Pathogenic mechanisms affecting mitochondrial DNA replication and transcription

Meenakshi Singh

Department of Medical Biochemistry and Cell Biology
Institute of Biomedicine
Sahlgrenska Academy, University of Gothenburg

Gothenburg, 2023

UNIVERSITY OF GOTHENBURG

Pathogenic mechanisms affecting mitochondrial DNA replication and transcription

Meenakshi Singh

© Meenakshi Singh, 2023

meenakshi.singh@gu.se

meenakshisingh@live.com

ISBN 978-91-8069-235-9 (PRINT)

ISBN 978-91-8069-236-6 (PDF)

http://hdl.handle.net/2077/75192

Printed in Borås, Sweden 2023

Printed by Stema Specialtryck AB

SVANENMÄRKET
Trycksak
3041 0234

‘’Ability is what you’re capable of doing, Motivation determines what you do and attitude determines how well you do it’’.

-Lou Holtz

To my beloved family

Pathogenic mechanisms affecting mitochondrial DNA replication and transcription

Meenakshi Singh

Department of Medical Biochemistry and Cell Biology
Institute of Biomedicine
Sahlgrenska Academy, University of Gothenburg

ABSTRACT

Mitochondria are cytoplasmic organelles fundamental to life and health. In mitochondria, energy from the food we eat is converted into adenosine triphosphate (ATP), which in turn is used as a source of chemical energy to drive a multitude of cellular reactions. Owing to its endosymbiotic origin, the mitochondrion contains its own genetic material, a circular double stranded DNA molecule (mtDNA) of about 16.6 kb. The mitochondrial genome contains 37 genes, which code for 13 protein components of the oxidative phosphorylation system, 2 ribosomal RNAs that are required for mitochondrial ribosome biogenesis, and a set of 22 transfer RNAs.

The enzymatic systems needed to replicate and transcribe mtDNA are distinct from those present in the nucleus. The objective of this thesis is to characterize pathogenic mechanisms affecting mtDNA replication and transcription in human cells. In collaboration with clinical colleagues, we report that mutations in the gene coding for mitochondrial RNA polymerase (POLRMT) can cause mitochondrial dysfunction and neurological disease. Using *in vitro* biochemistry and cell biology approaches, we find that the identified mutations

cause deleterious effects on mitochondrial transcription, which in turn impair biogenesis of the oxidative phosphorylation system.

Mammalian mitochondria lack systems for ribonucleotide excision repair and mtDNA therefore contains relatively high levels of embedded ribonucleotides, which can be even higher in patients suffering from genetic disorders associated with imbalanced nucleotide pools. We demonstrate that embedded ribonucleotides can cause problems for the transcribing POLRMT, causing premature termination of mitochondrial transcription. We suggest that these effects can contribute to the phenotypes associated with nucleotide pool disorders.

Others have demonstrated that nucleoside analogues used to treat retroviral infections can cause a progressive accumulation of somatic mtDNA mutations. We investigated the effects of two nucleoside analogues commonly used to treat childhood cancer, 5-Fluorouracil (5-FU) and 6-Thioguanine (6-TG). Using a reconstituted mtDNA replication system, with highly purified components, we demonstrate that both 5-FU and 6-TG impair the activity of mitochondrial DNA polymerase γ *in vitro*. We also find that these compounds can cause mtDNA replication stalling in cells. Taken together, our data suggest that 5-FU and 6-TG have the potential to cause mutations, but future studies of mtDNA isolated from patients treated with these compounds are required to validate this idea.

Keywords: mtDNA, ribonucleotides, nucleoside analogues, POLRMT

ISBN 978-91-8069-235-9 (PRINT)

ISBN 978-91-8069-236-6 (PDF)

SAMMANFATTNING PÅ SVENSKA

Mitokondrien är en cytoplasmatisk organell av central betydelse för eukaryota celler. I mitokondrien omvandlas energi från maten vi äter till adenosintrifosfat (ATP), som i sin tur används som en energikälla för att driva ett stort antal cellulära reaktioner. På grund av sitt endosymbiotiska ursprung innehåller mitokondrien sitt eget genetiska material, en cirkulär, dubbelsträngad DNA-molekyl (mtDNA) på cirka 16 600 baspar. Det mitokondriella genomet innehåller 37 gener, vilka kodar för 13 proteiner inblandade i oxidativ fosforylering, 2 ribosom-RNA som krävs för biogenes av mitokondriella ribosomer och en uppsättning av 22 transfer-RNA.

De enzymatiska system som replikerar och transkriberar mtDNA skiljer sig från de som finns i cellkärnan. Syftet med denna avhandling är att karakterisera patogena mekanismer som påverkar mtDNA-replikation och transkription i mänskliga celler. I samarbete med kliniska kollegor rapporterar vi att mutationer i genen som kodar för mitokondriellt RNA-polymeras (POLRMT) kan orsaka mitokondriell dysfunktion och neurologisk sjukdom. Genom att använda *in vitro*-biokemi och cellbiologiska tillvägagångssätt, finner vi att de identifierade mutationerna orsakar skadliga effekter på mitokondriell transkription, vilket i sin tur leder till nedsatt biogenes av det oxidativa fosforyleringssystemet.

I mitokondrien saknas enzymsystem som kan ta bort ribonukleotider som av misstag byggts in i mtDNA. Genomet innehåller därför höga nivåer av ribonukleotider och mängden kan stiga ytterligare hos patienter som lider av genetiska sjukdomar med obalanserade nukleotidnivåer. Vi visar att inbäddade ribonukleotider kan orsaka problem för POLRMT när enzymet använder

mtDNA som mall och resultera i oönskad terminerig av transkription. Vi föreslår att dessa effekter på transkription kan bidra till de symptom som patienter med förändrade nukleotidnivåer uppvisar.

Andra har visat att nukleosidanaloger som används för att behandla retrovirala infektioner kan orsaka en progressiv ackumulering av mtDNA-mutationer i somatiska celler. Vi undersökte om liknande effekter kunde orsakas av nukleosidanaloger som vanligtvis används för att behandla cancer hos barn, 5-fluorouracil (5-FU) och 6-tioguanin (6-TG). Genom att använda ett *in vitro*-system för mtDNA-replikation som återskapats med högt renade komponenter, visar vi att både 5-FU och 6-TG stör funktionen hos mitokondriellt DNA-polymeras γ. Vi visar också att dessa nukleosidanaloger kan terminera mtDNA-replikation i celler. Våra data demonstrerar att 5-FU och 6-TG har potential att orsaka mutationer, men framtida studier av mtDNA isolerat från patienter som genomgått behandling med dessa läkemedel krävs för att validera denna idé.

LIST OF PAPERS

The thesis is based on the following studies.

1. Ribonucleotides embedded in template DNA impair mitochondrial RNA polymerase progression.
   Meenakshi Singh, Viktor Posse, Bradley Peter, Maria Falkenberg and Claes M. Gustafsson. *Nucleic Acids Research*, 2022, Vol. 50, No. 2, 989–999. https://doi.org/10.1093/nar/gkab1251

2. POLRMT mutations impair mitochondrial transcription causing neurological disease.
   Monika Oláhová, Bradley Peter, Zsolt Szilagyi, Hector Diaz-Maldonado, **Meenakshi Singh**, Ewen W. Sommerville, Emma L. Blakely, Jack J. Collier, Emily Hoberg, Viktor Stránecký, Hana Hartmannová, Anthony J. Bleyer, Kim L. McBride, Sasigarn A. Bowden, Zuzana Korandová, Alena Pecinová, Hans-Hilger Ropers, Kimia Kahrizi, Hossein Najmabadi, Mark A. Tarnopolsky, Lauren I. Brady, K. Nicole Weaver, Carlos E. Prada, Katrin Õunap, Monica H. Wojcik, Sander Pajusalu, Safoora B. Syeda, Lynn Pais, Elicia A. Estrella, Christine C. Bruels, Louis M. Kunkel, Peter B. Kang, Penelope E. Bonnen, Tomáš Mráček, Stanislav Kmoch, Gráinne S. Gorman, Maria Falkenberg, Claes M. Gustafsson and Robert W. Taylor. *Nature Communications*, 2021, Vol. 12, No. 1, 1135. https://doi.org/10.1038/s41467-021-21279-0

3. 5-Fluorouracil and 6-Thioguanine impair mitochondrial DNA replication.
   Meenakshi Singh, Louise Jenninger, Emily Hoberg, Laleh Arabinian, Xuefeng Zhu, Claes M. Gustafsson and Maria Falkenberg. *Manuscript to be submitted.*

CONTENT

ABBREVIATIONS

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
bp	Base pair
CSB I-III	Conserved sequence blocks I-III
CTD	Carboxy-terminal domain
D-loop	Displacement loop
dNTP	Deoxyribonucleotide
dsDNA	Double-stranded deoxyribonucleic acid
HSP	Heavy-strand promoter
HMG	High mobility group
kDa	Kilo Dalton
LSP	Light-strand promoter
mtDNA	Mitochondrial DNA
mRNA	Messenger RNA
MTS	Mitochondrial targeting sequence
NCR	Non-coding region
NTD	Amino-terminal domain
OriH	Origin of heavy-strand
OriL	Origin of light-strand
OXPHOS	Oxidative phosphorylation
Pi	Inorganic phosphate

PPR	Pentatricopeptide repeat
POLRMT	Mitochondrial RNA Polymerase
rRNA	Ribosomal RNA
ssDNA	Single-stranded DNA
SSB	Single-stranded DNA-binding protein
tRNA	Transfer RNA
TAS	Termination associated sequences
TFAM	Mitochondrial transcription factor A
TFB2M	Mitochondrial transcription factor B2
TEFM	Mitochondrial transcription elongation factor

1 INTRODUCTION

1.1 THE ORIGIN AND FUNCTION OF MITOCHONDRIA

The mitochondrion is a membrane-enclosed organelle found in eukaryotic cells. These structures were first observed in 1850, when Albert von Köllicker described them as grains inside cells (Scheffler, 2008) and in 1898, Carl Benda named them "mitochondria''. Today, we know that mitochondria are of central importance for many aspects of cellular metabolism, including regulation of cell death (Alberts, 2015). Mitochondria are required for cellular energy transformation and produce the majority of adenosine triphosphate (ATP) used in eukaryotic cells. In the process of oxidative phosphorylation (OXPHOS), electrons are transported from nutrients to oxygen, creating water, carbon dioxide, and most importantly ATP. Enormous amounts of ATP are generated and used as an energy currency to drive processes within the cell. The chemical energy stored in ATP is liberated when the molecule is split into adenosine diphosphate (ADP) and inorganic phosphate (Pi) (Berg et al., 2012).

Mitochondria are believed to be descendants of α-proteobacteria, or a closely related lineage, that entered into an endosymbiotic relationship with an archaebacterium. During evolution, the majority of genes present in the original bacterial genome have been lost or transferred to the nucleus (Martin et al., 2015). Over time, a number of evolutionary changes have taken place, and many processes have been integrated between the symbiont and the host cell. The size of the symbiont's genome has been reduced, and genes involved in e.g., amino acid and nucleoside biosynthesis, anerobic glycolysis, and cellular regulation are now encoded by nuclear DNA (nDNA) in humans (Gray et al., 1999).

1.2 STRUCTURE AND DYNAMICS OF MITOCHONDRIA

In the 1950s, with the help of electron microscopy, morphological studies revealed that mitochondria are surrounded by two membranes, an inner membrane and an outer membrane, and that the inner membrane is folded into structures known as cristae. The compartment between the two membranes is referred to as the inter membrane space, whereas the space surrounded by the inner membrane is known as the matrix (Figure 1) (Palade, 1953). The OXPHOS machinery consists of complex I-IV of the electron transport chain (ETC) and ATP synthase, which are all located within the cristae of the inner membrane.

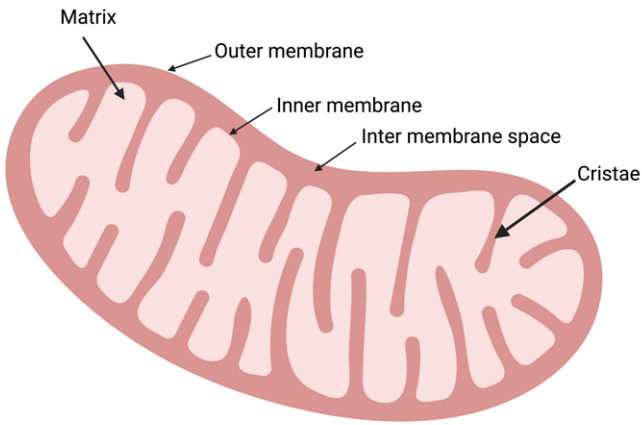

Figure 1. *The basic organization of the mitochondrion.*

The outer membrane separates the mitochondrion from the surrounding cytosol (Palmer & Hall, 1972). Ninety-nine percent of all mitochondrial proteins are nuclear encoded and protein import is therefore a critically important process. The mitochondrial genome only codes for 13 proteins, which are all key components of the OXPHOS machinery (Bolender et al.,

2008). Most mitochondrial proteins are instead coded in the nuclear genome and imported into mitochondria after translation in the cytoplasm (Dolezal et al., 2006). These proteins contain a mitochondrial targeting sequence (MTS), which is removed after import. The MTS varies in size but is about 30 amino acids long.

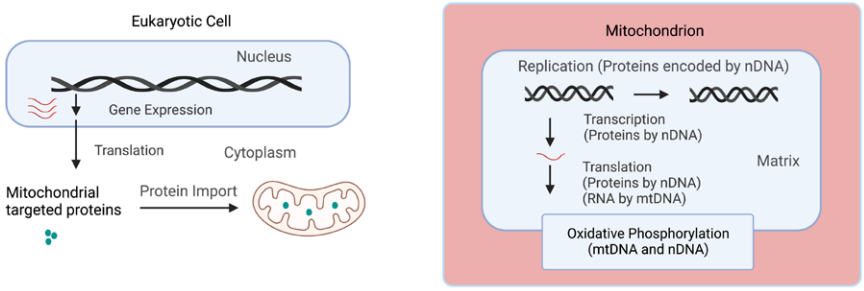

Figure 2. *The mitochondrial proteome has a dual origin. Most proteins are encoded by the nuclear genome (DNA is represented in black and RNA is represented in red). Their translation occurs in the cytoplasm and later they are targeted to mitochondria with a mitochondrial targeting sequence (MTS). When the proteins are imported to mitochondria, the MTS signal is removed. The proteins which are encoded by the human mt genome are membrane embedded subunits of the OXPHOS system. (Figure has been adapted from Viktor Posse, with permission and modified).*

Mitochondria are not static organelles, but part of a dynamic network within the cell that continuously undergoes fission and fusion (van der Bliek et al., 2013). The size of the network varies between cell types, with high mitochondrial density in tissues with a high energy demand, such as liver cells and cardiomyocytes, and smaller mitochondrial networks are in low-energy demanding cell types, e.g., keratinocytes (Alberts, 2015). Fusion of the outer membrane depends on an enzymatic machinery, which in mammals includes two large GTPases (mitofusins 1 and 2 (MFN1-2)) and loss of these enzymes causes excessive mitochondrial fragmentation. Fusion of the inner membrane

depends on the optical atrophy 1 (OPA1) protein, which is a dynamin related GTPase (Cipolat et al., 2006). Loss of OPA1 also causes mitochondrial fragmentation.

Mitochondrial fission is mediated by another set of factors, including a large GTPase known as dynamin-related protein 1 (Drp1). Drp1 is located in the cytosol and gets recruited to the outer mitochondrial membrane. The protein oligomerizes in a spiral-form around the OMM and GTP hydrolysis leads to the constriction of the spiral, which results in fission and leads to the formation of two separate organelles (Loson et al., 2013).

1.3 THE MITOCHONDRIAL GENOME

The first observations suggesting the existence of mtDNA came in the early 1960s from studies of chick embryo mitochondria. Later studies demonstrated that mtDNA was a small, circular, and double-stranded molecule that could exist both as a monomer and as catenated circles. The number of mtDNA molecules varies depending on cell type, ranging from a few hundred to more than 10,000 copies per cell (Bogenhagen & Clayton, 1974). Even if the copy number of mtDNA is high in animal cells, the size of the molecule is very small (16,569 bp in human cells) and mtDNA therefore constitutes less than one percent of the total cellular DNA.

In mammals, mtDNA is strictly maternally inherited (Hutchison et al., 1974). The copy number of mtDNA is very low (50-75) in sperm cells and sperm mitochondria are actively degraded during fertilization. In contrast, oocytes contain extremely high mtDNA levels (>100,000 copies per cell), which are segregated into daughter cells after fertilization (Michaels et al., 1982).

The mitochondrial genome is densely packed with coding information and the organization of genes is highly conserved in mammals (Wolstenholme, 1992). Human mtDNA encodes 37 genes that all lack introns, including 13 protein-coding genes that are transcribed into 11 messenger RNAs (mRNAs) coding for essential components of the OXPHOS system. The genome also contains genes for 2 ribosomal RNA (rRNA) molecules (12S and 16S rRNA), and 22 transfer RNAs (tRNAs) that are essential for mitochondrial protein synthesis. The two strands in mtDNA are referred to as the heavy (H) and light (L) strand. The H-strand contains more guanine and can be separated from the L-strand by cesium chloride alkaline gradient density centrifugation.

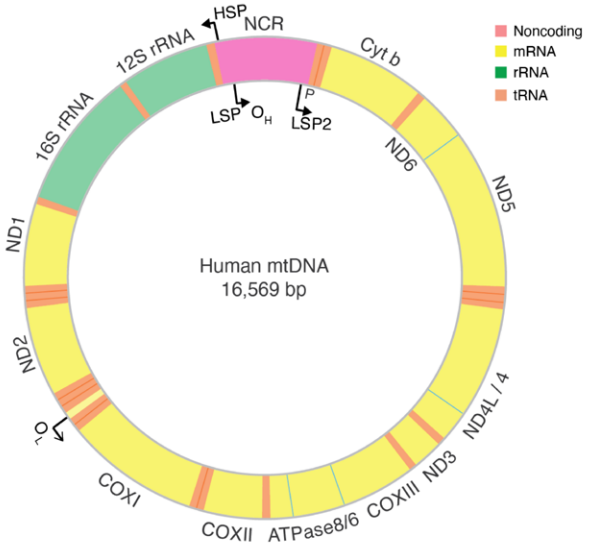

Figure 3. *Illustration of human mitochondrial genome. Indicated ND I to VI which occurs in complex I subunits, Cytb which occurs in complex III subunits, Cox which occurs in complex IV subunits and ATPase which occurs in complex V subunits. tRNAs are represented by orange lines. Replication and transcription sites are indicated as O_H (OriH) and O_L (OriL), and the promoter sites are reported as HSP, LSP and LSP2.*

Human mtDNA contains one long noncoding region (NCR), also known as the control region. The NCR of mammalian mtDNA contains two promoters, one on each strand, referred to as the H-strand and L-strand promoters (HSP and LSP). A third promoter, LSP2 was recently identified at the 3′-end of the NCR, but its physiological role in mitochondrial gene transcription remains to be established (Tan et al., 2022). The genome contains two origins of replication, the origins of H-strand and L-strand replication (OriH and OriL). OriH is located in the NCR, downstream of LSP, whereas OriL is located in a tRNA cluster, approximately 11,000 bp downstream of the NCR (Falkenberg, 2018).

D-Loop

In the majority of cases, mtDNA replication initiated at OriH is terminated already after about 650 nt, which leads to the formation of 7S DNA. The newly formed 7S DNA remains associated with the NCR and forms a triple-stranded displacement loop (D-loop) structure (Arnberg et al., 1971).

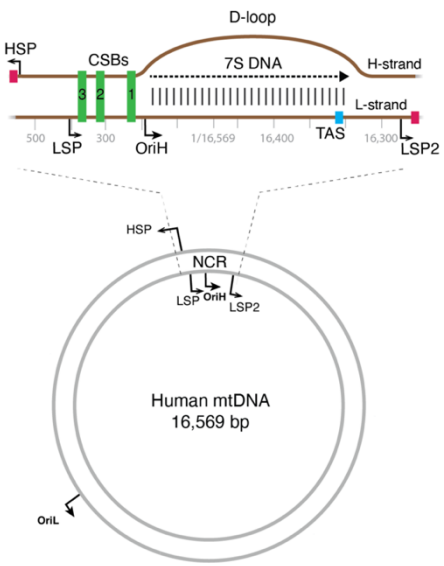

Figure 4. *Schematic representation of the non-coding region of mitochondrial DNA. The positions of the D-loop, 7S DNA and TAS region are indicated, Conserved sequence blocks I, II and III are represented in green colour. tRNAs are indicated in red. Mitochondrial promoter sites HSP, LSP and LSP2 are indicated. Colour coding is different from figure 2 for representing the D-loop in detail.*

How replication is terminated to form 7S DNA is not known, but termination coincides with a number of conserved sequence motifs, the so called "termination associated sequences" (TAS), located in the 3′-region of the NCR. The choice between 7S DNA synthesis and genome length mtDNA replication constitutes an important regulatory step, which helps to control mtDNA levels, but how this step is regulated is not well understood (Jemt et al., 2015).

Mitochondrial nucleoid

The mitochondrial genome is packaged together with proteins into compact nucleoprotein complexes, known as nucleoids. The average size of a mammalian nucleoid is about 100 nm in diameter and in most cases, a nucleoid contains only one single copy of mtDNA (Kukat et al., 2011). The most prominent nucleoid component is mitochondrial transcription factor A (TFAM), which is an abundant DNA-binding protein with the ability to bind non-sequence-specifically to DNA. TFAM can fully coat mtDNA and form bridges between two neighboring DNA duplexes (Kukat et al., 2015). In combination with the ability to induce DNA flexibility by local denaturation and bending, these effects allow TFAM to compact mtDNA (Farge et al., 2012; Fisher et al., 1992).

2 TRANSCRIPTION

A short introduction

In transcription, DNA is used as a template to form a complementary single-stranded RNA molecule. The process is performed by an RNA polymerase, which synthesizes the RNA strand in the 5′ to 3′ direction. The resulting RNA has the same sequence as that of the non-template strand of DNA (except thymine which is replaced by uracil (Alberts, 2015; Berg et al., 2012). The mechanisms of transcription have been studied in many different systems, but most relevant for our understanding of the process in mitochondria is previous work on transcription in bacteria and bacteriophages.

2.1 BACTERIAL TRANSCRIPTION

Transcription can be divided into three main steps: initiation, elongation, and termination. In *Escherichia coli*, the core RNA polymerase is composed of four different subunits (α (two copies), β, β', ω). A fifth subunit, known as the sigma factor (σ) helps the RNA polymerase to recognize promoters and initiate transcription. The core polymerase together with the sigma factor forms the holoenzyme. Bacteria produces a number of different sigma factors, which recognize and binds to different promoters. By expressing different sigma factors, the cell can change gene transcription patterns in response to environmental conditions. Sigma factor binds to two sequence motifs which are found upstream of the transcription start site at position -10 and at -35, with the first transcribed nucleotide denoted as +1.

Once transcription is initiated, the σ factor is released from the RNA polymerase holoenzyme and elongation begins. Transcription continues until the RNA polymerase encounters a transcription termination signal in the DNA template. There are two canonical ways to terminate transcription in bacteria,

rho dependent and rho independent termination. In the case of rho dependent termination, the nascent RNA strand contains an unstructured region rich in cytosine, which is recognized and bound by the rho protein. Rho is a ring-shaped, hexameric helicase that associates with and translocates along the nascent RNA strand. Once the RNA polymerase reaches a transcription pause site, the rho helicase dissociates the transcription machinery from the DNA template (Alberts, 2008; Berg et al., 2012).

In rho independent termination, a palindromic sequence followed by a stretch of adenine nucleotides in the template strand helps to terminate transcription. When the palindrome sequence is transcribed, nascent RNA forms a hairpin loop-like structure that disturbs its interactions with RNA polymerase. At the same time, the adenine stretch in the template strand is transcribed, which due to weak base interactions with the nascent RNA strand destabilizes the DNA-RNA hybrid. Together, these two effects cause bacterial RNA polymerase to dissociate, and transcription is terminated (Bebenek & Ziuzia-Graczyk, 2018).

2.2 NUCLEAR TRANSCRIPTION

The nuclear genome in eukaryotes is transcribed by three major RNA polymerases: RNA polymerase I, II, and III (Pol I-III), which are similar in structure and composition. Pol I transcribes rRNA genes, Pol II transcribes protein coding mRNA genes, whereas Pol III transcribes tRNA genes (Berg et al., 2012). As in bacteria, these RNA polymerases require the assistance of additional transcription factors to initiate transcription. Assembly of the Pol II preinitiation complex at promoters takes place in a stepwise fashion and involves at least six basal transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH). The first factor to bind promoters is TFIID, which recognizes the TATA-box located upstream of the transcription start site. The last factors to enter the complex are TFIIE and TFIIH. TFIIH has a protein

kinase activity that phosphorylates Pol II and stimulates the transition to active transcription (Vannini & Cramer, 2012). In addition to the basal transcription machinery, nuclear transcription is carefully regulated by a large number of gene specific factors that can activate or repress transcription from promoters. Also of importance is the epigenetic regulation of gene activity, which is conferred by histone packaging of nuclear DNA into nucleosomes and higher order chromatin structures. In epigenetic regulation of gene activity, posttranslational modifications and chromatin remodeling enzymes can regulate packaging of DNA and thereby govern RNA polymerase access and gene activity. As we will see, nuclear transcription factors are distinctly different from those found in mitochondria.

2.3 T7 BACTERIOPHAGE TRANSCRIPTION

Bacteriophage T7 RNA polymerase is a single-subunit protein of 99 kDa, which does not require additional factors for transcription initiation and elongation. The enzyme was first isolated in 1969 from infected *E. coli* cells and its primary sequence was determined in the 1980s. Later studies have provided detailed structural information about the T7 RNA polymerase, that also revealed significant structural similarities to DNA polymerases (Cheetham & Steitz, 1999, 2000; Sousa et al., 1993). In fact, T7 RNA polymerase is homologous to the DNA polymerase I family but shares few sequence similarities with other classes of RNA polymerases (Steitz et al., 1994).

The consensus promoter sequence recognized by T7 RNA polymerase contains 23 conserved bases which span a region around the transcription start site (position -17 to +6). The AT-rich recognition loop and the specificity loop are the two structural elements located at the N-terminus of the protein that are used by the polymerase to recognize and bind promoters (Dunn & Studier,

1983; Oakley et al., 1979). A third element, the intercalating hairpin is located in the C-terminal catalytic part of the protein. This later structure helps to separate the two DNA strands at position -4 and thereby promotes DNA melting, which leads to the formation of an open complex.

DNA polymerase I from the family of DNA polymerases has a steric gate at the active site that helps to discriminate deoxyribonucleotides (dNTPs) from ribonucleotides (NTPs), thus preventing misincorporation of NTPs into DNA. This effect depends on a glutamic acid side chain which sterically blocks nucleotides with a 2′-OH group (Cheetham et al., 1999; Cheetham & Steitz, 2000). In the T7 RNA polymerase, the glutamic acid has been replaced by glycine and a nearby histidine residue can form a hydrogen bond with the 2′-OH group. These structural changes make incorporation of ribonucleotides more energetically favorable as compared to the dNTPs and allows the enzyme to function as an RNA polymerase.

Once a promoter-T7 RNA polymerase complex forms, transcription is initiated by formation of the first phosphodiester bond between the +1 and +2 nucleotides (Cheetham & Steitz, 1999). Transcription initiation does not always lead to full-length transcripts. Instead, T7 RNA polymerase-dependent transcription is frequently aborted after only producing short RNA stretches between 2 to 8 nucleotides in length. Once the T7 RNA polymerases continues transcription and passes the eighth nucleotide it escapes the promoter and enters into the elongation phase. This transition to transcription elongation is linked to substantial confirmational changes in the N-terminal part of the protein (Durniak et al., 2008; Martin et al., 1988).

Transcription termination depends on rho dependent and rho independent termination mechanisms, identical to those in terminating transcription of host bacterial genes.

2.4 MITOCHONDRIAL TRANSCRIPTION

Transcription in human mitochondria depends on four proteins: the mitochondrial RNA polymerase (POLRMT); TFAM; mitochondrial transcription factor B2 (TFB2M); and mitochondrial transcription elongation factor (TEFM) (Gustafsson et al., 2016). Together these four proteins ensure correct initiation and elongation from mitochondrial promoters.

POLRMT

POLRMT belongs to the same family of single-subunit RNA polymerases as the bacteriophage T7 RNA polymerase. In contrast to its phage counterpart, POLRMT cannot recognize promoter sequences on its own and requires two additional transcription factors, TFAM and TFB2M, to initiate transcription (Falkenberg et al., 2002; Litonin et al., 2010; Shi et al., 2012).

The first mitochondrial RNA polymerase (Rpo41) was identified in *Saccharomyces cerevisiae* and later the mammalian counterpart was identified based on sequence similarities (Tiranti et al., 1997). In human cells, POLRMT is a 1230 amino acids long protein with a molecular weight of approximately 140 kDa. POLRMT contains an MTS at its N-terminus, spanning amino acids 1- 41 that is cleaved upon import into the mitochondria. The protein can be divided into a N-terminal extension (NTE) spanning amino acids 42 to 368, a N-terminal domain (NTD) spanning amino acids 368 to 647, and a C-terminal domain (CTD), spanning amino acids 648 to 1230. The NTE region is absent in bacteriophage polymerases, suggesting that it has evolved for mtDNA transcription in eukaryotic cells. An interesting feature of the NTE is the existence of pentatricopeptide repeats (PPRs). PPR domains are commonly present in a wide range of RNA-binding proteins and are needed in diverse RNA-related processes, such as RNA splicing, stability, and translation. The

precise role of the PPR domain in POLRMT has not been established, but newly synthesized RNA molecules exit towards the PPR domain (Schwinghammer et al., 2013). The NTE also contains the so called "tether helix" which interacts with TFAM during recruitment of POLRMT to promoters and helps to bring POLRMT's active site close to the transcription start site.

The NTD of POLRMT contains two loops corresponding to functional domains found in bacteriophage polymerases. The first is the AT-rich recognition loop, which binds to promoter elements during initiation. The second is the intercalating hairpin, which enables both POLRMT and the T7 RNA polymerase to melt DNA during transcription initiation. Structural studies of elongating POLRMT suggests that the intercalating hairpin also helps POLRMT to separate RNA from DNA during transcription elongation, a function not associated with the corresponding structure in T7 RNA polymerase (Hillen et al., 2018). The CTD contains the catalytic domain, which is also conserved in the T7 RNA polymerase. Similar to other members of the family A polymerases, this part of the protein adopts a right-hand fold, with three subdomains (the thumb, palm, and fingers) flanking the active center (Masters et al., 1987; Ringel et al., 2011; Tiranti et al., 1997).

TFAM

Mitochondrial transcription factor A was first discovered as Abf2, an abundant DNA binding-protein in budding yeast, primarily located to mitochondria (Diffley & Stillman, 1991). In contrast to its homologue in human cells, TFAM, the Abf2 protein is not required for initiation of transcription but is only involved in packaging of mtDNA (Xu & Clayton, 1992).

Human TFAM is a protein of 244 amino acids with a molecular weight of approximately 25 kDa. The protein contains an MTS at its N-terminus, spanning amino acids 1-42, which is cleaved off during import. TFAM contains two HMG boxes that are connected via a linker region. Similar to other HMG-box domain family members, TFAM can bind, bend, and partially unwind duplex DNA in a non-sequence-specific manner. In addition to its role in nucleoid formation, TFAM can also recognize specific sequence elements upstream of mitochondrial promoters, which is essential for promoter-specific transcription initiation (Fisher & Clayton, 1985; Shi et al., 2012). Binding of TFAM forces DNA to undergo a sharp U-turn, observed both with LSP and HSP containing DNA constructs (Ngo et al., 2011). The precise positioning of TFAM relative to the transcription start site is important for activation and changes in this distance can completely abolish promoter activity (Dairaghi et al., 1995a, 1995b). In mammalian cells, TFAM and POLRMT are both required to recognize specific promoter elements. However, transcription initiation can take place with only POLRMT under conditions that allow for extensive DNA unwinding, e.g., if transcription assays are performed using a negatively supercoiled template or at very low ionic strength (Gaspari et al., 2004; Shi et al., 2012).

TFB2M

TFB2M was identified based on sequence similarities to Mtf1, a mitochondrial transcription factor originally identified in yeast (Falkenberg et al., 2002; Schinkel et al., 1987). In mammalian cells, there is also a second Mtf1 homologue, TFB1M, and the two proteins are related to a large family of rRNA methyltransferases. TFB1M likely represents the original methyltransferase, whereas TFB2M resulted from a gene duplication event and only later evolved into a mitochondrial transcription factor (Shutt & Gray, 2006b). Later studies

of TFB1M, including a mouse knock-out model, have revealed that the protein is not a transcription factor *in vivo*, but is required for mitochondrial 12S rRNA adenine dimethylation (Metodiev et al., 2009).

TFB2M is 45 kDa protein consisting of 396 amino acids. TFB2M is absolutely required for initiation of transcription and acts to promote promoter melting (Falkenberg et al., 2002). Structural studies have revealed that the basic surface groove that binds RNA in structurally related rRNA methyltransferases has been repurposed in TFB2M to stabilize the open promoter complex during initiation of transcription.

TEFM

TEFM was originally identified based on similarities with transcription elongation factors identified in other systems. The protein interacts directly with POLRMT and helps to produce long transcripts. TEFM forms a ‘‘sliding clamp’’ which interacts with the DNA template downstream of POLRMT, which stabilize polymerase-template interactions, thereby increasing processivity. In support of these notions, a conditional knock out mice model revealed a dramatic down regulation of promoter-distal transcripts in *Tefm*-knockout hearts. This work also demonstrated that TEFM is required for the development of embryonic stages and that loss of TEFM causes mitochondrial cardiomyopathy with severe OXPHOS deficiency (Jiang et al., 2019).

In vitro transcription initiated from LSP is often terminated prematurely at conserved sequence block II (CSBII), a sequence element rich in guanines and located about 100 bp downstream of the promoter. Transcription termination is caused by formation of a G-quadruplex structure, that involves both the non-template DNA strand and the nascent RNA (Wanrooij et al., 2012). TEFM prevents transcription termination (Posse et al., 2015), by binding close to the

RNA exit channel of POLRMT and blocking the formation of G-quadruplex structure (Hillen, Parshin, et al., 2017)

Transcription termination at CSBII is linked to formation of primers required for initiation of mtDNA replication at OriH (Posse et al., 2019). *In vitro* work suggested that TEFM could influence this process and down-regulate levels of primer formation (Agaronyan et al., 2015), but characterization of transcription intermediates in mouse, formed in tissues lacking Tefm, refuted this idea and instead identified the protein as a transcription elongation factor needed for genome length transcription and the link to RNA processing (Jiang et al., 2019).

The sequential model of mitochondrial transcription

A high affinity binding site for TFAM is situated upstream of the transcription start site (position -15 to -35) at all three human promoters: HSP, LSP, and LSP2 (Fisher et al., 1987; Tan et al., 2022). Once bound, TFAM introduces a 180° bend in the DNA (Ngo et al., 2011; Rubio-Cosials et al., 2011) and recruits POLRMT, via direct interaction with a “tether” helix located in the NTE of the polymerase.

Next, TFB2M binds to POLRMT, which completes preinitiation complex formation. In the absence of TFB2M, POLRMT interactions with promoter DNA are restricted to sequences around position −50 to −60. Once TFB2M enters the complex, there is a pronounced structural change around the transcription start site (position -10 to +10) and the first phosphodiester bond can be formed (Posse & Gustafsson, 2017). Structural studies have shown that TFB2M induces structural changes in POLRMT, which helps to open duplex DNA at the promoter (Hillen, Morozov, et al., 2017).

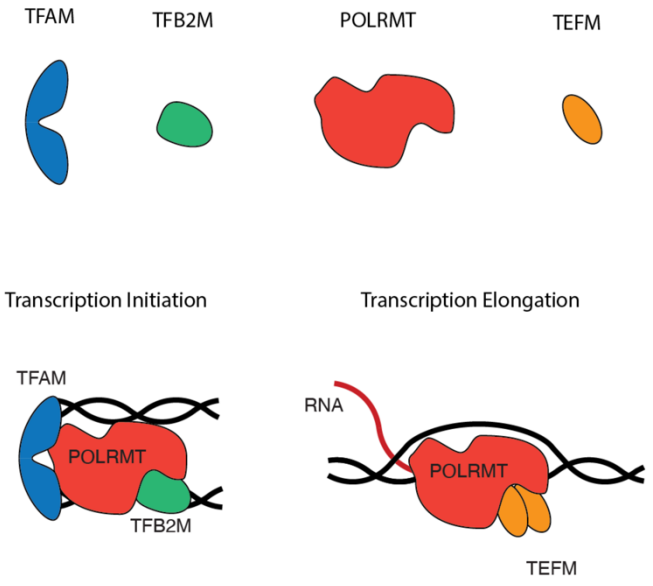

Figure 5. *The schematic diagram represents components involved in transcription initiation and elongation - TFAM is in blue, TFB2M is in green, POLRMT is in red and TEFM is a dimer in orange. The DNA is colour coded in black and RNA in red. (The figure has been adapted from Viktor Posse with permission and modified).*

After initiation of transcription, TFAM and TFB2M are no longer needed and are left behind as POLRMT initiates transcription elongation. At this stage, POLRMT instead associates with TEFM, which provides high processivity to the elongating polymerase, by stabilizing interactions between POLRMT and the template DNA (Agaronyan et al., 2015; Gaspari et al., 2004; Morozov et al., 2015; Posse et al., 2015). Mitochondrial transcription is polycistronic and once initiated from mitochondrial promoters, POLRMT continues to nearly genome-length transcription. Termination of L-strand transcription occurs just before the rRNA gene cluster and is directed by mitochondrial transcription termination factor 1 (MTERF1) (Shi et al., 2016; Terzioglu et al., 2013). Transcription from HSP is terminated in the TAS region, which is located near

the 3′-end of the D-loop, but the mechanisms of this process have not been established (Jemt et al., 2015).

MTERF1

MTERF1 was originally identified as a sequence-specific DNA-binding protein that associates with mtDNA at the 3′-end of the 16S rRNA gene and mediates transcription termination. The protein was initially believed to terminate transcription coming from the HSP direction, but later work has shown that it blocked transcription on the opposite strand (Asin-Cayuela et al., 2005). Based on analysis of transcription intermediates in *Tefm* knockout mice, the protein is now believed to prevent transcription from continuing over the rRNA transcription unit and to interfere with new rounds of LSP transcription. Recent evidence has also suggested that the protein can affect replication pausing, which may point to a role unrelated to transcription (Shi et al., 2016; Terzioglu et al., 2013).

3 DNA REPLICATION

A brief introduction

During DNA replication, two identical copies are produced from one original DNA molecule. The process is semiconservative, since one of the parental strands remains conserved in each of the two newly formed daughter molecules. The general principles of this process are similar in bacteria, archaea, and eukaryotes (Alberts, 2015; Berg et al., 2012). During DNA synthesis, a parental strand is used as a template by DNA polymerase to produce a nascent DNA strand in the 5′ to 3′ direction. Many DNA polymerases also have a 3′ to 5′ exonuclease activity, which helps to proofread and remove mis-incorporated nucleotides during active DNA synthesis (Alberts, 2008).

DNA polymerases are unable to start DNA synthesis *de novo*, but need a primer bound to the template with a free OH-group at the 3′-end. In most systems, these primers are produced by a class of enzymes known as primases, which are a type of RNA polymerase. Most DNA polymerases lack strand displacement activity and to work on a double-stranded DNA template they hence require a DNA helicase to unwind the duplex in an ATP-dependent process (Berg et al., 2012). A single-stranded DNA-binding protein (SSB) is required to stabilize unwound regions of DNA and to stimulate DNA replication, whereas DNA polymerase processivity factors help to anchor the moving enzyme to template DNA. There is also a need for topoisomerases that prevent and correct topological problems. Together with additional replication factors, these proteins form the enzymatic machinery, the replisome, which functions at the replication fork to ensure effective strand-separation and active DNA synthesis (Alberts, 2015; Berg et al., 2012). Because of the antiparallel nature of the double stranded DNA, only one strand (the leading strand) is

synthesized in a continuous manner at the replication fork. The second strand (the lagging strand) is made in small pieces (Okazaki fragments), each synthesized in the 5′ to 3′ direction and later ligated together by a DNA ligase (Alberts, 2015; Berg et al., 2012).

3.1 T7 BACTERIOPHAGE REPLICATION MACHINERY

The replication machinery of T7 bacteriophage is comparatively simple. The T7 replisome can be reconstituted with only four factors: T7 gp5 (DNA polymerase); *E. coli* thioredoxin (trx) (processivity factor), T7 gp2.5 (SSB protein), and T7 gp4 (helicase/primase) (Hamdan & Richardson, 2009). The T7 gp4 protein is a dual function protein that harbors both primase and helicase activities. The protein assembles onto DNA as a hexamer and unwinds the duplex in front of the gp5 protein. The primase activity is used to synthesize short RNA primers required by gp5 to initiate DNA synthesis (Crampton et al., 2006; Lee et al., 2006). The trx processivity factor interacts directly with T7 gp5 and stimulates processivity, enabling polymerization of very long stretches of DNA (Kato et al., 2003). The T7 gp2.5 protein binds to single-stranded DNA (ssDNA) regions to prevent reannealing. The protein also increases the primase activity gp4, thereby stimulating initiation of Okazaki fragment synthesis, needed for coordinated formation of leading and lagging strand DNA at the bacteriophage replication fork (Crampton et al., 2006; Hamdan et al., 2009; Kato et al., 2003; Lee et al., 2006).

3.2 BACTERIAL REPLICATION MACHINERY

E. coli has a circular, double-stranded DNA genome of about 4.6 Mbp and replication is initiated from a unique site, the bacterial origin of replication, oriC. From this site, DNA replication proceeds bi-directionally until the two

replication forks have progressed around the circular genome to meet at a specific termination sequence.

Initiation of DNA replication in *E. coli* starts when the DnaA proteins binds to oriC in an ATP-dependent process. DNA is wrapped around a core of multiple DnaA proteins and a region of oriC is unwound to expose ssDNA. At this point DnaB (DNA helicase) and DnaC (helicase loader) are directed to oriC and assist in further unwinding of the origin. In the next phase, DnaG (primase) and the *E. coli* SSB protein are recruited to oriC, followed by DNA polymerase III. At this point DnaA dissociation is triggered and DNA replication begins (Alberts, 2008; Berg et al., 2012; Costa et al., 2013).

3.3 MITOCHONDRIAL DNA REPLICATION

In mammalian cells, mtDNA replication is performed by a number of distinct replication factors that are related to the ones used for DNA replication in bacteria and bacteriophages (Shutt & Gray, 2006a). The core of this process is DNA polymerase γ (POLγ), which performs both H- and L-strand DNA synthesis. The polymerase is assisted by the TWINKLE DNA helicase, which unwinds dsDNA at the replication fork. Primers needed for initiation of mtDNA synthesis are formed by POLRMT, which thus also functions as a primase in mitochondria. Mitochondria also contains a dedicated, ssDNA binding protein (mtSSB), which blocks secondary structure formation and prevents random primer synthesis on ssDNA exposed during mtDNA synthesis.

3.4 MODELS REPPRTED FOR MITOCHONDRIAL DNA REPLICATION

According to the strand-displacement model of mtDNA replication, both L-strand and H-strand DNA synthesis are performed in a continuous manner, and

no Okazaki fragments are produced. DNA replication is first initiated at OriH, and H-strand DNA synthesis proceeds without simultaneous synthesis of a nascent L-strand. During this first phase, the parental H-strand is gradually exposed in a ssDNA conformation and covered by mtSSB (Clayton, 1982; Robberson et al., 1972; Tapper & Clayton, 1981). Once H-strand DNA synthesis has progressed about two-thirds of genome, the replication machinery passes OriL and the displaced parental H-strand forms a stem-loop structure at OriL. The stem blocks mtSSB binding, which leaves the loop region available for interactions with POLRMT, which starts to synthesize a primer. After approximately 25 nucleotides (nt), POLRMT leaves the template and POLγ uses the primer formed to initiate L-strand DNA synthesis. After initiation at OriL, the two strands continue to form in a continuous manner, until they both reach full circle. At this point, two complete daughter molecules have been formed, which are separated by topoisomerase 3α (TOP3A) (Fuste et al., 2010; Nicholls et al., 2018; Wanrooij et al., 2008).

Other, alternative modes of mtDNA replication have been suggested. The RITOLS model (ribonucleotide incorporation throughout the lagging strand) is related to the strand-displacement-model, with one important exception. Instead of mtSSB, processed RNA molecules are used to cover the displaced, parental H-strand during synthesis of the L-strand (Yasukawa et al., 2006).

3.5 THE MITOCHONDRIAL MACHINERY – REPLICATION ASSOCIATED PROTEINS AND THEIR FUNCTIONS

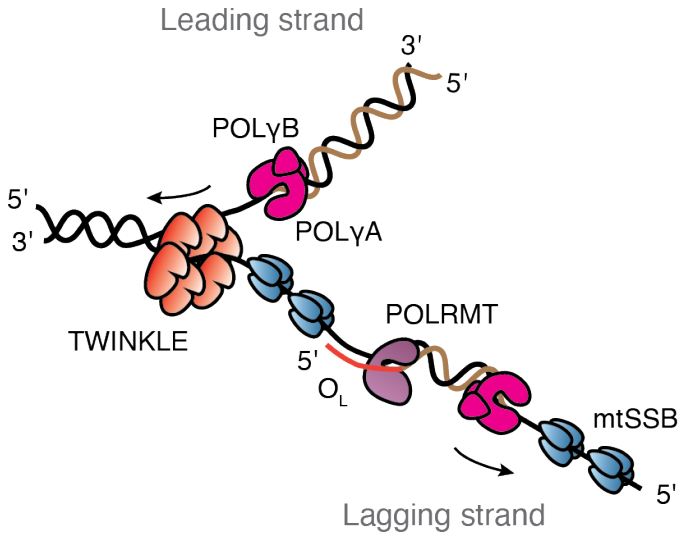

Figure 6. *The mitochondrial core replisome. Polγ (PolγA and PolγB subunits) represented in pink synthesizes DNA, mtSSB marked in blue binds to the single stranded DNA and TWINKLE helicase represented in light orange color unwinds the double stranded DNA ahead of Polγ. POLRMT shown in purple provides the RNA primer. (The figure has been adapted from Jay P. Uhler with permission and modified).*

DNA POLYMERASE γ

POLγ is a heterotrimeric protein complex composed of one catalytic POLγA subunit and two accessory POLγB subunits (Yakubovskaya et al., 2006). The molecular weight of POLγA and B are approximately 140 kDa and 55 kDa, respectively. In addition to its 5′ to 3′ polymerase activity, POLγA also harbors a 3′ to 5′ exonuclease domain used to remove misincorporated nucleotides during DNA replication. The exonuclease activity reduces the mutation rate

and is also needed for the formation of ligatable ends during mtDNA synthesis (Macao et al., 2015). POLγB is structurally related to amino-acyl tRNA synthetases and acts to enhance catalytic activity and processivity, via stabilization of POLγ interactions with template DNA (Lee et al., 2010).

TWINKLE

TWINKLE is a hexameric DNA helicase with a molecular mass of 420 kDa, which unwinds duplex DNA in the 5′ to 3′-direction (Korhonen et al., 2003). TWINKLE was first identified based on sequence similarity to the T7 gp4 protein in a screen for mutations causing mitochondrial disease (autosomal dominant progressive external ophthalmoplegia), but in contrast to T7 gp4, the primase domain is non-functional (Spelbrink et al., 2001).

mtSSB

Human mtSSB has a molecular weight of 15 kDa and forms homotetramers in solution, around which ssDNA is wrapped. The protein displays sequence similarity to *E. coli* SSB and stimulates both POLγ and TWINKLE activities (Li & Williams, 1997). mtSSB not only stimulates replication elongation, but also plays an essential role during initiation of replication at OriH. Loss of mtSSB causes an upregulation of LSP transcription, but abolishes primer formation (Posse et al., 2019). In part this effect could be due to mtSSB's ability to stabilize R-loops formed during primer synthesis and termination at CSBII (Jiang et al., 2021). mtSSB also prevents random transcription initiation by POLRMT on exposed ssDNA, which helps to restrict primer formation to OriH and OriL (Fuste et al., 2010; Posse et al., 2019).

MGME1

MGME1 belongs to the RecB family of nucleases. The enzyme prefers ssDNA and cleaves DNA with free ends. Disease-causing mutations that abolish MGME1 activity cause mtDNA depletion and various genomic rearrangements, including linear deletions (Kornblum et al., 2013; Nicholls et al., 2014). Biochemically, MGME1 processes DNA flaps formed during termination of DNA replication into ligatable nicks (Uhler et al., 2016).

RNASEH1

Ribonuclease H1 (RNASEH1) which belongs to the RNase H family of endonucleases, cleaves RNA that is hybridized to DNA with limited sequence specificity (Cerritelli & Crouch, 2009). Disease causing mutations in the *RNASEH1* gene impair primer removal from nascent DNA strands, a process required for correct termination of mtDNA replication (Reyes et al., 2015)**.** In support of these observations, knockout of the *Rnaseh1* gene in mouse causes an embryonically lethal phenotype and depletion of mtDNA levels. Analysis of mouse tissues also revealed that Rnaseh1 is needed for primer processing during replication termination and revealed impaired origin-specific initiation at OriH (Misic et al., 2022). In agreement with these findings, analysis of mouse embryonic fibroblasts lacking Rnaseh1 activity showed retention of RNA primers at the 5′ end of 7S DNA and at nascent DNA fragments initiated from OriL, supporting an important role for RNASEH1 in primer processing (Akman et al., 2016; Holmes et al., 2015).

The role for RNASEH1 in origin specific initiation of mtDNA replication was originally demonstrated by biochemical studies *in vitro* and a detailed analysis of replication intermediates in human fibroblasts with low RNASEH1 activity. This work revealed that RNASEH1 is required to process nascent R-loops

formed at CSBII, before LSP-transcripts can be used to prime OriH-specific initiation of mtDNA replication (Posse et al., 2019).

TOP3A

Topoisomerase 3α (TOP3A) helps to separate the two interconnected daughter molecules after completion of mtDNA replication. The enzyme belongs to the type 1A family of topoisomerases and disease-causing mutations in this gene lead to the accumulation of large catenated mtDNA molecules and multiple deletions. TOP3A also has a nuclear isoform, where it acts to resolve double Holliday junctions (Nicholls et al., 2018; Pommier et al., 2016).

4 MITOCHONDRIAL DNA MAINTENANCE DISORDERS

Mutations that cause defects in OXPHOS system biogenesis and other aspects of mitochondrial function can lead to diseases (Copeland, 2012; Reinecke et al., 2009). Even if individual mitochondrial diseases are rare, as a group they are surprisingly common. Studies from the UK suggest that the prevalence of mitochondrial disease in adults, due to mutations in either the mitochondrial or nuclear genome, is about 1 in 4,300 (Gorman et al., 2015)**.** The hallmark of these diseases is a lack of function of the OXPHOS machinery with consequent deficiency in ATP synthesis. The clinical picture is highly variable, but as a general rule, the mitochondrial diseases are progressive with symptoms in the muscle and nervous systems. Even if mitochondrial dysfunction is a relatively common cause of human disease, there is almost no effective treatment available.

A large group of mitochondrial diseases is caused by mutations in nuclear genes encoding factors needed for proper mtDNA maintenance. The most common nuclear cause of mitochondrial disease is mutations affecting POLγA (Rahman & Copeland, 2019), but mutations have also been identified in genes coding for other replication factors, including POLγB, TWINKLE, and mtSSB (Gustafson et al., 2020). These mutations impair mtDNA synthesis and cause effects such as mtDNA depletion, deletions, and duplications. Disease causing mutations can also disturb expression of mtDNA, e.g., by reducing mitochondrial transcription or impairing translation. As reported in Paper II in this thesis, disease causing mutations associated with the *POLRMT* gene reduces mitochondrial transcription, which leads to impaired OXPHOS system biogenesis and reduced respiration (Olahova et al., 2021). Recently, related

phenotypes have also been reported for patients with mutations in the *TEFM* gene (Van Haute et al., 2023).

Disease-causing mutations are also associated with genes needed to maintain the nucleotide pools. They are also required to produce and regulate dNTPs at proper levels (El-Hattab et al., 2017). Mutations that reduce dNTP levels can impair POLγ function and cause secondary effects on mtDNA. Changes in nucleotide pools can also increase the risk of incorporating higher levels of ribonucleotides (rNTPs) into mtDNA, which in turn can disturb later rounds of mtDNA replication and transcription (Wanrooij & Chabes, 2019). How embedded rNTPs affect mitochondrial transcription is investigated in Paper I of this thesis.

4.1 RIBONUCLEOTIDES IN MITOCHONDRIAL DNA

Already 50 years ago, the two strands of mtDNA were shown to contain unusually high levels of embedded ribonucleotides (rNMPs), and later work, demonstrated that the presence of these rNMPs were caused by misincorporation during mtDNA replication (Berglund et al., 2017; Wanrooij et al., 2017). rNMPs are also incorporated during nuclear DNA synthesis but in the nucleus, there are specialized repair pathways that remove these erroneously incorporated nucleotides and replace them using the proper dNTPs as substrates.

Removal of rNMPs from nuclear DNA depends on RNase H2 that nicks DNA on the 5′-side of an embedded rNMP. A part of the cleaved strand is next removed by nick translation synthesis, initiated from the 3′-end that is formed at the nick. The DNA flap created during translation synthesis is next processed by the flap endonuclease FEN1 or the exonuclease EXO1 and the nick is sealed by ligation (Lujan et al., 2013; Sparks et al., 2012). In this

manner, the embedded rNMP is replaced by a correct dNMP. In the absence of RNASEH2, a second RER also exists. This system depends on topoisomerase 1 (TOP1) and is more error prone.

The relative levels of rNMPs embedded in mtDNA depends on two factors. First, POLγ's ability to discriminate between nucleotides during DNA synthesis differs between the four ribonucleotides. Second, the levels of incorporation are dependent on the size of individual dNTP pools. In support of this notion, mutations that change nucleotide pools also cause changes in the relative levels of ribonucleotides incorporated, which has been seen in patient-derived fibroblasts with mutations in important enzymes needed for proper nucleotide pool maintenance, including thymidine kinase, deoxyguanosine kinase, and MPV17 (Berglund et al., 2017; Forslund et al., 2018; Kasiviswanathan & Copeland, 2011). This effect is also seen in yeast strains with mutations that alter their dNTP pools. Careful analysis of mtDNA from these strains demonstrated a direct correlation between dNTP concentrations and the risk of incorporating the corresponding rNMP into mtDNA (Wanrooij et al., 2017).

4.2 MUTATIONS IN MITOCHONDRIAL DNA AND AGEING

Mutations in mtDNA accumulate during the natural process of ageing. Already 70 years ago, studies suggested that the production of free radicals could have a deleterious effect on biological molecules and contribute to ageing (Harman, 1956). Later, the model was further refined, leading to the "mitochondrial free radical theory", which postulates that mtDNA mutations are the primary event in the biological ageing process (Miquel et al., 1980). Experimental support for the role of mtDNA mutations in mammalian ageing has come from studies of knock-in mice that expressed a version of PolγA that lacked exonuclease

activity and thus was unable to proofread during mtDNA synthesis. These mice displayed an increase in somatic mtDNA mutations, which was associated with shorter lifespan and a wide array of premature ageing phenotypes (Greaves et al., 2012; Trifunovic et al., 2004). Accumulation of mtDNA mutations with age has also been associated with age-related disorders, e.g., an accumulation of mtDNA deletions in the neurons of Parkinson's patients and point mutations in the hippocampus of Alzheimer's patients (Bender et al., 2006; Hoekstra et al., 2016). Depletion and mutations affecting mtDNA have also been linked to sarcopenia, heart failure, neurodegeneration, and childhood cancer (Chinnery et al., 2002; Hiona & Leeuwenburgh, 2008).

Later work has questioned the mitochondrial free radical theory of ageing and suggests that mtDNA mutations that accumulate during ageing are due to mtDNA replication errors rather than oxidative damage (Lagouge & Larsson, 2013).

4.3 NUCELOSIDE ANALOGUES AND THEIR EFFECTS

Nucleoside analogues are commonly used in the treatment of human diseases, including cancer and some viral infections (Zhu et al., 2000). These compounds are phosphorylated within the cell to form the corresponding triphosphates that may interact with different cellular targets and perturb many cellular processes ("Late Effects of Treatment for Childhood Cancer (PDQ(R)): Health Professional Version," 2002). Importantly, nucleoside analogues can be used as building blocks during DNA replication, causing termination of the nascent DNA chain, thereby blocking e.g., virus replication or cancer cell growth.

Some nucleoside analogues disturb mtDNA replication, resulting in mtDNA depletion and deletions. For instance, HIV patients treated with nucleoside

analogues progressively accumulate mtDNA mutations similar to those observed during normal ageing (Payne et al., 2011). This observation not only supports the notion that mtDNA mutations contribute to ageing, but also indicates that nucleoside analogues can cause serious, adverse side effects, long after treatment has been terminated.

Nucleoside analogues are also used to treat childhood cancer patients. In Paper III of this thesis, we explore how two of these compounds, 6-Thioguanine (6-TG) and 5-Flourouracil (5-FU), affect mtDNA replication. Our work is an extension of previous cellular work that demonstrated that 6-TG can be incorporated into mtDNA, which can cause mitochondrial dysfunction, eventually leading to cell death (Ott et al., 2007).

5 SPECIIFC AIMS

Work in this thesis aims to gain a deeper, molecular understanding of mitochondrial transcription and DNA replication in human cells. In the three papers of this thesis, we have investigated different aspects of these processes, with a special emphasis on processes of relevance for mitochondrial dysfunction in human health. The specific aims for each study are listed below:

In Paper I, we aim to investigate how ribonucleotides incorporated in mtDNA affect mitochondrial transcription. The levels of embedded ribonucleotides can change due to mutations affecting nucleotide pools and we analyze if disturbed mitochondrial transcription can be a possible downstream consequence of such pathological changes.

In Paper II, we aim to characterize the first known disease-causing mutations in POLRMT. These mutations are associated with neurological disease, and we investigate their effects on mitochondrial transcription *in vitro* and *in vivo*. We also aim to understand how specific mutations affect POLRMT structure and function, and how these changes cause defects in OXPHOS biogenesis and cellular respiration.

In Paper III, we aim to analyze how nucleoside analogues commonly used to treat childhood cancer patients affect mtDNA replication. We aim to elucidate if nucleoside analogues can be incorporated by POLγ and inhibit progression of the mtDNA replication machinery *in vitro* and *in vivo*.

6 RESULTS

6.1 PAPER I

Ribonucleotides embedded in template DNA impair mitochondrial RNA polymerase progression.

Mitochondrial DNA contains high levels of embedded ribonucleotides, which are misincorporated by POLγ during DNA replication. Ribonucleotides are evenly distributed between the two strands of mtDNA, and their relative levels depend on two factors. First, POLγ's ability to discriminate between nucleotides during DNA synthesis differs between the four ribonucleotides. Second, levels of incorporation are dependent on the size of the different nucleotide pools. In support of this notion, mutations that change nucleotide pools and cause human disease also changes the relative levels of ribonucleotides incorporated into mtDNA (Berglund et al., 2017). Adding to the problem, there is no known repair system in mitochondria that can remove embedded ribonucleotides.

Previous work had demonstrated that ribonucleotides embedded in mtDNA can disturb the function of POLγ. Whereas POLγ can efficiently bypass a single ribonucleotide, longer stretches in the template lead to stalling or even termination of DNA synthesis (Forslund et al., 2018; Kasiviswanathan & Copeland, 2011).

MtDNA is also used as a template for transcription and in Paper I, we analyzed how ribonucleotides embedded in template DNA affect mitochondrial transcription. To this end, we used a reconstituted mitochondrial transcription system with purified, recombinant proteins and defined DNA templates. To analyze the effects of ribonucleotides in the template strand, we first used an

RNA-primed DNA template that could support the production of a short transcript. Our experiments revealed that ribonucleotides in the template strand reduce the processivity of POLRMT. Addition of TEFM did not help POLRMT to overcome this effect.

POLRMT is not processive on ssDNA templates and we therefore repeated our experiments using dsDNA templates with various numbers of ribonucleotides embedded. We again observed that ribonucleotides in the template strand caused an increase in premature termination of transcription. The effect was accentuated when longer stretches of ribonucleotides were embedded in the template strand. Embedded ribonucleotides in the non-template strand did not affect transcription. On double-stranded templates, TEFM could help in removing some shorter, pre-maturely terminated products, but the elongation factor could not completely rescue the effect. We also investigated if embedded ribonucleotides could cause temporary transcriptional pausing. To this end, we performed an *in vitro* transcription pulse chase experiment. The data strongly suggested that the shorter transcription products observed in our experiments were due to true transcription termination and not merely pausing events.

In Paper 1, we also attempted to find direct interactions between POLRMT and embedded ribonucleotides, which could explain the effects we observed. Even if structural analysis suggested that POLRMT could form direct hydrogen bonds with the H at the 2′-position of deoxyribose in the template strand, we could not observe any effects on transcription elongation when the corresponding amino acids were mutated.

Based on our work, we suggested that the transcription termination caused by embedded ribonucleotides in the template strand is due to structural changes in the template, and not to specific interactions between embedded

ribonucleotides and POLRMT. Our data suggests that disease-causing processes that change the levels of embedded ribonucleotides in mtDNA can affect not only DNA replication, but also mitochondrial transcription.

6.2 PAPER II

POLRMT mutations impair mitochondrial transcription causing neurological diseases.

A number of genetic diseases with different clinical characteristics and patient prognoses have been linked to mtDNA (Gorman et al., 2016). Multi-organ disorders can be due to sporadic or inherited pathogenic variants in mtDNA or the genes required for mtDNA maintenance (Bonawitz et al., 2006; Wanrooij et al., 2008). Mutations affecting mitochondrial transcription have been more difficult to identify. In Paper II, we identify the first disease-causing mutations in the *POLRMT* gene. We use DNA sequencing and identify both recessive and dominant *POLRMT* variants, associated with a broad spectrum of neurological symptoms.

We also purified POLRMT and disease-causing mutant derivatives to analyze the effects on transcription *in vitro*. POLRMT mutations identified in affected subjects displayed negative effects on transcription. In patient-derived fibroblasts we observed lower levels of mitochondrial transcripts but observed no effects on mtDNA levels or mutation rate. In patient cells, we also observed lower levels of some OXPHOS subunits and reduced respiration.

Our study identified and characterized the first known pathogenic variants in *POLRMT* and connected them to mitochondrial disease. From our data, we concluded that *POLRMT* is a disease associated gene that should be included while examining patients that show a global delay in development, speech impairment, intellectual disabilities, decreased muscle tone, short stature, progessive external opthalmoplegia, and OXPHOS defects.

6.3 PAPER III

5-Fluorouracil and 6-Thioguanine impair mitochondrial DNA replication.

Previous cellular work had demonstrated that nucleoside analogues used to treat cancer can cause a reduction in mtDNA levels and mitochondrial dysfunction. We wanted to understand the molecular basis for these effects and selected to study two compounds which have previously been associated with mitochondrial dysfunction. The nucleoside analogues selected were 5-Flourouracil (5-FU) and 6-Thioguanine (6-TG), which are commonly used to treat childhood cancer patients.

In cells, 5-FU is converted into 5-fluorodeoxyuridine monophosphate (5-FdUMP) which inhibits thymidylate synthase (TYMS) (Christensen et al., 2019; Sommer & Santi, 1974). As a consequence, the conversion into dUMP from dTMP is inhibited, which leads to imbalances in nucleotide pools and disturbed DNA synthesis. 5-FU can also incorporate directly into genomic DNA and affect DNA replication (Huehls et al., 2016; Pettersen et al., 2011). 6-TG (also known as thioguanine) is used to treat acute and chronic myeloid leukemia. Systemic treatment with 6-TG leads to an accumulation of this purine analogue in DNA. 6-TG is vulnerable to oxidation, and the oxidized forms efficiently block transcription and DNA replication (Karran & Attard, 2008).

To address the ability of POLγ to incorporate 5-FU and 6-TG, we studied the effects of these compounds in a single-nucleotide incorporation assay. For this purpose, we used a template with a 21-bp duplex region and a 15-nt stretch of ssDNA. The 3′-end of the 21-mer acted as a starting point for DNA synthesis and the first nucleotide incorporated (position 22) was either dTTP or dGTP, ideal for investigating the effects of 5-FU and 6-TG, respectively. We found

that POLγ could incorporate the triphosphorylated forms 5-FU and 6-TG with an efficiency that was similar to that seen with the natural nucleotides, dTTP and dGTP.

We also studied how nucleoside analogues affect the ability of POLγ to synthesize longer stretches of DNA, using a circular single-stranded DNA template. Addition of 5-FU had no obvious effect on the rate of the reaction and could support the formation of long DNA products even in the complete absence of dTTP. In contrast, 6-TG inhibited POLγ activity, and when dGTP was omitted, the reaction was completely inhibited. We also investigated the effects of 6-TG and 5-FU by monitoring DNA synthesis on a dsDNA template, which required the entire core mitochondrial replication machinery (POLγ, TWINKLE, and mtSSB). In the presence of dTTP, the replication machinery tolerated lower levels of 5-FU, but when 5-FU was present in molar excess, DNA synthesis was significantly reduced. In contrast, 6-TG had a strong inhibitory effect already at very low concentrations.

To assess the effects *in vivo*, we used two-dimensional (2D) agarose gel electrophoresis and monitored the effects of 5-FU and 6-TG on mtDNA replication patterns. For both compounds, we observed clear effects on mtDNA synthesis in the form of replication stalling phenotypes, though the exact pattern differed between the two compounds. In parallel, we also monitored the effects on mtDNA and 7S DNA levels and observed a clear reduction.

Our findings support the idea that treatment with 5-FU or 6-TG can cause mtDNA depletion and reduced expression of mtDNA-encoded proteins required for proper OXPHOS system biogenesis and ATP production. The stalling effects on mtDNA synthesis observed *in vivo* may also stimulate

formation of mtDNA deletions, which can accumulate over time even after treatment with these analogues has been terminated.

7 CONCLUSION

Since its discovery more than 50 years ago, the importance of mtDNA in human health and disease has been increasingly appreciated. Significant progress has been made towards understanding how mtDNA is replicated and expressed, as well as its central importance in OXPHOS system biogenesis. Many different disease-causing mutations have been found both in mtDNA and in nuclear genes needed for mitochondrial function. In spite of this impressive progress, many questions remain unanswered, and the mitochondrial research field is expanding continuously. The main objective of this thesis was to characterize pathogenic mechanism underlying mitochondrial dysfunction. Specifically, we analyzed how disease-causing mutations can impair mitochondrial transcription and how nucleoside analogues used to treat diseases can cause adverse mitochondrial side effects.

For years, we have known that mutations in genes encoding the mtDNA replication machinery can lead to human disease. In fact, mutations affecting POLγ are the most common nuclear cause of mitochondrial disease, creating various phenotypes that can manifest from childhood to late adulthood. Some studies suggest that up to 2% of the population carry mutations in the gene encoding POLγA (*POLG*) and in recent years, these mutations have been linked to more common conditions, such as Parkinson's disease. Disease-causing mutations affecting POLRMT were identified only recently. In the current thesis, we have collaborated with clinical scientists and performed a detailed biochemical analysis of POLRMT variants that cause disease in affected patients. We showed that decreased POLRMT activity lead to lower levels of mitochondrial transcripts and impaired OXPHOS system biogenesis, which in turn caused defects in respiration and leads to various phenotypes. Most of the mutations identified are recessive and compound heterozygous,

meaning that the mutations vary between the two alleles in the same patient. This mixture of different mutations makes identification a bit more challenging and also explains the need for detailed biochemical analysis. By analyzing the defects of individual mutant variants, we can verify that they actually impair the activity of the mutant POLRMT, both in isolation and in combination with other mutant variants.

Interestingly, PORMT activity can also be affected indirectly, via changes in mtDNA. Previous studies have demonstrated that mtDNA contains high levels of incorporated ribonucleotides and that these levels can be increased in diseases affecting mitochondrial nucleotide pools. Previous work had demonstrated that POLγ activity can be impaired by high levels of embedded ribonucleotides in the mtDNA and in the current thesis we expanded on these previous observations and demonstrated that POLRMT activity can also be affected. Our work suggests that impaired mitochondrial transcription can contribute to disease phenotypes, even if the clinical relevance remains to be determined. So far, our analysis has been limited to *in vitro* characterization and further work in human cells and mammalian model systems are therefore needed. Mitochondrial disease can in some cases be iatrogenic, i.e., caused unintentionally by a medical treatment. It is extremely important to identify and understand such effects, since it allows us to develop new and better therapies without unwanted side effects. For instance, anti-retroviral therapy that is used to successfully treat HIV infection can also cause a clonal expansion of mtDNA mutations, which leads to phenotypes of premature ageing later in life. Nucleoside analogues are also used to treat various types of cancer, but how these compounds given in childhood affect patients later in life is not known. In the present thesis, we investigated the effects of two important nucleoside analogues, 5-FU and 6-TG, and demonstrated that they can impair mtDNA replication *in vitro* and in cells. Interestingly, the two

compounds caused different stalling phenotypes on mtDNA replication in cells, suggesting that they affect distinct steps in the replication process. Further work is needed to address the relevance of these findings for patients treated with these commonly used drugs.

ACKNOWLEDGEMENTS

Finally, the cumulation of a journey that started with a single step and gradually developed into one mighty task. My happiness would not be complete without making mention of everyone who offered help and support to complete my journey.

Thank you doesn't mean sufficient, but it is said with appreciation and respect to you for your support and encouragement.

Firstly, I would like to gratefully and sincerely thank my supervisor Prof. **Claes Gustafsson**. I cannot express my gratitude in words for him. You have been a great supervisor, mentor, and a role model for me. You have inspired me the most during the time I have spent under your guidance and helped me to become a better researcher. I would like to thank you for believing in me and providing all the support to reach a successful end to my studies. Whenever I had a hard time to take the right decision you were always there with your best support and guidance. It's hard for me to find words to describe a great supervisor like you and I could not find enough words of gratitude for your efforts to help me during the past few years. I feel it like an honor and truly a pleasure to be under your supervision. I would like to thank you again for your invaluable patience and feedback.

I would also like to thank my co-supervisor, **Eric Lekholm Larsson**. I have been always meeting to you at the coffee machine and we kept discussing the research I was involved in. It was very kind of you always sending me back with a cup of coffee together with some nice suggestions to put into my research.

My next thanks and gratitude which come directly from my heart is for my second supervisor for my other project, **Maria Falkenberg**. It was a privilege to work under your guidance. Thank you so much for providing all the support and guidance to finish the project and helping me to reach a successful end. Your door was always open for me to provide all kinds of support and guidance. You have always motivated me to become a successful researcher as well as a successful mother. Thank you very much **Mia** for always being there, and for understanding and helping me to pass the toughest situations.

Jay (Jennifer Uhler), Bertil, Viktor and Emily you all deserve a special thanks for providing all your support, motivation and guidance during my hard times of troubleshooting the experiments, I always received very good and intelligent suggestions**.** Thanks to all of you**.**

Next, I would like to thank my former colleagues **Bradley Peter, Anna-Karin, Thomas Nicholls, Yazh Muthukumar, Benedict Tan, Hector Diaz, Layal Ishak,** for always providing a cooperative and supportive environment. Also, thanks for sharing the knowledge and contributing to the projects I was involved in.

Örjan Persson, I was sharing the lab with you, and you have always helped me massively with a lot of suggestions and ideas to make the experiments successful. You have always been an inspiration to keep trying and never losing. Thank you so much Örjan for all your support.

Next are **Maru**, **Majda, and Louise** thanks to all of you for being such nice colleagues and very good friends. I really had a good time with you all in the lab and outside during our get-togethers for after-work dinner. Thanks for all the support you have provided, both professionally and personally. You all are close to my heart.

Anup Mishra, my present office mate. Thanks Anup, for always encouraging me to do good science and it was nice to share the office with you. Your suggestions were always motivating and thanks for answering my random questions.

Direnis Erdnic, you have been a very nice colleague and a good friend as well. There are a lot of unforgettable moments and sharing the stories with you. It was really a nice time spent with you during the past few years and now we are defending together. Hope we keep meeting after we finish our studies. All the best for your thesis.

Carlos Pardo Hernandez, you have always been a very nice and a polite lab mate. It was nice sharing the lab with you. I remember our chats during some incubation times and sharing stories. I will never forget your dog prank under my office table. I still get goosebumps. I wish you all the best for your future endeavors.

Mãdãlina Sãcultanu, the youngest PhD student of our lab and a very good friend of mine. I feel lucky to meet you before I defended. It has been really nice spending time with you in and outside the lab. The best stories ever. I hope to keep meeting you. All the best for your PhD, do well.

I would also like to thank **Xuefeng Zhu, Xie Xie, Yonghong Shi and Laleh Arabanian** for sharing their knowledge and suggestions.

A big thanks to my former supervisor **Kui Liu** for all his support and guidance. My previous lab mates Martina Devoti, Bilal and Jessie, I will never forget you guys. You are my good friends and we have spent so good time together.

Last but not the least my family – my parents, **Neelam Singh, and Suresh Singh**. Dad, you will be the happiest person when I get this doctoral degree as

it was your dream, now it is coming to fruition. Thank you for taxing yourself and motivating me to be a scientist. Mom, you have always been very supportive throughout my life whether it is personal or an experiment. You have been so patient and keen on listening to my silly experiment stories without having any scientific interests. You have always motivated and taught me to keep trying in life. You are my best teacher, my advisor and my friend since I was born. I have no suitable word that can fully describe your everlasting and unconditional love, that was the greatest strength to me. I remember many sleepless nights with you accompanying me whenever I needed you. You are my role model, thanks for being in my life.

Next is my husband **Avadhesh Kumar Singh**, you have been the candle of my darkest moments of life. You are a loving husband and with that also a mentor. You have stood by me through all my travails, my absences, my fits of pique and impatience. You have supported our family a lot during my doctoral studies, my late-night experiments by taking a good care of Advik. Thanks for being at my side. You have always been teaching me about the ups and downs whether it is my life, my experiments, or my studies. Thank you for making the life easier for me and being the important person in my life. Thanks for being there and all the love.

It's the turn of my little flourishing bouquet my son **Advik**. I remember you pulling my hand to come and play with you when I was busy in thesis writing. You are my motivation to become a successful person in life. I hope I could be a good mother and provide you everything what all you need or deserve in your beautiful life. Mumma loves you so much my little honey.

Next is my brother **Gaurav,** my sister-in-law **Nayan** and my niece **Shivanya,** a big thanks to all of you for being part of my life. I am lucky to have you all

in my life. Thanks for supporting me with all your love. You all mean a lot to me.

I cannot forget my school friends VS2UMARK who are still in contact Vaibhav, Sumit, Sakshi, Urvashi, Abhishek, Richa and Kanika. Thanks for being the part of my life.

Thanks Rakesh and Tulasi for all your love and support during the time we have spent together in Sweden. Thanks, Myra, for accompanying Advik and always missing him. He loves you the most. Hope to see you both growing up together.

Thank you, almighty God, for giving me this life and the loved ones who are so caring. Thanks for giving me the special things in my life which I never thought of. I will be always grateful.

REFERENCES

The figures are created with the scientific illustration software biorender.com. The illustration on the cover page is adapted from Explode/Shutterstock.com.

Agaronyan, K., Morozov, Y. I., Anikin, M., & Temiakov, D. (2015). Mitochondrial biology. Replication-transcription switch in human mitochondria. *Science*, *347*(6221), 548-551.

Akman, G., Desai, R., Bailey, L. J., Yasukawa, T., Dalla Rosa, I., Durigon, R., . . . Holt, I. J. (2016). Pathological ribonuclease H1 causes R-loop depletion and aberrant DNA segregation in mitochondria. *Proc Natl Acad Sci U S A*, *113*(30), E4276-4285.

Alberts, B. (2008). *Molecular Biology of Cell. New York; Garland Science.*

Alberts, B. (2015). *Molecular biology of the cell. 6th Edition, New York; Garland Science, Taylor and Francis Group.*

Arnberg, A., van Bruggen, E. F., & Borst, P. (1971). The presence of DNA molecules with a displacement loop in standard mitochondrial DNA preparations. *Biochim Biophys Acta*, *246*(2), 353-357.

Asin-Cayuela, J., Schwend, T., Farge, G., & Gustafsson, C. M. (2005). The human mitochondrial transcription termination factor (mTERF) is fully active in vitro in the non-phosphorylated form. *J Biol Chem*, *280*(27), 25499-25505.

Bebenek, A., & Ziuzia-Graczyk, I. (2018). Fidelity of DNA replication-a matter of proofreading. *Curr Genet*, *64*(5), 985-996.

Bender, A., Krishnan, K. J., Morris, C. M., Taylor, G. A., Reeve, A. K., Perry, R. H., . . . Turnbull, D. M. (2006). High levels of mitochondrial DNA

deletions in substantia nigra neurons in aging and Parkinson disease. *Nat Genet*, *38*(5), 515-517.

Berg, J. M., Tymoczko, J. L., & Stryer, L. (2012). Biochemistry. New York, W.H. Freeman.

Berglund, A. K., Navarrete, C., Engqvist, M. K., Hoberg, E., Szilagyi, Z., Taylor, R. W., . . . Clausen, A. R. (2017). Nucleotide pools dictate the identity and frequency of ribonucleotide incorporation in mitochondrial DNA. *PLoS Genet*, *13*(2), e1006628.

Bogenhagen, D., & Clayton, D. A. (1974). The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid. *J Biol Chem*, *249*(24), 7991-7995.

Bolender, N., Sickmann, A., Wagner, R., Meisinger, C., & Pfanner, N. (2008). Multiple pathways for sorting mitochondrial precursor proteins. *EMBO Rep*, *9*(1), 42-49.

Bonawitz, N. D., Clayton, D. A., & Shadel, G. S. (2006). Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol Cell*, *24*(6), 813-825.

Cerritelli, S. M., & Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS J*, *276*(6), 1494-1505.

Cheetham, G. M., Jeruzalmi, D., & Steitz, T. A. (1999). Structural basis for initiation of transcription from an RNA polymerase-promoter complex. *Nature*, *399*(6731), 80-83.

Cheetham, G. M., & Steitz, T. A. (1999). Structure of a transcribing T7 RNA polymerase initiation complex. *Science*, *286*(5448), 2305-2309.

Cheetham, G. M., & Steitz, T. A. (2000). Insights into transcription: structure and function of single-subunit DNA-dependent RNA polymerases. *Curr Opin Struct Biol*, *10*(1), 117-123.

Chinnery, P. F., Samuels, D. C., Elson, J., & Turnbull, D. M. (2002). Accumulation of mitochondrial DNA mutations in ageing, cancer, and mitochondrial disease: is there a common mechanism? *Lancet*, *360*(9342), 1323-1325.

Christensen, S., Van der Roest, B., Besselink, N., Janssen, R., Boymans, S., Martens, J. W. M., . . . Van Hoeck, A. (2019). 5-Fluorouracil treatment induces characteristic T>G mutations in human cancer. *Nat Commun*, *10*(1), 4571.

Cipolat, S., Rudka, T., Hartmann, D., Costa, V., Serneels, L., Craessaerts, K., . . . De Strooper, B. (2006). Mitochondrial rhomboid PARL regulates cytochrome c release during apoptosis via OPA1-dependent cristae remodeling. *Cell*, *126*(1), 163-175.

Clayton, D. A. (1982). Replication of animal mitochondrial DNA. *Cell*, *28*(4), 693-705.

Copeland, W. C. (2012). Defects in mitochondrial DNA replication and human disease. *Crit Rev Biochem Mol Biol*, *47*(1), 64-74.

Costa, A., Hood, I. V., & Berger, J. M. (2013). Mechanisms for initiating cellular DNA replication. *Annu Rev Biochem*, *82*, 25-54.

Crampton, D. J., Ohi, M., Qimron, U., Walz, T., & Richardson, C. C. (2006). Oligomeric states of bacteriophage T7 gene 4 primase/helicase. *J Mol Biol*, *360*(3), 667-677.

Dairaghi, D. J., Shadel, G. S., & Clayton, D. A. (1995a). Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. *J Mol Biol*, *249*(1), 11-28.

Dairaghi, D. J., Shadel, G. S., & Clayton, D. A. (1995b). Human mitochondrial transcription factor A and promoter spacing integrity are required for transcription initiation. *Biochim Biophys Acta*, *1271*(1), 127-134.

Diffley, J. F., & 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*(17), 7864-7868.

Dolezal, P., Likic, V., Tachezy, J., & Lithgow, T. (2006). Evolution of the molecular machines for protein import into mitochondria. *Science*, *313*(5785), 314-318.

Dunn, J. J., & Studier, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol*, *166*(4), 477-535.

Durniak, K. J., Bailey, S., & Steitz, T. A. (2008). The structure of a transcribing T7 RNA polymerase in transition from initiation to elongation. *Science*, *322*(5901), 553-557.

El-Hattab, A. W., Craigen, W. J., & Scaglia, F. (2017). Mitochondrial DNA maintenance defects. *Biochim Biophys Acta Mol Basis Dis*, *1863*(6), 1539-1555.

Falkenberg, M. (2018). Mitochondrial DNA replication in mammalian cells: overview of the pathway. *Essays Biochem*, *62*(3), 287-296.

Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N. G., & Gustafsson, C. M. (2002). Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat Genet*, *31*(3), 289-294.

Farge, G., Laurens, N., Broekmans, O. D., van den Wildenberg, S. M., Dekker, L. C., Gaspari, M., . . . Wuite, G. J. (2012). Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor A. *Nat Commun*, *3*, 1013.

Fisher, R. P., & 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*(20), 11330-11338.

Fisher, R. P., Lisowsky, T., Parisi, M. A., & Clayton, D. A. (1992). DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J Biol Chem*, *267*(5), 3358-3367.

Fisher, R. P., Topper, J. N., & Clayton, D. A. (1987). Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. *Cell*, *50*(2), 247-258.

Forslund, J. M. E., Pfeiffer, A., Stojkovic, G., Wanrooij, P. H., & Wanrooij, S. (2018). The presence of rNTPs decreases the speed of mitochondrial DNA replication. *PLoS Genet*, *14*(3), e1007315.

Fuste, J. M., Wanrooij, S., Jemt, E., Granycome, C. E., Cluett, T. J., Shi, Y., . . . Falkenberg, M. (2010). Mitochondrial RNA polymerase is needed for

activation of the origin of light-strand DNA replication. *Mol Cell*, *37*(1), 67-78.

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

Gorman, G. S., Chinnery, P. F., DiMauro, S., Hirano, M., Koga, Y., McFarland, R., . . . Turnbull, D. M. (2016). Mitochondrial diseases. *Nat Rev Dis Primers*, *2*, 16080.

Gorman, G. S., Schaefer, A. M., Ng, Y., Gomez, N., Blakely, E. L., Alston, C. L., . . . McFarland, R. (2015). Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Ann Neurol*, *77*(5), 753-759.

Gray, M. W., Burger, G., & Lang, B. F. (1999). Mitochondrial evolution. *Science*, *283*(5407), 1476-1481.

Greaves, L. C., Reeve, A. K., Taylor, R. W., & Turnbull, D. M. (2012). Mitochondrial DNA and disease. *J Pathol*, *226*(2), 274-286.

Gustafson, M. A., Sullivan, E. D., & Copeland, W. C. (2020). Consequences of compromised mitochondrial genome integrity. *DNA Repair (Amst)*, *93*, 102916.

Gustafsson, C. M., Falkenberg, M., & Larsson, N. G. (2016). Maintenance and Expression of Mammalian Mitochondrial DNA. *Annu Rev Biochem*, *85*, 133-160.

Hamdan, S. M., Loparo, J. J., Takahashi, M., Richardson, C. C., & van Oijen, A. M. (2009). Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature*, *457*(7227), 336-339.

Hamdan, S. M., & Richardson, C. C. (2009). Motors, switches, and contacts in the replisome. *Annu Rev Biochem*, *78*, 205-243.

Harman, D. (1956). Aging: a theory based on free radical and radiation chemistry. *J Gerontol*, *11*(3), 298-300.

Hillen, H. S., Morozov, Y. I., Sarfallah, A., Temiakov, D., & Cramer, P. (2017). Structural Basis of Mitochondrial Transcription Initiation. *Cell*, *171*(5), 1072-1081 e1010.

Hillen, H. S., Parshin, A. V., Agaronyan, K., Morozov, Y. I., Graber, J. J., Chernev, A., . . . Temiakov, D. (2017). Mechanism of Transcription Anti-termination in Human Mitochondria. *Cell*, *171*(5), 1082-1093 e1013.

Hillen, H. S., Temiakov, D., & Cramer, P. (2018). Structural basis of mitochondrial transcription. *Nat Struct Mol Biol*, *25*(9), 754-765.

Hiona, A., & Leeuwenburgh, C. (2008). The role of mitochondrial DNA mutations in aging and sarcopenia: implications for the mitochondrial vicious cycle theory of aging. *Exp Gerontol*, *43*(1), 24-33.

Hoekstra, J. G., Hipp, M. J., Montine, T. J., & Kennedy, S. R. (2016). Mitochondrial DNA mutations increase in early stage Alzheimer disease and are inconsistent with oxidative damage. *Ann Neurol*, *80*(2), 301-306.

Holmes, J. B., Akman, G., Wood, S. R., Sakhuja, K., Cerritelli, S. M., Moss, C., . . . Holt, I. J. (2015). Primer retention owing to the absence of RNase H1

is catastrophic for mitochondrial DNA replication. *Proc Natl Acad Sci U S A*, *112*(30), 9334-9339.

Huehls, A. M., Huntoon, C. J., Joshi, P. M., Baehr, C. A., Wagner, J. M., Wang, X., . . . Karnitz, L. M. (2016). Genomically Incorporated 5-Fluorouracil that Escapes UNG-Initiated Base Excision Repair Blocks DNA Replication and Activates Homologous Recombination. *Mol Pharmacol*, *89*(1), 53-62.

Hutchison, C. A., 3rd, Newbold, J. E., Potter, S. S., & Edgell, M. H. (1974). Maternal inheritance of mammalian mitochondrial DNA. *Nature*, *251*(5475), 536-538.

Jemt, E., Persson, O., Shi, Y., Mehmedovic, M., Uhler, J. P., Davila Lopez, M., . . . Falkenberg, M. (2015). Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. *Nucleic Acids Res*, *43*(19), 9262-9275.

Jiang, M., Xie, X., Zhu, X., Jiang, S., Milenkovic, D., Misic, J., . . . Falkenberg, M. (2021). The mitochondrial single-stranded DNA binding protein is essential for initiation of mtDNA replication. *Sci Adv*, *7*(27).

Jiang, S., Koolmeister, C., Misic, J., Siira, S., Kuhl, I., Silva Ramos, E., . . . Larsson, N. G. (2019). TEFM regulates both transcription elongation and RNA processing in mitochondria. *EMBO Rep*, *20*(6).

Karran, P., & Attard, N. (2008). Thiopurines in current medical practice: molecular mechanisms and contributions to therapy-related cancer. *Nat Rev Cancer*, *8*(1), 24-36.

Kasiviswanathan, R., & Copeland, W. C. (2011). Ribonucleotide discrimination and reverse transcription by the human mitochondrial DNA polymerase. *J Biol Chem*, *286*(36), 31490-31500.

Kato, M., Ito, T., Wagner, G., Richardson, C. C., & Ellenberger, T. (2003). Modular architecture of the bacteriophage T7 primase couples RNA primer synthesis to DNA synthesis. *Mol Cell*, *11*(5), 1349-1360.

Korhonen, J. A., Gaspari, M., & Falkenberg, M. (2003). TWINKLE Has 5' -> 3' DNA helicase activity and is specifically stimulated by mitochondrial single-stranded DNA-binding protein. *J Biol Chem*, *278*(49), 48627-48632.

Kornblum, C., Nicholls, T. J., Haack, T. B., Scholer, S., Peeva, V., Danhauser, K., . . . Prokisch, H. (2013). Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nat Genet*, *45*(2), 214-219.

Kukat, C., Davies, K. M., Wurm, C. A., Spahr, H., Bonekamp, N. A., Kuhl, I., . . . Larsson, N. G. (2015). Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc Natl Acad Sci U S A*, *112*(36), 11288-11293.

Kukat, C., Wurm, C. A., Spahr, H., Falkenberg, M., Larsson, N. G., & Jakobs, S. (2011). Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc Natl Acad Sci U S A*, *108*(33), 13534-13539.

Lagouge, M., & Larsson, N. G. (2013). The role of mitochondrial DNA mutations and free radicals in disease and ageing. *J Intern Med*, *273*(6), 529-543.

Late Effects of Treatment for Childhood Cancer (PDQ(R)): Health Professional Version. (2002). In *PDQ Cancer Information Summaries*. https://www.ncbi.nlm.nih.gov/pubmed/26389273

Lee, J. B., Hite, R. K., Hamdan, S. M., Xie, X. S., Richardson, C. C., & van Oijen, A. M. (2006). DNA primase acts as a molecular brake in DNA replication. *Nature*, *439*(7076), 621-624.

Lee, Y. S., Lee, S., Demeler, B., Molineux, I. J., Johnson, K. A., & Yin, Y. W. (2010). Each monomer of the dimeric accessory protein for human mitochondrial DNA polymerase has a distinct role in conferring processivity. *J Biol Chem*, *285*(2), 1490-1499.

Li, K., & Williams, R. S. (1997). Tetramerization and single-stranded DNA binding properties of native and mutated forms of murine mitochondrial single-stranded DNA-binding proteins. *J Biol Chem*, *272*(13), 8686-8694.

Litonin, D., Sologub, M., Shi, Y., Savkina, M., Anikin, M., Falkenberg, M., . . . 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*(24), 18129-18133.

Loson, O. C., Song, Z., Chen, H., & Chan, D. C. (2013). Fis1, Mff, MiD49, and MiD51 mediate Drp1 recruitment in mitochondrial fission. *Mol Biol Cell*, *24*(5), 659-667.

Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B., & Kunkel, T. A. (2013). Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol Cell*, *50*(3), 437-443.

Macao, B., Uhler, J. P., Siibak, T., Zhu, X., Shi, Y., Sheng, W., . . . Falkenberg, M. (2015). The exonuclease activity of DNA polymerase gamma is required for ligation during mitochondrial DNA replication. *Nat Commun*, *6*, 7303.

Martin, C. T., Muller, D. K., & Coleman, J. E. (1988). Processivity in early stages of transcription by T7 RNA polymerase. *Biochemistry*, *27*(11), 3966-3974.

Martin, W. F., Garg, S., & Zimorski, V. (2015). Endosymbiotic theories for eukaryote origin. *Philos Trans R Soc Lond B Biol Sci*, *370*(1678), 20140330.

Masters, B. S., Stohl, L. L., & Clayton, D. A. (1987). Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell*, *51*(1), 89-99.

Metodiev, M. D., Lesko, N., Park, C. B., Camara, Y., Shi, Y., Wibom, R., . . . Larsson, N. G. (2009). Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome. *Cell Metab*, *9*(4), 386-397.

Michaels, G. S., Hauswirth, W. W., & Laipis, P. J. (1982). Mitochondrial DNA copy number in bovine oocytes and somatic cells. *Dev Biol*, *94*(1), 246-251.

Miquel, J., Economos, A. C., Fleming, J., & Johnson, J. E., Jr. (1980). Mitochondrial role in cell aging. *Exp Gerontol*, *15*(6), 575-591.

Misic, J., Milenkovic, D., Al-Behadili, A., Xie, X., Jiang, M., Jiang, S., . . . Larsson, N. G. (2022). Mammalian RNase H1 directs RNA primer formation

for mtDNA replication initiation and is also necessary for mtDNA replication completion. *Nucleic Acids Res*, *50*(15), 8749-8766.

Morozov, Y. I., Parshin, A. V., Agaronyan, K., Cheung, A. C., Anikin, M., Cramer, P., & Temiakov, D. (2015). A model for transcription initiation in human mitochondria. *Nucleic Acids Res*, *43*(7), 3726-3735.

Ngo, H. B., Kaiser, J. T., & Chan, D. C. (2011). The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nat Struct Mol Biol*, *18*(11), 1290-1296.

Nicholls, T. J., Nadalutti, C. A., Motori, E., Sommerville, E. W., Gorman, G. S., Basu, S., . . . Gustafsson, C. M. (2018). Topoisomerase 3alpha Is Required for Decatenation and Segregation of Human mtDNA. *Mol Cell*, *69*(1), 9-23 e26.

Nicholls, T. J., Zsurka, G., Peeva, V., Scholer, S., Szczesny, R. J., Cysewski, D., . . . Minczuk, M. (2014). Linear mtDNA fragments and unusual mtDNA rearrangements associated with pathological deficiency of MGME1 exonuclease. *Hum Mol Genet*, *23*(23), 6147-6162.

Oakley, J. L., Strothkamp, R. E., Sarris, A. H., & Coleman, J. E. (1979). T7 RNA polymerase: promoter structure and polymerase binding. *Biochemistry*, *18*(3), 528-537.

Olahova, M., Peter, B., Szilagyi, Z., Diaz-Maldonado, H., Singh, M., Sommerville, E. W., . . . Taylor, R. W. (2021). POLRMT mutations impair mitochondrial transcription causing neurological disease. *Nat Commun*, *12*(1), 1135.

Ott, M., Gogvadze, V., Orrenius, S., & Zhivotovsky, B. (2007). Mitochondria, oxidative stress and cell death. *Apoptosis*, *12*(5), 913-922.

Palade, G. E. (1953). An electron microscope study of the mitochondrial structure. *J Histochem Cytochem*, *1*(4), 188-211.

Palmer, J. M., & Hall, D. O. (1972). The mitochondrial membrane system. *Prog Biophys Mol Biol*, *24*, 125-176.

Payne, B. A., Wilson, I. J., Hateley, C. A., Horvath, R., Santibanez-Koref, M., Samuels, D. C., . . . Chinnery, P. F. (2011). Mitochondrial aging is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations. *Nat Genet*, *43*(8), 806-810.

Pettersen, H. S., Visnes, T., Vagbo, C. B., Svaasand, E. K., Doseth, B., Slupphaug, G., . . . Krokan, H. E. (2011). UNG-initiated base excision repair is the major repair route for 5-fluorouracil in DNA, but 5-fluorouracil cytotoxicity depends mainly on RNA incorporation. *Nucleic Acids Res*, *39*(19), 8430-8444.

Pommier, Y., Sun, Y., Huang, S. N., & Nitiss, J. L. (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat Rev Mol Cell Biol*, *17*(11), 703-721.

Posse, V., Al-Behadili, A., Uhler, J. P., Clausen, A. R., Reyes, A., Zeviani, M., . . . Gustafsson, C. M. (2019). RNase H1 directs origin-specific initiation of DNA replication in human mitochondria. *PLoS Genet*, *15*(1), e1007781.

Posse, V., & Gustafsson, C. M. (2017). Human Mitochondrial Transcription Factor B2 Is Required for Promoter Melting during Initiation of Transcription. *J Biol Chem*, *292*(7), 2637-2645.

Posse, V., Shahzad, S., Falkenberg, M., Hallberg, B. M., & Gustafsson, C. M. (2015). TEFM is a potent stimulator of mitochondrial transcription elongation in vitro. *Nucleic Acids Res*, *43*(5), 2615-2624.

Rahman, S., & Copeland, W. C. (2019). POLG-related disorders and their neurological manifestations. *Nat Rev Neurol*, *15*(1), 40-52.

Reinecke, F., Smeitink, J. A., & van der Westhuizen, F. H. (2009). OXPHOS gene expression and control in mitochondrial disorders. *Biochim Biophys Acta*, *1792*(12), 1113-1121.

Reyes, A., Melchionda, L., Nasca, A., Carrara, F., Lamantea, E., Zanolini, A., . . . Zeviani, M. (2015). RNASEH1 Mutations Impair mtDNA Replication and Cause Adult-Onset Mitochondrial Encephalomyopathy. *Am J Hum Genet*, *97*(1), 186-193.

Ringel, R., Sologub, M., Morozov, Y. I., Litonin, D., Cramer, P., & Temiakov, D. (2011). Structure of human mitochondrial RNA polymerase. *Nature*, *478*(7368), 269-273.

Robberson, D. L., Kasamatsu, H., & Vinograd, J. (1972). Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. *Proc Natl Acad Sci U S A*, *69*(3), 737-741.

Rubio-Cosials, A., Sidow, J. F., Jimenez-Menendez, N., Fernandez-Millan, P., Montoya, J., Jacobs, H. T., . . . Sola, M. (2011). Human mitochondrial transcription factor A induces a U-turn structure in the light strand promoter. *Nat Struct Mol Biol*, *18*(11), 1281-1289.

Scheffler, I. E. (2008). Mitochondria. Hoboken, N.J., Wiley-Liss.

Schinkel, A. H., Koerkamp, M. J., Touw, E. P., & Tabak, H. F. (1987). Specificity factor of yeast mitochondrial RNA polymerase. Purification and interaction with core RNA polymerase. *J Biol Chem*, *262*(26), 12785-12791.

Schwinghammer, K., Cheung, A. C., Morozov, Y. I., Agaronyan, K., Temiakov, D., & Cramer, P. (2013). Structure of human mitochondrial RNA polymerase elongation complex. *Nat Struct Mol Biol*, *20*(11), 1298-1303.

Shi, Y., Dierckx, A., Wanrooij, P. H., Wanrooij, S., Larsson, N. G., Wilhelmsson, L. M., . . . Gustafsson, C. M. (2012). Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proc Natl Acad Sci U S A*, *109*(41), 16510-16515.

Shi, Y., Posse, V., Zhu, X., Hyvarinen, A. K., Jacobs, H. T., Falkenberg, M., & Gustafsson, C. M. (2016). Mitochondrial transcription termination factor 1 directs polar replication fork pausing. *Nucleic Acids Res*, *44*(12), 5732-5742.

Shutt, T. E., & Gray, M. W. (2006a). Bacteriophage origins of mitochondrial replication and transcription proteins. *Trends Genet*, *22*(2), 90-95.

Shutt, T. E., & Gray, M. W. (2006b). Homologs of mitochondrial transcription factor B, sparsely distributed within the eukaryotic radiation, are likely derived from the dimethyladenosine methyltransferase of the mitochondrial endosymbiont. *Mol Biol Evol*, *23*(6), 1169-1179.

Sommer, H., & Santi, D. V. (1974). Purification and amino acid analysis of an active site peptide from thymidylate synthetase containing covalently bound 5-fluoro-2'-deoxyuridylate and methylenetetrahydrofolate. *Biochem Biophys Res Commun*, *57*(3), 689-695.

Sousa, R., Chung, Y. J., Rose, J. P., & Wang, B. C. (1993). Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution. *Nature*, *364*(6438), 593-599.

Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., & Burgers, P. M. (2012). RNase H2-initiated ribonucleotide excision repair. *Mol Cell*, *47*(6), 980-986.

Spelbrink, J. N., Li, F. Y., Tiranti, V., Nikali, K., Yuan, Q. P., Tariq, M., . . . Larsson, C. (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*(3), 223-231.

Steitz, T. A., Smerdon, S. J., Jager, J., & Joyce, C. M. (1994). A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science*, *266*(5193), 2022-2025.

Tan, B. G., Mutti, C. D., Shi, Y., Xie, X., Zhu, X., Silva-Pinheiro, P., . . . Gustafsson, C. M. (2022). The human mitochondrial genome contains a second light strand promoter. *Mol Cell*, *82*(19), 3646-3660 e3649.

Tapper, D. P., & Clayton, D. A. (1981). Mechanism of replication of human mitochondrial DNA. Localization of the 5' ends of nascent daughter strands. *J Biol Chem*, *256*(10), 5109-5115.

Terzioglu, M., Ruzzenente, B., Harmel, J., Mourier, A., Jemt, E., Lopez, M. D., . . . Larsson, N. G. (2013). MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation. *Cell Metab*, *17*(4), 618-626.

Tiranti, V., Savoia, A., Forti, F., D'Apolito, M. F., Centra, M., Rocchi, M., & 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*(4), 615-625.

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., . . . Larsson, N. G. (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature*, *429*(6990), 417-423.

Uhler, J. P., Thorn, C., Nicholls, T. J., Matic, S., Milenkovic, D., Gustafsson, C. M., & Falkenberg, M. (2016). MGME1 processes flaps into ligatable nicks in concert with DNA polymerase gamma during mtDNA replication. *Nucleic Acids Res*, *44*(12), 5861-5871.

van der Bliek, A. M., Shen, Q., & Kawajiri, S. (2013). Mechanisms of mitochondrial fission and fusion. *Cold Spring Harb Perspect Biol*, *5*(6).

Van Haute, L., O'Connor, E., Diaz-Maldonado, H., Munro, B., Polavarapu, K., Hock, D. H., . . . Horvath, R. (2023). TEFM variants impair mitochondrial transcription causing childhood-onset neurological disease. *Nat Commun*, *14*(1), 1009.

Vannini, A., & Cramer, P. (2012). Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol Cell*, *45*(4), 439-446.

Wanrooij, P. H., & Chabes, A. (2019). Ribonucleotides in mitochondrial DNA. *FEBS Lett*, *593*(13), 1554-1565.

Wanrooij, P. H., Engqvist, M. K. M., Forslund, J. M. E., Navarrete, C., Nilsson, A. K., Sedman, J., . . . Chabes, A. (2017). Ribonucleotides incorporated by the yeast mitochondrial DNA polymerase are not repaired. *Proc Natl Acad Sci U S A*, *114*(47), 12466-12471.

Wanrooij, P. H., Uhler, J. P., Shi, Y., Westerlund, F., Falkenberg, M., & Gustafsson, C. M. (2012). A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop. *Nucleic Acids Res*, *40*(20), 10334-10344.

Wanrooij, S., Fuste, J. M., Farge, G., Shi, Y., Gustafsson, C. M., & Falkenberg, M. (2008). Human mitochondrial RNA polymerase primes lagging-strand DNA synthesis in vitro. *Proc Natl Acad Sci U S A*, *105*(32), 11122-11127.

Wolstenholme, D. R. (1992). Animal mitochondrial DNA: structure and evolution. *Int Rev Cytol*, *141*, 173-216.

Xu, B., & Clayton, D. A. (1992). Assignment of a yeast protein necessary for mitochondrial transcription initiation. *Nucleic Acids Res*, *20*(5), 1053-1059.

Yakubovskaya, E., Chen, Z., Carrodeguas, J. A., Kisker, C., & Bogenhagen, D. F. (2006). Functional human mitochondrial DNA polymerase gamma forms a heterotrimer. *J Biol Chem*, *281*(1), 374-382.

Yasukawa, T., Reyes, A., Cluett, T. J., Yang, M. Y., Bowmaker, M., Jacobs, H. T., & Holt, I. J. (2006). Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J*, *25*(22), 5358-5371.

Zhu, C., Johansson, M., & Karlsson, A. (2000). Incorporation of nucleoside analogs into nuclear or mitochondrial DNA is determined by the intracellular phosphorylation site. *J Biol Chem*, *275*(35), 26727-26731.