µM of each DNA primer, 200 µM of each dNTP and the standard concentration of Phusion High-Fidelity DNA polymerase and its respective reaction buffer. DNA amplification was performed in a Biometra T3000 Thermocycler over 30 cycles. Annealing temperatures and elongation times were adjusted to the respective primers and the length of the desired amplification product. In order to verify the success of the reaction, 5 µL of the PCR products were visualized by agarose gel electrophoresis. Remaining DNA was purified using the QIAquick PCR Purification Kit (QIAGEN).

2.2.1.2   Restriction digest and vector dephosphorylation

Enzymatic reactions were performed for 3 h at 37°C. 50 µL reaction contained DNA obtained by PCR or 1-5 µg vector DNA and respective amounts of restriction enzymes (New England Biolabs) as stated in the manufacturer's manual. Digested vectors were subsequently dephosphorylated by the addition of 1 u CIP enzyme according to the manufacturer's recommendations. Digested DNA was purified using the QIAquick Gel Extraction Kit (QIAGEN).

2.2.1.3   Enzymatic ligation

DNA amounts of insert and linearized plasmid were estimated by analyzing 1 µL each by agarose gel electrophoresis. The reaction mixture of 20 µL contained T4 DNA ligase, respective buffer, plasmid DNA and 2-10 fold molar excess of the insert. The ligation mixture was incubated for 1 h at 20°C.

2.2.1.4   Transformation into chemically competent *E.coli* cells

50 µL chemically competent cells were thawed on ice. After the addition of 1-10 ng plasmid DNA or 10 µL ligation product cells were incubated for 30 min on ice. A heat shock was applied for 30 sec at 42°C. Cells were incubated for 2 min on ice and mixed with 500 µL of LB medium. After incubation for 1 h at 37°C shaking vigorously the cells were plated on prewarmed LB-Ampicilin (Amp) plates for selection and incubated at 37°C overnight.

2.2.1.5   Transformation into electrocompetent *E.coli* cells

50 µL electrocompetent cells were thawed on ice. After the addition of 5 µL ligation product cells were transferred into a prechilled Gene Pulser cuvette (0.2 cm gap, Bio-Rad) and exposed to a 2.5 kV pulse using a MicroPulser electroporation apparatus (Bio-Rad). Immediately, 200 µL of LB medium were added and the cells suspension transferred to a 1.5 mL reaction tube, incubated at 37°C for 1 h and plated on prewarmed LB-Amp plates. The plates were incubated at 37°C for 20-24 h.

2.2.1.6   Plasmid amplification and isolation

A preculture of 5 mL LB-Amp was inoculated with a single XL1-Blue colony picked from a LB-Amp plate and incubated at 37°C overnight, shaking at 160 rpm for cell growth. For each protein three colonies were picked. Cells were harvested by centrifugation and purified using QIAprep Spin Miniprep Kit (QIAGEN).

2.2.1.7   Test restriction digest, colony PCR and sequencing

Only plasmids containing the desired DNA insert were relevant for further procedure. Therefore, a test restriction digest was performed in a total volume of 20 µL containing 1 µL plasmid DNA, 0.3 µL of each respective restriction enzymes and buffer (New England Biolabs). Test restriction mixture was incubated for 1 h at the respective temperature and analyzed by agarose gel electrophoresis.

Alternatively colony PCR was used to verify the sequence of a larger number of clones. A colony from the transformation plate was picked with a pipette tip and dipped into 50 µL PCR reaction mixture containing 0.64 µM of each primer, 150 µM dNTPs, 2.5 mM $MgCl_2$, 5% DMSO and 1.5 u Taq DNA polymerase (Fermentas) and 1 × Taq Pol buffer with KCl. The same tip was used to transfer cells on a LB-Amp plate which

was then incubated overnight at 37°C. PCR reactions were performed as described before and analyzed by agarose gel electrophoresis. A clone that was positively tested to contain the desired insert was further analyzed by sequencing at GATC Biotech.

2.2.1.8   Preparation of chemically competent *E.coli* cells

The selection of the antibiotic depends on the respective bacteria cells. XL1-Blue cells are resistant to tetracycline (Tet) and BL21 (DE3) CodonPlus® RIL to chloramphenicol. 400 mL LB medium with antibiotic were inoculated with an overnight culture of the desired strain and incubated at 37°C under shaking at 160 rpm, until an OD_{600} of ~ 0.5 was reached. Cells were cooled on ice and always kept on ice or at 4°C in the following. After centrifugation (10 min, 3,700 × g) the pellet was resuspended in 100 mL TFB-I. The bacteria were pelleted during a second centrifugation step and resuspended in 8 mL TFB-II. 50 µL aliquots of the cell solution were frozen in liquid nitrogen and stored at -80°C. The transformation competence was tested by test transformation. The transformation competence results from the number of colonies in correlation with to the amount of DNA added.

2.2.1.9   Preparation of electrocompetent *E.coli* cells

Electrocompetent XL1-blue cells were used to reach a better transformation efficiency after ligation. 400 mL LB-Tet were inoculated with an overnight culture and incubated at 37°C under shaking at 160 rpm, until an OD_{600} of ~ 0.5 was reached. Cells were cooled on ice and always kept on ice or at 4°C in the following. After centrifugation (10 min, 1,000 × g) the pellet was resuspended in 100 mL sterile H_2O. The bacteria were pelleted during a second centrifugation step and resuspended in 400 mL sterile H_2O. The centrifugation step was repeated and cells resuspended in 20 mL sterile 10% (v/v) glycerol. After another round of centrifugation (10 min, 5,000 × g) cells were resupended in 3 mL sterile 10% (v/v) glycerol. 50 µL aliquots of the cell solution were frozen in liquid nitrogen and stored at -80°C. transformation competence was tested by test transformation. The competence results from the number of colonies in correlation with to the amount of DNA added

2.2.2 General protein methods

2.2.2.1   Protein concentration determination

Protein concentrations were determined according to the Bradford assay (Bradford, 1976). A 1:5 dilution of Bio-Rad Protein Assay dye reagent (Bio-Rad) was used to measure the absorption at a wavelength of 595 nm in a BioPhotometer (Eppendorf). Protein concentration was calculated based on the standard absorption of each batch determined with bovine serum albumin. Alternatively protein concentrations were determined based on the absorption at a wavelength of 280 nm measured with a NanoDrop 1000 spectrophotometer (Peqlab).

2.2.2.2   Trichloroacetic acid (TCA) precipitation

In order to visualize low concentrated proteins within a sample by SDS-PAGE, TCA precipitation was used. For this purpose the sample was mixed with TCA to a final TCA concentration of 10 % (v/v). After a 30 min incubation on ice the solution was pelleted by centrifugation (20 min, 16,100 rcf, 4°C). The pellet washed with 1 mL of cold acetone and centrifuged for (5 min, 16,100 rcf, 4°C). The acetone was removed and the pellet dried by air. After the addition of 5-10 µL 1 × SDS loading dye the sample was incubated for 5 min at 95°C and analyzed by SDS-PAGE. In case the solution turned yellow, indicating an acidic pH, the solution was neutralized by exposing it with the gas phase of a 25 % NH3-solution until the sample turned blue.

2.2.2.3   SDS-PAGE for protein separation

Proteins were analyzed by vertical SDS-PAGE. Polyacrylamide gradient gels (NuPage Novex 4-12% Bis Tris Gel 1.0 mm, Invitrogen) were run in a Novex Mini Cell (Invitrogen) using MOPS or MES running buffers (Invitrogen). Before loading the samples onto the gel they were mixed with the appropriate amount 5 × SDS loading dye and incubated for 5 min at 95°C to denature proteins. Gels were stained with Instant Blue (Expedeon) for 1 h.

2.2.2.4   Mass spectrometry

To identify purified proteins mass spectroscopy was used. For this purpose a respective band of a coomassie-stained SDS gel was cut out with a clean scalpel and

analyzed by the Zentrallabor für Proteinanalytik of the Ludwig-Maximilians-University of Munich.

2.2.2.5   Edman sequencing

To determine a part of the primary sequence of a protein it was transferred to a membrane by Western blot and characterized by Edman sequencing. In order to assemble the blotting chamber the blotting frame was prepared with several layers that contain a sponge, three whatman papers, a polyvinylidene fluoride (PVDF) membrane, gel, three whatman papers and a sponge. All components were soaked with transfer buffer. Prior to its transfer into the buffer solution the PVDF membrane was incubated 1 min in ethanol. Air bubbles between the layers of the blotting frame were avoided to allow an optimal current flow and complete protein transfer. The blotting chamber was filled with transfer buffer. The blot was run at 100 mV for 1 h at 4°C. The membrane was stained with Ponceau S. solution and washed with water until only the protein bands were stained. The bands of interest were cut out with a scalpel, washed with 10 % ethanol and dried by air. Edman sequencing was performed at the core facility of the Max Planck Institute for Biochemistry in Martinsried (Germany).

2.2.2.6   Dynamic light scattering

Dynamic light scattering was used to determine size distribution of protein solutions. When monochromatic light hits small molecules that undergo Brownian motion in solution it is scattered and causes time-dependent fluctuations in the scattering intensity. Due to their higher average velocity small molecules cause a greater shift in light frequency. Therefore fluctuations are related to the size of the particles. 70 µL samples with a concentration of 1.6 µg/µL were transferred in a quartz cuvette and measured with a Viscotek 802 DLS (Malvern Instruments).

2.2.2.7   Limited proteolysis

Limited proteolysis was done using the endoproteinase ArgC (Sigma). The protein in the respective gelfiltration buffer G,H or I was optionally incubated with a 1.3-fold molar excess of elongation scaffold for 10 min at 20°C. The sample was mixed with ArgC in a protein:enzyme ratio of 1000:1 (w/w) and incubated at 23°C for 1 h. The enzymatic

reaction was stopped by the addition of SDS sample buffer and boiling for 5 min at 95°C. Protein fragments were analyzed by SDS-PAGE as described above.

2.2.2.8   Thermal shift assay

Buffer conditions were optimized via a thermal shift assay. Upon protein denaturation hydrophobic regions are exposed and become favored docking sites for the fluorophore SYPRO Orange. Therefore, the stabilizing effect of a buffer solution can be measured corresponding to the level of protein denaturation over a temperature gradient. The turning point between the folded and the unfolded protein state is defined as T_m and represents a comparative parameter.

The buffer screen was performed in 50 µL reactions comprising each 5 µg protein, 50 mM buffer and 1x SYPRO Orange. The screen covered a pH range from 5.6 to 9.0 (sodium citrate pH 5.6, MES pH 6.0, MES pH 6.5, HEPES pH 7.0, HEPES pH 7.5, Tris pH 8.0, Tris pH 8.5, CAPSO pH 9.0) and salt concentrations from 0 to 750 mM NaCl in a 96-well-plate. The samples were mixed, sealed and put into a Real-Time PCR cycler. Fluorescent detection was measured at 472 nm for each temperature from 20°C to 95°C in 1°C steps.

2.2.3   Recombinant protein purification

2.2.3.1   Human mitochondrial RNA polymerase

Cells were grown in LB medium at 37°C to an OD_{600} of 0.6. Expression was induced with 0.15 mM IPTG for 18 h overnight. Cells were harvested by centrifugation, resuspended in buffer A and lysed by sonification (Sonifier Cell Disrupter, Branson). Protein was incubated for 1 h with nickel-nitrilotriacetic acid agarose (Ni-NTA) beads equilibrated with buffer A. After washing the beads with 8 CV buffer B the protein was eluted with 4 CV buffer C. The sample was dialyzed against buffer D overnight at 4°C, centrifuged and loaded onto a HiPrep Heparin FF 16/30 cation exchange column (GE Healthcare) equilibrated with buffer E. Bound protein was eluted with a salt gradient from 7.5-60% buffer F in buffer E. Column fractions were checked via SDS-PAGE, pooled and concentrated using Amicon Ultra centrifugal filter devices with a cutoff of 50K (GE Healthcare). The sample was applied to a Superdex 200 10/300 GL size

exclusion column (GE Healthcare) equilibrated with buffer G, H or I. Resulting peak fractions were pooled and concentrated to a 5-9 mg/mL, aliquoted, flash frozen in liquid N_2 and stored at -80°C until usage.

2.2.3.2 Transcription factors

Expression and purification of TFAM and TFB2M for biochemical assays was carried out by the laboratory of Prof. Dmitry Temiakov (Rowan University, SOM, Stratford, NJ, USA) as described elsewhere (Sologub et al., 2009).

2.2.4 X-ray crystallographic analysis of mtRNAP elongation complexes

2.2.4.1 Nucleic acid scaffold formation

In order to form nucleic acid complexes, synthetic oligonucleotides of template DNA, non-template DNA and RNA were mixed in equimolar amounts to a final concentration of 0.5 mM. Annealing was performed in a Biometra T3000 Thermocycler. After a total volume of 20-40 µL mixture was heated to 95°C for 180 sec was reduced every 90 s by 1°C to 20°C final.

2.2.4.2 Binding study and protein-nucleic acid complex formation by gelfiltration

To observe a possible interaction between Δ150mtRNAP and the nucleic acid scaffold via gel filtration 1-1.8 nmol enzyme were mixed with a 1.3-fold molar excess of the scaffold of interest, diluted with buffer G, H or I to a total volume of 250 µL and incubated for 10 min at 20°C. A Superdex 200 10/300 GL size exclusion column (GE Healthcare) was equilibrated with the respective gelfiltration buffer (G, H or I). The sample was centrifuged (16,100 rcf, 4°C) for 10 min to remove possible particles and loaded on the column. Gelfiltration buffer G, H or I was used as running buffer.

2.2.4.3 Assembly of the mtRNAP elongation complex

The mtRNAP elongation complex was assembled by incubating Δ150mtRNAP (40 mM) with a 1.3-fold molar excess of nucleic acid scaffold for 10 min at 20°C. For

crystallization the mtRNAP elongation complex was digested *in situ* with ArgC protease from Sigma (1000:1, w/w) for 1 h at 23°C.

2.2.4.4   Crystallization screening

Crystallization drops with a total volume of 200 nL were set at room temperature or 8°C by a Phoenix nanodisperser robot at the Max Planck Institute of Biochemistry in Martinsried. Each screen was performed with 15 µL protein solution per plate and 10 µL protein solution in total excess. Optionally screens were supplemented with 10% glycerol and 120 mM DTT. In order to prevent possible crystallization seeds that cause early precipitation the protein solution was centrifuged (16,100 rcf, 4°C) 10 min.

2.2.4.5   Crystallization setup, crystal harvesting and freezing

Promising conditions obtained by screening were optimized by fine screening with varying pH, precipitate and salt concentrations in 24-well sitting drop plates. Therefore a drop of 1 µL mtRNAP elongation complex and 1 µl of reservoir solution (8% PEG 4000, 200 mM sodium acetate, 100 mM trisodium citrate (pH=5.5), 10% glycerol, 120 mM DTT) was incubated with 500 µL total reservoir solution at 20°C. Truncated rhombic dodecahedron crystals grew to a maximum size of approximately 0.2×0.2×0.2 mm within 4-6 days. Crystals were slowly transferred in cryo solution, mounted onto cryo loops and frozen in liquid N_2.

2.2.4.6   Soaking of substrate molecules

For soaking potential substrate molecules (ATP, AMPCPP or 3'dATP/PP_i) into the mtRNAP elongation complex, crystals were transferred into regular cryo solution as described above. Subsequently crystals were incubated in cryo solution supplemented with 20-50 µM substrate for 1 sec, 1 min, 5 min or longer time periods and frozen in liquid N_2.

2.2.4.7   X-ray diffraction measurement using synchrotron radiation

Diffraction data were collected in 0.25° increments at the protein crystallography beamline X06SA of the SLS in Villigen (Switzerland) using a Pilatus 6M pixel detector (Broennimann et al., 2006) and a wavelength of 0.91809 Å.

2.2.4.8   Data processing, refinement and model building

Raw data were integrated and scaled with XDS (Kabsch, 2010) and MOSFLM (Leslie, 2006). The structure was solved by molecular replacement using PHASER (McCoy et al., 2005) with the structure of human mtRNAP (PDB code 3SPA) (Ringel et al., 2011) as a search model. The molecular replacement solution was subjected to rigid-body refinement with phenix.refine (Afonine et al., 2005). The model was iteratively built with COOT (Emsley and Cowtan, 2004) and refined with phenix.refine (Afonine et al., 2005) and autoBuster (Global Phasing Limited). The structure and diffraction data of the human mtRNAP elongation complex have been deposited in the Protein Data Bank under the accession code 4BOC. All structural figures shown in this work were prepared using pymol (DeLano, 2002).

2.2.5   *In vitro* biochemical assays

All *in vitro* assays described in this chapter were performed by the laboratory of Prof. Dmitry Temiakov (Rowan University, SOM, Stratford, NJ, USA).

2.2.5.1   Primer extension assays

The catalytic activity of mtRNAP mutants was analyzed using a primer extension assay. An *in vitro* transcription system containing radioactively labeled scaffold (50 nM), mtRNAP (150 nM), TFAM (50 nM), TFB2M (150 nM), substrate NTPs (0.3 mM) were incubated for 2 min in primer extension buffer at 35°C. The reaction was stopped by the addition of an equal volume of 95% formamide in 0.05 M EDTA. The products were resolved using 20% PAGE containing 6 M urea and visualized by PhosphoImager (GE Health) (Temiakov et al., 2002).

2.2.5.2   Transcription run-off assays

Run-off transcription assays were performed using PCR DNA templates (50 nM) containing LSP promoter (nucleotides 338-478 in human mtDNA) and mRNAP (150 nM), TFAM (50 nM), TFB2M (150 nM), substrate NTPs (0.3 mM) in a transcription buffer containing 40 mM Tris (pH 7.9), 10 mM $MgCl_2$ and 10 mM DTT. Reactions were carried out at 35°C and stopped by the addition of an equal volume of 95% formamide

in 0.05 M EDTA. The products were resolved using 20% PAGE containing 6 M urea and visualized by PhosphoImager (GE Health) (Sologub et al., 2009).

2.2.5.3 Photo-cross-linking

RNA or DNA oligonucleotides containing photo reactive 4-thio-uridine monophosphate (Dharmacon Inc.) were used to assemble DNA-RNA scaffolds. For cross-linking of the RNA base at position -8, the elongation complex (1 mM) was assembled using the RS11sU-TS1-NT3 scaffold (Fig. 12a) and the RNA was labeled by incorporation of $[\alpha\text{-}^{32}P]$ uridine triphosphate (UTP) (800 Ci/mmol) for 5 min at room temperature. For cross-linking of the RNA base at position -13, the elongation complex (1 mM) was assembled using the R15mt_sU-TS02-NT02 scaffold in which the RNA primer was ^{32}P-labeled (Fig. 12b). For DNA-mtRNAP cross-linking the elongation complex (1 mM) was assembled using the YMRNA1-TS35sU-YMNT1 scaffold in which the TS35sU DNA was ^{32}P-labeled (Fig. 12c). The cross-linking was activated by ultraviolet (UV) irradiation at 312 nm for 10 min at room temperature as previously described (Temiakov et al., 2002).

2.2.5.4 Mapping of the cross-linking sites in mtRNAP

Mapping of the regions in mtRNAP that interact with RNA or DNA with cyanogen bromide (CNBr), 2-nitro-5-thiocyano-benzoic acid (NTCB), and hydroxylamine (NH$_2$OH) was performed as described previously (Sologub et al., 2009). Products of the cleavage reactions were resolved using a 4-12% Bis-Tris NuPAGE gel (Invitrogen) and visualized by PhosphorImagerTM (GE Health). Bands were identified by calculating their apparent molecular weights using protein standards (Mark 12, Invitrogen) and matched to the theoretical single-hit cleavage pattern for NTCB or CNBr (Fig. 13).

3 Results and Discussion

3.1 Structure of human mtRNAP elongation complex

Data presented in this chapter have been obtained during this thesis and have been published (see page V).

3.1.1 Structure of mtRNAP elongation complex

We co-crystallized human mtRNAP (residues 151-1230, Δ150 mtRNAP) with a nucleic acid scaffold that contained a 28-mer DNA duplex with a mismatched 'bubble' region and a 14-mer RNA with nine nucleotides that were complementary to the template strand in the bubble (Fig. 6a and chapter 2.2). The reconstituted elongation complex was active in a primer extension assay (Fig. 7). We solved the structure by molecular replacement and refined it to a free R-factor of 21% at 2.65 Å resolution (Table 12).

The structure reveals a new mtRNAP conformation, most of the DNA and RNA, and details of the polymerase-nucleic acid contacts (Figs. 6 and 8). The protein structure includes the previously mobile part of the thumb (residues 736-769), and only lacks two disordered loops, the terminal tip of the intercalating hairpin (residues 595-597), and a loop called specificity loop in T7 RNAP (residues 1086-1106). Compared to the clenched conformation of the free polymerase (Ringel et al., 2011), the active center is widened by rotations of the palm and fingers by 10° and 15°, respectively, and neatly accommodates a 9-base pair DNA-RNA hybrid (Fig. 6c).
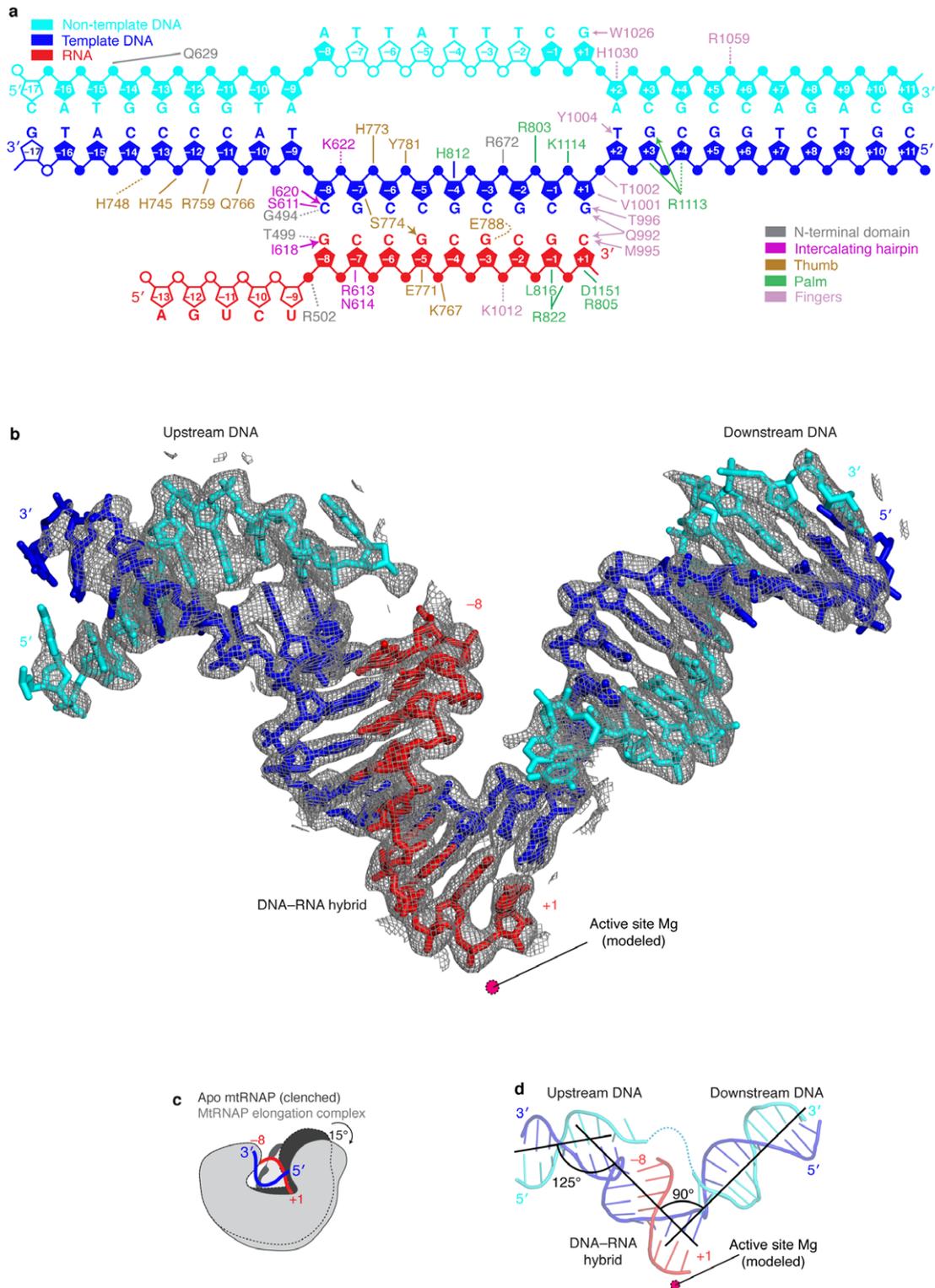
Figure 6 - Nucleic acid structure and mtRNAP interactions observed in the mtRNAP elongation complex crystal structure.

(**a**) Schematic overview of interactions between mtRNAP and nucleic acids. The nucleic acid scaffold contains template DNA (blue), non-template DNA (cyan) and RNA (red). Unfilled elements were not visible in the electron density map. Interactions with mtRNAP residues are indicated as lines (hydrogen bonds, ≤3.6 Å), dashed lines (electrostatic contacts, 3.6-4.2 Å), or arrows (stacking interactions).

(**b**) Refined nucleic acid structure with $2F_o$-F_c electron density omit map contoured at 1.5σ.

(**c**) Polymerase opening from the clenched conformation of free mtRNAP (PDB code 3SPA (Ringel et al., 2011), dark grey) to the elongation complex (light grey). Structures were superimposed based on their NTDs.

(**d**) Angles between duplex axes of upstream DNA, DNA-RNA hybrid, and downstream DNA.
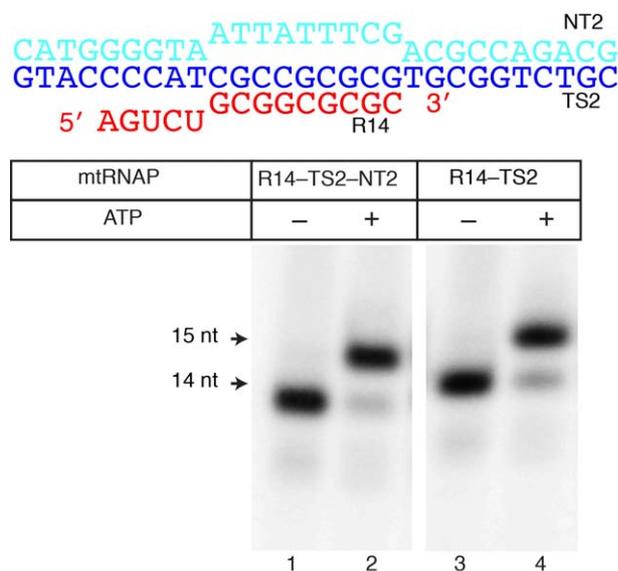
Figure 7 - Activity of mtRNAP elongation complex assembled on nucleic acid scaffolds.

MtRNAP (1 mM) was pre-incubated with the scaffolds indicated (1 mM) for 5 min at room temperature and the ^{32}P-labeled RNA primer extended by addition of 10 mM of adenosine triphosphate (ATP) for 2 min. The products of the reaction were resolved in 20% PAGE containing 6 M urea.

Table 12 - Data collection and refinement statistics (molecular replacement)

	mtRNAP elongation complex
Data collection[1]	
Space group	I23
Cell dimensions	
a=b=c (Å)	225.2
Resolution (Å)	39.8-2.65 (2.72-2.65)[2]
R_{sym} (%)	12 (229)
$I/\sigma I$	18.9 (1.7)
Completeness (%)	100.0 (100.0)
Redundancy	20.7 (20.2)
CC$_{(1/2)}$ (%)[3]	100 (42.5)
Refinement	
Resolution (Å)	39.81-2.65
No. reflections	54985
R_{work}/ R_{free} (%)	17.3/20.8
No. atoms	
Protein	7880
Ligand/ion	1265
Water	244
B-factors (Å^2)	
Protein	94.4
Ligand/ion	138.1
Water	83.5
RMSDs	
Bond lengths (Å)	0.010
Bond angles (º)	1.24

---

[1] Diffraction data were collected at beamline X06SA of the Swiss Light Source, Switzerland and processed with MOSFLM (Leslie, 2006).

[2] Numbers in parenthesis refer to the highest resolution shell.

[3] CC$_{1/2}$ = percentage of correlation between intensities from random half-datasets (Karplus and Diederichs, 2012).
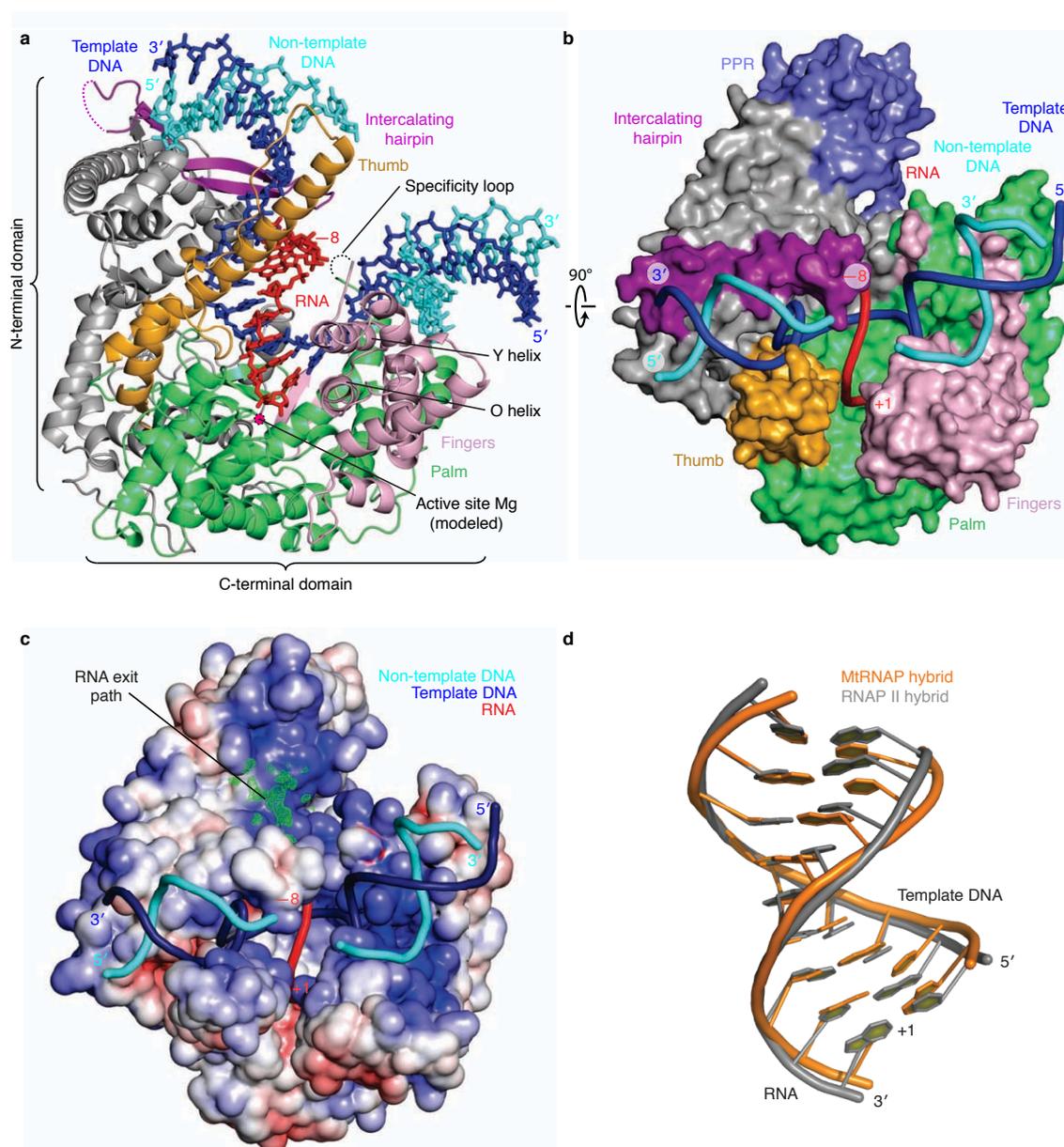
Figure 8 - Structure of mtRNAP elongation complex determined by X-ray crystallography.

(a) Overview with mtRNAP depicted as a ribbon (thumb, orange; palm, green; fingers, pink; intercalating hairpin, purple), and nucleic acids as sticks (color code as in Fig. 6). A Mg^{2+} ion (magenta) was placed according to a T7 RNAP structure (Yin and Steitz, 2004). The PPR domain was omitted for clarity.

(b) View of the structure rotated by 90° around a horizontal axis. The polymerase is depicted as a surface model and includes the PPR domain (slave). Nucleic acids are depicted as ribbons.

(c) Electrostatic surface representation of the mtRNAP elongation complex with template DNA (blue), non-template DNA (cyan) and RNA (red). The Fo-Fc electron density of the mobile 5'-RNA tail is shown as a green mesh (contoured at 2.5 σ).

(d) Superimposition of DNA-RNA hybrids in elongation complexes of mtRNAP (orange) and RNAP II (PDB CODE 1I6H (Gnatt et al., 2001a), grey).

3.1.2 Substrate selection and catalysis

The active site closely resembles that of T7 RNAP and harbors the RNA 3'-end at its catalytic residue D1151 (Arnold et al., 2012b; Steitz, 2009; Temiakov et al., 2004) (Fig. 9a). Comparison with phage RNAP structures that contain the NTP substrate (Basu and Murakami, 2013; Yin and Steitz, 2004) supports a conserved mechanism of substrate binding, selection, and catalysis. The location and relative arrangement of amino acid residues in the active center that bind catalytic metal ions and the NTP substrate are conserved in both enzymes. The trajectory of several side chains differs, but this was likely due to the absence of metal ions and NTP in our structure. In the mtRNAP elongation complex, the 3'-terminal RNA nucleotide occupies the NTP site and is paired with the DNA template base +1 (Fig. 9a). Thus the complex adopts the pre-translocation state (Steitz, 2009; Yin and Steitz, 2004), and this may be why we could not obtain a structure with NTP. Modeling suggested that translocation enables binding of the NTP between residues K853, R987 and K991 on one side and two metal ions coordinated by residues G923, D922 and D1151 on the other side (Fig. 9a). The NTP 2'-OH group may contact residue Y999 (Fig. 9a). This contact likely helps to discriminate NTP from dNTP substrates, as revealed by extensive biochemical (Kostyuk et al., 1995; Sousa and Padilla, 1995) and structural studies (Temiakov et al., 2004).
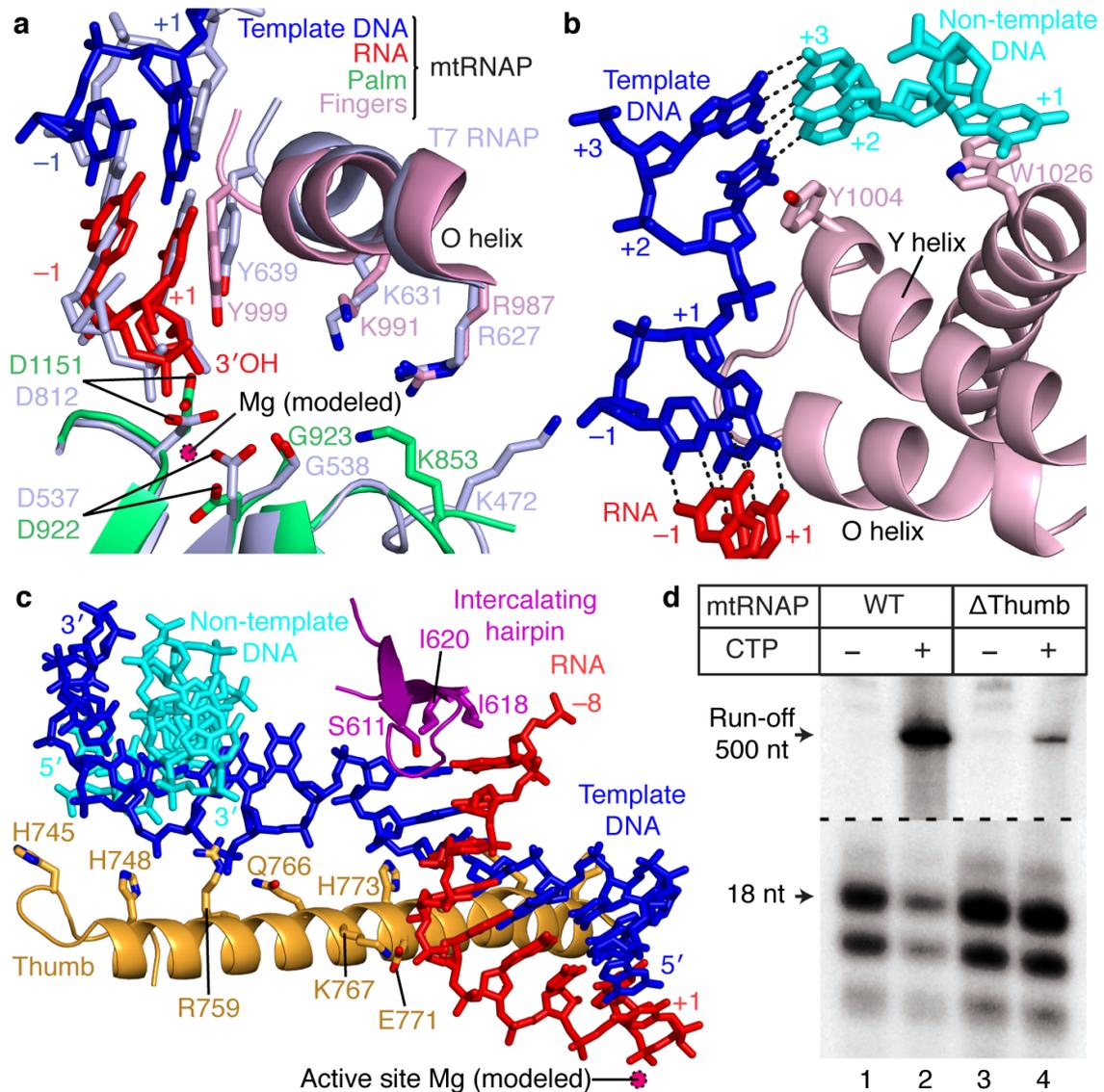
Figure 9 - Active center and nucleic acid strand separation observed in the crystal structure.

(**a**) Conservation of active centers in mtRNAP (color code as in Figs. 6 and 8) and T7 RNAP (PDB code 3E2E (Durniak et al., 2008), light blue). Structures were superimposed based on their palm subdomains and selected residues were depicted as stick models.

(**b**) Downstream DNA strand separation.

(**c**) RNA separation from DNA at the upstream end of the hybrid and thumb-hybrid interactions.

(**d**) Primer extension assays showed that a thumb subdomain plays a key role in elongation complex stability. Elongation complexes of wild-type (WT) (lanes 1 and 2) and Δthumb (lanes 3 and 4) mtRNAP variants were halted 18 nucleotides downstream of the light-strand promoter (LSP) by omitting cytidine triphosphate (CTP) (Sologub et al., 2009).

3.1.3 Polymerase-nucleic acid interactions

The active center is complementary to the hybrid duplex, which adopts A-form (Fig. 8d and Tab. 13), and could not accommodate a B-form duplex that would result from erroneous DNA synthesis. The DNA-RNA hybrid forms many contacts with the polymerase, including contacts to the thumb (Figs. 6a, 8a and 9c). Movement of the thumb was previously detected during different stages of the nucleotide addition cycle, implicating this domain in elongation complex stability, processivity, and translocation in the pol A family of polymerases (Brieba et al., 2001; Mentesana et al., 2000).

Table 13 - Base pair parameters of mtRNAP elongation complex DNA-RNA hybrid region

Register	Bp	Shear(Å)	Stretch(Å)	Stagger(Å)	Buckle(°)	Propeller(°)	Opening(°)
+1	G-C	-0.57	-0.13	-0.28	-13.85	-11.09	4.34
-1	C-G	-0.12	-0.23	0.43	-2.82	-11.09	-2.26
-2	G-C	0.01	-0.22	0.14	-8.82	-9.62	-2.76
-3	G-C	-0.3	-0.13	-0.19	-9.93	-16.22	2.08
-4	C-G	0.46	-0.18	0.02	-0.39	-11.15	0.54
-5	G-C	-0.06	-0.16	-0.02	-1.93	-12.26	-1.6
-6	C-G	0.24	-0.16	0.21	-0.48	-15.32	3
-7	G-C	-0.5	-0.1	-0.28	-21.38	-11.28	2.03
-8	C-G	-0.13	-0.13	0.18	-10.77	0.16	-2.17

Register	Step	Shift(Å)	Slide(Å)	Rise(Å)	Tilt(°)	Roll(°)	Twist(°)
+1/-1	GC/GC	-0.47	-0.48	3.16	-7.74	-0.55	32.53
-1/-2	CG/CG	0.4	-1.53	3.27	4.4	6.94	33.32
-2/-3	GG/CC	0.16	-1.18	3.31	3.38	11.58	32.11
-3/-4	GC/GC	0.47	-1.14	3.08	-1.28	7.37	29.52
-4/-5	CG/CG	-0.08	-1.85	3.3	-0.65	9.86	27.91
-5/-6	GC/GC	0.24	-1.69	3.24	-1.05	4.83	29.64
-6/-7	CG/CG	0.57	-1.16	3.65	9.92	10.14	35.42
-7/-8	GC/GC	-0.15	-0.64	3.16	-2.61	13.59	28.98

To test the functional role of the thumb domain, we carried out *in vitro* transcription assays. Deletion of thumb residues 734-773 in human mtRNAP did not result in any significant processivity defects, but we observed a markedly decreased stability of the elongation complex in salt-dependent primer extension assays (Fig. 10a,b). When we halted an elongation complex formed with the thumb deletion (Δthumb) mutant by withholding the substrate NTP, the polymerase was unable to resume elongation and dissociated during run-off transcription assays (Fig. 9d), suggesting a key role of thumb-hybrid interactions in maintaining complex stability during elongation.
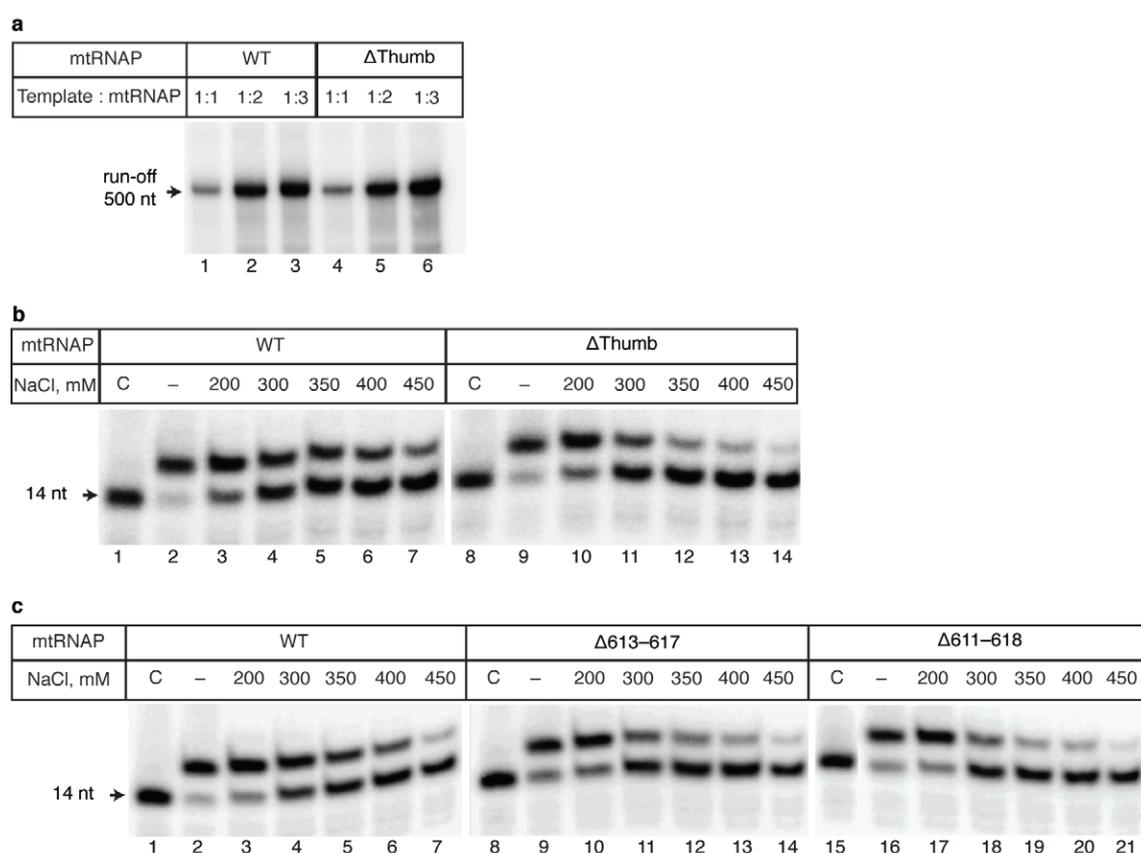
Figure 10 - Effects of mtRNAP variants on elongation complex stability.
(a,b) The thumb deletion mtRNAP mutant (Δthumb) is processive but forms unstable halted elongation complexes. (a) Processivity of the Δthumb mtRNAP. Run-off transcription assay was performed using PCR template containing the LSP promoter (50 nM) and the indicated amount of WT (lanes 1-3) and Δthumb (lanes 4-6) mtRNAPs and the products of the reactions resolved in 20% PAGE containing 6 M urea. (b) ΔThumb mutant forms an unstable halted elongation complex. The elongation complexes were assembled using DT1 scaffold and WT or Δthumb mtRNAP. As a control (C) only polymerase was loaded in lanes 1 and 8.

51

(c) Elongation complexes formed with mtRNAP variants that contain a deletion of the intercalating hairpin are sensitive to salt challenge. Elongation complexes were formed using DT1 scaffold and WT (lanes 1-7) or the intercalating hairpin deletion mutants Δ613-617 (lanes 8-14) and Δ611-618 (lanes 15-21). As a control (C) only polymerase was loaded in lanes 1,8 and 15.

We resolved both downstream and upstream duplexes in our structure. These DNA elements formed B-form duplexes near positively charged surfaces of the polymerase NTD and CTD, respectively (Fig. 8c). The downstream DNA runs perpendicular to the hybrid (Fig. 6d), as observed in elongation complex structures of T7 RNAP (Durniak et al., 2008; Steitz, 2009; Tahirov et al., 2002; Yin and Steitz, 2002, 2004) and the unrelated multisubunit RNAP II (Gnatt et al., 2001b; Kettenberger et al., 2004). Thus a 90° bend between downstream and hybrid duplexes is apparently a general feature of transcribing enzymes. The length and conformation of the hybrid are also very similar and apparently dictated by intrinsic nucleic acid properties (Fig. 8d and Tab. 13). The axes of upstream DNA and the hybrid encloses a 125° angle (Fig. 6d).

3.1.4 DNA strand separation

As the polymerase advances, the strands of downstream DNA must be separated before the active site. The structure showed that DNA strand separation involves the fingers domain (Fig. 9b). The side chain of tryptophan W1026 stacks onto the +1 base of the non-template DNA, directing it away from the template strand (Fig. 9b). The side chain of tyrosine Y1004 in the Y helix stacks onto the +2 DNA template base, stabilizing a 90° twist of the +1 template base and allowing its insertion into the active center (Fig. 9b). This is achieved by a 25° rotation of the Y helix compared to its position in free mtRNAP (Ringel et al., 2011). Whereas residue Y1004 has a structural counterpart in T7 RNAP, residue F644 (Cheetham and Steitz, 1999; Tahirov et al., 2002; Yin and Steitz, 2002, 2004), residue W1026 does not (Fig. 11), suggesting that the mechanisms of strand separation are likely conserved between the two polymerases.
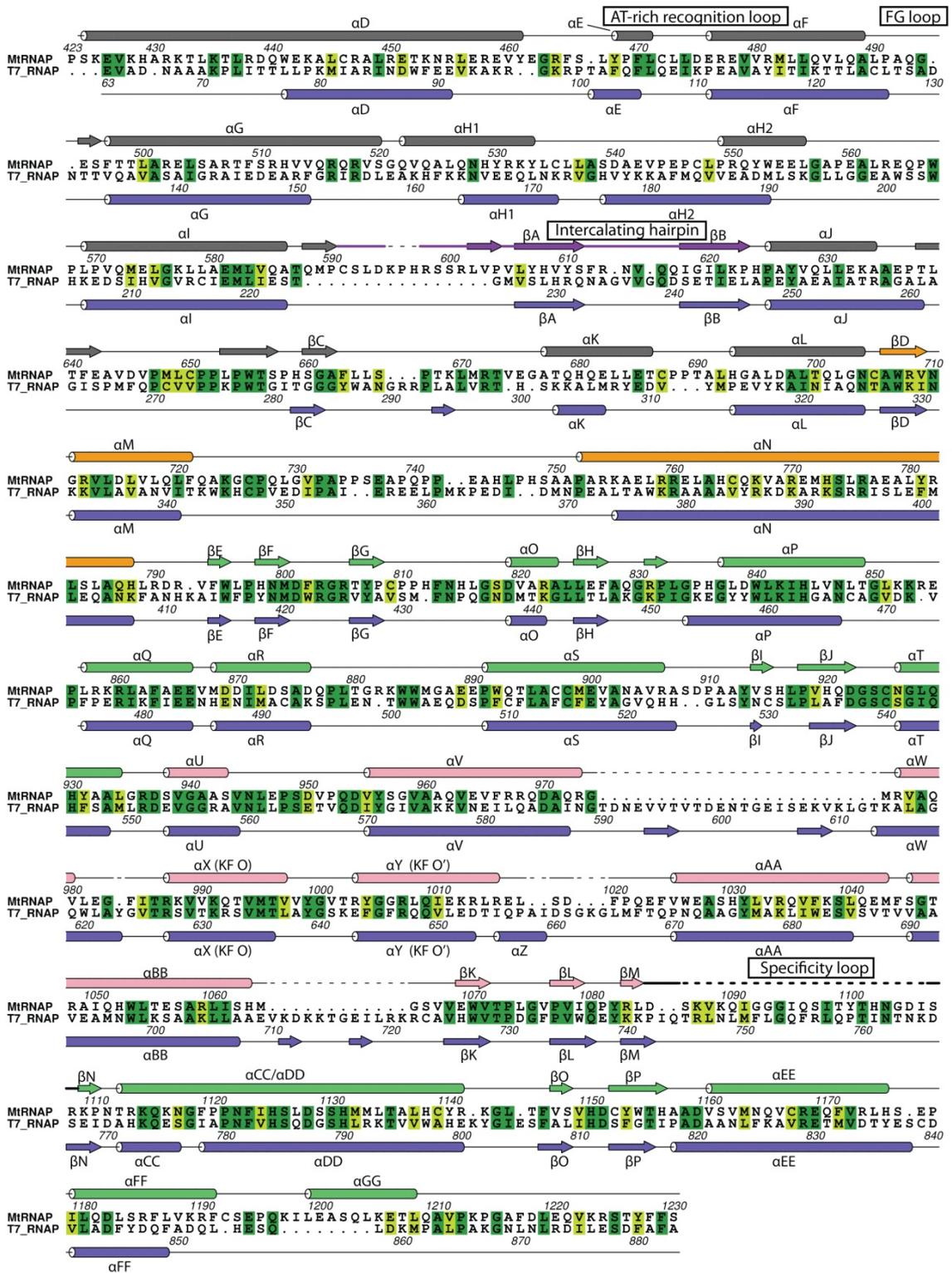
Figure 11 - Structure-based sequence alignment and conservation of human mtRNAP (residues 423-1230) and T7 RNAP (residues 63-883, PDB 1QLN (Cheetham et al., 1999)).

Secondary structure elements are consecutively labeled in alphabetical order (cylinders, α-helices; arrows, β-strands; lines, loops). Since helix X is commonly named helix O based on a corresponding helix in the *E.coli* Klenow (KF) fragment (Beese et al., 1993), we maintain this convention during this work. Identical residues are highlighted in dark green, conservative substitutions are shown light green. Color coding for mtRNAP secondary elements is as in Figs. 6-9.

3.1.5 RNA separation and exit

At the upstream end of the hybrid, RNA is separated from the DNA template by the intercalating hairpin, which protrudes from the NTD (Figs. 8a and 9c). The hairpin stacks with its exposed isoleucine residues I618 and I620 onto RNA and DNA bases, respectively, of the last hybrid base pair at the upstream position -8. Consistent with the role of the intercalating hairpin during elongation, elongation complexes assembled with the intercalating hairpin deletion, RNA extension assays revealed that variants of mtRNAP were considerably less stable than complexes with WT (wild-type) mtRNAP (Fig. 10c). This is in contrast to T7 RNAP (Brieba et al., 2001), where the intercalating hairpin is not important for RNA displacement and transcription bubble stability during T7 RNAP elongation.

RNA exits over a positively charged surface patch, but shows poor electron density that indicates mobility (Fig. 8c). To investigate whether the weak electron density reflects the RNA exit path, protein-RNA cross-linking experiments were applied. By replacing the first RNA base beyond the hybrid by a photo cross-linkable analogue, it was cross-linked to the specificity loop (Figs. 12a,d and 13). Thus the mobile specificity loop lines the RNA exit channel, as in the T7 RNAP elongation complex (Tahirov et al., 2002; Yin and Steitz, 2002). Exiting RNA at position -13 cross-linked to NTD helices I and G and thus the transcript emerges towards the PPR domain (Figs. 12b,d) that contains conserved RNA recognition motifs (Schmitz-Linneweber and Small, 2008).
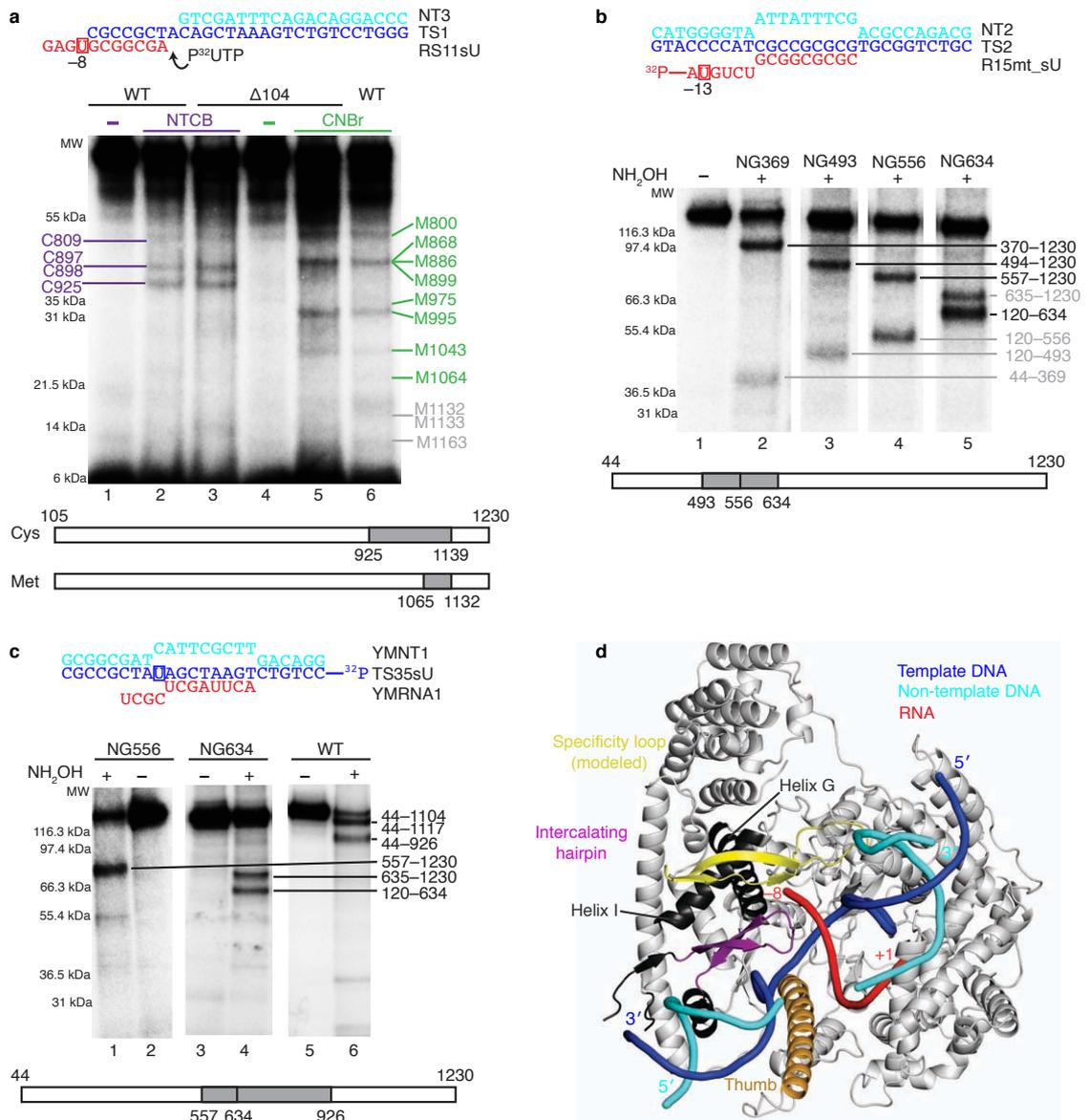
Figure 12 - Analysis of mtRNAP-nucleic acid contacts by cross-linking experiments.

(a) RNA nucleotide -8 cross-links to the specificity loop of mtRNAP. The cross-linked complexes were treated with 2-nitro-5-thiocyano-benzoic acid (NTCB, lanes 2 and 3) or cyanogen bromide (CNBr, lanes 5 and 6). Positions of the cysteine (Cys) and methionine (Met) residues that produced labeled peptides are indicated in purple and green, correspondingly. Grey numbers indicate methionine residues that did not produce labeled peptides and the expected migration of these peptides.

(b) Mapping of the RNA-mtRNAP cross-link at RNA nucleotide -13 with different mtRNAP variants having a single hydroxylamine cleavage site (NG) at a defined position. The cross-links were treated with hydroxylamine (NH_2OH). The major cross-linked peptides are highlighted in black, minor (less than 10%) cross-linking sites in grey.

(c) Mapping of the template strand DNA-mtRNAP cross-link at nucleotide -8. The cross-links were treated with NH_2OH as described above.

(d) Location of the cross-linked regions in mtRNAP elongation complex. The T7 RNAP specificity loop was built into the mRNAP structure by homology modeling. The structural elements that belong to the identified cross-linked regions and lie within 3-5 Å from the photo cross-linking probe include the modeled specificity loop (yellow, residues 1080-1108), part of the thumb (orange, residues 752-791) and part of the intercalating hairpin (purple, residues 605-623). Cross-linked regions that are not part of a defined structural element are shown in dark grey (e.g. helix G residues 587-571 and helix I residues 570-586).
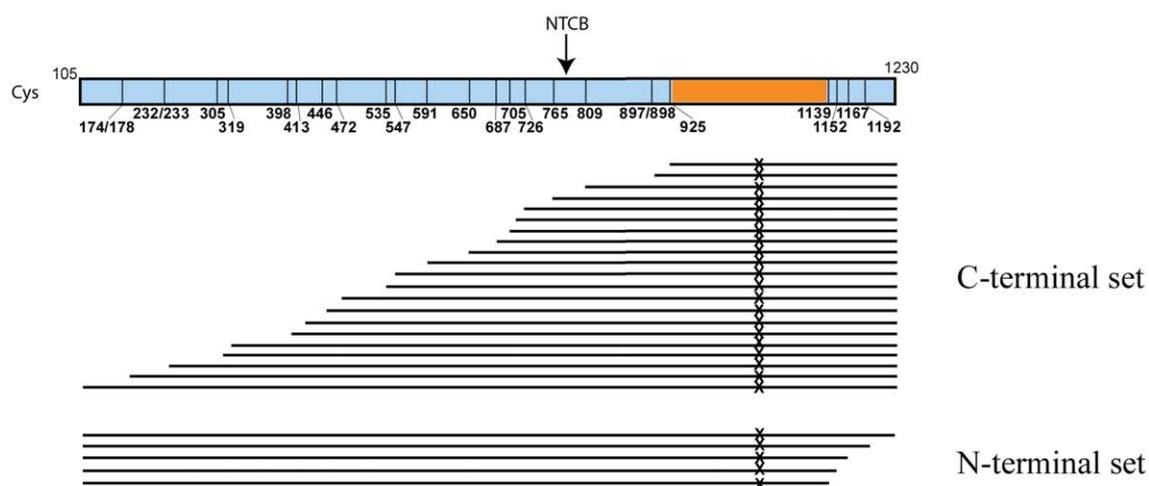
Figure 13 - Analysis of cross-linking mapping data.

Cross-linking mapping with NTCB and CNBr (Fig. 12a) was performed using the so-called "single-hit" conditions (Grachev et al., 1989; Korzheva et al., 2000) i.e. when every mtRNAP molecule is cleaved only once, on average. Thus, the single-hit conditions generate characteristic patterns of the N-terminal and C-terminal cleavage products. As an example, the theoretical pattern of mtRNAP cleavage by NTCB consistent with the position of the cross-link at the C-terminus is presented above. The size of the labeled fragments is identified by its mass (mobility in SDS PAGE) using SeeBlue protein standard markers (Invitrogen). To distinguish between the C-terminal and the N-terminal location of the cross-link two variants of mtRNAP were used, WT mtRNAP and Δ104 mtRNAP (Fig. 12a). No shift in bands migration was observed in SDS-PAGE (Fig. 12a, lanes 2 and 3) confirming the location of the cross-link site at the C-terminus of mtRNAP. The smallest labeled band visible on the SDS PAGE upon NTCB treatment corresponds to the 925-1230 peptide and thus positions the cross-linking site

between residues C925 and C1139. This interval was narrowed down even further by CNBr cleavage (Fig. 12a, lanes 5 and 6). The smallest band visible on the gel upon CNBr treatment corresponds to the 1064-1230 peptide and positions the cross-linking site between residues M1063 and M1132.

Cross-linking mapping of RNA at base -13 was performed using mtRNAP variants having a single NG at a defined position (Fig. 12b). The cleavage generates only two mtRNAP fragments simplifying identification of the labeled peptides. Thus the cleavage of the cross-link obtained with NG493 mutant results in appearance of a labeled fragment (83.2 kDa) representing the C-terminus of mtRNAP, while cleavage of NG634 mutant results in appearance of the N-terminal fragment (61.5 kDa). Taken together, these data suggest that the cross-linking site is between residues 494 and 634.

Mapping of cross-link at DNA template base at -8 (Fig. 12c) was performed using NH_2OH and WT, NG556 and NG634 mtRNAPs. WT mtRNAP contains four sites for NH_2OH cleavage at positions 710, 926, 1103 and 1117, however the most N-terminal site (710) is cleaved inefficiently and thus the resulting peptides are not visible. NH_2OH cleavage of the mtRNAP-DNA cross-link results in two major products corresponding to the intervals 44-926 and 44-1103 or 44-1117 (Fig. 12c, lane 6). Since no band was observed that corresponds to the interval 926-1103 or 926-1117 (about 28 kDa for peptide with the cross-linked DNA) we conclude that the cross-link is to the 44-926 interval of mtRNAP. Cleavage of the NG556 mutant results in appearance of the labeled C-terminal fragment (around 82 kDa), while cleavage of NG634 mutant generates two labeled fragments representing both the C- and the N-terminal parts of mtRNAP (Fig. 12c, lanes 1-4). Taken together these data suggest that the cross-link site of -8 base of DNA includes two adjacent mtRNAP regions: 557-634 and 635-926.

3.1.6 Lack of NTD refolding upon elongation

To initiate transcription, T7 RNAP binds promoter DNA with its NTD (Nayak et al., 2009; Steitz, 2009). The NTD then refolds during the transition from an initiation complex (Cheetham and Steitz, 1999) to an elongation complex (Yin and Steitz, 2002) via an intermediary state (Durniak et al., 2008). In contrast, the NTD of mtRNAP does not refold during the initiation-elongation transition (Fig. 14). The NTD fold observed in our mtRNAP elongation complex structure differs from that in T7 RNAP elongation complexes, but resembles that in the T7 initiation-elongation intermediate, and is partially related to that in the T7 initiation complex (Figs. 14a-c and Tab. 14).
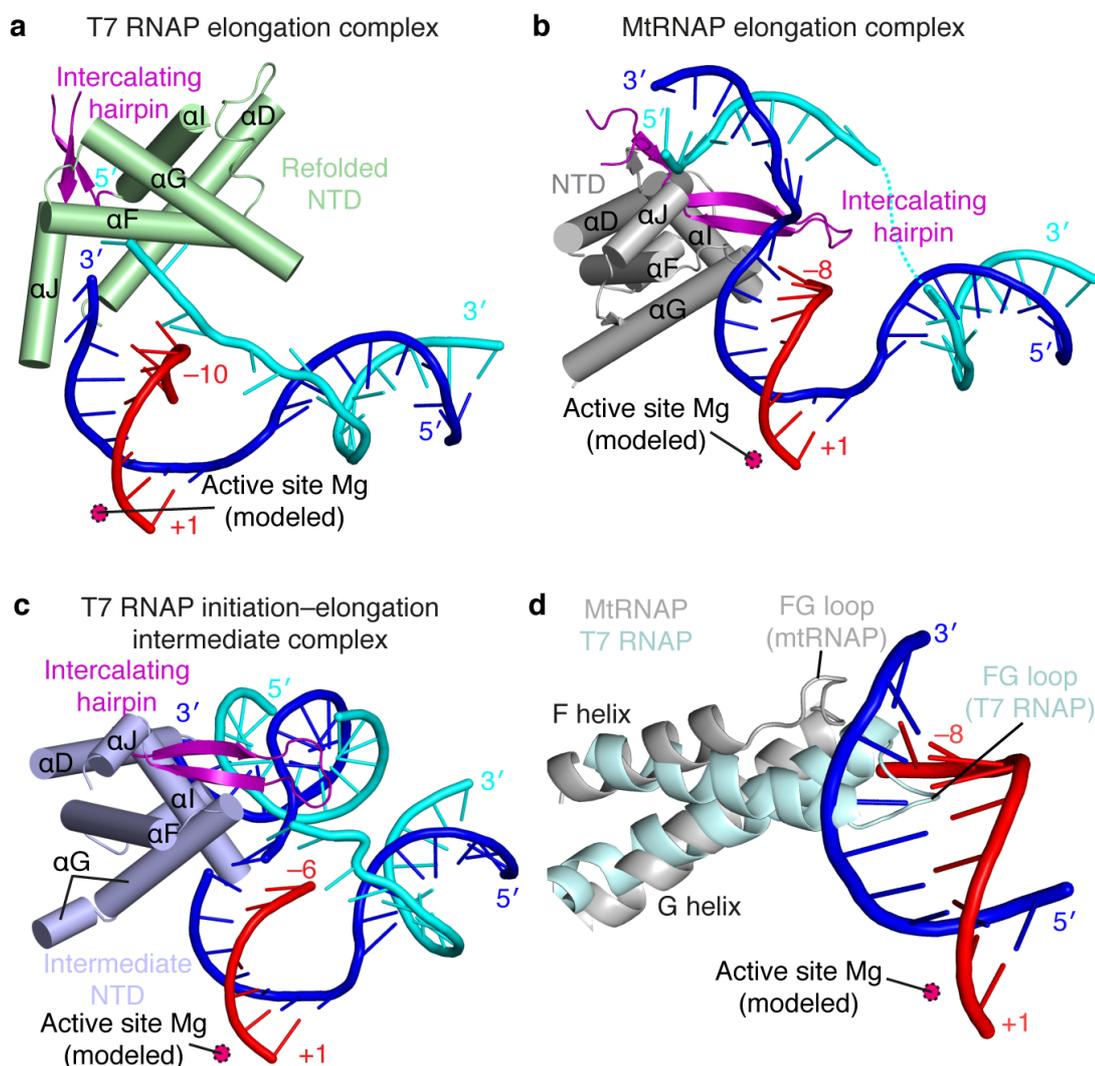
Figure 14 - Lack of NTD refolding upon mtRNAP elongation observed in the crystal structure.

(a-c) Structures of the NTD of T7 RNAP and mtRNAP. The NTD of T7 RNAP (a) is refolded in the elongation complex (PDB code 1MSW (Yin and Steitz, 2002), whereas the NTD of mtRNAP (b) is not, and resembles the NTD in the T7 intermediate (PDB code 3E2E (Durniak et al., 2008)) (c). Helices are depicted as cylinders and nucleic acids as ribbons with sticks for protruding bases.

(d) The FG loop of T7 RNAP (PDB code 1QLN (Cheetham and Steitz, 1999), pale cyan) protrudes into the hybrid-binding site but is shorter and positioned differently in mtRNAP (silver).

Our crystallized mtRNAP complex represents an elongation complex rather than an intermediate of the initiation-elongation transition because it shows full RNA-extension activity and comprises a mature 9-base pair DNA-RNA hybrid with a free 5'-RNA extension exiting the polymerase (Figs. 7 and 8c). Consistent with a lack of NTD refolding, the DNA template position -8 in the elongation complex could be cross-linked to a region that encompasses the intercalating hairpin (Ringel et al., 2011; Velazquez et al., 2012) (Figs. 12c,d). In striking contrast, NTD refolding in T7 RNAP moves the intercalating hairpin more than 40 Å away from the hybrid upon elongation (Figs. 14a-c).

Table 14 - Structural comparison of mtRNAP elongation complex NTD with different T7 NTD complexes by Cα root-mean-square deviation (RMSD) values.
Structures were aligned based on the sequence alignment (Fig. 12) and the RMSD calculated over all matching Cα pairs.

mtRNAP elongation complex NTD (residues 426-638) superimposed with:	RMSD (Å)
T7 initiation structure (PDB code 1QLN (Cheetham et al., 1999), residues 72-261)	6.4
T7 initiation-elongation intermediate (PDB code 3E2E (Durniak et al., 2008), residues 73-254)	4.7
T7 pre-translocated product structure (PDB code 1S77 (Yin and Steitz, 2004), residues 63-261)	8.3
T7 post-translocated structure (PDB code 1MSW (Yin and Steitz, 2002) residues 63-261)	8.0

3.1.7 Discussion

Transcription of the mitochondrial genome is essential for all eukaryotic cells, yet its mechanisms remain poorly understood. Thus far only the structure of free mtRNAP was reported, whereas structures of functional mtRNAP complexes were lacking. Here we present the structure of a functional mtRNAP complex, that of the human mtRNAP elongation complex. The structure showed that nucleic acid binding leads to an opening of the polymerase active center cleft, and an ordering of the thumb domain

and most of the intercalating hairpin. The structure revealed the arrangement of downstream and upstream DNA on the polymerase surface, and the DNA-RNA hybrid in the active center, as well as detailed nucleic acid-polymerase contacts.

The structure of the mtRNAP elongation complex also enabled a detailed comparison with the distantly related RNAP from bacteriophage T7. This indicated conserved mechanisms for substrate selection and binding, and for catalytic nucleotide incorporation into growing RNA. Downstream DNA strand separation is achieved by the fingers domain and at least partially resembles strand separation by T7 RNAP. Taken together, the polymerase CTD and mechanisms that rely on this domain were largely conserved during evolution of singlesubunit RNAPs (Gray, 2012).

Our results also revealed striking mechanistic differences between T7 RNAP and mtRNAP. In particular, the NTD does not refold during the transition from transcription initiation to elongation. In T7 RNAP (Cheetham and Steitz, 1999; Durniak et al., 2008), NTD refolding is triggered by a clash of the growing DNA-RNA hybrid with residues 127-133 in the FG loop. In contrast, this loop is two residues shorter in mtRNAP and positioned such that it allows for hybrid growth without NTD refolding (Figs. 11 and 14).

We suggest that during evolution of mtRNAP from an early bacteriophage-like RNAP the catalytic CTD and elongation mechanism remained highly conserved, whereas the NTD lost its capacity to adopt an initiation-specific fold with functions in promoter binding and opening, as initiation factors became available to take over these functions. A loss of NTD refolding and its intrinsic initiation functions in mtRNAP apparently went along with the evolution of initiation factors TFAM and TFB2M (Arnold et al., 2012b; Deshpande and Patel, 2012; Gaspari et al., 2004b; Litonin et al., 2010), which are responsible for promoter binding (Campbell et al., 2012; Ngo et al., 2011; Rubio-Cosials et al., 2011) and opening (Falkenberg et al., 2002; Sologub et al., 2009), respectively.

As a consequence, mtRNAP escapes the promoter by dissociating initiation factors (Mangus et al., 1994), whereas T7 RNAP release from the promoter involves NTD refolding, which destroys the promoter-binding site within the NTD and repositions the intercalating hairpin far away from the nucleic acids. In contrast, the intercalating hairpin in mtRNAP separates the RNA transcript from the DNA template at the upstream end of the hybrid during elongation. Thus, the mechanism of transcription initiation by mtRNAP is unique. In the future, the initiation mechanism of mtRNAP should be studied structurally and functionally.

3.2   Scaffold design and crystallization

Data presented in this chapter have been obtained during this thesis, but have not been published. The following section represents experimental (pre-)work that essentially contributed to solve the structure of human mtRNAP elongation complex but was not included in the published article.

As shown before the mtRNAP variant lacking residues 1-150 showed improved solubility compared to the full-length protein and was used for all experiments in this work (Ringel et al., 2011). In order to imitate the *in vivo* elongation phase as closely as possible, binding studies of Δ150mtRNAP with a variety of nucleic acid scaffolds were performed. Elongation complex formation was detected as a shift in the elution volume of mtRNAP and the ratio of absorption at 260 nm vs 280 nm during size exclusion chromatography (Fig. 15a). Consistent with the published results, mtRNAP was not able to bind linear or pre-melted DNA templates in the absence of TFB2M (Tab. 15) (Sologub et al., 2009). The presence of RNA, ideally a 9-base pair DNA-RNA hybrid and a free 5'-tail favored nucleic acid binding by mtRNAP, whereas the mismatched non-template DNA in the bubble region did not affect complex formation. Although mtRNAP relies on the presence of catalytic Mg^{2+} ions, the addition of $MgCl_2$ did not influence the interaction surface between mtRNAP and nucleic acids in this experimental set-up (data not shown). As shown in Fig. 15b and Tab. 15, decreased salt concentrations assisted in the assembly of the elongation complex probably due to

decreased ion-nucleic acid interactions. Therefore, the binding studies with different scaffolds (Tab. 15) were exclusively performed under low salt conditions (100 mM NaCl). For subsequent crystallization trials elongation complexes that also form under high salt condition (300 mM) were favored.
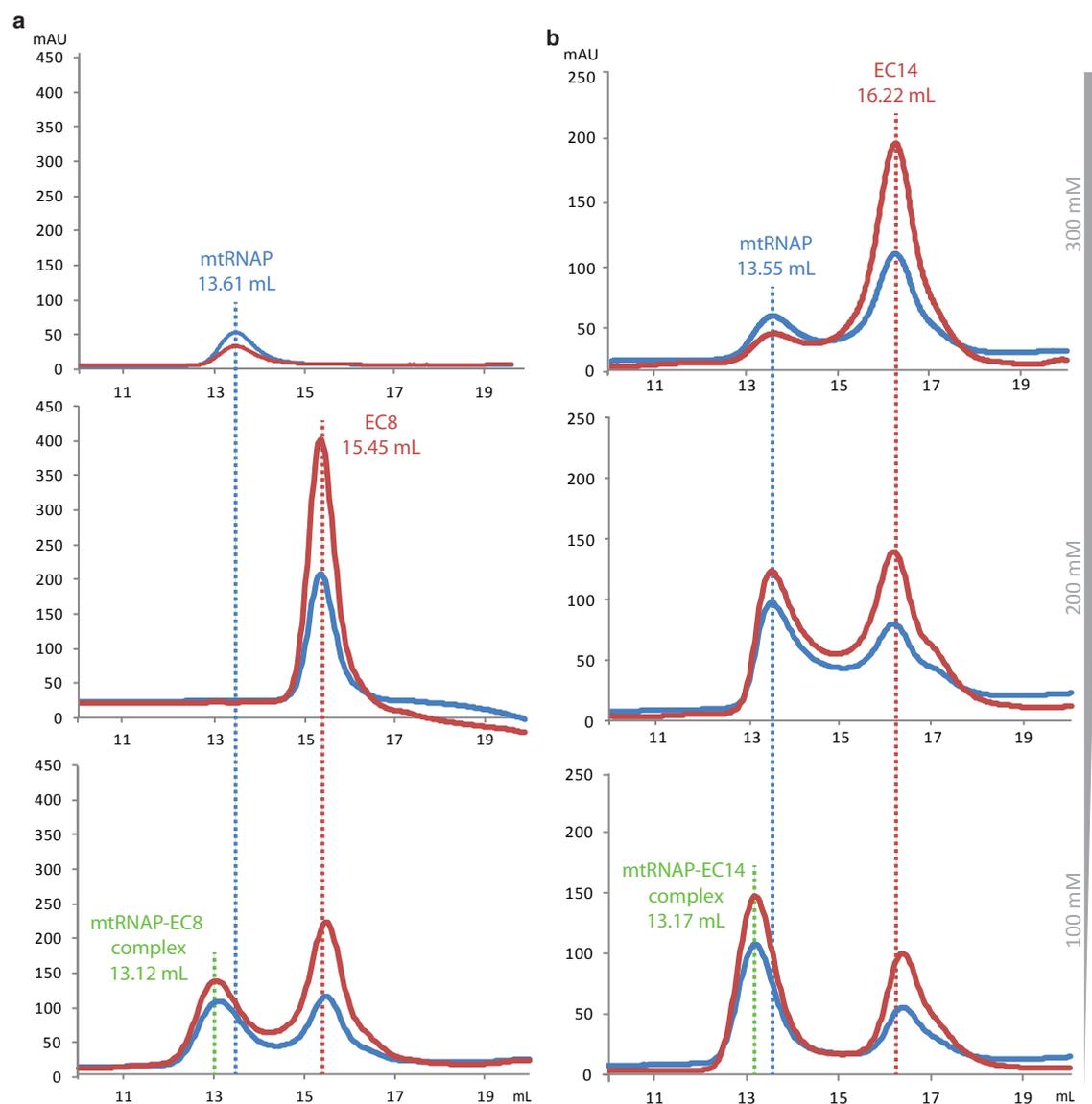
Figure 15 - Binding studies for mtRNAP elongation complex formation.
(a) Size exclusion chromatograms of a Superdex 200 10/300 GL (GE Healthcare) loaded with 900 pmol Δ150mtRNA (top), 1.1 nmol nucleic acid scaffold EC8 (middle) or with both Δ150mtRNAP and EC8 (bottom). Gelfiltration was performed in buffer G. Absorption at 280 nm

is shown in blue, absorption at 260 nm is shown in red. Retention volumes of Δ150mtRNAP, EC8 and the formed elongation complex are indicated with dashed lines.

(b) Size exclusion chromatograms of a Superdex 200 10/300 GL (GE Healthcare) loaded with each 900 pmol Δ150mtRNAP and 1.1 nmol nucleic acid scaffold EC14. Gelfiltration was performed in buffer I (top), buffer H (middle) and buffer G (bottom). Color code as in (a).

Table 15 - Summary of binding studies for the Δ150mtRNAP elongation complex

Binding studies of Δ150mtRNAP with different nucleic acid scaffolds were performed via gelfiltration on a Superdex 200 10/300 GL size exclusion column in the presence of 100 mM or 300 mM NaCl. The degree of complex formation was quantitatively measured based on the ratio of absorption at 260 nm vs 280 nm. Interactions between Δ150mtRNAP and nucleic acid scaffolds are indicated as ++ (strong complex formation, 260 nm / 280 nm ≥ 1.25), + (weak complex formation, 260 nm / 280 nm < 1.25) or — (no complex formation). Protein-scaffold combinations that have not been tested are indicated with n.d. (not determined). The nucleic acid scaffolds contain template DNA (blue), non-template DNA (cyan) and RNA (red). For a schematic overview of all scaffolds used in this work see also Fig. 5.

Scaffold	mM NaCl 300	mM NaCl 100
CC2	n.d.	—
OC1	—	n.d.
OC2	—	n.d.
OC3	—	n.d.
OC4	—	n.d.
EC1	—	n.d.
EC2	—	n.d.
EC3	—	n.d.

EC4
5′ GGGGTTGTA GCTTATGTCGA AGTAGGGAG 3′ | ++ n.d.
3′ CCCCAACATAACTACTCTAATCATCCCTC 5′ |
EC4 5′ UUUUUAG UUGAUGAGAU 3′ |

EC5
3′ CGGCGCGGTACGTAAGGG 5′ — n.d.
5′ UUUUGCCGCGCCA 3′

EC6
5′ GGGGTA GCTTATGTCGA AGTAGGGAG 3′ ++ n.d.
3′ CCCCATAACTACTCTAATCATCCCTC 5′
5′ UUAG UUGAUGAGAU 3′

EC7
5′ GGGGTA GCTTATGTCGA AGTGTG 3′ — n.d.
3′ CCCCATAACTACTCTAATCACAC 5′
5′ UUAG UUGAUGAGAU 3′

EC8
5′ CATGGGGTA ATTATTTCG ACTGACGCAG 3′ ++ n.d.
3′ GTACCCCATGGCCGGGCGTGACTGCGTC 5′
5′ CUGC CCGGCCCGC 3′

EC9
5′ CATGGGGTA ATTATTTCG ACTGACGCAG 3′ — n.d.
3′ GTACCCCATGGCCGGGCGTGACTGCGTC 5′
EC9 5′ CCGGCCCGC 3′

EC10
5′ CATGGGGTA ATTATTTCG ACTGACGCAG 3 — n.d.
3′ GTACCCCATGGCCGGGCGTGACTGCGTC 5′
5′ GCCCGC 3′

EC11
5′ GGGGTA ATTATTTCG ACTGACGCAG 3′ ++ n.d.
3′ CCCCATGGCCGGGCGTGACTGCGTC 5′
5′ CUGC CCGGCCCGC 3′

EC12
5′ GGGGTA ATTATTTCG ACTGACGCAG 3 — ++
3′ CCCCATGGCCGGGCGTGACTGCGTC 5
5′ CCGGCCCGC 3′

EC13
5′ GGGGTA ATTATTTCG ACTGACGCAG 3′ — n.d.
3′ CCCCATGGCCGGGCGTGACTGCGTC 5′
5′ GCCCGC 3′

EC14
5′ ATTATTTCG ACTGACGCAG 3′ — ++
3′ GGCCGGGCGTGACTGCGTC 5′
5′ CCGGCCCGC 3′

EC15
5′ ATTATTTCG ACTGACGCAG 3′ — +
3′ GGCCGGGCGTGACTGCGTC 5′
5′ GCCCGC 3′

EC16
5′ ACTGACGCAG 3′ — +
3′ GGCCGGGCGTGACTGCGTC 5′
5′ CCGGCCCGC 3′

EC17
5′ ACTGACGCAG 3′ — n.d.
3′ GGCCGGGCGTGACTGCGTC 5′
5′ GCCCGC 3′

EC18
5′ GGGGTA ATTATTTCG ACTGACGC 3′ — ++
3′ CCCCATGGCCGGGGCGTGACTGCG 5′
5′ CUGC CCGGCCCGC 3′

EC19
5′ GGGGTA ATTATTTCG ACTGACGC 3′ — ++
3′ CCCCATGGCCGGGGCGTGACTGCG 5′
5′ CCGGCCCGC 3′

64

EC20	—	n.d.
EC21	—	++
EC22	—	n.d.
EC23	—	+
EC24	—	—
EC25	—	++
EC26	—	+
EC27	—	—
EC28	—	—
EC29	—	n.d.
EC30	—	—
EC31	—	n.d.
EC32	n.d.	++
EC33	n.d.	+
EC34	n.d.	+
EC35	n.d.	+
EC36	n.d.	+

Figure 16 - Human mtRNAP elongation complex crystallization.

(a-c) Crystallization trials with *in situ* proteolysis. (a) Initial Δ150mtRNAP-EC8 crystals grew in a 96-well plate at the crystallization facility of the MPI at 8°C in a reservoir solution containing 8% PEG 4000, 150 mM sodium acetate, 80 mM trisodium citrate (pH 5.5), 10% glycerol and 120 mM DTT. Left: drop, right: close-up view of crystal. (b) Plate-like Δ150mtRNAP-DT1 crystals grown in a 24-well plate at 20°C in reservoir solution 8% PEG 4000, 200 mM sodium acetate, 100 mM trisodium citrate (pH 5.5), 10% glycerol and 120 mM DTT after 1-3 days. Left: drop, right: close-up view of crystal. (c) Truncated rhombic dodecahedron crystals (Δ150mtRNAP-DT1) grown from plate-like crystals in a 24-well plate at 20°C in the same reservoir solution as in (b) after 4-6 days (0.2 × 0.2 × 0.2 mm). Left: drop, right: close-up view of crystal.

(d) SDS-PAGE of limited proteolysis of Δ150mtRNAP (lane 3) and Δ150mtRNAP-DT1 complex (lane 4). Untreated Δ150mtRNAP is shown as a control (lane 2).

The optimization of initial co-crystals of human mtRNAP with the nucleic acid scaffold EC8 (Fig. 16a) was continued with the sequence optimized scaffold DT1 (Fig. 5). This resulted in a plate-like crystal morphology (Fig. 16b) that appeared after 1-3 days and finally transformed into truncated rhombic dodecahedron crystals (Fig. 16c) with a maximum size of approximately 0.2 × 0.2 × 0.2 mm within 4-6 days. The co-crystallized nucleic acid scaffold DT1 contained a 28-mer DNA duplex with a

mismatched bubble region and a 14-mer RNA with nine nucleotides that were complementary to the DNA template strand in the bubble (Figs. 5 and 6).

The final model was iteratively built and refined using autoBUSTER (Global Phasing Limited) and revealed R-factors of R_{work}=17.3% and R_{free}=20.8% (Tab. 12). Evaluation of protein and nucleic acid geometry of the crystal using MolProbility (Davis et al., 2004) identified 0.71% of the residues as Ramachandran outliers, 96.34% as Ramachandran favored residues and 4.67% as poor rotamers. Regions with weak or missing electron density caused 1.72% bad bonds and 0.52% bad angles.

3.3   Towards a human mtRNAP elongation substrate complex

Data presented in this chapter have been obtained during this thesis, but have not been published.

The highly conserved residues in the active center cleft of mtRNAP and T7 RNAP indicate a conserved catalytic mechanism of nucleotide addition (chapter 3.1.2). Nevertheless, the molecular mechanisms of the mitochondrial transcription cycle need to be studied further. The mitochondrial mtRNAP elongation complex obtained in this work represents an ideal starting.

Initially, various soaking and co-crystallization strategies of the human mtRNAP elongation complex (Δ150mtRNAP-DT1) with ATP, its non-hydrolysable analog α,β-methyleneadenosine 5′-triphosphate (AMPCPP) or 3'-deoxyadenosine-5'-triphosphate (3'dATP) and PP_i were tested. Since production of highly diffracting crystals has always been difficult (see also 3.1.7), obtaining a highly diffracting crystal of the mtRNAP elongation complex with an ATP or AMPCPP molecule bound in the active center was initially not successful.

However, through optimization of the DNA-RNA scaffold and the crystallization conditions, one highly diffracting crystal could eventually be obtained. Upon shortening the 3-'RNA end by one nucleotide (DT4, Fig. 5) truncated rhombic dodecahedron co-

crystals of Δ150mtRNAP-DT4-3'dATP-PPi grew in a reservoir solution containing 3.7% PEG 4000, 180 mM sodium acetate, 40 mM trisodium citrate (pH 5.5), 10% glycerol and 120 mM DTT within 5 days. Data processing and structure determination was performed as described in chapter 3.1 and methods section. As a search model for molecular replacement, the human mtRNAP elongation complex (PDB code 4BOC) missing all nucleic acid moiety was used (Tab. 16).
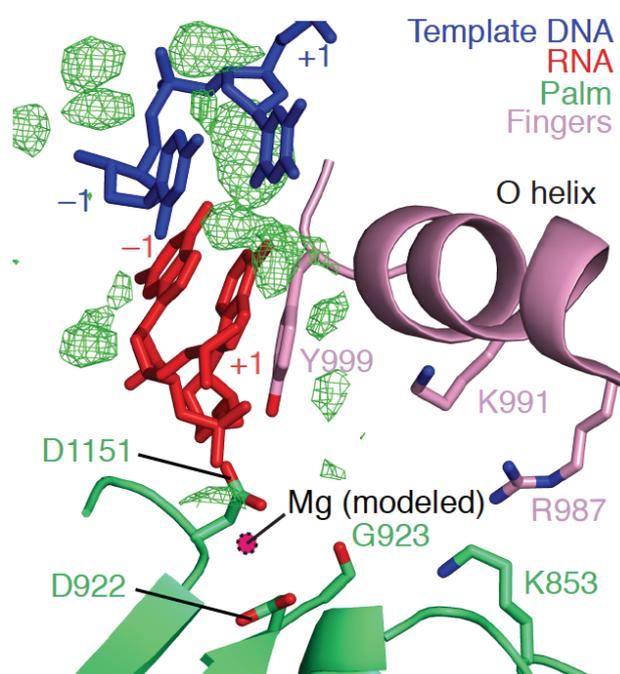
Figure 17 - Incorporated 3'dATP into the human mtRNAP elongation complex (Δ150mtRNAP-DT4).
Refined nucleic acid structure with F_o-F_c electron density omit map (green) contoured at 3 σ with a 3'dATP incorporated at the 3'end of the nascent RNA chain. (Color code as in Figs. 6 and 8.) Selected residues were depicted as stick models.

Through the described adaptions in the experimental set-up the crystal quality of the previously obtained mtRNAP elongation complex crystals was achieved (chapter 3.1). The crystal structure is similar to the human mtRNAP elongation complex except for an incorporated 3'dATP at the 3'-end of the nascent RNA chain (Fig. 17). In respect of mtRNAP elongation complex stability, the hybrid length seems to be a critical parameter that can overcome control mechanisms for NTP and dNTP discrimination *in*

vitro (Kostyuk et al., 1995; Sousa and Padilla, 1995). Although PP_i was present in the crystallization set-up it could not be detected in the crystal structure. Increasing PP_i concentration interfered with crystal growth. This indicates that the current crystallization condition is probably not able to complex PP_i. Thus an atomic structure of the mtRNAP elongation complex with both 3'dATP and PP_i bound in the active center requires further optimization of the crystallization parameters, such as the DNA-RNA scaffold sequence or the composition of the crystallization solution. Obtaining another crystal packing could allow proper substrate coordination in the active center of the mtRNAP elongation complex.

Table 16 - Data collection and refinement statistics (molecular replacement)

	3'dATP mtRNAP elongation complex
Data collection[4]	
Space group	I23
Cell dimensions	
$a=b=c$ (Å)	226.9
Resolution (Å)	48.4-3.15 (3.23-3.15)[5]
R_{sym} (%)	10 (157)
$I/\sigma I$	14.85 (1.9)
Completeness (%)	99.7 (100.0)
Redundancy	10.3 (10.7)
$CC_{(1/2)}$ (%)[6]	99.6 (68.4)
Refinement	
Resolution (Å)	48.4-3.15
No. reflections	33613
R_{work}/R_{free} (%)	18.3/23.8
No. atoms	
Protein	7880

[4] Diffraction data were collected at beamline X06SA of the Swiss Light Source, Switzerland and processed with XDS (Kabsch, 2010).

[5] Numbers in parenthesis refer to the highest resolution shell.

[6] $CC_{1/2}$ = percentage of correlation between intensities from random half-datasets (Karplus and Diederichs, 2012).

Ligand/ion	1177
Water	---
B-factors (Å^2)	
Protein	97.1
Ligand/ion	129.9
Water	---
RMSDs	
Bond lengths (Å)	0.01
Bond angles (º)	1.41

4 Conclusion and Outlook

The singlesubunit mtRNAP is the key player of transcription of the mitochondrial genome. This study applied a structure-function correlation, combining X-ray crystallography, transcription assays and cross-linking experiments to further characterize the mitochondrial transcription cycle. Comparisons of the mitochondrial system with the T7 system, helped to determine the degree of evolutionary conservation between certain protein domains. The results are an important step towards the understanding of the mitochondrial transcription cycle on a molecular level.

4.1   Functional studies of mtRNAP-specific mechanisms

This work highlights the lack of NTD refolding in mtRNAP as a significant difference to T7 RNAP. In order to delineate evolutionary adaptions between both RNAPs, new structure-based mtRNAP mutations need to be analyzed via biochemical assays.

The mobile fingers domain is a characteristic feature of pol A family polymerases and oscillations of O and Y helices are directly involved in the nucleotide addition cycle and the translocation of the nascent RNA (Doublie and Ellenberger, 1998; Steitz, 2009). Intriguingly, while the catalytic cores of mtRNAP and T7 RNAP are highly conserved the Y helices share a surprisingly low sequence similarity (Fig. 11). In mtRNAP, the Y helix is one turn shorter and has the three positively charged residues K1012, R1013 and R1015, whereas T7-like phage RNAPs feature negatively charged (E662 and D663) and aliphatic (I665) residues in the corresponding positions of the Y helix (Fig. 11). Mutation of these residues in combination with transcription assays could help to further investigate why a structurally conserved translocation element, such as the Y helix, displays significant sequence differences. In respect of its stabilizing contribution to the pre-translocated conformation, the Y helix may therefore have an intrinsic effect on the elongation rates of mtRNAP and T7 RNAP.

Although most domain functions of mtRNAP are described, the PPR domain still remains mostly uncharacterized. Previous investigations suggested a function in

binding promoter DNA or nascent RNA. This study revealed that the PPR domain does not interact with the nascent RNA chain, since the growing RNA chain exits the polymerase elsewhere (Fig. 8c). However, it needs to be investigated whether the PPR represents the positively charged trajectory for binding the promoter DNA during initiation. This question can be addressed by cross-linking experiments from different positions of the promoter DNA sequence to identify potential interacting regions in the PPR domain. Alternatively, mtRNAP surface mutations can be designed based on mtRNAP structures to weaken or reverse the positive charge along the PPR domain. Changes in the promoter binding ability of the surface mutants can be analyzed by transcription assays. Additionaly, obtaining the crystal structure of the mitochondrial IC will also help to elucidate this circumstance (see chapter 4.4).

Taken together, functional studies, based on the structure of the mtRNAP elongation complex can contribute to a better understanding of the complete cycle of mitochondrial transcription. Comparisons with the T7 system will help to integrate mtRNAP into the evolutionary context.

4.2 Towards crystallization of full length mtRNAP

Even though the previously unstructured thumb domain of mtRNAP could be solved by co-crystallization with nucleic acids, the structure of a major part of the NED (residues 1-217), the terminal tip of the intercalating hairpin (residues 595-597) and the specificity loop (residues 1086-1106) could not be determinded in the mtRNAP elongation complex structure. This could be due to the proteolytic digestion with ArgC as well as to a high flexibility of the respective regions. Since the reproduction of highly diffracting mtRNAP crystals has always been difficult, it must be a future concern to eliminate all experimental uncertainties. In order to obtain the full length mtRNAP crystal structure, research should concentrate on replacing the proteolytic treatment by the addition of regulatory cofactor proteins such as TFAM, TFB2M or TEFM in order to stabilize flexible domains through protein-protein interactions. Alternatively, sequence- and digestion-based structure predictions can be used to design new mtRNAP mutants that contain shortened flexible linkers between functional domains. The increased proximity

of these functional domains could affect crystal packing and reproducibility and could therefore help to solve the structure of so far disordered regions in the protein.

4.3 Extension of structural studies of the mtRNAP elongation complex

The availability of highly diffracting mtRNAP elongation complex crystals opens the door for various experimental set-ups towards an elongation complex structure containing a substrate molecule. As previously described, obtaining or reproducing highly diffracting mtRNAP crystals with different DNA-RNA scaffolds, substrates and additional factors is difficult. This is likely due to the proteolytic *in situ* digestion required for the established crystallization protocol. Trials to optimize critical parameters such as temperature, crystal age, DNA-RNA oligonucleotide quality, cryo solution composition or crystal freezing need to be continued. Additionally, a mtRNAP construct lacking the flexible specificity loop (residues 1086-1106) and parts of the unstructured NED can overcome this hurdle. Once an optimized mtRNAP construct is established, co-crystallization or (time-dependent) soaking experiments comprising NTPs or their non-hydrolysable analogs can be performed (Basu and Murakami, 2013). This could yield atomic resolution structures of different stages of the nucleotide addition cycle, including a substrate pre-insertion complex (Temiakov et al., 2004), a substrate insertion complex (Yin and Steitz, 2002), a pre-translocated complex (Yin and Steitz, 2004) or a post-translocated complex (Yin and Steitz, 2004) (see also Fig. 1). If the pre-translocated conformation of mtRNAP that was obtained in this work is not the appropriate starting point to crystallize a mtRNAP substrate elongation complex, sequence changes of the DNA-RNA scaffold might trigger a post-translocated polymerase conformation which might be more prone for substrate binding in the active center (Hein et al., 2011).

The recently identified elongation factor TEFM seems to play a significant role for mtRNAP processivity during RNA synthesis. A protocol for recombinant TEFM expression and purification, a direct interaction with mtRNAP *in vitro* as well as an enhancing effect on mtRNAP processivitiy *in vitro* have already been published

(Minczuk et al., 2011). Cross-linking and co-crystallization experiments of the mitochondrial mtRNAP elongation complex presented in this work together with recombinant TEFM will allow mapping of the underlying protein-protein and protein-nucleic acid interaction network. This would reveal the molecular basis of the stimulating effect of TEFM on mtRNAP activity observed *in vitro*.

Although it is well known that mitochondrial dysfunctions are the cause of a variety of human diseases, only little is known of how mtRNAP handles DNA damage that is introduced by the oxidative environment in mitochondria. Upon reaching an oxidatively damaged DNA site, mtRNAP pauses to activate either DNA repair mechanisms or translesion synthesis mechanisms (Nakanishi et al., 2013). In order to investigate the molecular changes that are responsible for factor recruitment or nucleotide incorporation, crystallographic approaches should comprise elongation scaffolds with a synthetic 8-oxoguanine - a typical oxidative DNA damage induced by ROS - placed at different positions in either the DNA template or non-template strand (Cline et al., 2010). Similar experiments have already been performed for multisubunit polymerases, such as RNAP II (Damsma and Cramer, 2009).

Besides damaged mtDNA, mtRNAP also has to deal with the presence of altered nucleotides derived from therapeutic nucleosides. Recent studies showed that a anti-hepatitis C virus ribonucleoside triphosphate known as ribavirin triphosphate is incorporated by both mtRNAP and nuclear RNAP II (Arnold et al., 2012a). Whereas RNAP II utilizes factor regulated proofreading activity to excise the incorrect nucleotide from the transcript, mtRNAP lacks this proofreading mechanism (Arnold et al., 2012a). Patient toxicity in clinical trials may be traced back to defects in mitochondrial transcription as an off target effect (Arnold et al., 2012a). The design of new, more agreeable anti-viral drugs implies the elucidation of the molecular mechanisms of therapeutic nucleotide incorporation. Therefore, the mtRNAP elongation complex could be expanded by soaking or co-crystallization experiments including ribavirin triphosphate or similar anti-viral ribonucleoside triphosphates.

The described extension of structural studies of the mtRNAP elongation complex will not only deepen our understanding of evolutionary developments in early eukaryotes but will also help modern medicine in developing better anti-viral drugs.

4.4 Crystallization of mtRNAP during different transcriptional phases

This work laid the foundation for future investigations of additional mtRNAP complexes comprising transcription initiation factors and regulatory factors.

Co-crystallization of mtRNAP with DNA-RNA scaffolds containing RNA oligonucleotides of different lengths could reveal an intermediate state between initiation and elongation phase with both promoter and downstream DNA duplexes bound (compare (Durniak et al., 2008)). This could reveal new insights into the interaction network and release mechanism of TFAM and/or TFB2M during the transition of mitochondrial initiation to elongation phase.

Another aim of future research will be the visualization of mtRNAP in other functional conformations to complete the picture of the molecular mechanisms during the mitochondrial transcription cycle. According to the current model, mitochondrial transcription initiation is induced by the sequential assembly of the pre-initiation complex (PIC) comprising mtRNAP, TFAM and double-stranded promoter DNA, that is then completed by binding of TFB2M forming the initiation complex (IC) (Morozov et al., 2014; Posse et al., 2014). Upon the availability of both complex structures, the individual steps towards transcription initiation comprising promoter recognition and binding by TFAM, recruitment of mtRNAP and TFB2M, promoter melting and binding of the priming nucleotide to the active center of the polymerase can be elucidated in more detail. Initial crystallization trials should concentrate on the strategy of co-crystallization of the respective components. Additionally, PIC or IC stability can be increased by the use of mtRNAP-TFAM or mtRNAP-TFB2M fusion constructs designed according to available cross-linking data and a low quality electron microscopy model (Morozov et al., 2014; Posse et al., 2014; Yakubovskaya et al., 2014).

Regulation of mitochondrial transcription is commonly due to the influence of protein cofactors that affect mtRNAP through protein-protein interactions. Therefore, crystallization of mtRNAP in complex with transcriptional activators such as LRPPRC, MRLP12 or terminating factors such as mTerf1 will contribute to a deeper understanding of regulatory processes during RNA synthesis.

Another remarkable but also little understood aspect of mtRNAP is its primase activity during mitochondrial replication. It needs to be further investigated how transcription and replication are linked on a molecular level and which regulatory proteins trigger this process. Especially the capability of mtRNAP to bind to single-stranded DNA stem-loop structures to initiate factor-independent RNA synthesis could be further probed by a combination of functional and structural approaches. For this purpose, transcription assays could be used to both gain kinetic data on origin-specific primase activity of mtRNAP as well as the design of DNA hairpin oligonucleotides for mtRNAP co-crystallization trials. A comparison with the crystal structure of virion RNAP of the bacteriophage N4 in complex with a single-stranded DNA hairpin (Gleghorn et al., 2008) might help in respect of the strategic approach.

Taken together, the structure of the human mtRNAP elongation complex presented in this work is an important step towards a molecular understanding of the mitochondrial transcription cycle. Additional insights into different mtRNAP complexes during different transcriptional phases will eventually reveal the regulatory network and molecular mechanisms dictating mitochondrial gene transcription.

REFERENCES# References

Adams, K.L., and Palmer, J.D. (2003). Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phylogenet Evol *29*, 380-395.

Afonine, P.V., Grosse-Kunstleve, R.W., and Adams, P.D. (2005). A robust bulk-solvent correction and anisotropic scaling procedure. Acta Crystallogr D Biol Crystallogr *61*, 850-855.

Alam, T.I., Kanki, T., Muta, T., Ukaji, K., Abe, Y., Nakayama, H., Takio, K., Hamasaki, N., and Kang, D. (2003). Human mitochondrial DNA is packaged with TFAM. Nucleic Acids Res *31*, 1640-1645.

Allen, J.F. (1993). Redox control of transcription: sensors, response regulators, activators and repressors. FEBS Lett *332*, 203-207.

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.

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F.*, et al.* (1981). Sequence and organization of the human mitochondrial genome. Nature *290*, 457-465.

Andersson, G.E., and Kurland, C.G. (1991). An extreme codon preference strategy: codon reassignment. Mol Biol Evol *8*, 530-544.

Andersson, S.G., Karlberg, O., Canback, B., and Kurland, C.G. (2003). On the origin of mitochondria: a genomics perspective. Philos Trans R Soc Lond B Biol Sci *358*, 165-177; discussion 177-169.

Arnold, J.J., Sharma, S.D., Feng, J.Y., Ray, A.S., Smidansky, E.D., Kireeva, M.L., Cho, A., Perry, J., Vela, J.E., Park, Y.*, et al.* (2012a). Sensitivity of Mitochondrial Transcription and Resistance of RNA Polymerase II Dependent Nuclear Transcription to Antiviral Ribonucleosides. PLoS Pathog *8*, e1003030.

Arnold, J.J., Smidansky, E.D., Moustafa, I.M., and Cameron, C.E. (2012b). Human mitochondrial RNA polymerase: structure-function, mechanism and inhibition. Biochim Biophys Acta *1819*, 948-960.

Asin-Cayuela, J., and Gustafsson, C.M. (2007). Mitochondrial transcription and its regulation in mammalian cells. Trends Biochem Sci *32*, 111-117.

Basu, R.S., and Murakami, K.S. (2013). Watching the bacteriophage N4 RNA polymerase transcription by time-dependent soak-trigger-freeze X-ray crystallography. J Biol Chem *288*, 3305-3311.

77

Becker, T., Bottinger, L., and Pfanner, N. (2012). Mitochondrial protein import: from transport pathways to an integrated network. Trends Biochem Sci 37, 85-91.

Beese, L.S., Friedman, J.M., and Steitz, T.A. (1993). Crystal structures of the Klenow fragment of DNA polymerase I complexed with deoxynucleoside triphosphate and pyrophosphate. Biochemistry 32, 14095-14101.

Bogenhagen, D., and Clayton, D.A. (1977). Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. Cell *11*, 719-727.

Bogenhagen, D.F., Rousseau, D., and Burke, S. (2008). The layered structure of human mitochondrial DNA nucleoids. J Biol Chem *283*, 3665-3675.

Bogenhagen, D.F., Wang, Y., Shen, E.L., and Kobayashi, R. (2003). Protein components of mitochondrial DNA nucleoids in higher eukaryotes. Mol Cell Proteomics *2*, 1205-1216.

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.

Borukhov, S., and Nudler, E. (2008). RNA polymerase: the vehicle of transcription. Trends Microbiol *16*, 126-134.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem *72*, 248-254.

Brieba, L.G., Gopal, V., and Sousa, R. (2001). Scanning mutagenesis reveals roles for helix n of the bacteriophage T7 RNA polymerase thumb subdomain in transcription complex stability, pausing, and termination. J Biol Chem *276*, 10306-10313.

Broennimann, C., Eikenberry, E.F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T.*, et al.* (2006). The PILATUS 1M detector. J Synchrotron Radiat *13*, 120-130.

Brookes, P.S., Levonen, A.L., Shiva, S., Sarti, P., and Darley-Usmar, V.M. (2002). Mitochondria: regulators of signal transduction by reactive oxygen and nitrogen species. Free Radic Biol Med *33*, 755-764.

Brown, T.A., Cecconi, C., Tkachuk, A.N., Bustamante, C., and Clayton, D.A. (2005). Replication of mitochondrial DNA occurs by strand displacement with alternative light-strand origins, not via a strand-coupled mechanism. Genes Dev *19*, 2466-2476.

Brown, W.M., George, M., Jr., and Wilson, A.C. (1979). Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci U S A *76*, 1967-1971.

Burgers, P.M., Koonin, E.V., Bruford, E., Blanco, L., Burtis, K.C., Christman, M.F., Copeland, W.C., Friedberg, E.C., Hanaoka, F., Hinkle, D.C.*, et al.* (2001). Eukaryotic DNA polymerases: proposal for a revised nomenclature. J Biol Chem *276*, 43487-43490.

Calvo, S.E., and Mootha, V.K. (2010). The mitochondrial proteome and human disease. Annu Rev Genomics Hum Genet *11*, 25-44.

Campbell, C.T., Kolesar, J.E., and Kaufman, B.A. (2012). Mitochondrial transcription factor A regulates mitochondrial transcription initiation, DNA packaging, and genome copy number. Biochim Biophys Acta *1819*, 921-929.

Canter, J.A., Kallianpur, A.R., Parl, F.F., and Millikan, R.C. (2005). Mitochondrial DNA G10398A polymorphism and invasive breast cancer in African-American women. Cancer Res *65*, 8028-8033.

Carafoli, E. (1970). Calcium ion transport in mitochondria. Biochem J *116*, 2P-3P.

Carter, R., and Drouin, G. (2010). The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. Mol Biol Evol *27*, 1035-1043.

Cermakian, N., Ikeda, T.M., Miramontes, P., Lang, B.F., Gray, M.W., and Cedergren, R. (1997). On the evolution of the single-subunit RNA polymerases. J Mol Evol *45*, 671-681.

Cerritelli, S.M., Frolova, E.G., Feng, C., Grinberg, A., Love, P.E., and Crouch, R.J. (2003). Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. Mol Cell *11*, 807-815.

Chamberlin, L.L., Mpanias, O.D., and Wang, T.Y. (1983). Isolation, properties, and androgen regulation of a 20-kilodalton protein from rat ventral prostate. Biochemistry *22*, 3072-3077.

Chang, D.D., and Clayton, D.A. (1985). Priming of human mitochondrial DNA replication occurs at the light-strand promoter. Proc Natl Acad Sci U S A *82*, 351-355.

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.

Cheetham, G.M., and Steitz, T.A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. Science *286*, 2305-2309.

Chen, C.T., Hsu, S.H., and Wei, Y.H. (2012). Mitochondrial bioenergetic function and metabolic plasticity in stem cell differentiation and cellular reprogramming. Biochim Biophys Acta *1820*, 571-576.

Chujo, T., Ohira, T., Sakaguchi, Y., Goshima, N., Nomura, N., Nagao, A., and Suzuki, T. (2012). LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. Nucleic Acids Res *40*, 8033-8047.

Claros, M.G., Perea, J., Shu, Y., Samatey, F.A., Popot, J.L., and Jacq, C. (1995). Limitations to in vivo import of hydrophobic proteins into yeast mitochondria. The case of a cytoplasmically synthesized apocytochrome b. Eur J Biochem *228*, 762-771.

Clayton, D.A. (1991). Replication and transcription of vertebrate mitochondrial DNA. Annu Rev Cell Biol *7*, 453-478.

Cline, S.D., Lodeiro, M.F., Marnett, L.J., Cameron, C.E., and Arnold, J.J. (2010). Arrest of human mitochondrial RNA polymerase transcription by the biological aldehyde adduct of DNA, M1dG. Nucleic Acids Res *38*, 7546-7557.

Connolly, B., Parsons, C.A., Benson, F.E., Dunderdale, H.J., Sharples, G.J., Lloyd, R.G., and West, S.C. (1991). Resolution of Holliday junctions in vitro requires the Escherichia coli ruvC gene product. Proc Natl Acad Sci U S A *88*, 6063-6067.

Cotney, J., McKay, S.E., and Shadel, G.S. (2009). Elucidation of separate, but collaborative functions of the rRNA methyltransferase-related human mitochondrial transcription factors B1 and B2 in mitochondrial biogenesis reveals new insight into maternally inherited deafness. Hum Mol Genet *18*, 2670-2682.

Cramer, P. (2002a). Common structural features of nucleic acid polymerases. Bioessays *24*, 724-729.

Cramer, P. (2002b). Multisubunit RNA polymerases. Curr Opin Struct Biol 12, 89-97.

Cramer, P., Armache, K.J., Baumli, S., Benkert, S., Brueckner, F., Buchen, C., Damsma, G.E., Dengl, S., Geiger, S.R., Jasiak, A.J., et al. (2008). Structure of eukaryotic RNA polymerases. Annu Rev Biophys 37, 337-352.

Crick, F. (1970). Central dogma of molecular biology. Nature *227*, 561-563.

Damsma, G.E., and Cramer, P. (2009). Molecular basis of transcriptional mutagenesis at 8-oxoguanine. J Biol Chem *284*, 31658-31663.

Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. Nucleic Acids Res *32*, W615-619.

Davydova, E.K., Santangelo, T.J., and Rothman-Denes, L.B. (2007). Bacteriophage N4 virion RNA polymerase interaction with its promoter DNA hairpin. Proc Natl Acad Sci U S A *104*, 7033-7038.

Delannoy, E., Stanley, W.A., Bond, C.S., and Small, I.D. (2007). Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochem Soc Trans *35*, 1643-1647.

DeLano, W. (2002). The PyMOL Molecular Graphics System. DeLano Scientific: San Carlos, CA, USA.

Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990). An attempt to unify the structure of polymerases. Protein Eng *3*, 461-467.

Deshpande, A.P., and Patel, S.S. (2012). Mechanism of transcription initiation by the yeast mitochondrial RNA polymerase. Biochim Biophys Acta *1819*, 930-938.

Diffley, J.F., and Stillman, B. (1991). A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. Proc Natl Acad Sci U S A *88*, 7864-7868.

Doublie, S., and Ellenberger, T. (1998). The mechanism of action of T7 DNA polymerase. Curr Opin Struct Biol *8*, 704-712.

Durniak, K.J., Bailey, S., and Steitz, T.A. (2008). The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. Science *322*, 553-557.

Ekstrand, M.I., Falkenberg, M., Rantanen, A., Park, C.B., Gaspari, M., Hultenby, K., Rustin, P., Gustafsson, C.M., and Larsson, N.G. (2004). Mitochondrial transcription factor A regulates mtDNA copy number in mammals. Hum Mol Genet *13*, 935-944.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.G., and Gustafsson, C.M. (2002). Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nat Genet *31*, 289-294.

Fisher, A.G., and Brown, G. (1980). A rapid method for determining whether monoclonal antibodies react with the same or different antigens on the cell surface. J Immunol Methods *39*, 377-385.

Fisher, R.P., and Clayton, D.A. (1985). A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. J Biol Chem *260*, 11330-11338.

Fisher, R.P., and Clayton, D.A. (1988). Purification and characterization of human mitochondrial transcription factor 1. Mol Cell Biol *8*, 3496-3509.

Fisher, R.P., Topper, J.N., and Clayton, D.A. (1987). Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. Cell *50*, 247-258.

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.

Gangelhoff, T.A., Mungalachetty, P.S., Nix, J.C., and Churchill, M.E. (2009). Structural analysis and DNA binding of the HMG domains of the human mitochondrial transcription factor A. Nucleic Acids Res *37*, 3153-3164.

Gaspari, M., Falkenberg, M., Larsson, N.G., and Gustafsson, C.M. (2004a). The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. EMBO J *23*, 4606-4614.

Gaspari, M., Larsson, N.G., and Gustafsson, C.M. (2004b). The transcription machinery in mammalian mitochondria. Biochim Biophys Acta *1659*, 148-152.

Geiduschek, E.P., and Ouhammouch, M. (2005). Archaeal transcription and its regulators. Mol Microbiol *56*, 1397-1407.

Gelfand, R., and Attardi, G. (1981). Synthesis and turnover of mitochondrial ribonucleic acid in HeLa cells: the mature ribosomal and messenger ribonucleic acid species are metabolically unstable. Mol Cell Biol *1*, 497-511.

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.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001a). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876-1882.

Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A., and Kornberg, R.D. (2001b). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. Science *292*, 1876-1882.

Grachev, M.A., Lukhtanov, E.A., Mustaev, A.A., Zaychikov, E.F., Abdukayumov, M.N., Rabinov, I.V., Richter, V.I., Skoblov, Y.S., and Chistyakov, P.G. (1989). Studies of the functional topography of Escherichia coli RNA polymerase. A method for localization of the sites of affinity labelling. Eur J Biochem *180*, 577-585.

Gray, M.W. (2012). Mitochondrial evolution. Cold Spring Harb Perspect Biol *4*, a011403.

Gray, M.W., and Doolittle, W.F. (1982). Has the endosymbiont hypothesis been proven? Microbiol Rev *46*, 1-42.

Gray, M.W., and Lang, B.F. (1998). Transcription in chloroplasts and mitochondria: a tale of two polymerases. Trends Microbiol *6*, 1-3.

Greaves, L.C., Reeve, A.K., Taylor, R.W., and Turnbull, D.M. (2012). Mitochondrial DNA and disease. J Pathol *226*, 274-286.

Green, D.R., and Reed, J.C. (1998). Mitochondria and apoptosis. Science *281*, 1309-1312.

Grummt, I. (2003). Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. Genes Dev 17, 1691-1702.

Guo, Q., Nayak, D., Brieba, L.G., and Sousa, R. (2005). Major conformational changes during T7RNAP transcription initiation coincide with, and are required for, promoter release. J Mol Biol 353, 256-270.

Hatefi, Y. (1985). The mitochondrial electron transport and oxidative phosphorylation system. Annu Rev Biochem 54, 1015-1069.

Hein, P.P., Palangat, M., and Landick, R. (2011). RNA transcript 3'-proximal sequence affects translocation bias of RNA polymerase. Biochemistry *50*, 7002-7014.

Holt, I.J., Harding, A.E., and Morgan-Hughes, J.A. (1988). Mitochondrial DNA polymorphism in mitochondrial myopathy. Hum Genet *79*, 53-57.

Holt, I.J., and Jacobs, H.T. (2003). Response: The mitochondrial DNA replication bubble has not burst. Trends Biochem Sci *28*, 355-356.

Hyvarinen, A.K., Pohjoismaki, J.L., Reyes, A., Wanrooij, S., Yasukawa, T., Karhunen, P.J., Spelbrink, J.N., Holt, I.J., and Jacobs, H.T. (2007). The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. Nucleic Acids Res *35*, 6458-6474.

Iyer, L.M., Koonin, E.V., and Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. BMC Struct Biol *3*, 1.

Jeruzalmi, D., and Steitz, T.A. (1998). Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. EMBO J *17*, 4101-4113.

Joyce, C.M., and Steitz, T.A. (1994). Function and structure relationships in DNA polymerases. Annu Rev Biochem *63*, 777-822.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kang, D., and Hamasaki, N. (2005). Mitochondrial transcription factor A in the maintenance of mitochondrial DNA: overview of its multiple roles. Ann N Y Acad Sci *1042*, 101-108.

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

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030-1033.

Kaufman, B.A., Durisic, N., Mativetsky, J.M., Costantino, S., Hancock, M.A., Grutter, P., and Shoubridge, E.A. (2007). The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. Mol Biol Cell *18*, 3225-3236.

Kettenberger, H., Armache, K.J., and Cramer, P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol Cell *16*, 955-965.

Kiefer, J.R., Mao, C., Hansen, C.J., Basehore, S.L., Hogrefe, H.H., Braman, J.C., and Beese, L.S. (1997). Crystal structure of a thermostable Bacillus DNA polymerase I large fragment at 2.1 A resolution. Structure *5*, 95-108.

Korhonen, J.A., Pham, X.H., Pellegrini, M., and Falkenberg, M. (2004). Reconstitution of a minimal mtDNA replisome in vitro. EMBO J *23*, 2423-2429.

Korzheva, N., Mustaev, A., Kozlov, M., Malhotra, A., Nikiforov, V., Goldfarb, A., and Darst, S.A. (2000). A structural model of transcription elongation. Science *289*, 619-625.

Kostyuk, D.A., Dragan, S.M., Lyakhov, D.L., Rechinsky, V.O., Tunitskaya, V.L., Chernov, B.K., and Kochetkov, S.N. (1995). Mutants of T7 RNA polymerase that are able to synthesize both RNA and DNA. FEBS Lett *369*, 165-168.

Kravchenko, J.E., Rogozin, I.B., Koonin, E.V., and Chumakov, P.M. (2005). Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. Nature *436*, 735-739.

Kruse, B., Narasimhan, N., and Attardi, G. (1989). Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. Cell *58*, 391-397.

Kucej, M., Kucejova, B., Subramanian, R., Chen, X.J., and Butow, R.A. (2008). Mitochondrial nucleoids undergo remodeling in response to metabolic cues. J Cell Sci *121*, 1861-1868.

Lahmy, S., Bies-Etheve, N., and Lagrange, T. (2010). Plant-specific multisubunit RNA polymerase in gene silencing. Epigenetics *5*, 4-8.

Larsson, N.G., Barsh, G.S., and Clayton, D.A. (1997). Structure and chromosomal localization of the mouse mitochondrial transcription factor A gene (Tfam). Mamm Genome *8*, 139-140.

Larsson, N.G., and Clayton, D.A. (1995). Molecular genetic aspects of human mitochondrial disorders. Annu Rev Genet *29*, 151-178.

Leslie, A.G. (2006). The integration of macromolecular diffraction data. Acta Crystallogr D Biol Crystallogr *62*, 48-57.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151.

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.

Liu, L., Sanosaka, M., Lei, S., Bestwick, M.L., Frey, J.H., Jr., Surovtseva, Y.V., Shadel, G.S., and Cooper, M.P. (2011). LRP130 protein remodels mitochondria and stimulates fatty acid oxidation. J Biol Chem *286*, 41253-41264.

Liu, P., and Demple, B. (2010). DNA repair in mammalian mitochondria: Much more than we thought? Environ Mol Mutagen *51*, 417-426.

Lodeiro, M.F., Uchida, A.U., Arnold, J.J., Reynolds, S.L., Moustafa, I.M., and Cameron, C.E. (2010). Identification of multiple rate-limiting steps during the human mitochondrial transcription cycle in vitro. J Biol Chem *285*, 16387-16402.

Lu, B., Lee, J., Nie, X., Li, M., Morozov, Y.I., Venkatesh, S., Bogenhagen, D.F., Temiakov, D., and Suzuki, C.K. (2013). Phosphorylation of human TFAM in mitochondria impairs DNA binding and promotes degradation by the AAA+ Lon protease. Mol Cell *49*, 121-132.

Mangus, D.A., Jang, S.H., and Jaehning, J.A. (1994). Release of the yeast mitochondrial RNA polymerase specificity factor from transcription complexes. J Biol Chem *269*, 26568-26574.

Martin, M., Cho, J., Cesare, A.J., Griffith, J.D., and Attardi, G. (2005). Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. Cell *123*, 1227-1240.

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

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

Matsunaga, M., and Jaehning, J.A. (2004a). Intrinsic promoter recognition by a "core" RNA polymerase. J Biol Chem *279*, 44239-44242.

Matsunaga, M., and Jaehning, J.A. (2004b). A mutation in the yeast mitochondrial core RNA polymerase, Rpo41, confers defects in both specificity factor interaction and promoter utilization. J Biol Chem *279*, 2012-2019.

Matsushima, Y., Goto, Y., and Kaguni, L.S. (2010). Mitochondrial Lon protease regulates mitochondrial DNA copy number and transcription by selective degradation of mitochondrial transcription factor A (TFAM). Proc Natl Acad Sci U S A *107*, 18410-18415.

McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C., and Read, R.J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr *61*, 458-464.

McCulloch, V., and Shadel, G.S. (2003). Human mitochondrial transcription factor B1 interacts with the C-terminal activation region of h-mtTFA and stimulates transcription independently of its RNA methyltransferase activity. Mol Cell Biol *23*, 5816-5824.

McFarland, R., Taylor, R.W., and Turnbull, D.M. (2010). A neurological perspective on mitochondrial disease. Lancet Neurol *9*, 829-840.

Mentesana, P.E., Chin-Bow, S.T., Sousa, R., and McAllister, W.T. (2000). Characterization of halted T7 RNA polymerase elongation complexes reveals multiple factors that contribute to stability. J Mol Biol *302*, 1049-1062.

Mili, S., and Pinol-Roma, S. (2003). LRP130, a pentatricopeptide motif protein with a noncanonical RNA-binding domain, is bound in vivo to mitochondrial and nuclear RNAs. Mol Cell Biol *23*, 4972-4982.

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.

Miquel, J., Economos, A.C., Fleming, J., and Johnson, J.E., Jr. (1980). Mitochondrial role in cell aging. Exp Gerontol *15*, 575-591.

MITOMAP (2013). MITOMAP: a human mitochondrial genome database. In http://wwwmitomaporg.

Mokranjac, D., and Neupert, W. (2005). Protein import into mitochondria. Biochem Soc Trans *33*, 1019-1023.

Montoya, J., Christianson, T., Levens, D., Rabinowitz, M., and Attardi, G. (1982). Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. Proc Natl Acad Sci U S A 79, 7195-7199.

Mooney, R.A., Darst, S.A., and Landick, R. (2005). Sigma and RNA polymerase: an on-again, off-again relationship? Mol Cell *20*, 335-345.

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 [Epub ahead of print].

Nakanishi, N., Fukuoh, A., Kang, D., Iwai, S., and Kuraoka, I. (2013). Effects of DNA lesions on the transcription reaction of mitochondrial RNA polymerase: implications for bypass RNA synthesis on oxidative DNA lesions. Mutagenesis *28*, 117-123.

Nass, M.M. (1966). The circularity of mitochondrial DNA. Proc Natl Acad Sci U S A *56*, 1215-1222.

Nayak, D., Guo, Q., and Sousa, R. (2007). Functional architecture of T7 RNA polymerase transcription complexes. J Mol Biol *371*, 490-500.

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.

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.

Ohniwa, R.L., Morikawa, K., Takeshita, S.L., Kim, J., Ohta, T., Wada, C., and Takeyasu, K. (2007). Transcription-coupled nucleoid architecture in bacteria. Genes Cells *12*, 1141-1152.

Ojala, D., Crews, S., Montoya, J., Gelfand, R., and Attardi, G. (1981). A small polyadenylated RNA (7 S RNA), containing a putative ribosome attachment site, maps near the origin of human mitochondrial DNA replication. J Mol Biol *150*, 303-314.

Ott, M., Gogvadze, V., Orrenius, S., and Zhivotovsky, B. (2007). Mitochondria, oxidative stress and cell death. Apoptosis *12*, 913-922.

Paratkar, S., Deshpande, A.P., Tang, G.Q., and Patel, S.S. (2011). The N-terminal domain of the yeast mitochondrial RNA polymerase regulates multiple steps of transcription. J Biol Chem *286*, 16109-16120.

Paratkar, S., and Patel, S.S. (2010). Mitochondrial transcription factor Mtf1 traps the unwound non-template strand to facilitate open complex formation. J Biol Chem *285*, 3949-3956.

Parisi, M.A., and Clayton, D.A. (1991). Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science *252*, 965-969.

Park, C.B., Asin-Cayuela, J., Camara, Y., Shi, Y., Pellegrini, M., Gaspari, M., Wibom, R., Hultenby, K., Erdjument-Bromage, H., Tempst, P*., et al.* (2007). MTERF3 is a negative regulator of mammalian mtDNA transcription. Cell *130*, 273-285.

Pedersen, P.L. (1978). Tumor mitochondria and the bioenergetics of cancer cells. Prog Exp Tumor Res *22*, 190-274.

Petros, J.A., Baumann, A.K., Ruiz-Pesini, E., Amin, M.B., Sun, C.Q., Hall, J., Lim, S., Issa, M.M., Flanders, W.D., Hosseini, S.H.*, et al.* (2005). mtDNA mutations increase tumorigenicity in prostate cancer. Proc Natl Acad Sci U S A *102*, 719-724.

Pfannschmidt, T., Nilsson, A., Tullberg, A., Link, G., and Allen, J.F. (1999). Direct transcriptional control of the chloroplast genes psbA and psaAB adjusts photosynthesis to light energy distribution in plants. IUBMB Life *48*, 271-276.

Pica-Mattoccia, L., and Attardi, G. (1972). Expression of the mitochondrial genome in HeLa cells. IX. Replication of mitochondrial DNA in relationship to cell cycle in HeLa cells. J Mol Biol *64*, 465-484.

Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M.A., Lerbs-Mache, S., Colot, V., and Lagrange, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev *19*, 2030-2040.

Ponting, C.P. (2002). Novel domains and orthologues of eukaryotic transcription elongation factors. Nucleic Acids Res *30*, 3643-3652.

Popot, J.L., and de Vitry, C. (1990). On the microassembly of integral membrane proteins. Annu Rev Biophys Biophys Chem *19*, 369-403.

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 [Epub ahead of print].

Rebelo, A.P., Dillon, L.M., and Moraes, C.T. (2011). Mitochondrial DNA transcription regulation and nucleoid organization. J Inherit Metab Dis *34*, 941-951.

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.

Roberti, M., Bruni, F., Loguercio Polosa, P., Manzari, C., Gadaleta, M.N., and Cantatore, P. (2006). MTERF3, the most conserved member of the mTERF-family, is a modular factor involved in mitochondrial protein synthesis. Biochim Biophys Acta *1757*, 1199-1206.

Rodeheffer, M.S., and Shadel, G.S. (2003). Multiple interactions involving the amino-terminal domain of yeast mtRNA polymerase determine the efficiency of mitochondrial protein synthesis. J Biol Chem *278*, 18695-18701.

Roeder, R.G. (1996). Nuclear RNA polymerases: role of general initiation factors and cofactors in eukaryotic transcription. Methods Enzymol *273*, 165-171.

Roeder, R.G., and Rutter, W.J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature *224*, 234-237.

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.

Ruhanen, H., Borrie, S., Szabadkai, G., Tyynismaa, H., Jones, A.W., Kang, D., Taanman, J.W., and Yasukawa, T. (2010). Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. Biochim Biophys Acta *1803*, 931-939.

Ruzzenente, B., Metodiev, M.D., Wredenberg, A., Bratic, A., Park, C.B., Camara, Y., Milenkovic, D., Zickermann, V., Wibom, R., Hultenby, K.*, et al.* (2012). LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. EMBO J *31*, 443-456.

Schmitz-Linneweber, C., and Small, I. (2008). Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci *13*, 663-670.

Seidel-Rogol, B.L., McCulloch, V., and Shadel, G.S. (2003). Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. Nat Genet *33*, 23-24.

Shadel, G.S., and Clayton, D.A. (1997). Mitochondrial DNA maintenance in vertebrates. Annu Rev Biochem *66*, 409-435.

Shock, L.S., Thakkar, P.V., Peterson, E.J., Moran, R.G., and Taylor, S.M. (2011). DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. Proc Natl Acad Sci U S A *108*, 3630-3635.

Shutt, T.E., and Gray, M.W. (2006). Bacteriophage origins of mitochondrial replication and transcription proteins. Trends Genet 22, 90-95.

Shutt, T.E., Lodeiro, M.F., Cotney, J., Cameron, C.E., and Shadel, G.S. (2010). Core human mitochondrial transcription apparatus is a regulated two-component system in vitro. Proc Natl Acad Sci U S A *107*, 12133-12138.

Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci *25*, 46-47.

Smidansky, E.D., Arnold, J.J., Reynolds, S.L., and Cameron, C.E. (2011). Human mitochondrial RNA polymerase: evaluation of the single-nucleotide-addition cycle on synthetic RNA/DNA scaffolds. Biochemistry *50*, 5016-5032.

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.

Sondheimer, N., Fang, J.K., Polyak, E., Falk, M.J., and Avadhani, N.G. (2010). Leucine-rich pentatricopeptide-repeat containing protein regulates mitochondrial transcription. Biochemistry *49*, 7467-7473.

Sousa, R. (1996). Structural and mechanistic relationships between nucleic acid polymerases. Trends Biochem Sci *21*, 186-190.

Sousa, R., and Padilla, R. (1995). A mutant T7 RNA polymerase as a DNA polymerase. EMBO J *14*, 4609-4621.

Spelbrink, J.N., Li, F.Y., Tiranti, V., Nikali, K., Yuan, Q.P., Tariq, M., Wanrooij, S., Garrido, N., Comi, G., Morandi, L.*, et al.* (2001). Human mitochondrial DNA deletions associated with mutations in the gene encoding Twinkle, a phage T7 gene 4-like protein localized in mitochondria. Nat Genet *28*, 223-231.

Steitz, T.A. (2009). The structural changes of T7 RNA polymerase from transcription initiation to elongation. Curr Opin Struct Biol *19*, 683-690.

Steitz, T.A., Smerdon, S.J., Jager, J., and Joyce, C.M. (1994). A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. Science *266*, 2022-2025.

Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci U S A *90*, 6498-6502.

Studier, F.W. (1972). Bacteriophage T7. Science *176*, 367-376.

Surovtseva, Y.V., and Shadel, G.S. (2013). Transcription-independent role for human mitochondrial RNA polymerase in mitochondrial ribosome biogenesis. Nucleic Acids Res *41*, 2479-2488.

Surovtseva, Y.V., Shutt, T.E., Cotney, J., Cimen, H., Chen, S.Y., Koc, E.C., and Shadel, G.S. (2011). Mitochondrial ribosomal protein L12 selectively associates with human mitochondrial RNA polymerase to activate transcription. Proc Natl Acad Sci U S A *108*, 17921-17926.

Sutovsky, P., Tengowski, M.W., Navara, C.S., Zoran, S.S., and Schatten, G. (1997). Mitochondrial sheath movement and detachment in mammalian, but not nonmammalian, sperm induced by disulfide bond reduction. Mol Reprod Dev *47*, 79-86.

Tahirov, T.H., Temiakov, D., Anikin, M., Patlan, V., McAllister, W.T., Vassylyev, D.G., and Yokoyama, S. (2002). Structure of a T7 RNA polymerase elongation complex at 2.9 A resolution. Nature *420*, 43-50.

Tapper, D.P., and Clayton, D.A. (1981). Mechanism of replication of human mitochondrial DNA. Localization of the 5' ends of nascent daughter strands. J Biol Chem *256*, 5109-5115.

Taylor, S.D., Zhang, H., Eaton, J.S., Rodeheffer, M.S., Lebedeva, M.A., O'Rourke T, W., Siede, W., and Shadel, G.S. (2005). The conserved Mec1/Rad53 nuclear checkpoint pathway regulates mitochondrial DNA copy number in Saccharomyces cerevisiae. Mol Biol Cell *16*, 3010-3018.

Temiakov, D., Anikin, M., and McAllister, W.T. (2002). Characterization of T7 RNA polymerase transcription complexes assembled on nucleic acid scaffolds. J Biol Chem *277*, 47035-47043.

Temiakov, D., Mentesana, P.E., Ma, K., Mustaev, A., Borukhov, S., and McAllister, W.T. (2000). The specificity loop of T7 RNA polymerase interacts first with the promoter and then with the elongating transcript, suggesting a mechanism for promoter clearance. Proc Natl Acad Sci U S A *97*, 14109-14114.

Temiakov, D., Patlan, V., Anikin, M., McAllister, W.T., Yokoyama, S., and Vassylyev, D.G. (2004). Structural basis for substrate selection by t7 RNA polymerase. Cell *116*, 381-391.

Tiranti, V., Savoia, A., Forti, F., D'Apolito, M.F., Centra, M., Rocchi, M., and Zeviani, M. (1997). Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. Hum Mol Genet *6*, 615-625.

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J.N., Rovio, A.T., Bruder, C.E., Bohlooly, Y.M., Gidlof, S., Oldfors, A., Wibom, R.*, et al.* (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature *429*, 417-423.

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.

von Heijne, G. (1986). Why mitochondria need a genome. FEBS Lett *198*, 1-4.

Wallace, D.C. (2007). Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. Annu Rev Biochem *76*, 781-821.

Wallace, D.C., Fan, W., and Procaccio, V. (2010). Mitochondrial energetics and therapeutics. Annu Rev Pathol *5*, 297-348.

Wallace, D.C., Singh, G., Lott, M.T., Hodge, J.A., Schurr, T.G., Lezza, A.M., Elsas, L.J., 2nd, and Nikoskelainen, E.K. (1988). Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science *242*, 1427-1430.

Wang, Y., and Bogenhagen, D.F. (2006). Human mitochondrial DNA nucleoids are linked to protein folding machinery and metabolic enzymes at the mitochondrial inner membrane. J Biol Chem *281*, 25791-25802.

Wang, Z., Cotney, J., and Shadel, G.S. (2007). Human mitochondrial ribosomal protein MRPL12 interacts directly with mitochondrial RNA polymerase to modulate mitochondrial gene expression. J Biol Chem *282*, 12610-12618.

Wanrooij, S., Fuste, J.M., Farge, G., Shi, Y., Gustafsson, C.M., and Falkenberg, M. (2008). Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. Proc Natl Acad Sci U S A *105*, 11122-11127.

Weinmann, R., and Roeder, R.G. (1974). Role of DNA-dependent RNA polymerase 3 in the transcription of the tRNA and 5S RNA genes. Proc Natl Acad Sci U S A *71*, 1790-1794.

Weiss, S., and Gladstone, L.A. (1959). A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. J Am Chem Soc *81:*, 4118.

Weissman, L., de Souza-Pinto, N.C., Stevnsner, T., and Bohr, V.A. (2007). DNA repair, mitochondria, and neurodegeneration. Neuroscience *145*, 1318-1329.

Wenz, T., Luca, C., Torraco, A., and Moraes, C.T. (2009). mTERF2 regulates oxidative phosphorylation by modulating mtDNA transcription. Cell Metab *9*, 499-511.

Werner, F., and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol *9*, 85-98.

Wong, T.W., and Clayton, D.A. (1985). In vitro replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. Cell *42*, 951-958.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.C., Dufour, M.E., Boulay, J., Regnault, B., Devaux, F., Namane, A., Seraphin, B., et al. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 121, 725-737.

Xu, B., and Clayton, D.A. (1996). RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. EMBO J *15*, 3135-3143.

Yang, M.Y., Bowmaker, M., Reyes, A., Vergani, L., Angeli, P., Gringeri, E., Jacobs, H.T., and Holt, I.J. (2002). Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. Cell *111*, 495-505.

Yin, Y., Chen, H., Hahn, M.G., Mohnen, D., and Xu, Y. (2010). Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. Plant Physiol *153*, 1729-1746.

Yin, Y.W., and Steitz, T.A. (2002). Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science *298*, 1387-1395.

Yin, Y.W., and Steitz, T.A. (2004). The structural mechanism of translocation and helicase activity in T7 RNA polymerase. Cell *116*, 393-404.

Yoshida, Y., Izumi, H., Torigoe, T., Ishiguchi, H., Itoh, H., Kang, D., and Kohno, K. (2003). P53 physically interacts with mitochondrial transcription factor A and differentially regulates binding to damaged DNA. Cancer Res *63*, 3729-3734.

Zeviani, M., Bonilla, E., DeVivo, D.C., and DiMauro, S. (1989). Mitochondrial diseases. Neurol Clin *7*, 123-156.

Zhang, H., Barcelo, J.M., Lee, B., Kohlhagen, G., Zimonjic, D.B., Popescu, N.C., and Pommier, Y. (2001). Human mitochondrial topoisomerase I. Proc Natl Acad Sci U S A *98*, 10608-10613.

Zylber, E.A., and Penman, S. (1971). Products of RNA polymerases in HeLa cell nuclei. Proc Natl Acad Sci U S A *68*, 2861-2865.

Abbreviations

3'dATP	3'-deoxyadenosine-5'-triphosphate
°C	degree Celsius
aa	amino acids
AMPCPP	α,β-methyleneadenosine-5'-triphosphate
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumine
CAPSO	N-cyclohexyl-2-hydroxyl-3-aminopropanesulfonic acid
CNBr	cyanogen bromide
CTD	carboxy-terminal domain
CV	column volumes
cys	cysteine
D-loop	displacement loop
Da	Dalton
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	1,4-dithio-D,L-threitol
EC	elongation complex
E.coli	Escherichia coli
EDTA	ethylene diamine tetraacetic acid
EtOH	ethanol
h	hour
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethane sulfonic acid
HhH	helix-hairpin-helix
HMG	high mobility group
HSP	heavy-strand promoter
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	kilo Dalton

KF	Klenow fragment
LB	Luria-Bertani
LRPPRC	mitochondrial leucine-rich pentatricopeptide repeat containing protein
LSP	light-strand promoter
M	molar (mole/litre)
MCS	multiple cloning site
MDa	mega Dalton
MES	2-N-morpholino-ethanesulfonic acid
Met	methionine
min	minutes
MLS	mitochondrial localization signal
MOPS	4-morpholine-propanesulfonic acid
MRLP12	mitochondrial ribosomal protein L12
mRNA	messenger RNA
mTerf	mitochondrial termination factor
Mtf1	mitochondrial transcription factor 1
mtRNAP	mitochondrial DNA-dependent RNA polymerase
mtSSB	mitochondrial single-stranded DNA binding proteins
MW	molecular weight
NED	N-terminal extension domain
NG	hydroxylamine clevage site
NH_2OH	hydroxylamine
Ni-NTA	Nickel-nitrilotriacetic acid
nt	nucleotides
NTCB	2-nitro-5-thiocyano-benzoic acid
NTD	amino-terminal domain
NTP	nucleoside triphosphate
O_H	origin of replication in the heavy strand
OD_{600}	optical density at a wavelength of 600 nm
O_L	origin of replication in the light strand
OXPHOS	oxidative phosphorylation system
o.n.	over night

PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	poly(ethylene glycol)
PI	protease inhibitor
PIC	pre-initiation complex
PMSF	phenylmethylsulfonyl fluoride
PP_i	pyrophosphate
PPR	pentratricopeptide repeat
PVDF	polyvinylidene fluoride
R-factor	normalized linear residual between observed and calculated structure factor amplitudes
RNA	riboculeic acid
RNAP	DNA-dependent RNA polymerase
ROS	reactive oxygen species
rRNA	ribosomal RNA
tRNAP	transfer RNA
RMSD	root means square deviation
ROS	reactive oxygen species
rpm	rounds per minute
Rpo41	DNA dependent RNA polymerase of S.c.
S.c.	*Saccharomyces cerivisiae*
SDS	sodium dodecylsulfate
sec	seconds
SLS	Swiss Light Source
TBE	tris-borate/-DTA
TBP	TATA-binding protein
TCA	trichloroacetic acid
TEFM	transcription elongation factor of mitochondria
TERM1	termination region of HSP1-dependent trascription
TFAM	transcription factor A, mitochondrial
TFB	transcription factor B

TFB2M	transcription factor B2, mitochondrial
TPR	tetratricopeptide
Tris	Tris-(hydroxymethyl)-aminomethane
tRNA	transfer RNA
U	unit
UV	ultraviolet
v/v	volume per volume
vs	versus
w/v	weight per volume
WT	wild-type

List of figures

List of tables